



FORSCHUNGSDATENZENTRUM
der Bundesagentur für Arbeit im Institut für
Arbeitsmarkt- und Berufsforschung

FDZ-DATENREPORT

Dokumentation zu Arbeitsmarktdaten

05|2019 DE Linked-Employer-Employee-Daten des IAB: LIAB-Längsschnittmodell (LIAB LM) 1975-2017

Lisa Schmidlein, Stefan Seth, Matthias Umkehrer



Bundesagentur für Arbeit

Linked-Employer-Employee-Daten des IAB: LIAB-Längsschnittmodell (LIAB LM) 1975-2017

Lisa Schmidtlein, Stefan Seth, Matthias Umkehrer
(Institut für Arbeitsmarkt- und Berufsforschung (IAB))

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

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

Inhaltsverzeichnis

1	Einleitung und Kurzbeschreibung.....	6
1.1	Einleitung	6
1.2	Datennutzung.....	8
1.2.1	Datenzugang	8
1.2.2	Datenhaltung.....	8
1.2.3	Dateinamen im LIAB LM 7517	11
1.3	Veränderungen zum LIAB LM 9314	11
1.3.1	Beobachtungszeitraum.....	11
1.3.2	Stichprobendesign	11
1.3.3	Korrektur der Fehlbuchungen in der Gemeinde Lahn	12
1.3.4	Merkmalskanon.....	12
1.4	Steckbrief	14
1.5	Merkmalsliste	15
1.6	Mengengerüst	18
2	Datenquellen.....	20
2.1	Beschäftigtenhistorik (BeH).....	20
2.2	Leistungsempfängerhistorik (LeH).....	22
3	Datenaufbereitung und Stichproben-ziehung.....	22
3.1	Bereinigungen und Aufbereitungen	22
3.1.1	Beschäftigtenhistorik (BeH).....	23
3.1.2	Leistungsempfängerhistorik (LeH)	23
3.2	Episodensplitting.....	23
3.3	Stichprobenziehung.....	25
3.4	Fehlende Werte	26
4	Datenqualität und Probleme	26
4.1	Gesamte IEB	26
4.1.1	Lücken in den Erwerbsbiografien.....	26
4.1.2	Einführung des SGB II und spätere Trägerwechsel.....	27
4.2	Beschäftigtenhistorik (BeH).....	28
4.3	Leistungsempfängerhistorik (LeH).....	31
5	Variablenbeschreibung.....	31
5.1	Identifikatoren	31
5.1.1	Systemfreie Personennummer (persnr)	31
5.1.2	Systemfreie Betriebsnummer (betnr).....	31

5.1.3	Betriebsnummer nach Betriebspanel (idnum)	32
5.2	Generierte technische Merkmale	33
5.2.1	Satzzähler pro Konto (spell)	33
5.2.2	Quelle des Satzes (quelle)	33
5.2.3	Jahr (jahr)	33
5.2.4	Status der Betriebsnummer (betr_st)	33
5.3	Gültigkeitszeitraum	34
5.3.1	Beginndatum Originalbeobachtung (begorig)	34
5.3.2	Enddatum Originalbeobachtung (endorig)	34
5.3.3	Beginndatum Episode (begepi)	35
5.3.4	Enddatum Episode (endepe)	35
5.4	Informationen zur Person	35
5.4.1	Geschlecht (frau)	35
5.4.2	Geburtsjahr (gebjahr)	35
5.4.3	Geburtsmonat (gebmon)	36
5.4.4	Staatsangehörigkeit (nation)	36
5.4.5	Staatsangehörigkeit vergrößert (nation_gr)	36
5.4.6	Familienstand (famst)	37
5.4.7	Kinderzahl (kind)	37
5.4.8	Ausbildung (ausbildung)	37
5.4.9	Ausbildung (imputiert) (ausbildung_imp)	38
5.4.10	Schulausbildung (schule)	38
5.5	Informationen zu Beschäftigung und Leistungsbezug	39
5.5.1	Tagesentgelt/ täglicher Leistungssatz (tentgelt)	39
5.5.2	Beruf – ausgeübte/letzte Tätigkeit (KldB 1988) (beruf)	40
5.5.3	Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_3)	41
5.5.4	Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_4)	41
5.5.5	Anforderungsniveau – ausgeübte/letzte Tätigkeit (KldB 2010) (niveau)	42
5.5.6	Teilzeit (teilzeit)	43
5.5.7	Stellung im Beruf und Arbeitszeit (stib)	43
5.5.8	Erwerbsstatus (erwstat)	44
5.5.9	Gleitzone (gleitz)	45
5.5.10	Leiharbeit (leih)	45
5.5.11	Befristung (befrist)	45
5.5.12	Abmeldegrund / Abgabegrund / Beendigungsgrund (grund)	46
5.5.13	Restanspruch/geplante Dauer (restanspruch)	46
5.5.14	Beginndatum der Arbeitslosigkeit (alo_beg)	47
5.5.15	Dauer der Arbeitslosigkeit (alo_dau)	47
5.6	Ortsangaben	48
5.6.1	Wohnort - Kreis (wo_kreis)	48
5.6.2	Wohnort - Bundesland (wo_bula)	49
5.6.3	Wohnort - Arbeitsagentur (wo_aa)	49
5.6.4	Wohnort - Regionaldirektion (wo_rd)	50

5.7	Betriebsmerkmale	50
5.7.1	WZ73 3-Steller (w73_3)	50
5.7.2	WZ93 5-Steller (w93_5)	50
5.7.3	WZ93 3-Steller (w93_3)	51
5.7.4	WZ03 5-Steller (w03_5)	51
5.7.5	WZ03 3-Steller (w03_3)	52
5.7.6	WZ08 5-Steller (w08_5)	52
5.7.7	WZ08 3-Steller (w08_3)	53
5.7.8	w73_3 vervollständigt durch Extrapolation/Imputation (w73_3_gen).....	53
5.7.9	Art der Vervollständigung w73_3 (group_w73_3).....	53
5.7.10	w93_3 vervollständigt durch Extrapolation/Imputation (w93_3_gen).....	54
5.7.11	Art der Vervollständigung w93_3 (group_w93_3).....	54
5.7.12	w08_3 vervollständigt durch Extrapolation/Imputation (w08_3_gen).....	54
5.7.13	Art der Vervollständigung w08_3 (group_w08_3).....	55
5.7.14	Jahr des 1. Auftretens der Betriebsnummer (grd_jahr).....	55
5.7.15	Erstes Auftreten Betriebsnummer (grd_dat).....	55
5.7.16	Jahr des letzten Auftretens der Betriebsnummer (lzt_jahr).....	56
5.7.17	Letztes Auftreten Betriebsnummer (lzt_dat)	56
5.7.18	Anzahl Beschäftigte gesamt (az_ges)	57
5.7.19	Anzahl in Vollzeit (Normalbeschäftigte + sonstige) (az_vz)	57
5.7.20	Anzahl geringfügig Beschäftigte (az_gf)	57
5.7.21	Mittelwert imp. Bruttotagesentgelt Vollzeitbeschäftigte (te_imp_mw).....	57
5.7.22	Arbeitsort - Kreis (ao_kreis)	58
5.7.23	Arbeitsort - Bundesland (ao_bula)	58
6	Literatur	60
7	Anhang	62
7.1	Auszählungen.....	62
7.2	Abkürzungsverzeichnis.....	62

Zusammenfassung

Dieser Datenreport beschreibt das LIAB-Längsschnittmodell 1975 – 2017 (LIAB LM 7517).

Abstract

This data report describes the LIAB Longitudinal Model 1975 – 2017 (LIAB LM 7517).

Keywords

German administrative micro data, labour market data, linked employer-employee data, data manual

Danksagung

Für ihre Mitarbeit und Unterstützung danken wir den Kolleginnen und Kollegen des Forschungszentrums (FDZ), des Bereichs Betriebe und Beschäftigung (D1) sowie des Geschäftsbereichs Daten- und IT-Management (DIM) des Instituts für Arbeitsmarkt- und Berufsforschung. Der vorliegende Datenreport ist eine Überarbeitung des FDZ-Datenreports 10/2016 (Autoren: Heining, Klosterhuber, Lehnert, Seth). Darüber hinaus wurden einzelne Passagen aus IAB-internen Datendokumentationen von IAB-DIM übernommen.

Datenverfügbarkeit

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

1 Einleitung und Kurzbeschreibung

1.1 Einleitung

Das LIAB Längsschnittmodell (LM) 7517 ist einer der Linked-Employer-Employee-Datensätze des Instituts für Arbeitsmarkt- und Berufsforschung (IAB), die über das Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit (BA) für wissenschaftliche Auswertungen zur Verfügung stehen. Es verknüpft Informationen zu Betrieben aus dem IAB-Betriebspanel, einer jährlich stattfindenden Betriebsbefragung, mit Information zu deren Beschäftigten aus den Prozessdaten der BA.

Das IAB-Betriebspanel ist eine jährliche, repräsentative Arbeitgeberbefragung zu betrieblichen Determinanten der Beschäftigung. Die Erhebung wird seit 1993 in Westdeutschland und seit 1996 in Ostdeutschland durchgeführt. Sie stellt die zentrale Quelle für Analysen zur Arbeitskräftenachfrage auf dem deutschen Arbeitsmarkt dar. **Dieser Datenreport konzentriert sich auf die Beschreibung der Personendaten im LIAB LM 7517. Nähere Informationen zum IAB-Betriebspanel finden sich auf den Internetseiten des FDZ (https://fdz.iab.de/de/FDZ_Establishment_Data/IAB_Establishment_Panel.aspx).**

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

- sozialversicherungspflichtige Beschäftigung in Deutschland (erfasst ab 1975),
- geringfügige Beschäftigung in Deutschland (erfasst ab 1999),
- Bezug von Leistungen nach dem Rechtskreis SGB III (erfasst ab 1975) oder SGB II (erfasst ab 2005),
- bei der Bundesagentur für Arbeit (BA) bzw. den Grundsicherungsträgern als arbeitsuchend gemeldet (erfasst ab 1997),
- Teilnahme an arbeitsmarktpolitischer Maßnahme (erfasst ab 2000).

Diese aus unterschiedlichen Datenquellen stammenden Informationen werden in den IEB zusammengeführt und tagesgenau abgebildet.

Während die Beschäftigtenhistorik (BeH) Ursprung der Informationen zu sozialversicherungspflichtiger und geringfügiger Beschäftigung ist, wird der Bezug von Leistungen nach den Rechtskreisen des SGB III und des SGB II in der Leistungsempfängerhistorik (LeH) bzw. in der Leistungshistorik Grundsicherung (LHG) erfasst. Die Arbeitsuchendenhistoriken (ASU und XASU) stellen die Datenquelle für die durch die BA bzw. die durch die kommunalen Grundsicherungsträger erfassten Zeiten der Arbeitsuche dar, während die Teilnahme an arbeitsmarktpolitischen Maßnahmen in den Maßnahmeteilnahmehistoriken (MTH und XMTH) erfasst wird. **Die Personendaten im LIAB LM 7517 umfassen nur Informationen aus den Quellen Beschäftigtenhistorik (BeH) und Leistungsempfängerhistorik (LeH). Die übrigen Quellen sind nicht in den Personendaten des LIAB LM 7517 enthalten.**

Das IAB-Betriebspanel bildet den Ausgangspunkt für die Ziehung der Personendaten des LIAB LM 7517. Abbildung 1 bietet einen Überblick über die einzelnen Schritte (nähere Details finden sich in Kapitel 3). Zuerst werden die Betriebe aus dem IAB-Betriebspanel ausgewählt, die Teil des Panel-

Zeitraums 2009 – 2016 sind (sogenannte *Panelfälle*). Diese Einschränkung ist aufgrund von Beschränkungen bei der Datensatzgröße notwendig. Anschließend werden sämtliche Personen ausgewählt, die laut den Prozessdaten mindestens einen Tag zwischen 2008 und 2017 bei zumindest einem dieser Betriebe beschäftigt waren. Letztlich werden die administrativen Erwerbsbiographien dieser Personen aus den IEB im Zeitraum 1975 – 2017 dem LIAB LM 7517 zugespielt.

Trotz der genannten Einschränkung der Betriebe bei der Ziehung der Personendaten umfasst das LIAB LM 7517 die Befragungsdaten zu allen Betrieben, die im IAB-Betriebspanel zwischen 1993 und 2017 befragt wurden (also auch die Befragungsdaten zu den Betrieben, die keinen *Panelfall* darstellen). Allerdings können lediglich für die gewählten *Panelfälle* alle Beschäftigten eines Betriebes in den Personendaten beobachtet werden. Zudem bilden die Personendaten auch Beschäftigungszeiten bei Betrieben ab, die nicht befragt wurden. Für diese Betriebe kann abermals nicht die Gesamtheit aller Beschäftigten in den Personendaten beobachtet werden. Allerdings können administrative Betriebsinformationen aus dem Betriebs-Historik-Panel (BHP) allen im LIAB LM 7517 enthaltenen Betrieben zugespielt werden (siehe https://fdz.iab.de/de/FDZ_Establishment_Data/Establishment_History_Panel.aspx).

Das LIAB LM 7517 wird im Forschungsdatenzentrum der Bundesagentur für Arbeit im IAB produziert. Der Datensatz umfasst die Erwerbsverläufe von 1.692.611 Individuen, deren Erwerbsbiographien in insgesamt 56.536.054 Datenzeilen dokumentiert sind. Der vorliegende Datenreport beschreibt die einzelnen Merkmale, welche weitgehend auf den Originaldaten, d.h. nicht vergrößerte Daten, basieren. Einige Merkmale sind allerdings als besonders sensibel eingestuft und werden nur auf gesonderten Antrag (siehe Abschnitt 1.2) weitergegeben.

1. Ausgangspunkt: Betriebsbefragung (IAB-Betriebspanel)

IAB-Betriebspanel

2. Auswahl der Betriebe aus dem IAB-Betriebspanel, die Teil des Panel-Zeitraums 2009 – 2016 sind (*Panelfälle*)

Panelfälle

3. Auswahl der Personen, die in den *Panel-Betrieben* zwischen 2008 und 2017 beschäftigt sind

Beschäftigte aus Panel-Betrieben

4. Zuspielen der tagesgenauen administrativen Erwerbsbiographien von 1975 bis 2017 zu den in Schritt 3 gewählten Personen

Erwerbsbiographien

Fiktives Beispiel:

persnr	begepi	endept	Betrieb ein <i>Panelfall</i> ? (1=Ja)	...
1	01.01.1975	31.12.1975	0	
1	01.01.1976	30.06.1976	0	
...				
1	01.01.2017	31.12.2017	1	

Abbildung 1: Aufbau des LIAB LM 7517. Für *Panelfälle* lassen sich in den Personendaten alle Beschäftigten eines Betriebes beobachten, für alle anderen Betriebe sind jeweils nur einzelne Beschäftigte beobachtbar. Jedoch lassen sich für alle im LIAB LM 7517 enthaltenen Betriebe administrative Informationen auf Betriebsebene aus dem Betriebs-Historik-Panel (BHP) zuspielen.

In Abgrenzung zum Querschnittmodell des LIAB umfasst das Längsschnittmodell die gesamten zeitraumbezogenen Erwerbsbiographien von Personen einer Unterstichprobe der im IAB-Betriebspanel wiederholt befragten Betriebe. Für diese Untergruppe lassen sich Erwerbsbiographien somit auch bei Betriebswechseln und Übergängen in Arbeitslosigkeit verfolgen. Im Querschnittmodell hingegen werden die Erwerbsbiographien jeweils immer nur zum Stichtag 30. Juni abgebildet. Allerdings werden zu allen im IAB-Betriebspanel befragten Betrieben die Erwerbsbiographien aller Beschäftigten zugespielt, die im Befragungsjahr zum Stichtag 30. Juni bei einem dieser Betriebe beschäftigt waren. **Eine Übersicht zu den verschiedenen LIAB-Datensätzen findet sich auf den Internetseiten des FDZ** (https://fdz.iab.de/de/Integrated_Establishment_and_Individual_Data/LIAB.aspx).

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

1.2 Datennutzung

1.2.1 Datenzugang

Die LIAB-Daten liegen in schwach anonymisierter Form vor und können somit ausschließlich im Rahmen von Gastaufenthalten am FDZ und anschließender Datenfernverarbeitung ausgewertet werden.

Für die Datennutzung ist ein Antrag beim FDZ einzureichen. Über die Genehmigung des Forschungsprojekts entscheidet das FDZ im Auftrag und ggf. unter Abstimmung mit dem Bundesministerium für Arbeit und Soziales (BMAS). Nach erfolgter Genehmigung wird ein Nutzungsvertrag mit der wissenschaftlichen Einrichtung der Forschenden geschlossen. Details zur Beantragung des Datensatzes und Datenverarbeitungsmöglichkeiten können der FDZ-Website entnommen werden.

1.2.2 Datenhaltung

Die LIAB-Daten, die sowohl deutsch- als auch englischsprachige Labels enthalten,¹ sind modular gegliedert und in mehreren Dateien abgelegt (vgl. Abbildung 2). Ein erstes Modul umfasst alle Wellen des IAB-Betriebspanels, die jeweils als separater Datensatz abgelegt sind. In den Wellen finden sich jahresspezifische Informationen zu Beschäftigungsentwicklung, Geschäftspolitik und Geschäftsentwicklung, betriebliche Investitionen, Innovationen im Betrieb, öffentliche Förderung, Personalstruktur, Berufsausbildung und Ausbildungsstellen, Personaleinstellungen und -abgänge, Personalsuche, Löhne und Gehälter, betriebliche Arbeitszeiten, Fort- und Weiterbildung oder auch allgemeine Angaben zum Betrieb. Das LIAB LM 7517 enthält die gesamten Befragungsdaten des IAB-Betriebspanels, unabhängig davon, ob ein Betrieb für die Ziehung der Personendaten ausgewählt wurde oder nicht.

¹ Mit den Stata-Befehlen `label language en` oder `label language de` kann auf englisch- bzw. deutschsprachige Labels umgestellt werden.

Ein zweites Modul, das im weiteren Verlauf Personendatei genannt wird, enthält Identifikatoren (systemfreie Personen- und Betriebsnummer), persönliche Merkmale, Informationen zu Beschäftigung und Leistungsbezug, Merkmale zum Wohnort sowie technische Merkmale.

Ein drittes Modul, die Basis-Betriebsdatei, enthält die Betriebsnummer, das Jahr und Merkmale zu Arbeitsort und Wirtschaftszweigen sowie weitere Betriebsinformationen zum Stichtag 30. Juni. Für Ostdeutschland liegen diese administrativen Betriebsdaten erst ab 1992 vor. Die Basis-Betriebsdatei stammt aus dem Betriebs-Historik-Panel (BHP). Das BHP umfasst alle Betriebe in Deutschland, die zum 30. Juni eines Jahres mindestens einen sozialversicherungspflichtigen oder seit 1999 auch geringfügig Beschäftigten aufweisen. Quelle des BHP ist die Beschäftigten-Historik (BeH) des IAB. Im BHP werden die Personendaten der BeH mit Hilfe der Betriebsnummern auf die Betriebs-Jahrebene aggregiert. Nähere Informationen zum BHP finden sich auf den Internetseiten des FDZ (https://fdz.iab.de/de/FDZ_Establishment_Data/Establishment_History_Panel.aspx). Die in der Merkmalsliste (vgl. Tabelle 4) mit einem „*“ gekennzeichneten Merkmale sind in der Basis-Betriebsdatei enthalten. Auf begründeten Antrag können aus dem BHP weitere **Betriebsmerkmale** in Variablenblöcken bereitgestellt werden (vgl. Abbildung 2). Zudem können die Erweiterungsmodule Beschäftigungsströme (Eintritte/Austritte) und Betriebsdynamik (Gründungen/Schließungen) mit gesonderter Begründung beantragt werden.

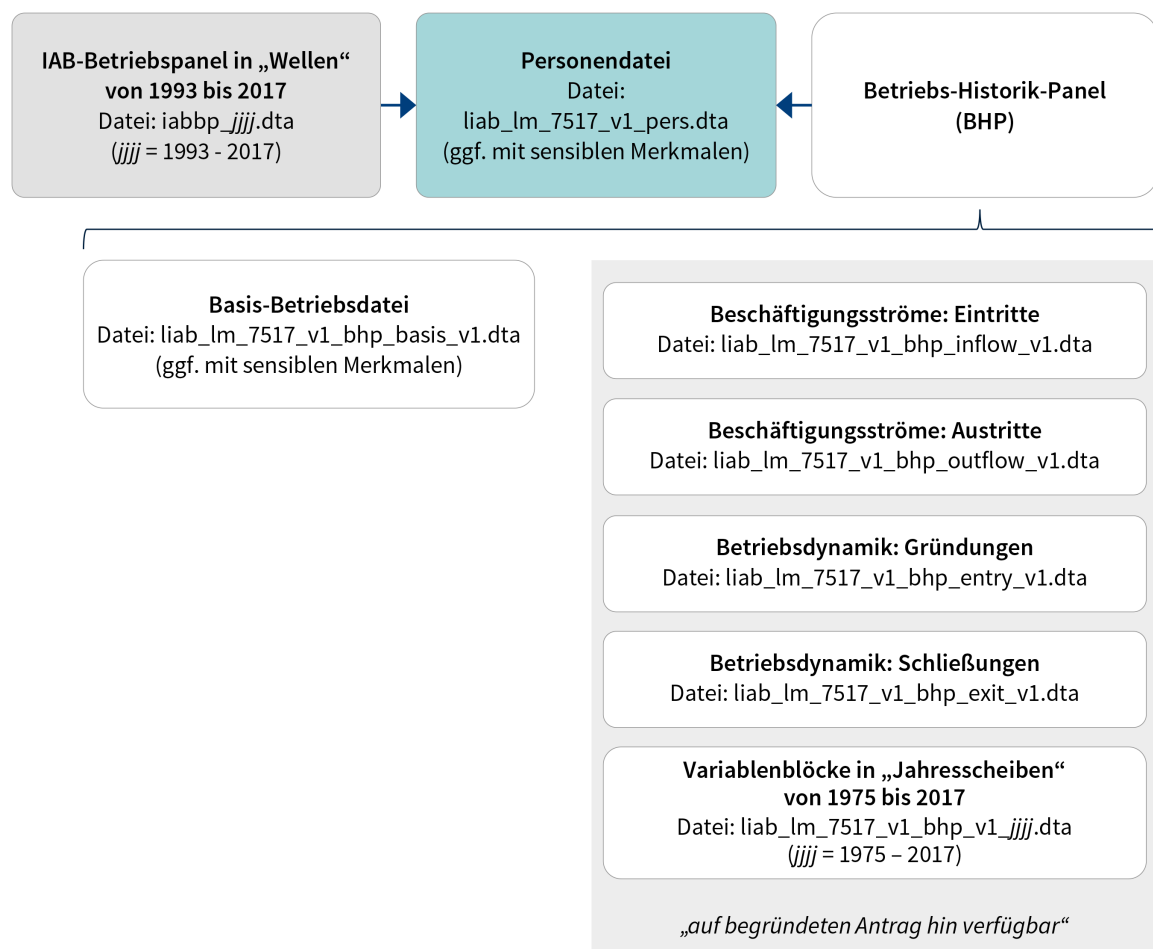


Abbildung 2: Datenhaltung der Linked-Employer-Employee-Daten im LIAB LM 7517

Betriebs- und Personendaten sind organisatorisch voneinander getrennt, was die Struktur der Daten deutlich macht und Speicherplatz spart. Die Verknüpfung der einzelnen Dateien erfolgt über die programmspezifischen Kommandos der zur Datenaufbereitung und -analyse genutzten Softwarepakete. So können Verknüpfungen beispielsweise in Stata unter Verwendung des „merge“-Befehls durchgeführt werden. Die Verknüpfung der Wellen des IAB-Betriebspanels und der Personendatei erfolgt über die in beiden Modulen enthaltene Betriebsnummer des IAB-Betriebspanels (*idnum*; siehe Kasten 1). Basis-Betriebsdatei und Personendatei werden dagegen über die systemfreie Betriebsnummer (*betnr*) und das Jahr (*jahr*) verknüpft (siehe Kasten 2). Gleiches gilt für die Erweiterungsmodule aus dem BHP.

```
use liab_lm_7517_v1_pers.dta
sort idnum
merge m:1 idnum using iabbp_2000.dta
```

Kasten 1: Beispielcode Stata 14; Verknüpfung der Personendatei mit einer Welle des IAB-Betriebspanels

```
use liab_lm_7517_v1_pers.dta
sort betnr jahr
merge m:1 betnr jahr using liab_lm_7517_v1_bhp_basis_v1.dta
```

Kasten 2: Beispielcode Stata 14; Verknüpfung der Personendatei mit der Basis-Betriebsdatei

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:

Personendatei:

- Staatsangehörigkeit (*nation*)
- Geburtsmonat (*gebmon*)
- Berufsuntergruppe (*beruf2010_4*)
- Wohnort - Arbeitsagentur (*wo_aa*)
- Wohnort - Kreis (*wo_kreis*)

Basis-Betriebsdatei:

- Arbeitsort - Kreis (*ao_kreis*)
- Wirtschaftszweig – Wirtschaftsunterklasse (5-Steller) (*w93_5*)
- Wirtschaftszweig – Wirtschaftsunterklasse (5-Steller) (*w03_5*)

- Wirtschaftszweig - Wirtschaftsunterklasse (5-Steller) (*w08_5*)
- Erstes Auftreten Betriebsnummer (*grd_dat*)
- Letztes Auftreten Betriebsnummer (*lzt_dat*)

Sofern sie beantragt und genehmigt wurden, befinden sich die hier aufgeführten sensiblen Merkmale in den entsprechenden Dateien (siehe Abbildung 2).

1.2.3 Dateinamen im LIAB LM 7517

Wellen des IAB-Betriebspanels

iabbp_####.dta, *####* = 1993 – 2017

Personendatei

liab_lm_7517_v1_pers.dta

Basis-Betriebsdatei

liab_lm_7517_v1_bhp_basis_v1.dta

Erweiterungsmodule BHP

Bestandsdaten mit Variablenblöcken

liab_lm_7517_v1_bhp_v1_####.dta, *####* = 1975 – 2017

Beschäftigtenströme

liab_lm_7517_v1_bhp_inflow_v1.dta

liab_lm_7517_v1_bhp_outflow_v1.dta

Betriebsdynamik

liab_lm_7517_v1_bhp_entry_v1.dta

liab_lm_7517_v1_bhp_exit_v1.dta

1.3 Veränderungen zum LIAB LM 9314

1.3.1 Beobachtungszeitraum

Das LIAB LM 7517 unterscheidet sich vom LIAB LM 9314 im abgebildeten Zeitraum für die Personen- und Betriebsdaten.

1.3.2 Stichprobendesign

Für die Auswahl der *Panelfälle* aus dem IAB-Betriebspanel wurde ein aktuellerer Panel-Zeitraum verwendet. Während im LIAB LM 9314 der Panel-Zeitraum 2003 – 2011 als Grundlage der Ziehung diente, wird im LIAB LM 7517 der Panel-Zeitraum 2009 – 2016 verwendet. Außerdem verschiebt

sich der Zeitraum, in dem die Erwerbsbiographien der ausgewählten Personen abgebildet werden, von 1993 – 2014 auf 1975 – 2017 (vgl. Abschnitt 3.3).

1.3.3 Korrektur der Fehlbuchungen in der Gemeinde Lahn

Für die Jahre 1975 – 1977 gab es in der Quelle BeH eine historische Fehlzurordnung der Betriebe und Beschäftigten der Gemeinde Lahn (Niedersachsen). Diese Fehlzurordnungen wurden soweit wie möglich korrigiert. Weitere Informationen finden sich in Abschnitt 4.2.

1.3.4 Merkmalskanon

Das LIAB LM 7517 bietet gegenüber dem LIAB LM 9314 eine Reihe von zusätzlichen Merkmalen. Das in der vorherigen Version entfernte Merkmal „Stellung im Beruf und Arbeitszeit“ (*stib*) ist nun wieder enthalten. Es gilt jedoch die Einschränkung, dass das Merkmal nur für Meldungen gefüllt ist, die vor der Einführung des neuen Tätigkeitsschlüssels abgegeben wurden (siehe Abschnitt 4.2). Neu aufgenommen wurde zudem das Merkmal „Ausbildung (imputiert)“ (*ausbildung_imp*, siehe Abschnitt 5.4.9). Diese Variable enthält für Beobachtungen mit ursprünglich fehlender Angabe zum beruflichen Ausbildungsniveau imputierte Angaben. Außerdem enthält das LIAB LM 7517 nun das sensible Merkmal „Geburtsmonat“ (*gebmon*, siehe Abschnitt 5.4.3) als separate Variable. Zudem wurden zeitkonsistente Wirtschaftszweige der Klassifikation WZ08 durch Extrapolation und Imputation generiert. Weitere Informationen dazu finden sich in den Abschnitten 5.7.12 und 5.7.13. Mit den Variablen „Beginndatum der Arbeitslosigkeit“ (*alo_beg*) und „Dauer der Arbeitslosigkeit“ (*alo_dau*) sind zudem zwei Variablen neu im LIAB LM 7517 enthalten, die Informationen zur individuellen Arbeitslosigkeit liefern (siehe Abschnitte 5.5.14 und 5.5.15). Diese Variablen werden bereits in den IEB erstellt, da das LIAB LM 7517 nicht alle hierzu benötigten Quellen beinhaltet.

Die individuellen Erwerbsbiographien sind im LIAB LM 7517 ab einschließlich 1975 enthalten. Das Vorgängermodell hat diese Biographien vor 1993 abgeschnitten. Um den Teil der Biographien vor 1993 dennoch zumindest teilweise abbilden zu können, wurden folgende biographische Merkmale für das Vorgängermodell generiert:

- Eintritt Erwerbsleben (*ein_erw*)
- Anzahl der Tage im Erwerbsleben (*tage_erw*)
- Eintritt Betrieb (*ein_bet*)
- Anzahl der Tage im Betrieb (*tage_bet*)
- Eintritt Job (*ein_job*)
- Anzahl der Tage im Job (*tage_job*)
- Anzahl der Tage mit Leistungsbezug (*tage_ist*)

Diese Merkmale können nun von den Forschenden selbst gebildet werden, weshalb sie nicht länger im LIAB 7517 als Variablen enthalten sind. Entsprechende Programme für Stata finden sich in Eberle und Schmucker (2019).

Tabelle 1 gibt einen Überblick über die Änderungen im Merkmalskanon über alle Quellen hinweg.

Tabelle 1: Änderungen des Merkmalskanons

Merkmal	Art der Änderung	Erläuterung
Ausbildung imputiert (<i>ausbildung_imp</i>)	N	Neu aufgenommen, jedoch mit weniger Ausprägungen als die Variable „Ausbildung“ (<i>ausbildung</i>)
Geburtsmonat (<i>gebmon</i>)	N	Sensibles Merkmal
Stellung im Beruf und Arbeitszeit (<i>stib</i>)	N	Wieder aufgenommen, jedoch nur für Meldungen gefüllt, die vor der Einführung des neuen Tätigkeitsschlüssels abgegeben wurden
w08_3 vervollständigt durch Extrapolation/Imputation (<i>w08_3_gen</i>)	N	Neu aufgenommen in Basis-Betriebsdatei
Art der Vervollständigung w08_3 (<i>group_w08_3</i>)	N	Neu aufgenommen in Basis-Betriebsdatei
Beginndatum der Arbeitslosigkeit (<i>alo_beg</i>)	N	Das Merkmal wird in den IEB unter Einbezug mehrerer Quellen generiert und kann somit nicht exakt anhand der Personendaten des LIAB nachgebildet werden.
Dauer der Arbeitslosigkeit (<i>alo_dau</i>)	N	Das Merkmal wird in den IEB unter Einbezug mehrerer Quellen generiert und kann somit nicht exakt anhand der Personendaten des LIAB nachgebildet werden.
Geburtsdatum (<i>gebdat</i>)	E	Anstelle der kombinierten Variable <i>gebdat</i> werden die Angaben zu Geburtsmonat und Geburtsjahr nun in zwei Variablen ausgewiesen (<i>gebmon</i> , <i>gebjahr</i>).
Eintritt Erwerbsleben (<i>ein_erw</i>)	E	Kann anhand der Erwerbsbiographien ab 1975 nun selbst gebildet werden
Anzahl der Tage im Erwerbsleben (<i>tage_erw</i>)	E	Kann anhand der Erwerbsbiographien ab 1975 nun selbst gebildet werden
Eintritt Betrieb (<i>ein_bet</i>)	E	Kann anhand der Erwerbsbiographien ab 1975 nun selbst gebildet werden
Anzahl der Tage im Betrieb (<i>tage_bet</i>)	E	Kann anhand der Erwerbsbiographien ab 1975 nun selbst gebildet werden
Eintritt Job (<i>ein_job</i>)	E	Kann anhand der Erwerbsbiographien ab 1975 nun selbst gebildet werden
Anzahl der Tage im Job (<i>tage_job</i>)	E	Kann anhand der Erwerbsbiographien ab 1975 nun selbst gebildet werden
Anzahl der Tage mit Leistungsbezug (<i>tage_lst</i>)	E	Kann anhand der Erwerbsbiographien ab 1975 nun selbst gebildet werden

N = neu, E = entfällt

1.4 Steckbrief

Tabelle 2: Steckbrief

Themen/ Merkmalsgruppen	Beschäftigtenhistorik (BeH): Jahres- und Abmeldungen zur Sozialversicherung für sozialversicherungs-pflichtig und geringfügig Beschäftigte Leistungsempfängerhistorik (LeH): Informationen zu Leistungsbezug nach SGB III für Bezieher von Arbeitslosen-geld, Arbeitslosenhilfe und Unterhaltsgeld IAB-Betriebspanel: Informationen aus jährlicher Betriebsbefragung mit wechselnden Schwer-punktt Themen
Untersuchungs-einheit	Betriebe, sozialversicherungspflichtige Beschäftigte (ab 1999 auch geringfügig Beschäftigte), Leistungsempfänger
Fallzahlen	Verknüpfte Betriebe: 2 187 bis 12 261 pro Jahr, 41 777 insgesamt Personen: 135 985 bis 1 399 924 pro Jahr, 1 692 611 insgesamt Überschneidungsfreie Sätze (nach Episodensplitting): 56 586 054
Zeitraum	Der Zeitraum hängt von der Datenquelle ab. BeH 01.01.1975 – 31.12.2017 (2014: 36-Monats-Datei 2015 und 2016: 18-Monats-Datei, 2017: 6-Monats-Datei) LeH 01.01.1975 – 31.12.2017
Zeitbezug	Betriebe: Stichtag 30. Juni Personen: Tagesgenau
Regionale Gliederung	Bundesländer, Kreise
Gebietsstand	Gebietsstandbereinigt zum 31.12.2017
Erhebungsdesign	Linked-Employer-Employee-Daten; Auswahl bestimmter Betriebe aus Betriebsbefragung und Zuspätschicken der Erwerbsbiographien ihrer Beschäftigten aus Prozessdaten.
Beteiligte Institutionen	Sozialversicherungsträger, Bundesagentur für Arbeit, Kantar Public (ehemals TNS Infratest Sozialforschung GmbH) in Zusammenarbeit mit dem Institut für Sozialökonomische Strukturanalysen GmbH (SÖSTRA).
Frequenz der Datensammlung	Administrative Daten: laufend Daten des IAB-Betriebspanels: einmal jährlich
Dateiformat/-größe	Stata Betriebsdaten (ohne sensible Merkmale): IAB-Betriebspanel: ca. 1,7 bis 11 MB pro Jahr Basis-Betriebsdatei: ca. 455 MB Personendatei (ohne sensible Merkmale): ca. 3,5 GB

Dateiorganisation	Die Personendaten werden in einer Datei abgelegt. Hinzu kommt eine Datei mit den betriebsbezogenen Informationen aus dem BHP, die über <i>betnr</i> und <i>jahr</i> an die Personendaten gespielt werden kann. Auf begründeten Antrag werden weitere Dateien mit zusätzlichen Betriebsinformationen aus dem BHP bereitgestellt. Die Dateien des IAB-Betriebspanels werden unverändert als eine Datei pro Welle übernommen und können über die <i>idnum</i> an die Personendaten gespielt werden.
Datenzugang	Gastaufenthalt und anschließende Datenfernverarbeitung
Anonymisierungsgrad	Schwach anonymisiert
Sensible Merkmale	Personendatei: Staatsangehörigkeit (<i>nation</i>), Geburtsmonat (<i>gebmon</i>), Berufsuntergruppe (<i>beruf2010_4</i>), Wohnort – Kreis (<i>wo_kreis</i>), Wohnort – Arbeitsagentur (<i>wo_aa</i>) Basis-Betriebsdatei: Arbeitsort – Kreis (<i>ao_kreis</i>), Wirtschaftsunterklasse (<i>w93_5</i>), Wirtschaftsunterklasse (<i>w03_5</i>), Wirtschaftsunterklasse (<i>w08_5</i>), Erstes Auftreten der Betriebsnummer (<i>grd_dat</i>), Letztes Auftreten der Betriebsnummer (<i>lzt_dat</i>)
Zitierung der Daten und Datendokumentation	Datensatz: Linked-Employer-Employee-Daten des IAB, Längsschnittmodell 1975 – 2017 (LIAB LM 7517), Nürnberg 2019 Daten: „Die Datengrundlage dieses Beitrags bildet das Längsschnittmodell 1975 – 2017 der Linked-Employer-Employee-Daten des IAB. Der Datenzugang erfolgte über einen Gastaufenthalt am Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (FDZ) und mittels kontrollierter Datenfernverarbeitung beim FDZ.“ DOI: 10.5164/IAB.LIABLM7517.de.en.v1 Datendokumentation: Schmidtlein, Lisa; Seth, Stefan; Umkehrer, Matthias (2019): Linked-Employer-Employee-Daten des IAB: LIAB Längsschnittmodell (LIAB LM) 1975 – 2017. FDZ-Datenreport, 05/2019 (de), Nürnberg. DOI: 10.5164/IAB.FDZD.1905.de.v1
Datensatzversion	Linked-Employer-Employee-Daten des IAB, Längsschnittmodell 1975 – 2017 (LIAB LM 7517) – Version 7517 v1; DOI: 10.5164/IAB.LIABLM7517.de.en.v1

1.5 Merkmalsliste

Die Merkmalsübersicht in Tabelle 4 listet die Variablennamen und die längeren Merkmalsbezeichnungen auf. Weiterhin bietet sie eine Übersicht darüber, ob und wie gut Merkmale in den einzelnen Datenquellen gefüllt sind. Tabelle 3 macht die Bedeutung der Schattierungen deutlich, die in Tabelle 4 den Füllgrad pro Merkmal und Quelle angeben.

Beispiel: Das Merkmal „Beruf – ausgeübte/letzte Tätigkeit“ (*beruf*) ist nur für BeH-Sätze gefüllt. Die Sätze aus der Datenquelle LeH enthalten in dieser Variablen den Missing-Wert „n“.

Tabelle 3: Füllgrad der Merkmale

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

Eine weitere Besonderheit liegt darin, dass einige Merkmale je nach Datenquelle unterschiedliche Inhalte haben. So enthält der „Erwerbsstatus“ (*erwstat*) für BeH-Sätze die Personengruppe des Meldeverfahrens und für LeH-Sätze die Leistungsart. Nicht bei jedem Merkmal sind diese Unterschiede direkt aus dem Variablennamen ersichtlich.

Tabelle 4: Merkmalsliste mit Füllgrad

Merkmalsliste	In Basis-Betriebsdatei	Seite	BeH	LeH
Identifikatoren		31		
Systemfreie Personennummer (<i>persnr</i>)		31	h	h
Systemfreie Betriebsnummer (<i>betnr</i>)	*	31	h	l
Betriebsnummer nach Betriebspanel (<i>idnum</i>)		32	m	l
Generierte technische Merkmale		33		
Satzzähler pro Konto (<i>spell</i>)		33	h	h
Quelle des Satzes (<i>quelle</i>)		33	h	h
Jahr (<i>jahr</i>)	*	33	h	h
Status der Betriebsnummer (<i>betr_st</i>)		33	h	
Gültigkeitszeitraum		34		
Beginndatum Originalbeobachtung (<i>begorig</i>)		34	h	h
Enddatum Originalbeobachtung (<i>endorig</i>)		34	h	h
Beginndatum Episode (<i>begepi</i>)		35	h	h
Enddatum Episode (<i>endeipi</i>)		35	h	h
Informationen zur Person		35		
Geschlecht (<i>frau</i>)		35	h	h
Geburtsjahr (<i>gebjahr</i>)		35	h	h
Geburtsmonat (<i>gebmon</i>)		36	h	h
Staatsangehörigkeit (<i>nation</i>)		36	h	h
Staatsangehörigkeit vergrößert (<i>nation_gr</i>)		36	h	h
Familienstand (<i>famst</i>)		37	l	h
Kinderzahl (<i>kind</i>)		37	l	h

Ausbildung (<i>ausbildung</i>)		37	m	l
Ausbildung (imputiert) (<i>ausbildung_imp</i>)		38	h	l
Schulausbildung (<i>schule</i>)		38	m	l
Informationen zu Beschäftigung und Leistungsbezug		39		
Tagesentgelt/ täglicher Leistungssatz (<i>tentgelt</i>)		39	h	h
Beruf – ausgeübte/letzte Tätigkeit (KldB 1988) (<i>beruf</i>)		40	h	l
Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (<i>beruf2010_3</i>)		41	h	m
Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (<i>beruf2010_4</i>)		41	h	m
Anforderungsniveau – ausgeübte/letzte Tätigkeit (KldB 2010) (<i>niveau</i>)		42	h	m
Teilzeit (<i>teilzeit</i>)		43	h	l
Stellung im Beruf und Arbeitszeit (<i>stib</i>)		43	h	l
Erwerbsstatus (<i>erwstat</i>)		44	h	h
Gleitzone (<i>gleitz</i>)		45	h	l
Leiharbeit (<i>leih</i>)		45	h	l
Befristung (<i>befrist</i>)		45	h	l
Abmeldegrund / Abgabegrund / Beendigungsgrund (<i>grund</i>)		46	h	h
Restanspruch/geplante Dauer (<i>restanspruch</i>)		46	l	h
Beginndatum der Arbeitslosigkeit (<i>alo_beg</i>)		47	m	m
Dauer der Arbeitslosigkeit (<i>alo_dau</i>)		47	h	h
Ortsangaben		48		
Wohnort - Kreis (<i>wo_kreis</i>)		48	h	h
Wohnort - Bundesland (<i>wo_bula</i>)		49	h	h
Wohnort - Arbeitsagentur (<i>wo_aa</i>)		49	h	h
Wohnort - Regionaldirektion (<i>wo_rd</i>)		50	h	h
Betriebsmerkmale		50		
WZ73 3-Steller (<i>w73_3</i>)	*	50	h	l
WZ93 5-Steller (<i>w93_5</i>)	*	50	h	l
WZ93 3-Steller (<i>w93_3</i>)	*	51	h	l
WZ03 5-Steller (<i>w03_5</i>)	*	51	h	l
WZ03 3-Steller (<i>w03_3</i>)	*	52	h	l
WZ08 5-Steller (<i>w08_5</i>)	*	52	h	l
WZ08 3-Steller (<i>w08_3</i>)	*	53	h	l
w73_3 vervollständigt durch Extrapolation/Imputation (<i>w73_3_gen</i>)	*	53	h	l
Art der Vervollständigung w73_3 (<i>group_w73_3</i>)	*	53	h	l
w93_3 vervollständigt durch Extrapolation/Imputation (<i>w93_3_gen</i>)	*	54	h	l

Art der Vervollständigung w93_3 (<i>group_w93_3</i>)	*	54	h	l
w08_3 vervollständigt durch Extrapolation/Imputation (<i>w08_3_gen</i>)	*	54	h	l
Art der Vervollständigung w08_3 (<i>group_w08_3</i>)	*	55	h	l
Jahr des 1. Auftretens der Betriebsnummer (<i>grd_jahr</i>)	*	55	h	l
Erstes Auftreten Betriebsnummer (<i>grd_dat</i>)	*	55	h	l
Jahr des letzten Auftretens der Betriebsnummer (<i>lzt_jahr</i>)	*	56	h	l
Letztes Auftreten Betriebsnummer (<i>lzt_dat</i>)	*	56	h	l
Anzahl Beschäftigte gesamt (<i>az_ges</i>)	*	57	h	l
Anzahl in Vollzeit (Normalbeschäftigte + sonstige) (<i>az_vz</i>)	*	57	h	l
Anzahl geringfügig Beschäftigte (<i>az_gf</i>)	*	57	h	l
Mittelwert imp. Bruttotagesentgelt Vollzeitbeschäftigte (<i>te_imp_mw</i>)	*	57	h	l
Arbeitsort - Kreis (<i>ao_kreis</i>)	*	58	h	l
Arbeitsort - Bundesland (<i>ao_bula</i>)	*	58	h	l

1.6 Mengengerüst

In Tabelle 5 gibt Spalte (2) die Anzahl der Personen pro Jahr im LIAB LM 7517 an. In Spalte (3) werden die in den Personendaten enthaltenen Betriebe pro Jahr abgebildet.

Im Gegensatz dazu enthalten die Spalten (4) bis (7) die Zahl an verknüpften Betrieben im Längsschnittmodell. Spalte (4) gibt die Gesamtzahl an Betrieben aus dem IAB-Betriebspanel an, die mit den Personendaten verknüpft werden können. Dagegen unterscheiden die Spalten (5) bis (7) die Anzahl der Verknüpfungen anhand des Merkmals „Status der Betriebsnummer“ (*betr_st*, siehe Abschnitt 5.2.4).

In Spalte (5) befinden sich alle originär für das Längsschnittmodell ausgewählten Betriebe. Diese sind als sogenannte *Panelfälle* für den Zeitraum 2009 – 2016 aus dem IAB-Betriebspanel gezogen worden (vgl. Abschnitt 3.3). Die dargestellten Zahlen in Spalte (5) beziehen sich dabei auf Verknüpfungen von Betrieben, bei denen

- zum Stichtag der Befragung (30.06. des jeweiligen Jahres) ein gültiges Interview vorliegt und
- dieser Stichtag vom Beginn- und Enddatum der Episode eingeschlossen wird.

Bei der Anzahl in Spalte (6) handelt es sich dagegen um Verknüpfungen, bei denen

- zum Stichtag der Befragung (30.06. des jeweiligen Jahres) ein gültiges Interview vorliegt,
- dieser Stichtag jedoch nicht vom Beginn- und Enddatum der Episode eingeschlossen wird.

Spalte (7) beinhaltet schließlich die Zahl an Betrieben aus dem IAB-Betriebspanel, für die zwar eine Verknüpfung mit den Personendaten hergestellt werden kann, die jedoch keine *Panelfälle* für den Zeitraum 2009 bis 2016 darstellen.

Tabelle 5: Mengengerüst

Jahr	Personen	Betriebe	Betriebe im IAB-Betriebspanel mit Verknüpfung zu den Personendaten			
	Insgesamt	Insgesamt	Insgesamt	Betrieb für LIAB-Modell ausgewählt		Betrieb nicht für LIAB-Modell ausgewählt
				<i>betr_st</i> == 1	<i>betr_st</i> == 2	<i>betr_st</i> == 3
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1975	135 985	70 735				
1976	150 451	78 959				
1977	167 835	84 347				
1978	182 942	89 467				
1979	203 935	97 216				
1980	222 369	102 529				
1981	239 335	104 380				
1982	254 273	104 940				
1983	268 892	108 173				
1984	288 445	113 315				
1985	305 704	115 323				
1986	329 271	119 919				
1987	348 979	123 477				
1988	366 464	126 622				
1989	390 956	132 856				
1990	424 651	143 858				
1991	571 840	184 054				
1992	671 553	212 978				
1993	686 299	219 461	2 432	172	126	2 258
1994	702 476	226 405	2 293	185	123	2 104
1995	728 138	230 554	2 187	204	143	1 979
1996	749 684	232 710	4 958	698	466	4 250
1997	777 219	236 379	4 643	790	534	3 843
1998	816 908	241 145	5 100	970	662	4 114
1999	879 748	306 509	5 592	1 254	899	4 322
2000	953 230	321 842	7 795	1 740	1 265	6 016
2001	988 326	319 758	8 680	2 052	1 525	6 588
2002	1 015 304	310 463	8 412	2 368	1 855	6 012
2003	1 047 107	315 020	8 535	2 986	2 530	5 508
2004	1 084 647	323 266	8 863	3 321	2 652	5 499
2005	1 085 709	324 397	9 062	3 710	2 939	5 310
2006	1 127 706	334 092	9 076	4 213	3 431	4 826

2007	1 190 587	342 093	9 354	4 806	4 012	4 509
2008	1 257 950	348 111	9 789	5 494	4 701	4 250
2009	1 275 709	358 149	10 204	6 176	5 359	3 938
2010	1 299 344	370 501	10 269	6 460	5 630	3 628
2011	1 339 255	382 583	10 388	6 562	5 701	3 646
2012	1 357 859	386 260	10 911	6 750	5 730	3 966
2013	1 374 278	388 351	11 180	7 049	6 048	3 920
2014	1 389 022	392 233	11 348	7 483	6 363	3 646
2015	1 396 878	392 862	11 823	7 964	6 693	3 622
2016	1 399 924	393 460	12 261	8 672	7 345	3 303
2017	1 393 830	391 475	11 136	7 304	6 152	3 578

Zu beachten: Der Wert in Spalte (4) ergibt sich nicht aus der Summe der Spalten (5), (6) und (7).

2 Datenquellen

Die administrativen Personendaten werden aus den Integrierten Erwerbsbiografien (IEB) des IAB gezogen. Diese vereinen Daten aus verschiedenen Datenquellen, die jeweils Informationen aus unterschiedlichen EDV-Verfahren enthalten können. Zusätzlich werden einige ergänzende Merkmale aus diesen Datenquellen in die administrativen Personendaten aufgenommen, die nicht Teil der IEB sind. Abbildung 3 veranschaulicht die Datenflüsse, die zum LIAB LM 7517 führen und stellt deren Beziehung zu anderen Datenprodukten des FDZ dar.

In den Personendaten des LIAB LM 7517 werden nur die Datensätze aus den dargestellten Quellen der Beschäftigtenhistorik (BeH) und Leistungsempfängerhistorik (LeH) aus der IEB übernommen.

2.1 Beschäftigtenhistorik (BeH)

Die Quelle für Informationen über Beschäftigung ist die Beschäftigtenhistorik (BeH) des IAB. Die Grundlage der Daten ist das mit Wirkung vom 01.01.1973 eingeführte (und seit dem 01.01.1991 auf Ostdeutschland ausgedehnte) integrierte Meldeverfahren zur Kranken-, Renten- und Arbeitslosenversicherung (DEÜV; früher DEVO/DÜVO; vgl. für weitere Details: Bender et al., 1996, S. 4ff.; Wermter/Cramer, 1988). Es verlangt von den Arbeitgebern mindestens einmal jährlich Meldungen für alle sozialversicherungspflichtig Beschäftigten an die zuständigen Sozialversicherungsträger abzugeben.

Die BeH erfasst alle Arbeiter und Angestellten sowie alle Auszubildenden, soweit sie nicht von der Sozialversicherungspflicht befreit sind. Dies bedeutet, dass verbeamtete und selbstständige Tätigkeiten sowie Studierende² (vgl. Cramer, 1985) grundsätzlich nicht in der BeH erfasst werden. Seit der Änderung des Meldeverfahrens zum 01.01.1999 werden auch geringfügig Beschäftigte und mithelfende Familienangehörige erfasst (enthalten erst ab 01.04.1999). Die Daten werden von den Krankenkassen aufgenommen, von der Bundesagentur für Arbeit gesammelt, aufbereitet und anschließend vom IAB in dessen Historikdatei integriert.

² Studierende können dennoch in der BeH erscheinen, wenn sie parallel zum Studium z.B. geringfügig beschäftigt waren.

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

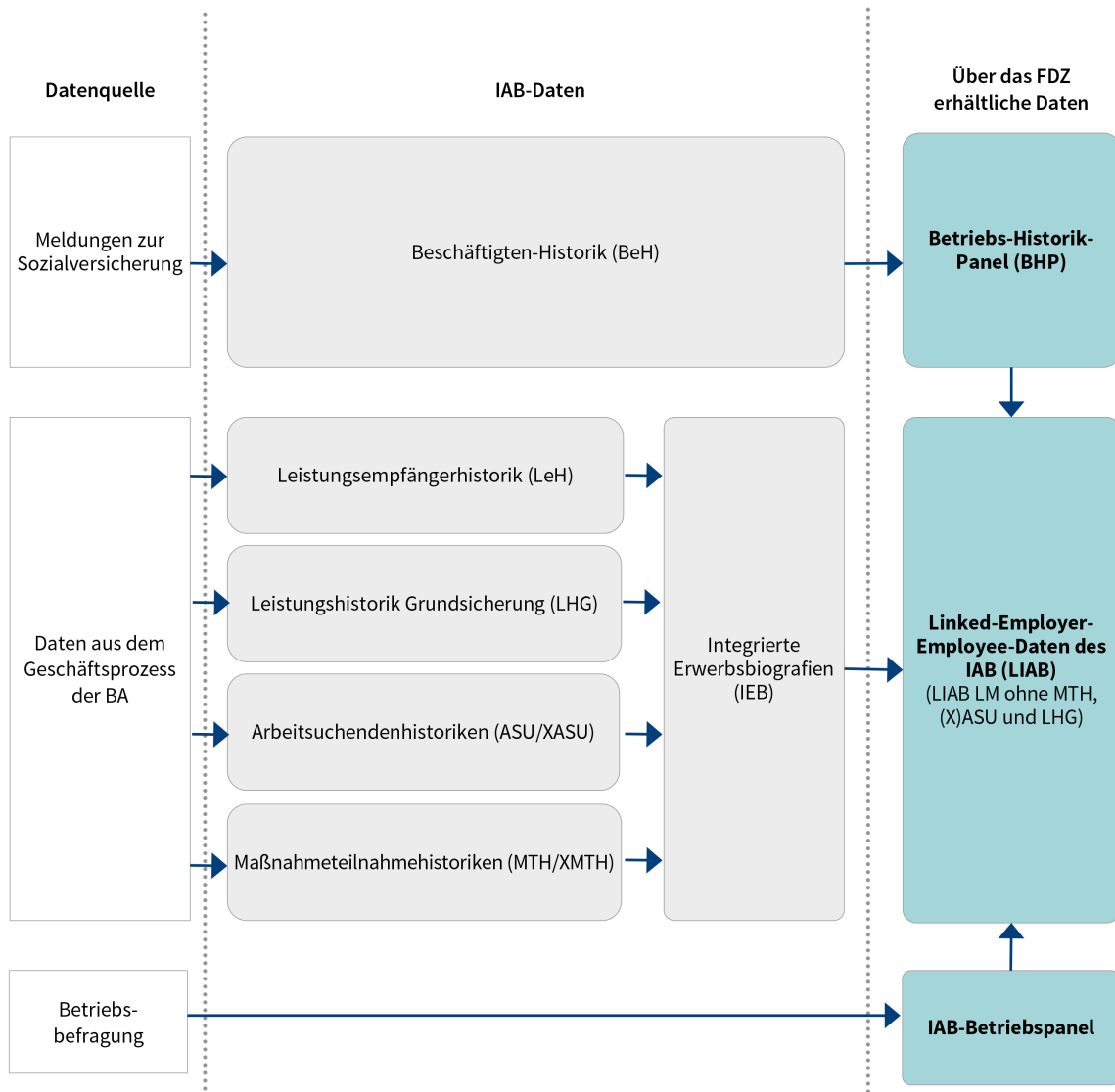


Abbildung 3: Quellen des LIAB

Bei der Verknüpfung von Personen- mit Betriebsdaten aus dem BHP muss berücksichtigt werden, dass die Merkmale in der Basis-Betriebsdatei sowie in den BHP-Erweiterungsmodulen zum 30.6. eines Jahres aggregiert werden. Die Betriebsmerkmale sind deshalb, im Unterschied zu den Personenangaben, keine zeitraumbezogenen Daten, sondern nur am 30.6. jeden Jahres exakt gültig.³

³ 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.

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 (vgl. Kasten 2).

2.2 Leistungsempfängerhistorik (LeH)

Die Leistungsempfängerhistorik (LeH) erfasst Zeiträume, in denen Personen Lohnersatzleistungen der Bundesagentur für Arbeit beziehen (Rechtskreis SGB III). Die Leistungen beinhalten Arbeitslosengeld, Arbeitslosenhilfe, Unterhaltsgeld und von der Bundesagentur für Arbeit übernommene Beiträge zur privaten Kranken- bzw. Pflegeversicherung während des Leistungsbezugs. Leistungen nach dem Rechtskreis SGB II (z.B. Arbeitslosengeld II) sind nicht enthalten. Da der Leistungsanspruch von der Erfüllung gesetzlicher Voraussetzungen abhängt, werden Arbeitslosigkeitsperioden, in denen die Voraussetzungen nicht vorliegen (z.B. fehlende Bedürftigkeit im Falle der Arbeitslosenhilfe oder Nichterfüllung der Anwartschaftszeit beim Arbeitslosengeld) nicht in der Leistungsempfängerhistorik erfasst. Der linke Datenrand der LeH ist der 01.01.1975.

3 Datenaufbereitung und Stichprobenziehung

3.1 Bereinigungen und Aufbereitungen

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

- Sätze, bei denen das Alter an ihrem Ende kleiner 13 oder zu Beginn größer 75 ist, werden gelöscht.
- Datensätze, bei denen das Enddatum kleiner als das Beginndatum ist, werden gelöscht.
- Widersprüchliche Angaben zu Geschlecht oder Geburtsdatum innerhalb eines Kontos werden bereinigt.
- Datensätze ohne Angabe von Geburtsdatum und Geschlecht nach Bereinigung werden gelöscht.

Weitergehende Bereinigungen, wie z.B. das Ergänzen von vermutlich fehlenden Meldungen, Streikbereinigungen, etc., erfolgen nicht.

3.1.1 Beschäftigtenhistorik (BeH)

- Um einen im Zeitablauf möglichst konstanten Personenkreis zu erfassen, werden einige Personengruppen ausgeschlossen, deren Daten nicht für den gesamten Beobachtungszeitraum zur Verfügung stehen. Ab dem Meldejahr 2011 stammen die Daten der BeH aus neu konzipierten Quelldaten. Dadurch wurde eine Reihe von Personengruppen reaktiviert bzw. neu eingeführt, da sie von der BA-Statistik als sozialversicherungspflichtig eingestuft werden. Damit sind fortan die Personengruppen 101 - 107, 111 - 114, 118, 119, 120, 140, 141, 142, 143, 149, 201 und 203 - 205 enthalten sowie die beiden Gruppen 109 und 209, die geringfügig Beschäftigte kennzeichnen. Nicht enthalten sind unter anderem kurzfristig Beschäftigte, d. h. Personengruppen 110, 202 und 210.
- Die Personengruppen 123, 124 und 127 wurden ab 2011 neu eingeführt.
- Aus datenschutzrechtlichen Gründen werden die Personengruppen 107, 111, 113, 114, 127 und 204 zur Personengruppe "Sonstige Erwerbstätige" (599) zusammengefasst.
- Ab dem Meldejahr 2012 wurden Auszubildende mit den Personengruppen 121 und 122 neu aufgenommen.
- Sätze mit fehlendem Entgelt sowie der Ausprägung 101 beim Merkmal Erwerbsstatus und dem Wert 50 beim Abgabegrund (Jahresmeldung) werden nicht in die IEB übernommen.
- Geschlecht und Geburtsdatum werden aus dem Data Warehouse (DWH) der Statistik der BA übernommen. Dort liegen diese Angaben quellenübergreifend harmonisiert vor.
- Der Gebietsstand für Arbeitsort und Wohnort entspricht dem 31.12.2017.

3.1.2 Leistungsempfängerhistorik (LeH)

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

3.2 Episodensplitting

Die administrativen Personendaten liegen „gesplittet“ vor. Falls sich innerhalb eines Kontos Sätze überlappen, werden diese durch künstliche Sätze mit neuen Datumsangaben derart ersetzt, dass vollständig parallele oder überschneidungsfreie Zeiträume entstehen (Episodensplitting). Die Anzahl der Sätze erhöht sich dadurch (siehe Abbildung 4).

Die ursprünglichen Datumsvariablen zum 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 wiederherzustellen bzw. die durch das Episodensplitting künstlich erzeugten Datensätze zu löschen, selektiert man alle Sätze, bei denen der Beginn des Originalsatzes gleich dem Beginn der gesplitteten Episode ist (*begepi* == *begorig*).

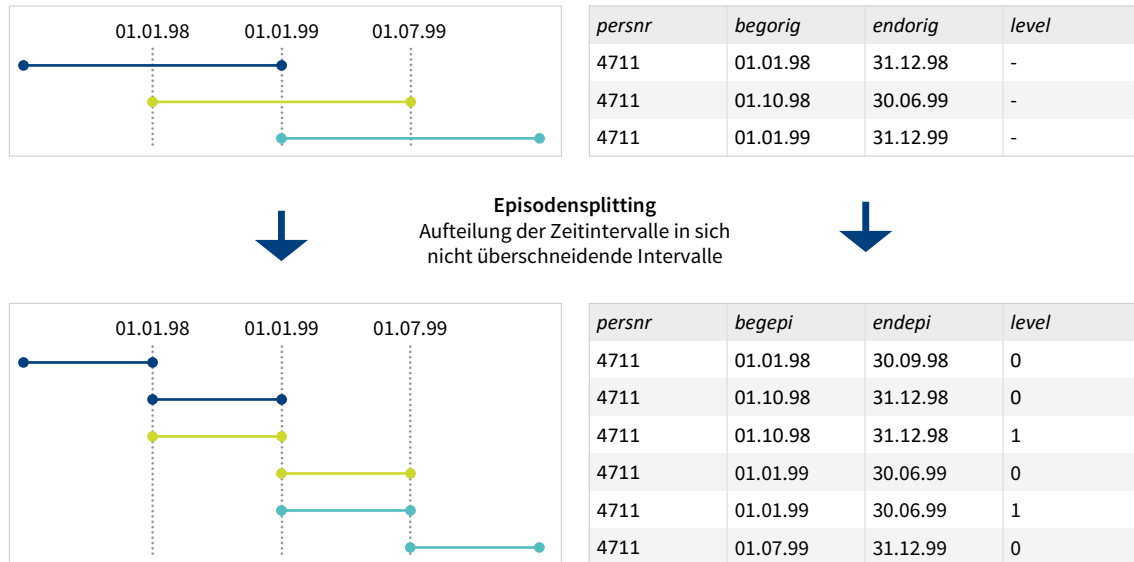


Abbildung 4: Episodensplitting

Es ist ratsam, die durch das Splitting erzeugten, vollständig parallelen Sätze einheitlich zu sortieren. Die früher in den LIAB Modellen enthaltenen Variablen „Satzzähler pro Episode“ (*level2*) und „Satzzähler pro Episode und Quelle“ (*level1*) können bei Bedarf mit folgenden Stata-Befehlen selbst erzeugt werden:

```
bysort persnr begepi quelle (spell): gen byte level1 = _n-1
bysort persnr begepi (spell): gen byte level2 = _n-1
```

Kasten 3: Beispielcode zum Erstellen weiterer Satzzähler in Stata 14

3.3 Stichprobenziehung

Grundlage der Stichprobenziehung in den LIAB sind die Betriebe im IAB-Betriebspanel. Das IAB-Betriebspanel ist eine jährliche Betriebsbefragung. Die Stichprobe der zu befragenden Betriebe wird aus der Grundgesamtheit aller deutschen Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten gezogen. Dabei wird nach Branche, Betriebsgröße und Bundesland geschichtet.

Die Personendaten des LIAB LM 7517 werden nach folgendem Prinzip aus den IEB gezogen. Im ersten Schritt werden aus dem IAB-Betriebspanel alle *Panelfälle* für den Zeitraum 2009 bis 2016 ausgewählt. *Panelfälle* umfassen dabei

- den Bestand an Betrieben im gewählten Startjahr 2009,
- in den Folgejahren neu hinzugekommene Betriebe sowie
- im Panelzeitraum erloschene Betriebe.

Voraussetzung für die Auswahl als *Panelfall* ist, dass eine gültige Befragung in den Folgejahren vorliegt oder der Betrieb in den Folgejahren erlischt.⁴ Es ist zu beachten, dass nicht alle Querschnittsfälle einer Welle des IAB-Betriebspanels auch gleichzeitig *Panelfälle* darstellen. Wurde ein Betrieb im Rahmen der Ersatzstichprobe, die aus dem Bestand gezogen wird, hinzugefügt, oder wurde nicht die gleiche Betriebseinheit wie im Vorjahr befragt, so ist der betreffende Fall kein *Panelfall* (vgl. Fischer et al., 2008, S.23f).

Im zweiten Schritt werden aus den IEB alle Personen gezogen, die im Zeitraum 2008 – 2017 mindestens einen Tag in einem dieser ausgewählten Panelbetriebe beschäftigt waren.

Für diese Personen werden im dritten Schritt die Erwerbsbiographien für die Quellen BeH und LeH aus den IEB gezogen. Die Erwerbsbiographien umfassen dabei den Zeitraum 1975 – 2017. Abbildung 5 verdeutlicht die Zeitdimension im LIAB LM 7517.

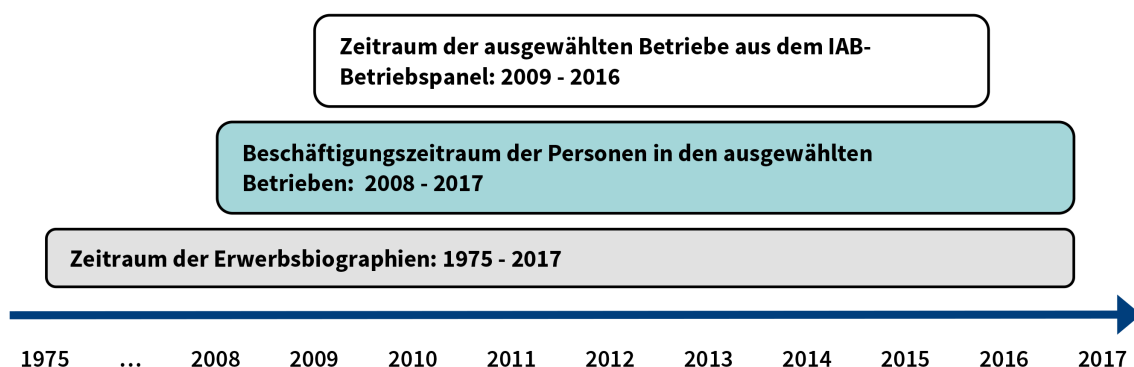


Abbildung 5: Zeitdimension im LIAB LM 7517

⁴ Vgl. Fischer et al. (2008), S. 29f.

3.4 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 (weiße Felder der Merkmalsübersicht in Tabelle 4) oder für einen gewissen Zeitraum nicht gefüllt.

4 Datenqualität und Probleme

4.1 Gesamte IEB

4.1.1 Lücken in den Erwerbsbiografien

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

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

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

Diese Lücken wurden identifiziert anhand der Ausprägungen der Merkmale "Abmeldegrund / Abgabegrund / Beendigungsgrund" (*grund*) und „Erwerbsstatus“ (*erwstat*) in den verschiedenen Quellen. Die Auflistung in Tabelle 6 erhebt keinen Anspruch auf Vollständigkeit.

Tabelle 6: Biographische 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	XASU
Ungeförderte Selbstständige	LeH, ASU
Studierende, Personen in schulischer Weiterbildung	LeH, LHG, ASU, XASU

Personen, die länger als 6 Wochen krank / arbeitsunfähig sind (Krankheit während Arbeitslosigkeit ist aber unter bestimmten Bedingungen in der Quelle ASU abgebildet)	BeH, LeH, ASU
Personen in Altersrente ohne Beschäftigung, wenn nicht Mitglied in einer BG	LeH, LHG, ASU
Personen in Mutterschutz / Elternzeit	XASU
Bezieher von Vorruhestandsgeld	LeH, ASU
Hausgewerbetreibende	
Kurzfristig Beschäftigte	ASU
Personen in Einrichtungen der Jugendhilfe, Berufsbildungswerken, anerkannten Werkstätten oder ähnlichen Einrichtungen für behinderte Menschen	ASU
Teilnehmende an Leistungen zur Teilhabe am Arbeitsleben (Rehabilitanden)	ASU
(Nebenerwerbs-)Landwirte	
Pflegepersonen i.S. von § 19 SGB XI	
Grundwehrdienstleistende	BeH, LeH, LHG, ASU, XASU
Wehrübungsleistende	BeH, LeH, LHG, ASU, XASU
Zivildienstleistende	BeH, LeH, LHG, ASU, XASU
Ableistende eines freiwilligen sozialen bzw. ökologischen Jahres anstelle des Zivildienstes	
Sonstige Personen, die nicht bei der Rentenversicherung oder der BA registriert sind (z.B. Sabbatjahr, Finanzierung aus privatem Vermögen oder Rentenbezügen, Emigration, Beschäftigung im Ausland, ehrenamtliche Tätigkeit, etc.)	BeH, LeH, ASU
Streikende bei Streikdauer über einem Monat	LeH
Sozialhilfe-Bezieher (vor Einführung des SGB II 2005), Sozialgeld-Bezieher (im Sinne des SGB II)	
SGB-II-Leistungsberechtigte, bei deren Trägern Lieferausfälle vorkamen	
Ausgleichsempfänger nach dem FELEG (Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit)	

4.1.2 Einführung des SGB II und spätere Trägerwechsel

Mit der Einführung des SGB II zum 01.01.2005 wurden auch die Zuständigkeiten für die Betreuung von Arbeitsuchenden neu gestaltet, sodass die Aufgabenwahrnehmung im SGB II von unterschiedlichen Trägertypen übernommen werden kann:

- In der Regel übernahmen die Arbeitsgemeinschaften (ARGE) aus BA und Kreis die Aufgaben im Bereich des SGB II. Diese wurde 2011 von den gemeinsamen Einrichtungen (gE) – auch „Job-center“ genannt – abgelöst.
- Zudem war es bis Ende 2011 möglich, dass BA und Kreis die die ihnen übertragenen Aufgaben in eigener (getrennter) Zuständigkeit – in sogenannten getrennten Trägerschaften (gT) oder auch getrennten Aufgabenwahrnehmungen (gAw) – erbrachten.

- Seit 2005 besteht zudem die Möglichkeit, dass zugelassene kommunale Träger (zkT; auch: optierende Kommunen) die Aufgaben übernehmen. Zunächst erhielten 69 Landkreise oder kreisfreie Städte die alleinige Trägerschaft der Grundsicherung für Arbeitsuchende bis 31.12.2010. Mit der Verfassungsänderung wurde die Optionsmöglichkeit ab 2012 auf 110 Träger ausgedehnt.

Während die ARGEn/gE das Leistungs- und Fallmanagement über die BA-Verfahren pflegen und die erfassten Daten an die Statistik der BA übermitteln, verwenden die kommunalen Grundsicherungsträger jeweils eigene Softwaresysteme. Die Übermittlung an die Statistik der BA erfolgt über den Standard XSozial. Für getrennte Träger lag eine Aufteilung der Zuständigkeiten vor. Der jeweilige Trägerteil war nur für die Erfassung und Lieferung der Daten verantwortlich, die in seinen Zuständigkeitsbereich fallen. Daher gab es reduzierte Datenanforderungen für die kommunale Seite der getrennten Träger. Die Übermittlung über den Standard XSozial erfolgt einmal monatlich zu einem festgelegten Zeitfenster. In der Woche darauf besteht an einem zweiten Meldetag die Möglichkeit, missglückte Lieferungen zu wiederholen. Im Lauf der Jahre wurden zahlreiche Qualitätssicherungsinstrumente entwickelt. Dennoch erhöhen Versionswechsel des Standards oder der vor Ort genutzten Software regelmäßig das Risiko für Datenqualitätsprobleme.

Im Rahmen der Neuorganisation von SGB-II-Trägern jeweils zum 01.01. der Jahre 2011 bis 2014 gingen diverse ARGEn/gE und getrennte Träger in kommunale Trägerschaft über bzw. wurden von zkT in gE umgewandelt. Diese Trägerwechsel und der damit einhergehende Wechsel der vor Ort verwendeten Software führen teilweise auch zu Brüchen in den Daten der SGB-II-Quellen. Detaillierte Informationen zu den daraus resultierenden Qualitätsproblemen in den Daten finden sich in den einzelnen Kapiteln zu den Datenquellen.

4.2 Beschäftigtenhistorik (BeH)

- Die Übermittlung von Informationen zur Ausbildung, ausgeübten Tätigkeit und Stellung im Beruf erfolgt durch Meldungen der Arbeitgeber nach DEÜV (siehe Abschnitt 2.1) über einen sogenannten Tätigkeitsschlüssel. Für Zeiträume mit einem Enddatum nach dem 30.11.2011 erfolgte der Wechsel zum neuen Tätigkeitsschlüssel 2010 (siehe Bertat et al., 2013, für mehr Details). Dieser Wechsel wurde von den Spitzenorganisationen der Sozialversicherungsträger beschlossen, da der Tätigkeitsschlüssel 2003 eine Reihe von Sachverhalten nicht mehr zeitgemäß und nicht mehr realitätsnah erfassen konnte. Da die Meldungen der Arbeitgeber nach DEÜV in die BeH einfließen, hat die Umstellung des Tätigkeitsschlüssels ausschließlich Auswirkungen für Beobachtungen aus dieser Quelle. Die Messung folgender bisher enthaltener Angaben ist von dieser Umstellung betroffen: Arbeitszeit, Beruf, Stellung im Beruf sowie schulisches und berufliches Qualifikationsniveau. Darüber hinaus gibt es seit der Umstellung auch Angaben zur Befristung eines Beschäftigungsverhältnisses sowie dazu, ob eine Person als Verleihpersonal in einem Zeitarbeitsunternehmen angestellt ist. Die wichtigste Auswirkung ist der Umstieg auf eine neue Berufsklassifikation. Statt der bisherigen Klassifikation der Berufe 1988 (KldB 1988) wird mit dem neuen Tätigkeitsschlüssel die differenziertere KldB 2010 gemeldet.⁵

⁵ Siehe Paulus/Matthes, 2013, für Details zur Klassifikation der Berufe 2010.

- Anders als im LIAB LM 9314 ist das Merkmal „Stellung im Beruf und Arbeitszeit“ (*stib*) im LIAB LM 7517 wieder enthalten. Die Variable ist allerdings nur für Meldungen gefüllt, die vor Einführung des neuen Tätigkeitsschlüssels abgegeben wurden, und das FDZ nimmt keine Extrapolation oder Imputation für spätere Meldungen vor. Die Kategorien der Variable *stib*, die über den gesamten Beobachtungszeitraum konsistent beobachtet werden können (Erwerbsstatus als Auszubildender, Unterscheidung von Teil- und Vollzeit), sind in den Variablen *erwstat* und *teilzeit* nachzuvollziehen.
- Die Einführung des neuen Tätigkeitsschlüssels hat zu einer Reihe von Problemen geführt. So kam es während des Übergangszeitraums, der Arbeitgebern im Meldeverfahren zur Sozialversicherung eingeräumt wurde,⁶ zu einer temporären Zunahme fehlender Angaben. Analysen der Statistik der BA (Bertat et al., 2013, S. 10) zeigen, dass die neuen bzw. umgestellten Merkmale „ausgeübte Tätigkeit“, „Arbeitszeit“ und „Ausbildungsabschluss“ nach der Umstellung zu etwa 20 bis 30 Prozent keine Angaben enthielten. Eine deutliche Verbesserung stellt sich im ersten Halbjahr 2013 ein. Um die Qualität der Arbeitszeitangabe im Übergangszeitraum zu verbessern, haben Ludsteck/Thomsen (2016) ein Imputationsverfahren entwickelt, das die fehlenden Werte durch imputierte Werte ersetzt. Die imputierten Daten werden im LIAB LM 7517 verwendet. Für die Lücken in den anderen Merkmalen werden keine Imputationen vorgenommen.
- 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. Ausbildung, Erwerbsstatus und Tagesentgelt) fehlende Werte auf.
- Der Anstieg der Zahl der BeH-Sätze ab 1999 folgt aus der Einführung der Meldepflicht für geringfügig Beschäftigte ab dem 01.04.1999.
- Insbesondere im Jahr 1999 lässt sich ein deutlicher Anstieg bei den Meldungen zu Teilzeitbeschäftigung feststellen. Verursacht wird dieser Befund sowohl vom tatsächlich zu beobachtenden Anstieg der Teilzeitbeschäftigung als auch von der Tatsache, dass seit 1999 die Beschäftigtenmeldungen generell besser ausgefüllt werden.
- Innerhalb des Meldeverfahrens ist eine gewisse Zeitverzögerung nicht vermeidbar. Zwar müssen Änderungen von Beschäftigungsverhältnissen sofort gemeldet und bestehende Beschäftigungsverhältnisse jährlich bis zum April (bis Jahresmeldung 2012) bzw. Mitte Februar (ab Jahresmeldung 2013) des folgenden Jahres bestätigt werden. Faktisch treffen Meldungen jedoch zum Teil um Jahre verspätet ein. Die Historikdatei des IAB wird aber nicht kontinuierlich fortgeschrieben, sondern in bestimmten Abständen aktualisiert. Dies geschieht unter Verwendung von Dateien mit den Beschäftigten-Meldungen jeweils eines Jahres, die 36, 18, 12 oder 6 Monate nach Ablauf des Berichtsjahres eingetroffen sind (frühestens im Juli 2015 z.B. kann die 18-Monats-Datei für das Jahr 2013 erstellt werden). Mehr als 3 Jahre verspätete Meldungen werden im IAB nicht berücksichtigt, so dass eine 36-Monats-Datei per Definition einen Füllgrad von 100% aufweist.

⁶ Die im Rahmen des Meldeverfahrens eingesetzten Prüfprogramme ließen bis einschließlich Mai 2012 leere Angaben beim Tätigkeitsschlüssel 2010 zu.

- In der zugrundeliegenden Version der IEB ist 2014 das letzte Jahr mit einem Füllgrad der BeH-Meldungen von 100%.⁷ Für die Jahre 2015 und 2016 wurden 18-Monats-Dateien verwendet, für 2017 eine 6-Monats-Datei. Es ist daher davon auszugehen, dass Beschäftigungsmeldungen für die Jahre 2015 bis 2016 leicht und für 2017 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 et al., 1996).
- Für die Jahre 1992 bis 2000 wurden in den IEB auffällige Einbrüche und Anstiege in der Anzahl der Meldungen festgestellt. Es lassen sich v.a. Einbrüche bei folgenden zehn Kreisen beobachten: Braunschweig (03101), Wolfsburg (03103), Emden (03402), Kassel (06633), Essen (05113), Neuss (05162), Erftkreis (05362), Hersfeld-Rotenburg (06632), Miltenberg (09676) und Kempten (Allgäu) (09763). Grund hierfür sind Meldeprobleme bei einzelnen bzw. mehreren Betrieben in diesen Regionen.
- Nennenswerte Einbrüche verzeichnen auch die Kreise Salzgitter (03102) und Hoyerswerda (14264).
- Bei den Meldungen zu Vollzeitbeschäftigung in den IEB sind besonders die Kreise Main-Taunus-Kreis (06436) und Alzey-Worms (07331) erwähnenswert. Hier sind überdurchschnittliche Anstiege zu beobachten. Als Grund sind ebenfalls Meldeprobleme bei einzelnen bzw. mehreren Betrieben in diesen Regionen anzuführen.
- In den Jahren 1996 bis 1998 kommen die Werte 841-844 (Ärzte und Apotheker) in der Variable Beruf – ausgeübte Tätigkeit (*beruf*) im Vergleich zu benachbarten Jahren sehr selten vor. Die Gründe dafür sind nicht bekannt.
- In den Jahren 1975 bis 1977 gab es bisher deutlich zu viele Beschäftigte mit einem Arbeitsort in der niedersächsischen Gemeinde Lahn (Kreis Emsland). Statt im zu erwartenden zweistelligen Bereich gab es pro Jahr bis zu 90.000 Beschäftigtenmeldungen mit diesem Arbeitsort in der Grundgesamtheit der BeH. Ursächlich ist eine historische Fehlzurordnung der Beschäftigtenmeldungen aus der Stadt Lahn. Hierbei handelte es sich um einen nach kurzer Zeit wieder aufgelösten Zusammenschluss der hessischen Gemeinden Wetzlar, Gießen, Heuchelheim, Wettenberg und Lahna (Kreise Gießen und Lahn-Dill-Kreis). Auf Basis dieser Erkenntnisse und unter der Annahme, dass es keine realen Betriebsumzüge zwischen diesen Regionen gab, wurde auf Ebene der Betriebsnummer folgende Bereinigungsregel umgesetzt: Sobald der Betriebsort in den Jahren 1975 bis 1978 von der Gemeinde Lahn auf eine der aufgeführten hessi-

⁷ Durch ein Redesign der Datengrundlage der BeH stand zum Zeitpunkt der Erstellung der BeH ausnahmsweise nur das 30- statt des 36-Monatsmaterials zur Verfügung. Analysen mit früheren Datenversionen haben jedoch gezeigt, dass i.d.R. bereits die 18-Monats-Datei einen Füllgrad von rund 99% aufweist. Demnach kommt es in den weiteren 18 Monaten nur noch zu extrem wenigen Nach- oder Korrekturmeldungen.

schen Gemeinden wechselt, wird die frühere Angabe Gemeinde Lahn mit der späteren korrekten Angabe überschrieben. Über diese Regel wird der Überhang deutlich reduziert und die niedersächsische Gemeinde Lahn weist danach in den kritischen Jahren zumindest nur noch unter 2.000 Beschäftigtenmeldungen in der Grundgesamtheit der BeH auf.

4.3 Leistungsempfängerhistorik (LeH)

- Für die neuen Bundesländer werden die LeH-Sätze erst ab 1992 vollständig erfasst.
- Die Leistungsbezugsdaten wurden früher auf Magnetbändern gesichert. Aufgrund eines Schadens an einem Magnetband sind die Leistungsbezugsdaten bis einschließlich des Jahres 1980 nur teilweise enthalten. Daher ist für das vorliegende Datenprodukt auch davon auszugehen, dass die Informationen zum Leistungsbezug im genannten Zeitraum nur zum Teil enthalten sind.
- Durch einen internen Systemwechsel kommt es 2004 zu einem Bruch in der Erfassung von Sperr- und Säumniszeiten. Bis 01.07.2004 können Sperr- und Säumniszeiten nur durch den Abmeldegrund im vorhergehenden LeH-Satz identifiziert werden. Danach existiert ein eigener Satz mit täglichem Leistungssatz = 0 für Versagens-, Entziehens- und Ruhezeiten.

5 Variablenbeschreibung

Auszählungen sowie Übersichten über die einzelnen Ausprägungen und Labels der Merkmale finden sich in den Arbeitshilfen zum LIAB unter <https://fdz.iab.de>.

5.1 Identifikatoren

5.1.1 Systemfreie Personennummer (persnr)

Variablenlabel	Systemfreie Personennummer
Variablenname	persnr
Kategorie	Identifikatoren
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	Die systemfreie Personennummer zeigt an, welche Sätze zur gleichen Person gehören. Systemfrei bedeutet, dass von dieser Personennummer nicht auf Eigenschaften der Person oder Original-Identifikatoren geschlossen werden kann. Da es in den verschiedenen Datenquellen keinen einheitlichen Personen-Identifikator gibt, ist die Zuordnung der Informationen aus den unterschiedlichen Datenquellen zu Personen(-Konten) nicht immer eindeutig. In solchen Fällen kann es zu unplausiblen Erwerbshistoriken kommen. Die Bildung des quellenübergreifenden Personen-Identifikators basiert auf einer von der BA entwickelten Heuristik.

5.1.2 Systemfreie Betriebsnummer (betnr)

Variablenlabel	Systemfreie Betriebsnummer
Variablenname	betnr

Kategorie	Identifikatoren
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Die systemfreie Betriebsnummer zeigt an, welche Sätze zum gleichen Betrieb gehören. Sie basiert auf der von der BA vergebenen Betriebsnummer, die durch eine systemfreie Nummer ersetzt wurde (weitere Informationen zur Betriebsnummernvergabe der BA finden sich in Bender et al., (1996, S. 15f. und S. 27-30) sowie direkt beim Betriebsnummern-Service der BA unter https://www.arbeitsagentur.de/betriebsnummern-service/alles-wichtige). Die Betriebsnummer kann zusammen mit der Jahresangabe genutzt werden, um Personen- und Betriebsinformationen zusammenzuspielen. Für die Betriebsnummer sollte grundsätzlich Folgendes beachtet werden: 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. Hat das Unternehmen mehrere Niederlassungen in einer 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 nicht derselben Wirtschaftsklasse angehören. Hat das Unternehmen mehrere Niederlassungen in mehreren Gemeinden, so ist jede dieser Niederlassungen ein Betrieb und erhält je eine Betriebsnummer. In diesem Zusammenhang sind auch folgende Definitionen im Sinne der Betriebsnummernvergabe im Rahmen des Meldeverfahrens zur Sozialversicherung zu beachten: Ein Betrieb ist eine regional und wirtschaftsfachlich abgegrenzte Einheit, in der Beschäftigte tätig sind und für die nach o.a. Grundsätzen eine Betriebsnummer vergeben wird. Eine Arbeitsstätte ist eine Einheit, in der Beschäftigte tätig sind und für die nach o.a. Grundsätzen keine Betriebsnummer zu vergeben ist. Ein Unternehmen als Begriff fasst Betriebs- und Arbeitsstätten desselben Arbeitgebers zusammen. Ein Arbeitgeber ist jede natürliche oder juristische Person, für die mindestens ein sozialversicherungspflichtig bzw. geringfügig Beschäftigter tätig ist. 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 nur sehr wenigen BeH-Meldungen. Bei diesen Meldungen handelt es sich in den meisten Fällen um Meldungen der Personengruppe „205“ (Entgeltmeldungen für unständig Beschäftigte). Da Betriebsmerkmale (Arbeitsort, Wirtschaftszweig, Betriebsgröße usw.) über die Betriebsnummer zugespielt werden, fehlen sie in diesen Meldungen.

5.1.3 Betriebsnummer nach Betriebspanel (idnum)

Variablenlabel	Betriebsnummer gemäß Betriebspanel
Variablenname	idnum
Kategorie	Identifikatoren
Herkunft	IAB-Betriebspanel

Datentyp	Numerisch
Detailbeschreibung	Die Betriebsnummer gemäß Betriebspanel wird aus dem IAB-Betriebspanel übernommen. Sie ist in der Personendatei nur gefüllt, falls die Person am Stichtag in einem Betrieb mit gültiger Befragung im IAB-Betriebspanel beschäftigt war.

5.2 Generierte technische Merkmale

5.2.1 Satzzähler pro Konto (spell)

Variablenlabel	Satzzähler pro Konto
Variablenname	spell
Kategorie	Generierte technische Merkmale
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	Der Satzzähler pro Konto nummeriert aufsteigend die Sätze einer Person. Das Merkmal wird beim Episodensplitting erzeugt und bezieht sich auf die gesplitteten Sätze. Mit dem Merkmal „Satzzähler pro Konto“ kann die ursprüngliche Sortierung auf einfache Art und Weise wiederhergestellt werden. Die Sortierung erfolgt grundsätzlich zunächst nach dem Beginndatum der Episode und dann nach der Quelle. Innerhalb der Beschäftigungsmeldungen werden sozialversicherungspflichtige vor geringfügige Beschäftigungsmeldungen und höhere vor niedrigere Tagesentgelte sortiert. Eine Ausnahme bilden Einmalzahlungen, die nach hinten sortiert werden.

5.2.2 Quelle des Satzes (quelle)

Variablenlabel	Quelle des Satzes
Variablenname	quelle
Kategorie	Generierte technische Merkmale
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt die Datenquelle an (siehe Kapitel 2).

5.2.3 Jahr (jahr)

Variablenlabel	Jahr
Variablenname	jahr
Kategorie	Generierte technische Merkmale
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Diese Variable gibt das Kalenderjahr zu Beginn der Episode (also nach dem Episodensplitting, siehe Abschnitt 3.2) an. Mit Hilfe dieses Merkmals und der Betriebsnummer können Personen- und Betriebsdateien verknüpft werden. Siehe Kasten 2 auf Seite 10 für einen Beispielcode mit dem „merge“-Befehl in Stata 14.

5.2.4 Status der Betriebsnummer (betr_st)

Variablenlabel	Status der Betriebsnummer
Variablenname	betr_st
Kategorie	Generierte technische Merkmale

Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt für die gegenwärtige Episode in den Personendaten an, ob für diese Episode Informationen aus dem IAB-Betriebspanel vorliegen. Die folgenden Ausprägungen existieren: 1. Am 30.6. eines Befragungsjahrs beschäftigt, Betrieb für LIAB-Modell ausgewählt 2. Nicht am 30.6. eines Befragungsjahrs beschäftigt, Betrieb für LIAB-Modell ausgewählt 3. Beschäftigung in Betriebspanelbetrieb, Betrieb nicht für LIAB-Modell ausgewählt 4. Beschäftigung in Nicht-Betriebspanelbetrieb Die Ausprägungen 1-3 kennzeichnen Beschäftigungen in Betrieben, die irgendwann einmal am IAB-Betriebspanel teilgenommen haben, während die Ausprägung 4 Beschäftigungen in Betrieben kennzeichnet, die nie am IAB-Betriebspanel teilgenommen haben. Mit 1 und 2 sind Beschäftigungen in IAB-Betriebspanelbetrieben markiert, die für das jeweilige LIAB Modell ausgewählt wurden, mit 3 Beschäftigungen in IAB-Betriebspanelbetrieben, die nicht für das jeweilige LIAB Modell ausgewählt wurden. 1 und 2 unterscheiden schließlich, ob die Episode den 30.6. eines Jahres einschließt, für das der Betrieb befragt wurde, oder nicht. Die Ausprägung 3 tritt definitionsgemäß im Querschnittmodell des LIAB nicht auf, da alle Betriebe des IAB-Betriebspanels irgendwann einmal befragt wurden und somit per Konstruktion für das LIAB Querschnittmodell ausgewählt werden.

5.3 Gültigkeitszeitraum

5.3.1 Beginndatum Originalbeobachtung (begorig)

Variablenlabel	Beginndatum Originalbeobachtung
Variablenname	begorig
Kategorie	Gültigkeitszeitraum
Herkunft	BeH, LeH
Datentyp	Datum
Detailbeschreibung	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). In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe von Jahresmeldungen). Ein durchgehendes Beschäftigungsverhältnis kann also auf mehrere Meldungen verteilt sein.

5.3.2 Enddatum Originalbeobachtung (endorig)

Variablenlabel	Enddatum Originalbeobachtung
Variablenname	endorig
Kategorie	Gültigkeitszeitraum

Herkunft	BeH, LeH
Datentyp	Datum
Detailbeschreibung	Das Enddatum des Originalsatzes entspricht dem ursprünglichen Enddatum der Meldung. Dieses kann vom Enddatum der betreffenden Datenzeile, dem sogenannten Episoden-Enddatum, abweichen (vgl. auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2). In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe von Jahresmeldungen). Ein durchgehendes Beschäftigungsverhältnis kann also auf mehrere Meldungen verteilt sein.

5.3.3 Beginndatum Episode (begepi)

Variablenlabel	Beginndatum Episode
Variablenname	begepi
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	BeH, LeH
Datentyp	Datum
Detailbeschreibung	Das Beginndatum der Episode ist immer gleich dem oder größer als das Beginndatum der Originalbeobachtung (vgl. auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2).

5.3.4 Enddatum Episode (endepe)

Variablenlabel	Enddatum Episode
Variablenname	endepe
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	BeH, LeH
Datentyp	Datum
Detailbeschreibung	Das Enddatum der Episode ist immer gleich dem oder kleiner als das Enddatum der Originalbeobachtung (vgl. auch die Erläuterungen zum Episodensplitting in Abschnitt 3.2).

5.4 Informationen zur Person

5.4.1 Geschlecht (frau)

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

5.4.2 Geburtsjahr (gebjahr)

Variablenlabel	Geburtsjahr
Variablenname	gebjahr
Kategorie	Persönliche Merkmale

Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	Das Geburtsjahr ist innerhalb eines Personenkontos konstant.
Qualitätshinweis	In den Originaldaten kann es vorkommen, dass das Geburtsdatum zwischen den Quellen wechselt. Dies wird bei der Datenaufbereitung bereinigt. Hierbei kommt Angaben aus dem Leistungsbezug die höchste Priorität zu.

5.4.3 Geburtsmonat (*gebmon*)

Variablenlabel	Geburtsmonat
Variablenname	gebmon
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	Der Geburtsmonat ist innerhalb eines Personenkontos konstant. Um aus den Angaben zum „Geburtsjahr“ (<i>gebjahr</i>) und „Geburtsmonat“ (<i>gebmon</i>) eine Variable im Datumsformat JJJJmM (z.B. 1984m6) zu generieren, kann folgender Befehl in Stata verwendet werden: <code>gen int gebdat = ym(gebjahr, gebmon)</code> <code>format gebdat %tm</code>
Qualitätshinweis	In den Originaldaten kann es vorkommen, dass das Geburtsdatum zwischen den Quellen wechselt. Dies wird bei der Datenaufbereitung bereinigt. Hierbei kommt Angaben aus dem Leistungsbezug die höchste Priorität zu.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt. Ansonsten wird lediglich das „Geburtsjahr“ (<i>gebjahr</i>) bereitgestellt.

5.4.4 Staatsangehörigkeit (*nation*)

Variablenlabel	Staatsangehörigkeit
Variablenname	nation
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält als Ausprägungen die Staatenschlüssel des Statistischen Bundesamtes (Statistisches Bundesamt, 2019).
Qualitätshinweis	Diese Variable ist in der LeH vor 1983 nicht gut gefüllt.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird lediglich die „Staatsangehörigkeit vergrößert“ (<i>nation_gr</i>) bereitgestellt.

5.4.5 Staatsangehörigkeit vergrößert (*nation_gr*)

Variablenlabel	Staatsangehörigkeit vergrößert
Variablenname	nation_gr
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält als Ausprägungen die zu Gruppen zusammengefassten Staatenschlüssel des Statistischen Bundesamtes (Statistisches Bundesamt, 2019).

Qualitätshinweis	Diese Variable ist in der LeH vor 1983 nicht gut gefüllt.
------------------	---

5.4.6 Familienstand (famst)

Variablenlabel	Familienstand
Variablenname	famst
Kategorie	Persönliche Merkmale
Herkunft	LeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal beschreibt den Familienstand. In der Quelle LeH umfasst das Merkmal zwei Ausprägungen (0 - nicht verheiratet, 1 - verheiratet).
Qualitätshinweis	Die Qualität des Merkmals aus der LeH wird bis 2005 als schlecht eingestuft, da nicht ganz geklärt ist, wie der Familienstand gepflegt wurde. Auch ab Beginnjahr 2014 ist die Qualität fraglich. So sinkt der Anteil der Verheirateten z.B. zwischen 2013 und 2017 von 37% auf 14%.

5.4.7 Kinderzahl (kind)

Variablenlabel	Kinderzahl
Variablenname	kind
Kategorie	Persönliche Merkmale
Herkunft	LeH
Datentyp	Numerisch
Detailbeschreibung	In der Quelle LeH gibt das Merkmal die Anzahl der Kinder unter 16 Jahren zum Zeitpunkt der Antragstellung an. Es wird nur unterschieden, ob die Anzahl an Kindern gleich Null oder mindestens eins ist. Folgende Ausprägungen kommen daher vor: 0 Keine Kinder 100 Ein oder mehrere Kinder
Qualitätshinweis	In der LeH wird das Merkmal bei Änderungen der Leistungsart oder neuen Bewilligungen nicht aktualisiert, sondern erst, wenn ein erneuter Leistungsfall auftritt. Dadurch ist die Datenqualität stark eingeschränkt.

5.4.8 Ausbildung (ausbildung)

Variablenlabel	Ausbildung
Variablenname	ausbildung
Kategorie	Persönliche Merkmale
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal enthält für BeH-Sätze die von den Arbeitgebern im Rahmen des Meldeverfahrens anzugebende Berufsausbildung. Folgende Ausprägungen kommen vor: 1 ohne abgeschlossene Berufsausbildung 2 betriebliche/außerbetriebliche Ausbildung 11 FH ohne nähere Spezifikation 12 Hochschulabschluss ohne nähere Spezifikation

	Bei Meldungen nach dem neuen Tätigkeitsschlüssel (siehe Abschnitt 4.2) ist es nicht mehr möglich, FH-AbsolventInnen wie bisher eindeutig zu identifizieren, da der neue Tätigkeitsschlüssel die separate Erfassung dieses Bildungsabschlusses nicht länger erlaubt. Sie werden der Kategorie 12 zugeordnet.
Qualitätshinweis	„Zusammen mit einem Betriebswechsel treten häufig auch Änderungen im Ausbildungsstatus auf. Das liegt daran, dass beim neuen Betrieb die Meldedaten neu zusammengestellt werden. Wenn ein Beschäftigter z.B. durch berufsbegleitende Weiterbildung einen höheren Abschluss erlangt hat, dann wird dieser Statuswechsel wahrscheinlich erst von einem neuen Beschäftigungsbetrieb erfasst. Es ist in der Regel davon auszugehen, dass bei längeren Beschäftigungszeiten innerhalb eines Betriebs die einmal erhobenen personenbezogenen Daten fortgeschrieben werden.“ (Meinken/Koch, 2004, S. 63). Der Anteil der fehlenden Werte steigt in der BeH im Zeitverlauf fast kontinuierlich an. Durch die Einführung des neuen Tätigkeitsschlüssels gab es in 2011 einen stark erhöhten Anteil von rund 51%. Seit 2014 pendelt sich der Anteil der fehlenden Werte in der BeH bei rund 40% ein. Fehlende Werte treten besonders häufig bei folgenden Personengruppen auf: geringfügig Beschäftigte, Personen in Teilzeit, ausländische Beschäftigte und ostdeutsche Arbeitnehmer. Hintergrund ist, dass das Merkmal versicherungsrechtlich keine besondere Bedeutung hat (siehe Meinken/Koch, 2004, S. 63). Für das Merkmal „Ausbildung (imputiert)“ (<i>ausbildung_imp</i>, siehe Abschnitt 5.4.9) wurde eine Methode angewendet, um fehlende Werte oder inkonsistente zeitliche Verläufe der Bildungsangaben zu bereinigen. Diese Variable ist jedoch nur in der Quelle BeH gefüllt und hat andere Ausprägungen als die Variable <i>ausbildung</i>.

5.4.9 Ausbildung (imputiert) (*ausbildung_imp*)

Variablenlabel	Ausbildung (imputiert)
Variablenname	<i>ausbildung_imp</i>
Kategorie	Persönliche Merkmale
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal „Ausbildung (imputiert)“ (<i>ausbildung_imp</i>) ist ein Zusatzmerkmal zum Merkmal „Ausbildung“ (<i>ausbildung</i>) und enthält für BeH-Meldungen ergänzte und harmonisierte Angaben zur Berufsausbildung von Beschäftigten. Das Merkmal bietet damit einen Lösungsansatz für die in Abschnitt 5.4.8 beschriebenen Probleme der Variable <i>ausbildung</i> bei BeH-Meldungen. Das Imputationsverfahren wird in Thomsen et al. (2018) beschrieben. Da das Merkmal nur Bildungsangaben der BeH-Meldungen nutzt, und weil die Bildungskategorien des alten und des neuen Tätigkeitsschlüssels für das Merkmal <i>ausbildung</i> harmonisiert werden mussten, bietet das Merkmal <i>ausbildung_imp</i> andere Kategorien als das Merkmal <i>ausbildung</i>.

5.4.10 Schulausbildung (*schule*)

Variablenlabel	Schulausbildung
Variablenname	<i>schule</i>
Kategorie	Persönliche Merkmale
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Mit der Umstellung des Tätigkeitsschlüssels ändern sich die möglichen Ausprägungen (siehe Abschnitt 4.2). Davor gelten für die BeH:

	5 Volks-/Hauptschule o./m. Abschluss, Mittlere Reife oder gleichwertige Schulbildung 8 Fachhochschulreife oder Abitur/Hochschulreife 9 Abitur/Hochschulreife Danach gelten: 1 Kein Hauptschulabschluss 4 Hauptschulabschluss/Volksschulabschluss 6 Mittlere Reife 8 Fachhochschulreife oder Abitur/Hochschulreife
Qualitätshinweis	Der Füllgrad nimmt in der BeH über die Jahre kontinuierlich ab und liegt in jüngeren Jahren unter zwei Dritteln.

5.5 Informationen zu Beschäftigung und Leistungsbezug

5.5.1 Tagesentgelt/ täglicher Leistungssatz (tentgelt)

Variablenlabel	Tagesentgelt/täglicher Leistungssatz
Variablenname	tentgelt
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	1) BeH In BeH-Sätzen weist das Merkmal jeweils das durchschnittliche Bruttotagesentgelt einer beschäftigten Person 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 01.04.1999 werden auch Entgelte unterhalb der Geringfügigkeitsgrenze erfasst. Die Beitragsbemessungsgrenze als obere Kappungsgrenze gilt jedoch weiterhin. In einigen Fällen übersteigen die gemeldeten Entgelte die Beitragsbemessungsgrenze allerdings trotzdem. Dies dürfte in der Regel auf Jahressonderzahlungen zurückzuführen sein, die der Arbeitgeber bei den Jahres-, Unterbrechungs- oder Abmeldungen zum normalen Entgelt hinzurechnen kann. Dabei ist es dann unerheblich, ob durch diese Addition die für den Meldezeitraum maßgebliche Beitragsbemessungsgrenze in der Rentenversicherung überschritten wird. Es könnte sich allerdings auch um fehlerhafte Angaben bei der Beschäftigungszeit handeln. (Die Entgelt-Informationen hingegen können auf Grund ihrer Versicherungsrelevanz als weniger fehleranfällig gesehen werden.) Die Geringfügigkeits- und Beitragsbemessungsgrenzen unterscheiden sich von Jahr zu Jahr und in West- und Ostdeutschland (maßgeblich ist jeweils der Standort des Betriebs). Eine Übersicht über diese Grenzen findet sich unter https://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. Ab 2013 nimmt die Anzahl der Meldungen mit Abmeldegrund 54 (Meldung eines einmalig gezahlten Entgelts) sprunghaft zu (siehe Abschnitt 5.5.12). Es ist davon auszugehen, dass Sonderzahlungen, die vor 2013 mit den Jahresmeldungen angegeben wurden, nun extra gemeldet werden. Daher empfiehlt es sich bei Analysen der Entgelte im Zeitverlauf, die Sonderzahlungen zu den entsprechenden Entgelten von zeitgleichen Beschäftigungsepisoden im gleichen Betrieb dazu zu addieren. 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-Enddatum, das vor dem 01.01.1998 liegt, für Werktage (also ohne Sonn- und Feiertage, jedoch inklusive Samstage) und bei Datensätzen mit einem Original-Enddatum ab 01.01.1998 für Kalendertage gilt. Die Angabe von 0 Euro ist seit 01.01.2005 auf Sperrzeiten oder Leistungsunterbrechungen zurückzuführen. Ist für einen Satz mit Leistungssatz gleich 0 ein Beendigungsgrund angegeben, so handelt es sich um eine Unterbrechungsmeldung. Bei Sätzen, die eine Sperrzeit widerspiegeln, ist der Anspruch derselbe wie vor Beginn der Sperrzeit.
--	---

5.5.2 Beruf – ausgeübte/letzte Tätigkeit (KldB 1988) (beruf)

Variablenlabel	Beruf – ausgeübte/letzte Tätigkeit (KldB 1988)
Variablenname	beruf
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	In der BeH ist die Berufsbezeichnung der im Meldezeitraum von der beschäftigten Person ausgeübten Tätigkeit Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für eine beschäftigte Person zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. Bundesagentur für Arbeit, 2005, S. V). Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifizierung der Berufe 1988“ (Bundesanstalt für Arbeit, 1988), in der ca. 25.000 Berufsbezeichnungen zu finden sind. Die Berufsordnung besteht aus einer 3-stelligen Kennziffer und umfasst rund 330 Ausprägungen. Beschäftigungsmeldungen mit einem Enddatum nach dem 30.11.2011 werden nach dem neuen Tätigkeitsschlüssel 2010 (KldB2010) gemeldet (siehe Abschnitt 4.2). Diese Werte werden durch einen Schwerpunktumstieg auf die KldB1988 zurückgeschlüsselt. Dabei kommt es zu Unschärfen.
Qualitätshinweis	Im Jahr 2011 kommt es in der BeH aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf.

5.5.3 Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_3)

Variablenlabel	Berufsgruppe – ausgeübte/letzte Tätigkeit (KldB 2010), 3-Steller
Variablenname	beruf2010_3
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	1) BeH Die Berufsbezeichnung der im Meldezeitraum von der beschäftigten Person ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für eine beschäftigte Person zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. Bundesagentur für Arbeit, 2005, S. V). Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2011). Die Berufsgattung besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Die weniger detaillierte Berufsgruppe wird durch die ersten drei Stellen der Kennziffer erfasst. Das in der KldB2010 an der fünften Stelle erfasste Anforderungsniveau einer Tätigkeit wird gesondert in der Variable „Anforderungsniveau“ (<i>niveau</i>) zur Verfügung gestellt. Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitsschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 4.2). Diese Werte werden durch einen Schwerpunktumstieg auf die KldB2010 umgeschlüsselt. Da der neue Tätigkeitsschlüssel wesentlich detaillierter als der alte ist, kommt es dabei zu erheblichen Unschärfen. Dies ist bei Analysen zu beachten. 2) LeH Das Merkmal enthält den Beruf der letzten Tätigkeit. Zum Schlüssel siehe 1).
Qualitätshinweis	1) BeH Im Jahr 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf. 2) LeH Das Merkmal wurde früher für die Quelle LeH nicht gefüllt, da die Berufsangaben in der LeH eine schlechte Qualität aufwiesen. Mittlerweile wird der Beruf in der LeH aus der Arbeitsuchendenhistorik zugespielt, so dass die Qualität gestiegen ist.

5.5.4 Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010) (beruf2010_4)

Variablenlabel	Berufsuntergruppe – ausgeübte/letzte Tätigkeit (KldB 2010), 4-Steller
Variablenname	beruf2010_4
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH, LeH
Datentyp	Numerisch

Detailbeschreibung	1) BeH Die Berufsbezeichnung der im Meldezeitraum von der beschäftigten Person ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für eine beschäftigte Person zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. Bundesagentur für Arbeit, 2005, S. V). Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2011). Die Berufsgattung besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Die weniger detaillierte Berufsuntergruppe wird durch die ersten vier Stellen der Kennziffer erfasst. Das in der KldB2010 an der fünften Stelle erfasste Anforderungsniveau einer Tätigkeit wird gesondert in der Variable „Anforderungsniveau“ (<i>niveau</i>) zur Verfügung gestellt. Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitsschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 4.2). Diese Werte werden durch einen Schwerpunktumstieg auf die KldB2010 umgeschlüsselt. Da der neue Tätigkeitsschlüssel wesentlich detaillierter als der alte ist, kommt es dabei zu erheblichen Unschärfen. Dies ist bei Analysen zu beachten. 2) LeH Das Merkmal enthält den Beruf der letzten Tätigkeit. Zum Schlüssel siehe 1).
Qualitätshinweis	1) BeH Im Jahr 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf. 2) LeH Das Merkmal wurde früher für die Quelle LeH nicht gefüllt, da die Berufsangaben in der LeH eine schlechte Qualität aufwiesen. Mittlerweile wird der Beruf in der LeH aus der Arbeitsuchendenhistorik zugespielt, so dass die Qualität gestiegen ist.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen unvergrößert bereitgestellt. Ansonsten wird als Berufsangabe gemäß KldB2010 lediglich die Berufsgruppe <i>beruf2010_3</i> bereitgestellt.

5.5.5 Anforderungsniveau – ausgeübte/letzte Tätigkeit (KldB 2010) (*niveau*)

Variablenlabel	Anforderungsniveau - ausgeübte/letzte Tätigkeit (KldB 2010)
Variablenname	<i>niveau</i>
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	1) BeH Die Berufsbezeichnung der im Meldezeitraum von der beschäftigten Person ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für eine beschäftigte Person zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. Bundesagentur für Arbeit, 2005, S. V). Der Arbeitgeber verschlüsselt die Tätigkeit der/des Beschäftigten hierbei gemäß der „Klassifikation der Berufe 2010“ (Bundesagentur für Arbeit, 2011). Die Berufsgattung

	besteht aus einer 5-stelligen Kennziffer und umfasst rund 1.300 Ausprägungen. Das an der fünften Stelle erfasste Anforderungsniveau einer Tätigkeit wird gesondert in der Variable „Anforderungsniveau“ (<i>niveau</i>) zur Verfügung gestellt. Beschäftigungsmeldungen mit einem Enddatum vor dem 30.11.2011 werden nach dem alten Tätigkeitsschlüssel 1988 (KldB 1988) gemeldet (siehe Abschnitt 4.2). Diese Werte werden durch einen Schwerpunktumstieg auf die KldB2010 umgeschlüsselt. Da der neue Tätigkeitsschlüssel wesentlich detaillierter als der alte ist, kommt es dabei zu erheblichen Unschärfen. Dies ist bei Analysen zu beachten. 2) LeH Das Merkmal enthält den Beruf der letzten Tätigkeit. Zum Schlüssel siehe 1).
Qualitätshinweis	1) BeH Im Jahr 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf. 2) LeH Das Merkmal wurde früher für die Quelle LeH nicht gefüllt, da die Berufsangaben in der LeH eine schlechte Qualität aufwiesen. Mittlerweile wird der Beruf in der LeH aus der Arbeitsuchendenhistorik zugespielt, so dass die Qualität gestiegen ist.

5.5.6 Teilzeit (teilzeit)

Variablenlabel	Teilzeit
Variablenname	teilzeit
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal „Teilzeit“ (<i>teilzeit</i>) unterscheidet zwischen Vollzeit- und Teilzeitbeschäftigten. Maßgeblich ist dabei das Verhältnis zwischen der arbeitsvertraglich vereinbarten und der betriebsüblichen Arbeitszeit. Für Teilzeitbeschäftigte wird nur erfasst, 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.
Qualitätshinweis	Im Jahr 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Um dieses Problem abzumildern, wurde am IAB eine Imputation der Arbeitszeit für den fraglichen Zeitraum vorgenommen. Nähere Informationen zum Verfahren können Ludsteck/Thomsen (2016) entnommen werden. Eine ähnliche Häufung tritt 1991 aufgrund der Wiedervereinigung auf. Hier wird keine Imputation vorgenommen.

5.5.7 Stellung im Beruf und Arbeitszeit (stib)

Variablenlabel	Stellung im Beruf und Arbeitszeit
Variablenname	stib
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Die berufliche Stellung eines Beschäftigten im Meldezeitraum wird vom Arbeitgeber im Rahmen der „Angaben zur Tätigkeit“ übermittelt. Das Merkmal „Stellung im Beruf und Arbeitszeit“ (<i>stib</i>) unterscheidet zwischen Vollzeit- und Teilzeitbeschäftigten.

	ten. Maßgeblich ist dabei das Verhältnis zwischen der arbeitsvertraglich vereinbarten und der betriebsüblichen Arbeitszeit. Für Teilzeitbeschäftigte wird nur erfasst, 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. Nur für Vollzeitbeschäftigte liefert die Variable tatsächlich Angaben zur beruflichen Stellung, wobei unter anderem zwischen vollzeitbeschäftigten Arbeitern und Angestellten sowie Auszubildenden unterschieden wird. Die Unterscheidung zwischen vollzeitbeschäftigten Arbeitern und Angestellten richtet sich ausschließlich nach der Art der Rentenversicherungspflicht (BfA für Angestellte und LVA für Arbeiter). Meister und Poliere werden nur dann in einer eigenen Kategorie erfasst, wenn sie in der Arbeiterrentenversicherung pflichtversichert sind. Die Zuordnung eines Meisters oder Poliers zu den Arbeitern oder Angestellten kann nur über den jeweiligen Rententräger erfolgen. 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. 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. BA 2005, S. VI). Mit der Einführung des neuen Tätigkeitsschlüssels (siehe Abschnitt 4.2) fiel diese Unterscheidung jedoch weg. Die Variable <i>stib</i> ist daher nur für Meldungen gefüllt, die noch vor Einführung des neuen Tätigkeitsschlüssels abgegeben wurden.
Qualitätshinweis	Im Jahr 1991 kommt es aufgrund der Wiedervereinigung zu einer deutlichen Häufung der fehlenden Angaben.

5.5.8 Erwerbsstatus (erwstat)

Variablenlabel	Erwerbsstatus
Variablenname	erwstat
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal nimmt für jede Datenquelle unterschiedliche Werte mit unterschiedlichen Bedeutungen an. 1) BeH Für BeH-Sätze entspricht das Merkmal „Erwerbsstatus“ (<i>erwstat</i>) der im neuen Meldeverfahren (DEÜV) ab dem 01.01.1999 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 es sich dabei um sozialversicherungspflichtige Beschäftigungsverhältnisse ohne Besonderheiten, die unter der 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. 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 01.04.1999 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“ (<i>erwstat</i>) 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.
--	---

5.5.9 Gleitzone (gleitz)

Variablenlabel	Gleitzone
Variablenname	gleitz
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal ist erst ab 2003 und nur für BeH-Datensätze gefüllt. Es zeigt an, ob es sich bei der Beschäftigungsmeldung um eine Beschäftigung im Niedriglohnbereich innerhalb der sogenannten Gleitzone mit einem monatlichen Bruttoarbeitsentgelt von 400,01 bis 800,00 € handelt (sog. Midi-Job), für die die die/der Arbeitnehmer nur einen reduzierten Beitrag zum Gesamtsozialversicherungsbeitrag leisten muss. Da ArbeitnehmerInnen 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.

5.5.10 Leiharbeit (leih)

Variablenlabel	Leiharbeit
Variablenname	leih
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt an, ob es sich bei der Beschäftigung der Person um ein Leiharbeitsverhältnis handelt. Das Merkmal entstammt dem Tätigkeitsschlüssel 2010 und ist nur für Meldungen mit einem Enddatum nach dem 30.11.2011 gefüllt.
Qualitätshinweis	Im Jahr 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Im Jahr 2012 beträgt die Missingquote nur noch rund 3%.

5.5.11 Befristung (befrist)

Variablenlabel	Befristung
Variablenname	befrist
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH

Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt an, ob es sich bei der Beschäftigung der Person um ein befristetes oder unbefristetes Arbeitsverhältnis handelt. Das Merkmal entstammt dem Tätigkeitsschlüssel 2010 und ist nur für Meldungen mit einem Enddatum nach dem 30.11.2011 gefüllt.
Qualitätshinweis	Im Jahr 2011 kommt es aufgrund der Meldeumstellung zu einer deutlichen Häufung der fehlenden Angaben. Im Jahr 2012 beträgt die Missingquote nur noch rund 3%.

5.5.12 Abmeldegrund / Abgabegrund / Beendigungsgrund (grund)

Variablenlabel	Abmeldegrund / Abgabegrund / Beendigungsgrund
Variablenname	grund
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	BeH, LeH
Datentyp	Numerisch
Detailbeschreibung	1) BeH 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 möglichen 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 01.01.1999 geltenden Meldeverfahrens (gemäß DEÜV) kodiert. 2) LeH In LeH-Meldungen gibt das Merkmal den Grund für das Ende des Bezugs von Arbeitslosengeld, -hilfe oder Unterhaltsgeld an. Über die Gründe für den Beginn des Leistungsbezugs finden sich in der LeH keine Informationen, da die LeH mit den Meldungen der Arbeitsagenturen an die Krankenkassen über abgeschlossene Leistungsbezugsdauern gefüllt wird.
Qualitätshinweis	In der BeH nimmt ab 2013 die Anzahl der Meldungen mit Abmeldegrund 54 (Meldung eines einmalig gezahlten Entgelts) sprunghaft zu. Solange ein Beschäftigungsverhältnis besteht, können Sonderzahlungen, die bis März des Folgejahres ausgezahlt werden, mit der normalen Meldung (v.a. Jahresmeldung) für das Vorjahr angegeben werden. Eine gesonderte Meldung mit Abmeldegrund 54 ist dann nicht erforderlich. Bis 2012 konnten die Jahresmeldungen bis Mitte April abgegeben werden, seit 2013 müssen sie nun bis spätestens Mitte Februar erfolgen. Für Sonderzahlungen, die im Februar und März ausgezahlt werden, muss nun eine extra Meldung erfolgen.

5.5.13 Restanspruch/geplante Dauer (restanspruch)

Variablenlabel	Restanspruch/geplante Dauer
Variablenname	restanspruch
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	LeH
Datentyp	Numerisch

Detailbeschreibung	In der LeH enthält das Merkmal den Anspruch, der nach Ende des aktuellen Leistungsbezugs verbleibt. Endet der Leistungsbezug vor dem Ende der maximalen Anspruchsdauer (z. B. durch Arbeitsaufnahme), bleibt ein Restanspruch mit der Dauer des nicht ausgeschöpften Anspruchs bestehen. Entsteht innerhalb von fünf Jahren ein neuer Anspruch, wird die Dauer des Restanspruchs zur neuen Anspruchsdauer addiert. Allerdings ist hierbei die Höchstdauer für das jeweilige Lebensalter die Obergrenze. Entsteht kein neuer Anspruch, so kann der Restanspruch auf Antrag innerhalb von vier Jahren für Leistungen verwendet werden. Liegt das Enddatum des Leistungsempfangs vor dem 01.01.1998, wird der Restanspruch in Werktagen angegeben. Danach erfolgt die Angabe in Kalendertagen. Die Angabe bezieht sich nicht auf das Enddatum der Episode, sondern auf das Ende des Originalzeitraums.
Qualitätshinweis	In der LeH ist im Verlauf einer Leistungsbiografie grundsätzlich mit einer systematischen Entwicklung des Restanspruchs zu rechnen. Das bedeutet, dass der Restanspruch am Ende einer Leistungsperiode minus der Dauer der Folgeperiode den Restanspruch nach der Folgeperiode ergeben sollte, wenn in der Zwischenzeit keine Ansprüche entstanden oder verfallen sind. Es kommt jedoch teilweise zu unerwarteten Erhöhungen oder Reduzierungen des Restanspruchs. Hierbei handelt es sich vermutlich überwiegend um Korrekturen bei der Bearbeitung von Leistungsfällen. Solche Korrekturen treten zum Beispiel auf, wenn ein Anspruch zu Beginn des Leistungsfalls aufgrund unvollständiger Informationen nicht korrekt festgelegt oder ein noch nicht verfallener Restanspruch nicht berücksichtigt wurde. Diese Korrekturen erfolgen dann aber in der Regel nur in der aktuellen Buchung, eine Bereinigung des kompletten Falls findet nicht statt. Darüber hinaus kann die Angabe zur Anspruchsdauer in Einzelfällen die aus den gesetzlichen Vorgaben abzuleitende individuelle Obergrenze übersteigen.

5.5.14 Beginndatum der Arbeitslosigkeit (alo_beg)

Variablenlabel	Beginndatum der Arbeitslosigkeit
Variablenname	alo_beg
Kategorie	Informationen zu Beschäftigung und Leistungsbezug
Herkunft	LeH, LHG, ASU, XASU, MTH, XMTH
Datentyp	Numerisch
Detailbeschreibung	Das Merkmal gibt das Beginndatum einer ununterbrochenen Folge von Arbeitslosigkeitsperioden an und ist zu Beginn des Satzes gültig. Folgende Lücken führen nicht zu einer Unterbrechung der Arbeitslosigkeit: • jede Lücke kleiner oder gleich 7 Tage • Krankheitslücken bis zu 42 Tagen Das Merkmal wird in den IEB unter Einbezug mehrerer Quellen generiert und kann somit nicht exakt anhand der Personendaten des LIAB nachgebildet werden.

5.5.15 Dauer der Arbeitslosigkeit (alo_dau)

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

Detailbeschreibung	Das Merkmal gibt die Dauer (in Tagen) einer ununterbrochenen Folge von Arbeitslosigkeitsperioden an und ist zu Beginn des Satzes gültig. Folgende Lücken führen nicht zu einer Unterbrechung der Arbeitslosigkeit: • jede Lücke kleiner oder gleich 7 Tage • Krankheitslücken bis zu 42 Tagen Bei der Berechnung der Dauer werden diese Lücken jedoch nicht addiert. Daher kann sich die Dauer in diesem Merkmal von dem Ergebnis der Dauerberechnung mit dem "Beginndatum der Arbeitslosigkeit" (<i>alo_beg</i>) unterscheiden. Vor 1997 bedeutet der Wert 0 nicht, dass keine Arbeitslosigkeit vorliegt, da hier die Quellen ASU/XASU nicht gefüllt sind. Das Merkmal wird in den IEB unter Einbezug mehrerer Quellen generiert und kann somit nicht exakt anhand der Personendaten des LIAB nachgebildet werden.
--------------------	---

5.6 Ortsangaben

5.6.1 Wohnort - Kreis (wo_kreis)

Variablenlabel	Wohnort - Kreis
Variablenname	wo_kreis
Kategorie	Ortsangaben
Herkunft	BeH, LeH
Datentyp	Numerisch
Hierarchie	Bundesland Kreis
Detailbeschreibung	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 die/der Versicherte wohnt. Der 5-stellige Kreisschlüssel enthält in den ersten beiden Stellen den Schlüssel des Bundeslandes (NUTS-1) und in der ersten bis dritten Stelle den Regierungsbezirk (NUTS-2). Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null. In der Systematik NUTS (Nomenclature des unités territoriales statistiques) der Europäischen Union entsprechen die Kreise der Hierarchieebene NUTS 3. In der BeH wird der Wohnort jeweils zum Jahresende ermittelt und dann einheitlich für alle Sätze eines Jahres zugespielt. In der LeH enthält das Merkmal den Wohnort zu Beginn des Originalzeitraums. Mit der Dauer des Spells steigt also das Risiko, dass der Wohnort veraltet ist und die Angaben zu späteren Stichtagen falsch sind. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Kreisangaben für alle Quellen auf den Gebietsstand vom 31.12.2017 umgeschlüsselt, d. h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2017 hatten. Da sich die Grenzen der Kreise im Zeitverlauf geändert haben, würden ohne Gebietsstandbereinigung Fälle auftreten, bei denen die Kreiskennziffer des Wohnortes wechselt, ohne dass der/die betreffende Versicherte umgezogen ist.
Qualitätshinweis	Hinsichtlich der Wohnortangaben bestehen für einzelne Beschäftigte Erhebungsungenauigkeiten. Die Meldevorschrift stellt nicht klar, welcher Wohnsitz – Haupt- oder Nebenwohnsitz mit überwiegendem Aufenthaltsort – vom Arbeitgeber zu melden

	ist. Dies kann in der Beschäftigungsstatistik zum Nachweis von "Fernpendlern" zwischen gemeldetem Hauptwohnsitz und Arbeitsort führen, obwohl die beschäftigte Person ihrer Beschäftigung am Nebenwohnsitz nachgeht, also faktisch nicht pendelt. Im Jahr 2015 kommt es in den Daten zu einer Reduktion der Wohnortwechsel von ca. 10% bis 15% gegenüber dem üblichen Niveau der umliegenden Jahre, wobei regionale Unterschiede auftreten. Der genaue Grund für diese Abweichung konnte bisher nicht ermittelt werden.
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 Merkmal „Wohnort - Bundesland“ (<i>wo_bula</i>) bereitgestellt.

5.6.2 Wohnort - Bundesland (*wo_bula*)

Variablenlabel	Wohnort - Bundesland
Variablenname	<i>wo_bula</i>
Kategorie	Ortsangaben
Herkunft	BeH, LeH
Datentyp	Numerisch
Hierarchie	Bundesland Kreis
Detailbeschreibung	Dieses Merkmal ist eine Aggregation aus dem Merkmal „Wohnort – Kreis“ (<i>wo_kreis</i>) auf die 16 deutschen Bundesländer. In BeH- und LeH-Sätzen liegt das Bundesland des Wohnortes erst für die Jahre ab 1999 vor. Weitere Informationen zum Kreis des Wohnortes finden sich in Abschnitt 5.6.1.

5.6.3 Wohnort - Arbeitsagentur (*wo_aa*)

Variablenlabel	Wohnort - Arbeitsagentur
Variablenname	<i>wo_aa</i>
Kategorie	Ortsangaben
Herkunft	BeH, LeH
Datentyp	Numerisch
Hierarchie	Regionaldirektion Arbeitsagentur
Detailbeschreibung	Das Merkmal enthält ab 1999 den Agenturbezirk der am Wohnort des Beschäftigten/Kunden zuständigen Arbeitsagentur. Sie wird aus der Wohnadresse ermittelt. In der LeH enthält das Merkmal den Wohnort zu Beginn der Meldung. Mit der Dauer des Spells steigt also das Risiko, dass der Wohnort veraltet ist und die Angaben zu späteren Stichtagen falsch sind. Es ist für die BeH und LeH erst ab 1999 vorhanden. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Angaben für alle Quellen auf den Gebietsstand vom 31.12.2017 umgeschlüsselt, d. h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einem Agenturbezirk nach den Grenzen, welche die Bezirke zum 31.12.2017 hatten. Da sich die Grenzen im Zeitverlauf geändert haben, würden ohne diese Gebietsstandbereinigung Fälle auftreten, bei denen der Bezirk einer Person wechselt, ohne dass diese Person umgezogen ist. Einen Problemfall stellt hinsichtlich der Gebietsstandbereinigung 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. Berlin sollte möglichst als Einheit und nicht getrennt nach Ost und West analysiert werden, da seit der Gebietsreform der Arbeitsämter in Berlin zum 01.07.1997 deren

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

5.6.4 Wohnort - Regionaldirektion (wo_rd)

Variablenlabel	Wohnort - Regionaldirektion
Variablenname	wo_rd
Kategorie	Ortsangaben
Herkunft	BeH, LeH
Datentyp	Numerisch
Hierarchie	Regionaldirektion Arbeitsagentur
Detailbeschreibung	Dieses Merkmal ist eine Aggregation aus dem Merkmal „Wohnort – Arbeitsagentur“ (wo_aa) auf die Ebene der Regionaldirektionen. Weitere Informationen finden sich in Abschnitt 5.6.3.

5.7 Betriebsmerkmale

5.7.1 WZ73 3-Steller (w73_3)

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

5.7.2 WZ93 5-Steller (w93_5)

Variablenlabel	WZ93 5-Steller
Variablenname	w93_5
Kategorie	Betriebsmerkmale
Herkunft	BHP

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 5-Steller aus und ist für die Jahre 1999 bis inklusive 2003 gefüllt. WZ93 steht für die „Klassifikation der Wirtschaftszweige, Ausgabe 1993“ (Statistisches Bundesamt, 2002). Die WZ93 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1 („Nomenclature générale des activités économiques dans les communautés européennes“) auf, die vier Gliederungsebenen vorsieht und bezüglich der ersten beiden wiederum auf dem internationalen Standard ISIC Rev.3 („International Standard Industrial Classification of All Economic Activities“) basiert. 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 (w93_3) bereitgestellt.

5.7.3 WZ93 3-Steller (w93_3)

Variablenlabel	WZ93 3-Steller
Variablenname	w93_3
Kategorie	Betriebsmerkmale
Herkunft	BHP
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 und ist für die Jahre 1999 bis inklusive 2003 gefüllt. WZ93 steht für die „Klassifikation der Wirtschaftszweige, Ausgabe 1993“ (Statistisches Bundesamt, 2002). Die WZ93 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1 („Nomenclature générale des activités économiques dans les communautés européennes“) auf, die vier Gliederungsebenen vorsieht und bezüglich der ersten beiden wiederum auf dem internationalen Standard ISIC Rev.3 („International Standard Industrial Classification of All Economic Activities“) basiert. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widerspiegelt werden.

5.7.4 WZ03 5-Steller (w03_5)

Variablenlabel	WZ03 5-Steller
Variablenname	w03_5
Kategorie	Betriebsmerkmale
Herkunft	BHP
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" (Statistisches Bundesamt, 2003). Die WZ03 baut ebenso wie die WZ93 auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.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.

5.7.5 WZ03 3-Steller (w03_3)

Variablenlabel	WZ03 3-Steller
Variablenname	w03_3
Kategorie	Betriebsmerkmale
Herkunft	BHP
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" (Statistisches Bundesamt, 2003). Die WZ03 baut ebenso wie die WZ93 auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.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.

5.7.6 WZ08 5-Steller (w08_5)

Variablenlabel	WZ08 5-Steller
Variablenname	w08_5
Kategorie	Betriebsmerkmale
Herkunft	BHP
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 WZ08 als 5-Steller aus und ist für die Jahre ab 2008 gefüllt. WZ08 steht für die "Klassifikation der Wirtschaftszweige, Ausgabe 2008" (Statistisches Bundesamt, 2008). Die WZ08 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.2 auf. 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 (<i>w08_3</i>) bereitgestellt.
----------------	--

5.7.7 WZ08 3-Steller (*w08_3*)

Variablenlabel	WZ08 3-Steller
Variablenname	<i>w08_3</i>
Kategorie	Betriebsmerkmale
Herkunft	BHP
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 WZ08 als 3-Steller aus und ist für die Jahre ab 2008 gefüllt. WZ08 steht für die "Klassifikation der Wirtschaftszweige, Ausgabe 2008" (Statistisches Bundesamt, 2008). Die WZ08 baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.2 auf. Jedem Betrieb ist nur eine Kennziffer zugeordnet. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, soll die wirtschaftliche Haupttätigkeit widerspiegelt werden.

5.7.8 *w73_3* vervollständigt durch Extrapolation/Imputation (*w73_3_gen*)

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

5.7.9 Art der Vervollständigung *w73_3* (*group_w73_3*)

Variablenlabel	Art der Vervollständigung <i>w73_3</i>
Variablenname	<i>group_w73_3</i>
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal weist die Art der Vervollständigung für das Merkmal <i>w73_3_gen</i> aus. Es ist angegeben, ob der jeweilige Wert in <i>w73_3_gen</i> dem Originalwert von <i>w73_3</i> entspricht, weiterhin fehlend ist bzw. extrapoliert oder auf Basis von Umschlüsselungstabellen imputiert wurde.

	Eine ausführliche Beschreibung zur Vorgehensweise findet sich bei Eberle et al., 2011.
--	--

5.7.10 w93_3 vervollständigt durch Extrapolation/Imputation (w93_3_gen)

Variablenlabel	w93_3 vervollständigt durch Extrapolation/Imputation
Variablenname	w93_3_gen
Kategorie	Betriebsmerkmale
Herkunft	BHP
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- oder fortgeschrieben oder mit Hilfe von Umschlüsselungstabellen ersetzt, sodass das Merkmal eine zeitkonsistente Wirtschaftszweigangabe auf Basis der Wirtschaftszweigklassifikation WZ93 enthält. Eine ausführliche Beschreibung dazu findet sich bei Eberle et al., 2011. Weitere Informationen zur Klassifikation WZ93 finden sich in der Variablenbeschreibung zu w93_3.

5.7.11 Art der Vervollständigung w93_3 (group_w93_3)

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

5.7.12 w08_3 vervollständigt durch Extrapolation/Imputation (w08_3_gen)

Variablenlabel	w08_3 vervollständigt durch Extrapolation/Imputation
Variablenname	w08_3_gen
Kategorie	Betriebsmerkmale
Herkunft	BHP
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 WZ08 als 3-Steller aus. Von 2008 bis 2017 enthält das Merkmal die Originalwerte aus w08_3. Vor 2008 wird die Information rückgeschrieben oder mit Hilfe von Umschlüsselungstabellen ersetzt, sodass das Merkmal eine zeitkonsistente Wirtschaftszweigangabe auf Basis der Wirtschaftszweigklassifikation WZ08 enthält. Eine ausführliche Beschreibung dazu findet sich bei Eberle et al., 2011.

	Weitere Informationen zur Klassifikation WZ08 finden sich in der Variablenbeschreibung zu <i>w08_3</i> .
--	--

5.7.13 Art der Vervollständigung *w08_3* (*group_w08_3*)

Variablenlabel	Art der Vervollständigung <i>w08_3</i>
Variablenname	<i>group_w08_3</i>
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Dieses Merkmal weist die Art der Vervollständigung für das Merkmal <i>w08_3_gen</i> aus. Es ist angegeben, ob der jeweilige Wert in <i>w08_3_gen</i> dem Originalwert von <i>w08_3</i> entspricht, weiterhin fehlend ist bzw. extrapoliert oder auf Basis von Umschlüsselungstabellen imputiert wurde. Eine ausführliche Beschreibung zur Vorgehensweise findet sich bei Eberle et al., 2011.

5.7.14 Jahr des 1. Auftretens der Betriebsnummer (*grd_jahr*)

Variablenlabel	Jahr des 1. Auftretens der Betriebsnummer
Variablenname	<i>grd_jahr</i>
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Bei dieser Variablen handelt es sich um das erste Auftreten der Nummer des Betriebes im Datensatz. Wird eine Betriebsnummer in Westdeutschland erst nach 1975 bzw. in Ostdeutschland nach 1992 erstmalig nachgewiesen, könnte es sich um das Gründungsdatum des betreffenden Betriebes handeln. Allerdings könnte es sich auch um einen bereits länger existierenden Betrieb handeln, dem nach einem Wechsel des Inhabers oder einer Änderung der Rechtsform eine neue Betriebsnummer zugeordnet wurde (zur Betriebsnummernvergabe siehe Bundesagentur für Arbeit, 2007, S. 9-11). Ebenfalls könnte es sein, dass der Betrieb bereits vorher existierte, aber keine sozialversicherungspflichtig bzw. ab 1999 auch keine geringfügig Beschäftigten hatte. Ein Betrieb muss nicht zwingend in dem Jahr des ersten Auftretens im BHP enthalten sein, da für das BHP nur der Stichtag 30.06. relevant ist. Hat der Betrieb in seinem Gründungsjahr an diesem Stichtag keine Beschäftigten, erscheint er auch nicht im BHP.

5.7.15 Erstes Auftreten Betriebsnummer (*grd_dat*)

Variablenlabel	Erstes Auftreten Betriebsnummer
Variablenname	<i>grd_dat</i>
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Datum
Detailbeschreibung	Bei dieser Variablen handelt es sich um das tagesgenaue erste Auftreten der Nummer des Betriebes in der BeH. Wird eine Betriebsnummer in Westdeutschland erst nach 1975 bzw. in Ostdeutschland nach 1992 erstmalig nachgewiesen, könnte es sich um das Gründungsdatum des betreffenden Betriebes handeln. Allerdings könnte es sich auch um einen bereits länger existierenden Betrieb handeln, dem nach einem Wechsel des Inhabers oder einer Änderung der Rechtsform eine neue

	Betriebsnummer zugeordnet wurde (zur Betriebsnummernvergabe siehe Bundesagentur für Arbeit, 2007, S. 9-11). Ebenfalls könnte es sein, dass der Betrieb bereits vorher existierte, aber keine sozialversicherungspflichtig bzw. ab 1999 auch keine geringfügig Beschäftigten hatte. Ein Betrieb muss nicht zwingend in dem Jahr des ersten Auftretens im BHP enthalten sein, da für das BHP nur der Stichtag 30.06. relevant ist. Hat der Betrieb in seinem Gründungsjahr an diesem Stichtag keine Beschäftigte, erscheint er auch nicht im BHP.
Besonderheiten	Dieses Merkmal wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen tagesgenau bereitgestellt. Ansonsten wird lediglich das Jahr des ersten Auftretens der Betriebsnummer (<i>grd_jahr</i>) bereitgestellt.

5.7.16 Jahr des letzten Auftretens der Betriebsnummer (lzt_jahr)

Variablenlabel	Jahr des letzten Auftretens der Betriebsnummer
Variablenname	lzt_jahr
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Bei dieser Variablen handelt es sich um das letztmalige Auftreten der Nummer des Betriebes im Datensatz (vgl. Bender et al., 1996). Endet die Existenz einer Betriebsnummer im BHP bereits vor dem Jahr 2017, 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 neuen Nummer“ oder andere verwaltungstechnische Änderungen in Betracht (vgl. Bender et al., 1996 oder Bundesagentur für Arbeit, 2007, S. 9-11). Ein Betrieb muss nicht zwingend in dem Jahr des letzten Auftretens im BHP enthalten sein, da für das BHP nur der Stichtag 30.06. relevant ist. Hat der Betrieb in seinem Schließungsjahr an diesem Stichtag keine Beschäftigte, erscheint er auch nicht im BHP.

5.7.17 Letztes Auftreten Betriebsnummer (lzt_dat)

Variablenlabel	Letztes Auftreten Betriebsnummer
Variablenname	lzt_dat
Kategorie	Betriebsmerkmale
Herkunft	BHP
Datentyp	Datum
Detailbeschreibung	Bei dieser Variablen handelt es sich um das tagesgenaue letztmalige Auftreten der Nummer des Betriebes im Datensatz (vgl. Bender et al., 1996). Endet die Existenz einer Betriebsnummer im BHP bereits vor dem Jahr 2017, 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 neuen Nummer“ oder andere verwaltungstechnische Änderungen in Betracht (vgl. Bender et al., 1996 oder Bundesagentur für Arbeit, 2007, S. 9-11). Ein Betrieb muss nicht zwingend in dem Jahr des letzten Auftretens im BHP enthalten sein, da für das BHP nur der Stichtag 30.06. relevant ist. Hat der Betrieb in seinem Schließungsjahr an diesem Stichtag keine Beschäftigte, erscheint er auch nicht im BHP.

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 (<i>lzt_jahr</i>) bereitgestellt.
----------------	---

5.7.18 Anzahl Beschäftigte gesamt (az_ges)

Variablenlabel	Anzahl Beschäftigte gesamt
Variablenname	az_ges
Kategorie	Generierte Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Die Variable enthält die Anzahl aller zum 30.06. eines Jahres zur Sozialversicherung gemeldeten Mitarbeitenden eines Betriebes. Seit Einführung der neuen Meldevorschriften im Jahr 1999 sind auch die geringfügig Beschäftigten erfasst. Ruhende Beschäftigungsverhältnisse (Tagesentgelt von Null) sind nicht enthalten.

5.7.19 Anzahl in Vollzeit (Normalbeschäftigte + sonstige) (az_vz)

Variablenlabel	Anzahl Vollzeit (Normalbeschäftigte + sonstige)
Variablenname	az_vz
Kategorie	Generierte Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch
Detailbeschreibung	Die Variable enthält die Anzahl der Personen im Betrieb, die am 30.06. eines Jahres mit den Personengruppenschlüsseln 101, 140, 143, 105, 106, 112, 113, 114, 118, 119, 120, 149, 201, 203, 205, 999 und YYY und als Vollzeitbeschäftigte gemeldet sind. Auszubildende, geringfügig Beschäftigte und Personen in Altersteilzeit werden nicht berücksichtigt.

5.7.20 Anzahl geringfügig Beschäftigte (az_gf)

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

5.7.21 Mittelwert imp. Bruttotagesentgelt Vollzeitbeschäftigte (te_imp_mw)

Variablenlabel	Mittelwert imp. Bruttotagesentgelt Vollzeitbeschäftigte
Variablenname	te_imp_mw
Kategorie	Generierte Betriebsmerkmale
Herkunft	BHP
Datentyp	Numerisch

Detailbeschreibung	Diese Variable enthält den Mittelwert des imputierten Bruttotagesentgeltes der Vollzeitbeschäftigten eines Betriebes. Nicht enthalten sind die Löhne der geringfügig Beschäftigten, der Auszubildenden sowie der Personen in Altersteilzeit. Die Werte sind für alle Jahre in Euro angegeben. Gemäß den Meldevorschriften zur Sozialversicherung müssen Arbeitgeber angeben, wie hoch das beitragspflichtige Bruttoarbeitsentgelt der/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. Zur Berechnung der Mittelwerte wurden diese zensierten Löhne imputiert (vgl. Abschnitt 3.1.3.3 in Schmucker et al. 2018). Diese Daten wurden dann auf Betriebsebene aggregiert. Die Werte sind auf zwei Dezimalstellen gerundet. Aufgrund des "storage types" in Stata werden jedoch weitere Dezimalstellen angezeigt, die nicht korrekt sind.
--------------------	---

5.7.22 Arbeitsort - Kreis (ao_kreis)

Variablenlabel	Arbeitsort - Kreis
Variablenname	ao_kreis
Kategorie	Ortsangaben
Herkunft	BHP
Datentyp	Numerisch
Hierarchie	Bundesland Kreis
Detailbeschreibung	Angegeben ist jeweils der Kreis (kreisfreie Stadt oder Landkreis), in dem der Betrieb der beschäftigten Person angesiedelt ist. Der 5-stellige Kreisschlüssel enthält in den ersten beiden Stellen den Schlüssel des Bundeslandes (NUTS-1), in der ersten bis dritten Stelle den Regierungsbezirk (NUTS-2). Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null. In der Systematik NUTS (Nomenclature des unités territoriales statistiques) der Europäischen Union entsprechen die Kreise der Hierarchieebene NUTS 3. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Kreisangaben auf den Gebietsstand vom 31.12.2017 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Betriebsortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2017 hatten. Da sich die Grenzen der Kreise im Zeitverlauf geändert haben, würden ohne Gebietsstandbereinigung Fälle auftreten, bei denen die Kreiskennziffer des Betriebsortes wechselt, ohne dass dieser Betrieb seinen Standort verlagert hat.
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 (<i>ao_bula</i>) bereitgestellt.

5.7.23 Arbeitsort - Bundesland (ao_bula)

Variablenlabel	Arbeitsort - Bundesland
Variablenname	ao_bula
Kategorie	Ortsangaben
Herkunft	BHP

Datentyp	Numerisch
Hierarchie	Bundesland Kreis
Detailbeschreibung	Angegeben ist jeweils das Bundesland, in dem der Betrieb angesiedelt ist. Dieses Merkmal wird aus der Kreiskennziffer (<i>ao_kreis</i>) generiert. Die ersten beiden Stellen der Kreiskennziffer bezeichnen das Bundesland (NUTS-1).

6 Literatur

- Bender, Stefan / Hilzendegen, Jürgen / Rohwer, Götz / Rudolph, Helmut (1996): Die IAB-Beschäftigtenstichprobe 1975-1990. Beiträge zur Arbeitsmarkt- und Berufsforschung 197, Nürnberg.
- Bertat, Thomas / Dundler, Agnes / Grimm, Christopher / Kiewitt, Jochen / Schomaker, Christine / Schridde, Dr. Henning / Zemann, Dr. Christian (2013): Neue Erhebungsinhalte 'Arbeitszeit', 'ausgeübte Tätigkeit' sowie 'Schul- und Berufsabschluss' in der Beschäftigungsstatistik. Methodenbericht, Bundesagentur für Arbeit – Statistik, URL: <https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Methodenberichte/Beschaeftigungsstatistik/Generische-Publikationen/Methodenbericht-Neue-Erhebungsinhalte-Arbeitszeit-ausgeuebte-Taetigkeit-sowie-Schul-und-Berufsabschluss-in-der-Beschaeftigungsstatistik.pdf> (Zugriff: 12.06.2019).
- Bundesagentur für Arbeit (Hrsg.) (2005): Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Ausgabe Januar 2005, Nürnberg.
- Bundesagentur für Arbeit (Hrsg.) (2007): Handbuch für die Betriebsnummernvergabe und –pflege im Rahmen des Meldeverfahren zur Sozialversicherung. Ausgabe Dezember 2007, Nürnberg.
- Bundesagentur für Arbeit (2009): Klassifikation der Wirtschaftszweige 1973, Nürnberg. URL: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Wirtschaftszweige/Klassifikation-der-Wirtschaftszweige-1973-2003/Klassifikationen-der-Wirtschaftszweige-1973-2003-Nav.html> (Zugriff: 12.06.2019).
- Bundesagentur für Arbeit (2011): Klassifikation der Berufe 2010. Band 1: Systematischer und alphabetischer Teil mit Erläuterungen, Nürnberg. URL: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB2010/Printausgabe-KldB2010/Printausgabe-KldB-2010-Nav.html> (Zugriff: 12.06.2019).
- Bundesanstalt für Arbeit (1988): Klassifikation der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennung, Nürnberg. URL: <http://statistik.arbeitsagentur.de/Navigation/Statistik/Grundlagen/Klassifikation-der-Berufe/KldB1975-1992/KldB1975-1992-Nav.html> (Zugriff: 12.06.2019).
- Cramer, Ulrich (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv 69: S. 56-68.
- Drews, Nils (2006): Qualitätsverbesserung der Bildungsvariable in der IAB-Beschäftigtenstichprobe 1975-2001. FDZ Methodenreport, 05/2006 (de), Nürnberg.
- Eberle, Johanna/ Jacobebbinghaus, Peter/ Ludsteck, Johannes/ Witter, Julia (2011): Generation of time-consistent industry codes in the face of classification changes * Simple heuristic based on the Establishment History Panel (BHP). FDZ Methodenreport, 05/2011 (en), Nürnberg.
- Eberle, Johanna/ Schmucker, Alexandra (2019): Creating cross-sectional data and biographical variables with the Sample of Integrated Labour Market Biographies 1975-2017 * Programming examples for Stata. FDZ Methodenreport, 04/2019 (en), Nürnberg.

- Fischer, Gabriele/ Janik, Florian/ Müller, Dana/ Schmucker, Alexandra (2008): Das IAB-Betriebspanel - von der Stichprobe über die Erhebung bis zur Hochrechnung. FDZ Methodenreport, 01/2008 (de), 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, Jg. 126, H. 3, S. 405-436.
- Ludsteck, Johannes/ Thomsen, Ulrich (2016): Imputation of the Working Time Information for the Employment Register Data. FDZ Methodenreport 01/2016 (en), Nürnberg.
- Meinken, Holger/ Koch, Iris (2004): BA-Beschäftigtenpanel 1998-2002. Codebuch, Nürnberg.
- Paulus, Wiebke/ Matthes, Britta (2013): Klassifikation der Berufe * Struktur, Codierung und Umsteigeschlüssel. FDZ-Methodenreport, 08/2013 (de), Nürnberg.
- Schmucker, Alexandra/Eberle, Johanna/ Ganzer, Andreas/Stegmaier, Jens/ Umkehrer, Matthias (2018): Betriebs-Historik-Panel 1975-2016. FDZ-Datenreport, 01/2018 (de), Nürnberg.
- Schmucker, Alexandra/ Seth, Stefan/ Ludsteck, Johannes/ Eberle, Johanna/ Ganzer, Andreas (2016): Betriebs-Historik-Panel 1975-2014. FDZ Datenreport, 03/2016 (de), Nürnberg.
- Statistisches Bundesamt (2002): Klassifikation der Wirtschaftszweige, Ausgabe 1993 (WZ 93), Wiesbaden. URL: <https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/Downloads/klassifikation-wz-1993.html> (Zugriff: 12.06.2019).
- Statistisches Bundesamt (2003): Klassifikation der Wirtschaftszweige, Ausgabe 2003 (WZ 2003), Wiesbaden. URL: <https://www.destatis.de/DE/Methoden/Klassifikationen/GueterWirtschaftsklassifikationen/Content75/KlassifikationWZ2003.html> (Zugriff: 12.06.2019).
- Statistisches Bundesamt (2008): Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), Wiesbaden. URL: <https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/klassifikation-wz-2008.html> (Zugriff: 12.06.2019).
- Statistisches Bundesamt (2019): Staats- und Gebietssystematik, Wiesbaden. URL: https://www.destatis.de/DE/Methoden/Klassifikationen/Staat-Gebietsystematik/Staatsangehoerigkeitsgebietsschluesel_pdf.pdf?__blob=publicationFile&v=5 (Zugriff: 12.06.2019).
- Thomsen, Ulrich/ Ludsteck, Johannes/ Schmucker, Alexandra (2018): Skilled or unskilled - Improving the information on qualification for employee data in the IAB Employee Biography. FDZ-Methodenreport, 09/2018 (en), 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 4/88, S. 468-482.

7 Anhang

7.1 Auszählungen

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

7.2 Abkürzungsverzeichnis

AA	Agentur für Arbeit / Arbeitsamt
ALG	Arbeitslosengeld
ALH	Arbeitslosenhilfe
ARGE	Arbeitsgemeinschaft
ASU	Arbeitsuchendenhistorik
A2LL	Arbeitslosengeld II – Leistungen zum Lebensunterhalt
BA	Bundesagentur für Arbeit
BeH	Beschäftigtenhistorik
BfA	Bundesversicherungsanstalt für Angestellte
BG	Bedarfsgemeinschaft
BHP	Betriebs-Historik-Panel
BMAS	Bundesministerium für Arbeit und Soziales
coArb	Computerunterstützte Arbeitsvermittlung (operatives Verfahren zur Verwaltung der Vermittlung (Altverfahren))
DEÜV	Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und -übermittlungsverordnung –
DEVO	Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung –
DIM	Daten- und IT-Management
DÜVO	Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung
EDV	Elektronische Datenverarbeitung
FDZ	Forschungsdatenzentrum der Bundesagentur für Arbeit am IAB
FELEG	Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit
gAw	Träger mit getrennter Aufgabenwahrnehmung
gE	Gemeinsame Einrichtungen
gT	Getrennte Trägerschaft
IAB	Institut für Arbeitsmarkt- und Berufsforschung
IEB	Integrierte Erwerbsbiografien

ISIC	International Standard Industrial Classification of All Economic Activities
KldB	Klassifikation der Berufe
LeH	Leistungsempfängerhistorik
LIAB	Linked-Employer-Employee-Daten des IAB
LHG	Leistungshistorik Grundsicherung
LIAB	Linked-Employer-Employee-Daten des IAB
LM	Längsschnittmodell
LVA	Landesversicherungsanstalt
MTH	Maßnahmeteilnahmehistorik
NACE	Nomenclature générale des activités économiques dans les communautés européennes
NUTS	Nomenclature des unités territoriales statistiques
QM	Querschnittmodell
SGB	Sozialgesetzbuch
UHG	Unterhaltsgeld
VerBIS	Vermittlungs- und Beratungsinformationssystem
XASU	Arbeitsuchendenhistorik aus XSozial-BA-SGB II
XLHG	Leistungshistorik Grundsicherung aus XSozial-BA-SGB II
zkT	Zugelassene kommunale Träger

Impressum

FDZ-Datenreport 05|2019

Veröffentlichungsdatum

01. Oktober 2019

Herausgeber

Forschungsdatenzentrum
der Bundesagentur für Arbeit
im Institut für Arbeitsmarkt- und Berufsforschung
Regensburger Straße 104
90478 Nürnberg

Rechte

Nachdruck - auch auszugsweise - nur mit Genehmigung des FDZ gestattet

Bezugsmöglichkeit dieses Dokuments

http://doku.iab.de/fdz/reporte/2019/DR_05-19.pdf

Dokumentation Version

LIAB_LM_7517_DE_v1_dok1, DOI: 10.5164/IAB.FDZD.1905.de.v1

Datensatz Version

LIAB LM 7517 v1, DOI: 10.5164/IAB.LIABLM7517.de.en.v1

Bezugsmöglichkeit aller Veröffentlichungen der Reihe „FDZ-Datenreport“

https://fdz.iab.de/de/FDZ_Publications/FDZ_Publication_Series/FDZ-Datenreporte.aspx

Website

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

Rückfragen zum Inhalt

Matthias Umkehrer

Telefon: 0911 179-6211

E-Mail: Matthias.Umkehrer@iab.de