

## LIAB-Datenhandbuch

Version 3.0

*Peter Jacobebbinghaus*

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung .....</b>	<b>5</b>
<b>2</b>	<b>Veränderungen des Datensatzes im Vergleich zu früheren Versionen.....</b>	<b>6</b>
2.1	Änderungen beim Update auf 2005.....	6
2.2	Änderungen beim Update auf 2006.....	9
<b>3</b>	<b>Kurzbeschreibungen .....</b>	<b>11</b>
<b>4</b>	<b>Datenquellen.....</b>	<b>15</b>
4.1	Betriebsdaten .....	15
4.2	Personendaten.....	17
<b>5</b>	<b>Datenaufbereitung.....</b>	<b>19</b>
5.1	Grundsätzliche Verknüpfungsmöglichkeiten.....	19
5.2	LIAB-Daten am FDZ.....	21
5.2.1	LIAB-Querschnittmodell, Version 1.....	22
5.2.2	LIAB-Längsschnittmodell, Version 1.....	22
5.2.3	LIAB-Längsschnittmodell, Version 2.....	23
5.2.4	LIAB-Längsschnittmodell, Version 3.....	23
5.2.5	Vor- und Nachteile von Quer- und Längsschnittmodellen.....	23
5.3	Mengengerüste .....	24
<b>6</b>	<b>Datenqualität: Verknüpfung von Betrieben und Personen .....</b>	<b>26</b>
<b>7</b>	<b>Variablenbeschreibung zu den Personendaten .....</b>	<b>29</b>
7.1	Allgemeiner Teil.....	29
7.1.1	Merkmalsliste .....	29
7.1.2	Sensible Merkmale.....	31
7.2	Beschreibung der einzelnen Variablen.....	31
7.2.1	Identifikatoren.....	31
7.2.1.1	Personen-ID (id).....	31
7.2.1.2	Betriebs-ID nach Betriebspanel (idnum) .....	32
7.2.1.3	Betriebs-ID nach BA-BNR (betnr) .....	32
7.2.2	Technische Merkmale.....	32
7.2.2.1	Spellbeginn original (begorig) .....	32
7.2.2.2	Spellende original (endorig) .....	32
7.2.2.3	Spellbeginn Episode (begepi).....	32
7.2.2.4	Spellende Episode (endepi).....	33
7.2.2.5	Quelle (quelle).....	33
7.2.2.6	Spellzähler der Person (spell_nr).....	33
7.2.2.7	Spellzähler pro Episode und Quelle (level1) .....	33
7.2.2.8	Status der Betriebsnummer (betr_st).....	34
7.2.2.9	Datumsbereinigung (dat_ber).....	34
7.2.3	Merkmale zur Person.....	35
7.2.3.1	Geschlecht (geschl).....	35
7.2.3.2	Geburtsjahr (gebjahr) .....	35
7.2.3.3	Staatsangehörigkeit, gruppiert (nation) .....	35
7.2.3.4	Staatsangehörigkeit, original (nationor).....	35
7.2.3.5	Ausbildung (ausbild).....	36

7.2.3.6	Familienstand (fam).....	36
7.2.3.7	Kinder (kind) .....	37
7.2.4	Beschäftigung und Leistungsbezug.....	37
7.2.4.1	Personengruppe (pers_gr) .....	37
7.2.4.2	Abgabegrund der Meldung (abg_gr).....	38
7.2.4.3	Stellung im Beruf (berstell) .....	40
7.2.4.4	Beruf (beruf).....	41
7.2.4.5	Tagesentgelt/Leistungssatz (tag_entg) .....	41
7.2.4.6	Leistungsart (leistart).....	42
7.2.4.7	Bewilligungsart (bew_art) .....	43
7.2.4.8	Abmeldegrund Leistungsbezug (abm_gr).....	43
7.2.4.9	Wirtschaftszweig WZ73, 3-steller (wz73).....	43
7.2.4.10	Wirtschaftszweig WZ93, 5-steller (wz93).....	44
7.2.4.11	Wirtschaftszweig WZ03, 5-steller (wz03).....	44
7.2.5	Generierte biographische Merkmale .....	44
7.2.5.1	Eintrittsdatum Erwerbsleben (ein_erw).....	44
7.2.5.2	Eintrittsdatum Betrieb (ein_bet) .....	44
7.2.5.3	Dauer Betrieb (tage_bet) .....	45
7.2.5.4	Anzahl bisheriger Leistungsbezüge (anz_ist).....	45
7.2.5.5	Dauer bisheriger Leistungsbezüge (tage_ist).....	45
7.2.5.6	Erwerbsstatus 5 Jahre nach Stichtag (estat5jn).....	46
7.2.5.7	Erwerbsstatus 8 Tage vor Eintritt in den Betrieb (estat8tv).....	46
7.2.5.8	Lohn 5 Jahre nach Stichtag (lohn5jn).....	48
7.2.5.9	Lohn 8 Tage vor Eintritt in den Betrieb (lohn8tv) .....	48
7.2.6	Ortsangaben.....	48
7.2.6.1	Arbeitsort Arbeitsagentur (arb_aa).....	48
7.2.6.2	Arbeitsort Kreis (arb_kr).....	49
7.2.6.3	Wohnort Arbeitsagentur (wohn_aa) .....	49
7.2.6.4	Wohnort Kreis (wohn_kr).....	49
7.2.6.5	Ost-West-Kennzeichen (west) .....	50
<b>8</b>	<b>Datenanonymisierung .....</b>	<b>51</b>
<b>9</b>	<b>Datennutzung .....</b>	<b>52</b>
9.1	<i>Datenzugang</i> .....	52
9.2	<i>Hinweise für die Datenanalyse</i> .....	52
9.2.1	Testdaten.....	52
9.2.2	Überprüfung der Verknüpfungsqualität .....	53
9.2.3	Weitere Tipps zur Datenaufbereitung.....	54
<b>10</b>	<b>Abkürzungsverzeichnis.....</b>	<b>55</b>
<b>11</b>	<b>Literatur .....</b>	<b>56</b>

---

### **Datenverfügbarkeit**

Die in diesem Beitrag beschriebenen Datensätze sind für Forschungsprojekte zugänglich. Nähere Informationen dazu finden Sie auf der Internetseite: <http://fdz.iab.de> unter der Rubrik „Integrierte Betriebs- und Personendaten“.

### **Danksagung**

Für die Mitwirkung bei der Erstellung der ersten LIAB-Datenmodelle und Datenhandbücher ist Herrn Holger Alda gedankt. Das Team Datenmanagement des Bereichs ITM am Institut für Arbeitsmarkt- und Berufsforschung (IAB) hat die für die Erstellung der LIAB-Personendaten erforderliche Dateninfrastruktur des IAB bereitgestellt. Zudem wird für die in diesem Rahmen erfolgte datentechnische Beratung und Unterstützung gedankt. Dem Forschungsbereich „Betriebe und Beschäftigung“ des IAB wird für die Bereitstellung des IAB-Betriebspanels und die Unterstützung bei inhaltlichen Fragen gedankt. Schließlich danken wir den Nutzern des LIAB für ihre Anregungen, insbesondere für die Verbesserungsvorschläge zu den Testdaten und für die Erweiterungsvorschläge zu den biographischen Merkmalen.

---

## 1 Einleitung

Dieses LIAB-Datenhandbuch beschreibt die Linked Employer-Employee Daten des Instituts für Arbeitsmarkt- und Berufsforschung (IAB), wie sie über das Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit (BA) im IAB für wissenschaftliche Auswertungen zur Verfügung stehen. Die Linked Employer-Employee Daten im **IAB** (LIAB) bestehen aus den Betriebsdaten der jährlichen Erhebungswellen des IAB-Betriebspanels und Personendaten aus den Prozessdaten der BA. Durch die Verknüpfung beider Datenquellen entstehen Linked Employer-Employee Daten.

Das LIAB-Datenhandbuch stellt die wichtigsten Informationen zur Auswertung der LIAB-Daten zusammen. Diese überarbeitete Neuauflage entstand vor dem Hintergrund der Vereinheitlichung der FDZ-Datenreporte und deren Übersetzung ins Englische. Anlagen zum Handbuch sind eine Datei mit Merkmalsausprägungen, wenn die Liste der Ausprägungen besonders lang ist und eine Datei mit Auszählungen zu allen Merkmalen. Die bisherigen Dokumentationen werden nicht weiter gepflegt, bleiben aber dem deutschsprachigen Leser als reichhaltige Nachschlagewerke erhalten. Weitere Literatur ist in Abschnitt 11 aufgeführt. Wer schon mit älteren LIAB-Daten gearbeitet hat, sollte die Beschreibung der Erweiterungen und Änderungen beim Update auf 2005 und 2006 in Abschnitt 2 lesen.

Bezogen auf das IAB-Betriebspanel wird in diesem LIAB-Datenhandbuch nur auf Aspekte eingegangen, die im Zusammenhang mit der Verknüpfung mit Personendaten von Bedeutung sind. Vertiefende Informationen für das IAB-Betriebspanel stellt das FDZ auf seinen Internetseiten zur Verfügung.

---

## 2 Veränderungen des Datensatzes im Vergleich zu früheren Versionen

Bei einem Update werden immer alle Wellen neu generiert. Gründe dafür sind die Kontenbereinigungen der Rentenversicherung, die auch die Qualität der älteren Daten verbessern, die Gebietsstandsbereinigungen der Ortsangaben und die Sicherstellung der konsistenten Datenaufbereitung über die Wellen. Damit bestehende Programme weiter funktionieren, sollten bei Updates so wenige Änderungen wie möglich erfolgen.

### 2.1 Änderungen beim Update auf 2005

Die meisten Änderungen beim Update auf 2005 sind Erweiterungen: es gibt mehr Betriebe und mehr Merkmale. Daneben gibt es aber auch Änderungen an bereits bestehenden Merkmalen. Diese dienen zum einen der Vereinheitlichung der LIAB-Versionen und zum anderen der Vereinheitlichung mit anderen Personendaten des FDZ. Die folgende Aufstellung soll die Anpassung bestehender Programme erleichtern:

Erweiterungen:

- Auswahl der Betriebe im Querschnittmodell: Bisher waren in LIAB-Daten nur so genannte Querschnittsbetriebe (quer`jahr' = "Q") des IAB-Betriebspanels enthalten. Nun gilt die allgemeinere Regel: alle Fälle mit Interview in der laufenden Welle (Wellenkennzeichen well`jahr' = A, B, C, D, E oder G). Neu sind also Beobachtungen mit Wellenkennzeichen C. C bedeutet, dass die gleiche Betriebseinheit wie im Vorjahr befragt wurde, am Stichtag für die Hochrechnung (30.6. des Vorjahres) aber keine sozialversicherungspflichtig Beschäftigten im Betrieb angestellt waren.
  - Neue zusätzliche Betriebs-ID nach BA-BNR (betnr): In den ersten Jahren des Betriebspanels wurden Betriebe, die einmal teilgenommen hatten, dann aber nicht mehr, bei Ziehungen von Betrieben für Aufstockungen nicht berücksichtigt. Dies führte dazu, dass in manchen Zellen (Betriebsgrößen-/Branchenkombinationen) schließlich keine Betriebe für Aufstockungen mehr
-

blieben. Daher wanderten "gebrauchte" Betriebsnummern nach einer Karenzzeit wieder in den Topf und wurden wieder angeschrieben. Nehmen sie (erneut) an der Befragung teil, erhalten sie eine neue idnum. In vielen Fällen wird es sich aber um den gleichen Betrieb handeln, der vor Jahren schon mal teilgenommen hat. Da wir dies nicht eindeutig klären können, stellen wir neben der idnum auch das neue Merkmal Betriebs-ID nach BA-BNR (betnr) bereit, das kennzeichnet, welche Betriebe die gleiche von der BA vergebene Betriebsnummer haben. Der Forscher muss durch einen Vergleich im Zeitablauf beurteilen, ob es sich um den gleichen Betrieb handelt wie zuvor.

Bei den Längsschnittmodellen wird eine eindeutige Zuordnung von idnum und betnr verwendet, so dass die betnr nicht relevant ist.

- Auswahl der Beobachtungen in den Personendaten im Querschnittmodell: Bisher waren Spells zu Beschäftigungen in IAB-Betriebspanel-Betrieben enthalten. Informationen zu weiteren Beschäftigungen und Leistungsbezug waren in den Merkmalen mjob und mjob\_tag\_entg enthalten. Nun werden alle parallelen Spells der Personen zum Stichtag aufgenommen, also auch Spells zu weiteren Beschäftigungen und Spells zum Leistungsempfang. Dadurch kann der Anwender selbst entscheiden, wie er diese Informationen verwenden möchte. Wer diese Information nicht benötigt, kann die Datenzeilen über die Variablen Quelle und Status der Betriebsnummer leicht löschen. Die Variablen mjob und mjob\_tag\_entg werden nicht mehr angeboten, da sie anhand der Daten berechnet werden können.
- Die Testdaten sind neu erstellt worden, wobei versucht wurde, möglichst viele Nutzerwünsche umzusetzen (vgl. Abschnitt 9.2.1). Weitere Anregungen sind natürlich willkommen.

#### Änderungen:

- Die Auswahl, Bezeichnung und Aufbereitung der Variablen in LIAB-Querschnitt- und Längsschnittmodellen wurde vereinheitlicht. Dadurch ergeben sich einzelne Änderungen.
  - Fehlende Werte wurden einheitlich kodiert (vgl. Abschnitt 7.1).
  - Alle Strings wurden in numerische Merkmale umgesetzt.
-

## Übersicht der Erweiterungen und Änderungen bei einzelnen Merkmalen

	Variablen-name	Querschnittmodell	Längsschnittmodelle
<b>Identifikatoren</b>			
Personen-ID	id	–	–
Betriebs-ID nach Betriebspanel	idnum	–	–
Betriebs-ID nach BA-BNR	betnr	neu	neu
<b>Technische Merkmale</b>			
Spellbeginn original	begorig	ersetzt beginn	neu
Spellende original	endorig	ersetzt ende	neu
Spellbeginn Episode	begepi	neu	ersetzt beginn
Spellende Episode	endepe	neu	ersetzt ende
Quelle	quelle	neu	–
Spellzähler der Person	spell_nr	–	–
Spellzähler pro Episode und Quelle	level1	neu	–
Status der Betriebsnummer	betr_st	neu	verändert, keine Unterscheidung von Ost und West
Datumsbereinigung	dat_ber	neu	–
<b>Merkmale zur Person</b>			
Geschlecht	geschl	–	–
Geburtsjahr	gebjahr	ersetzt alter	ersetzt alter
Staatsangehörigkeit, gruppiert	nation	analog zu FDZ-Daten kodiert	nation_neu in nation
Staatsangehörigkeit, original (sensibel)	nationor	Original als sensibles Merkmal	Original als sensibles Merkmal
Ausbildung	ausbild	0 auf -7 umgesetzt	0 auf -7 umgesetzt
Familienstand	fam	neu	–
Kinder	kind	neu	–
<b>Beschäftigung und Leistungsbezug</b>			
Personengruppe	pers_gr	–	–
Abgabegrund der Meldung	abg_gr	LEH: rekodiert	numerisch LeH: rekodiert
Stellung im Beruf	berstell	5, 6 -> -9	5, 6 -> -9
Beruf	beruf	BeH: 3-steller LeH: Missing	BeH: 3-steller LeH: Missing
Tagesentgelt/Leistungssatz	tag_entg	immer in Euro	immer in Euro taglohn in tag_entg umbenannt
Leistungsart	leistart	neu	–
Bewilligungsart	bew_art	neu	–
Abmeldegrund Leistungsbezug	abm_gr	neu	–
Wirtschaftszweig WZ73, 3-steller (sensibel)	wz73	–	–

Wirtschaftszweig WZ93, 5-steller (sensibel)	wz93	–	–
Wirtschaftszweig WZ03, 5-steller (sensibel)	wz03	numerisch	neu
<b>Generierte biographische Merkmale</b>			
Eintrittsdatum Erwerbsleben	ein_erw	–	–
Eintrittsdatum Betrieb	ein_bet	–	–
Dauer Betrieb	tage_bet	umbenannt in tage_bet (vorher dauer_`jahr')	–
Anzahl bisheriger Leistungsbezüge	anz_lst	neu	neu
Dauer bisheriger Leistungsbezüge	tage_lst	neu	neu
Erwerbsstatus 5 Jahre nach Stichtag	estat5jn	neu	nicht aufgenommen
Erwerbsstatus 8 Tage vor Eintritt in den Betrieb	estat8tv	neu	neu
Lohn 5 Jahre nach Stichtag	lohn5jn	neu	nicht aufgenommen
Lohn 8 Tage vor Eintritt in den Betrieb	lohn8tv	neu	neu
<b>Ortsangaben</b>			
Arbeitsort Arbeitsagentur (sensibel)	arb_aa	–	–
Arbeitsort Kreis (sensibel)	arb_kr	–	–
Wohnort Arbeitsagentur (sensibel)	wohn_aa	–	–
Wohnort Kreis (sensibel)	wohn_kr	LeH auf -5 numerisch	numerisch
Ost-West-Kennzeichen, BeH: Arbeitsort, LeH: Wohnort	west	–	–

### Nicht mehr enthalten sind die Variablen:

Alter der Person in Jahren	alter	ersetzt durch gebjahr
Währungskennzeichen	euro	alle Entgelte in Euro
Dauer der Meldung in Tagen	dauer	kann selber berechnet werden
Parallele Zustände, ja/nein	mjob	kann selber berechnet werden
Parallele Zustände, Geld	mjob_tag_entg	kann selber berechnet werden
Zweites Ost-West-Kennzeichen im Längsschnittmodell	ost_west	Personendaten werden nicht mehr nach Ost- und West getrennt

## 2.2 Änderungen beim Update auf 2006

Mit dem Update auf 2006 gibt es folgende Änderungen:

- Die Welle 2006 des IAB-Betriebspanels wurde aufgenommen.
- Personendaten wurden aus der BLH 6.01 gezogen, der aktuelle Datenrand ist Ende 2006.
- Jedem LIAB-Nutzer wird nun das vollständige IAB-Betriebspanel zur Verfügung gestellt. In der Vergangenheit waren im Querschnittmodell nur Betriebe

mit gültiger Befragung enthalten, in den Längsschnittmodellen Teilstichproben dieser Betriebe. Durch die vollständige Bereitstellung kann nun auch der LIAB-Nutzer über das Wellenkennzeichen ermitteln, warum ein Betrieb im nächsten Jahr nicht enthalten ist. Die Variablen `in_QM1`, `in_LM1`, `in_LM2` und `in_LM3` zeigen an, ob es sich um einen LIAB-Betrieb der entsprechenden Version handelt. Alle Nicht-LIAB-Betriebe können über folgende Schleife leicht wieder gelöscht werden:

```
forvalues y = 1993/2006 {  
    use orig/LIAB_2006_estab_`y', clear  
    keep if in_LM1 = 1  
    sort idnum  
    save data/LIAB_LM1_2006_estab_`y', replace  
}
```

- Die Identifikatoren `id` und `betnr` wurden in der Vergangenheit bei jedem Update neu generiert, so dass neue Daten nicht mit alten verknüpft werden konnten. Beim Update 2006 wurden `id` und `betnr` erstmalig konsistent für alle LIAB-Versionen und konsistent mit den Längsschnittmodellen von 2005 erstellt. D.h. die neuen Daten für das Jahr 2006 können bei den Längsschnittmodellen den alten von 2005 auch auf Personenebene zugespielt werden. Auf Betriebsebene war das anhand der `idnum` schon immer möglich.
- Bei den Längsschnittmodellen aus 2005 wurden bei der Berechnung der Variablen Eintritt ins Erwerbsleben irrtümlich auch Ausbildungszeiten in Betrieben berücksichtigt. Dies widerspricht der Variablenbeschreibung (siehe Abschnitt 7.2.5.1), ist also ein Fehler, der beim Update auf 2006 behoben wurde.
- Bei den Längsschnittmodellen wurde das Zeitfenster der gezogenen Personendaten nach hinten um ein Jahr erweitert. Für die ausgewählten Betriebe wurden alle Personen gezogen, die in den folgenden Zeitfenstern mindestens einen Tag dort beschäftigt waren:
  - LM1: 01.01.1996 bis 31.12.2002
  - LM2: 01.01.1997 bis 31.12.2003
  - LM3: 01.01.1993 bis 31.12.2006

Man kann also für diese Zeiträume für jeden Tag die gesamte sozialversicherungspflichtige Belegschaft beobachten, mit diesem Update also auch ein Jahr nach dem eigentlichen Auswahlfenster der Längsschnittversion.

### 3 Kurzbeschreibungen

#### LIAB-Querschnittmodell, Version 1

Inhaltliche Charakteristika	
Themen/Merkmalgruppen	<p>Betriebsdaten: 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, allgemeine Angaben zum Betrieb</p> <p>Personendaten: Soziodemographische Merkmale: systemfreie Personennummer, Geschlecht, Geburtsjahr, Staatsangehörigkeit, Schul- und Berufsausbildung Beschäftigungsbezogene Merkmale: u.a. Beginn und Ende der Beschäftigung, Art der Beschäftigung (insbesondere Unterscheidung zwischen sozialversicherungspflichtiger und geringfügiger Beschäftigung), sozialversicherungspflichtiges Bruttoentgelt, Beruf, Stellung im Beruf (einschließlich Voll- oder Teilzeitbeschäftigung), Grund für die Abgabe der Beschäftigungsmeldung (z.B. Beschäftigungsende oder -unterbrechung) Leistungsbezugsbezogene Merkmale: Beginn und Ende einer Leistungsbezugsepisode, Art der Leistung (Arbeitslosengeld, -hilfe oder Unterhaltsgeld), Grund für das Ende des Leistungsbezugs</p>
Untersuchungseinheit	Betriebe und sozialversicherungspflichtig Beschäftigte (ab 1999 auch geringfügig Beschäftigte)
Fallzahlen	Betriebe: 4.000 bis 16.000 pro Jahr Beschäftigte: 1,8 bis 2,5 Mio. pro Jahr
Zeitraum	1993 bis 2006
Zeitbezug	jährlich zum 30. Juni
Regionale Gliederung	Betriebsdaten: Bundesland Personendaten: Ost-West (tiefere Gliederung durch sensible Merkmale)
Gebietsstand	BeH-Meldungen sind gebietsstandsaktualisiert zum 31.12.2005
Methodische Charakteristika	
Erhebungsdesign	<p>Betriebsdaten: alle im betreffenden Jahr befragten Betriebe des IAB-Betriebspanels (disproportional geschichtete Zufallsstichprobe)</p> <p>Personendaten: alle Beschäftigten- und Leistungsmeldungen zum 30.6. von Personen, die am 30.6. in einem Betriebspanel-Betrieb beschäftigt sind</p>
An der Erhebung beteiligte Institutionen	<p>IAB-Betriebspanel: Auftraggeber: Forschungsbereich "Betriebe und Beschäftigung" des IAB, Ministerien der Bundesländer bzw. Regionaldirektionen Durchführung: Infratest Sozialforschung München, SÖSTRA Berlin Partnerinstitute: BAW, IAW, IWAK, INIFES Personendaten: Sozialversicherungsträger</p>
Frequenz der Datensammlung	<p>Betriebsdaten: einmal jährlich</p> <p>Personendaten: laufend</p>

Dateiformat/-größe	Stata Personendaten: <270 MB pro Jahr in Stata Betriebsdaten: <8 MB pro Jahr in Stata
Dateiorganisation	Es gibt je einen Personen- und einen Betriebsdatensatz pro Jahr: LIAB_QM1_pers_*jahr*.dta LIAB_QM1_estab_*jahr*.dta
<b>Datenzugang</b>	
Zugangsweg	Gastaufenthalt nach der Bewilligung des Antrages auf Datenzugang gemäß § 75 SGB X; nach erfolgreichem Aufbau des Analysedatensatzes im Rahmen eines Gastaufenthalts auch Datenfernverarbeitung (Nachbereitung des Gastaufenthalts)
Anonymisierungsgrad	Schwach anonymisiert
Sensible Merkmale	Personendaten: Staatsangehörigkeit, original; Wirtschaftszweig Z73, 3-steller; Wirtschaftszweig WZ93, 5-steller; Wirtschaftszweig WZ03, 5-steller; Arbeitsort Arbeitsagentur; Arbeitsort Kreis; Wohnort Arbeitsagentur; Wohnort Kreis

### LIAB-Längsschnittmodell, Version 1

<b>Inhaltliche Charakteristika</b>	
Themen/Merkmalgruppen	Personendaten: Merkmale aus dem Verfahren der Beschäftigten-Leistungsempfänger-Historik Betriebe: Merkmale des IAB-Betriebspanel
Untersuchungseinheit	Beschäftigte / Betriebe
Fallzahlen	Betriebe: je ca. 2.100 in West- und Ostdeutschland Beschäftigte: 25.100.000 Meldungen (entspricht ca. 1,9 Mio Personen)
Zeitraum	Betriebe: mindestens 1999 bis 2001 Beschäftigte: 1990 bis 2006
Zeitbezug	Betriebe: jährlich zum 30. Juni Beschäftigte: mindestens eintägige Beschäftigung zwischen 1996 und 2002 in einem der je etwa 2.100 Betriebe
Regionale Gliederung	Ost- und Westdeutschland
Gebietsstand	laufende Aktualisierung wie in BLH
<b>Methodische Charakteristika</b>	
Erhebungsdesign	Die Personendatenbanken enthalten alle Beschäftigungs- und Leistungsmeldungen von 1991 bis 2006, wenn die entsprechende Person mindestens einen Tag in einem der je 2.100 Betriebe zwischen 1996 und 2002 gearbeitet hat. Diese IAB-Betriebspanelbetriebe haben mindestens zwischen 1999 und 2001 in jedem Jahr ein gültiges Interview.
An der Erhebung beteiligte Institutionen	IAB
Frequenz der Datensammlung	einmalig mit jährlicher Aktualisierung
Dateiformat/-größe	STATA (2.650 MB), SAS (4.200 MB)
<b>Datenzugang</b>	

Zugangsweg	Gastaufenthalt, Datenfernverarbeitung (Nachbereitung des Gastaufenthalts)
Anonymisierungsgrad	schwach anonymisiert
Sensible Merkmale	Wohnort (Kreis, Arbeitsagentur), Arbeitsort (Kreis, Arbeitsagentur), Staatsangehörigkeit (original), WZ73 3-Steller, WZ93 5-Steller, WZ03 5-Steller

## LIAB-Längsschnittmodell, Version 2

<b>Inhaltliche Charakteristika</b>	
Themen/Merkmalgruppen	Personendaten: Merkmale aus dem Verfahren der Beschäftigten-Leistungsempfänger-Historik Betriebe: Merkmale des IAB-Betriebspanel
Untersuchungseinheit	Beschäftigte / Betriebe
Fallzahlen	Betriebe: ca. 5.300 (West) und 4.300 (Ost) Beschäftigte: West: ca. 28,5 Mio. Meldungen (entspricht 2.358.451 Personen) Ost: ca. 15,8 Mio. Meldungen (entspricht 1.100.955 Personen)
Zeitraum	Betriebe: mindestens 1999 bis 2001 und mindestens von 2000 bis 2002 Beschäftigte: 1993 bis 2006
Zeitbezug	Betriebe: jährlich zum 30. Juni Beschäftigte: mindestens eintägige Beschäftigung zwischen 1997 und 2002 in einem der ausgewählten Betriebe
Regionale Gliederung	Ost- und Westdeutschland
Gebietsstand	laufende Aktualisierung wie in BLH
<b>Methodische Charakteristika</b>	
Erhebungsdesign	Die Personendatenbanken enthalten alle Beschäftigungs- und Leistungsmeldungen von 1993 bis 2006, wenn die entsprechende Person mindestens einen Tag in einem der ausgewählten Betriebe zwischen 1997 und 2003 gearbeitet hat. Diese IAB-Betriebspanelbetriebe haben mindestens zwischen 1999 und 2001 bzw. von 2000 bis 2002 in jedem Jahr ein gültiges Interview
An der Erhebung beteiligte Institutionen	IAB
Frequenz der Datensammlung	einmalig mit jährlicher Aktualisierung
Dateiformat/-größe	STATA (West: 2.420 MB; Ost: 1.346 MB) SAS (West: 3.870 MB; Ost: 2.153 MB)
<b>Datenzugang</b>	
Zugangsweg	Gastaufenthalt, Datenfernverarbeitung (Nachbereitung des Gastaufenthalts)
Anonymisierungsgrad	schwach anonymisiert
Sensible Merkmale	Wohnort (Kreis, Arbeitsagentur), Arbeitsort (Kreis, Arbeitsagentur), Staatsangehörigkeit (original), WZ73 3-Steller, WZ93 5-Steller, WZ03 5-Steller

**LIAB-Längsschnittmodell, Version 3**

<b>Inhaltliche Charakteristika</b>	
Themen/Merkmalgruppen	Personendaten: Merkmale aus dem Verfahren der Beschäftigten-Leistungsempfänger-Historik Betriebe: Merkmale des IAB-Betriebspanel
Untersuchungseinheit	Beschäftigte / Betriebe
Fallzahlen	Betriebe: je ca. 2.100 in West- und Ostdeutschland Beschäftigte: 25.100.000 Meldungen (entspricht ca. 1,9 Mio Personen)
Zeitraum	Betriebe: mindestens 1996 bis 2005 Beschäftigte: 1993 bis 2006
Zeitbezug	Betriebe: jährlich zum 30. Juni Beschäftigte: mindestens eintägige Beschäftigung zwischen 1996 und 2002 in einem der je etwa 2.100 Betriebe
Regionale Gliederung	Ost- und Westdeutschland
Gebietsstand	laufende Aktualisierung wie in BLH
<b>Methodische Charakteristika</b>	
Erhebungsdesign	Die Personendatenbanken enthalten alle Beschäftigungs- und Leistungsmeldungen von 1993 bis 2006, wenn die entsprechende Person mindestens einen Tag in einem der je 2.100 Betriebe zwischen 1993 und 2006 gearbeitet hat. Diese IAB-Betriebspanelbetriebe haben mindestens zwischen 1996 und 2005 in jedem Jahr ein gültiges Interview.
An der Erhebung beteiligte Institutionen	IAB
Frequenz der Datensammlung	einmalig mit jährlicher Aktualisierung
Dateiformat/-größe	STATA (2.650 MB), SAS (4.200 MB)
<b>Datenzugang</b>	
Zugangsweg	Gastaufenthalt, Datenfernverarbeitung (Nachbereitung des Gastaufenthalts)
Anonymisierungsgrad	schwach anonymisiert
Sensible Merkmale	Wohnort (Kreis, Arbeitsagentur), Arbeitsort (Kreis, Arbeitsagentur), Staatsangehörigkeit (original), WZ73 3-Steller, WZ93 5-Steller, WZ03 5-Steller

## 4 Datenquellen

### 4.1 Betriebsdaten

Die Betriebsseite der LIAB-Daten bildet das IAB-Betriebspanel.<sup>1</sup> Das IAB-Betriebspanel ist eine seit 1993 jährlich durchgeführte Arbeitgeberbefragung, zunächst in etwa 4000 westdeutschen, seit 1996 zusätzlich in etwa 4300 ostdeutschen Betrieben. Ab dem Jahr 2000 sind für jedes Bundesland repräsentative Auswertungen möglich, denn seitdem finanzieren auch westdeutsche Bundesländer entsprechende Aufstockungsstichproben. Eine genaue Beschreibung des IAB-Betriebspanels findet sich in Bellmann (2002)<sup>2</sup> und auf unseren Internetseiten (<http://fdz.iab.de/pageText.asp?PageID=19>), hier folgt eine zusammenfassende Darstellung.

Das IAB-Betriebspanel enthält u.a. die folgenden Informationen:

- Beschäftigtenzahl
- Beschäftigtenstruktur (Qualifikation, Geschlecht, Befristung, Leiharbeit,...)
- Wochenarbeitszeit für Vollzeitbeschäftigte und Überstunden
- Tarifbindung des Betriebs, Existenz eines Betriebsrats
- Umsatz, Vorleistungen und Exportanteil
- Lohnsumme
- Investitionssumme
- technischer Stand der Betriebsanlagen
- Wirtschaftszweig, Betriebsalter, Rechtsform und Unternehmensstellung
- Beurteilung der betrieblich-wirtschaftlichen Gesamtsituation
- Reorganisationsmaßnahmen und betriebliche Weiterbildungsaktivitäten (in mehrjährigen Abständen)

Die Grundgesamtheit des IAB-Betriebspanels sind alle Betriebe in Deutschland mit mindestens einem sozialversicherungspflichtig Beschäftigten. Aus diesen wird eine disproportionale Stichprobe gezogen, in der Großbetriebe überrepräsentiert sind. Die Ziehung erfolgt geschichtet nach 10 Betriebsgrößenklassen und 20 (bis 1999: 16)

---

<sup>1</sup> Das Betriebs-Historik-Panel (BHP) basierend auf dem Meldeverfahren zur Sozialversicherung ist eine ergänzende Datenbasis für zukünftige LIAB-Versionen. Zum BHP siehe FDZ Datenreport 3/2006.

<sup>2</sup> Bellmann, L. (2002): *Das IAB-Betriebspanel - Konzeption und Anwendungsbereiche*, Allgemeines Statistisches Archiv 86, 177 – 188.

---

Wirtschaftszweigen (vgl. Tabelle 1). Da die Ziehung der Betriebe ein Jahr vor der Befragung erfolgt und es zu dieser Zeit noch keine Beschäftigendaten für das Befragungsjahr gibt, basiert die Hochrechnungsmatrix für das Befragungsjahr auf der betrieblichen Vorjahresbeschäftigung.

### Ziehungswahrscheinlichkeiten und Rücklaufquoten in der ersten Befragungswelle 1993 (nur westdeutsche Betriebe)

Beschäftigte	gezogene Betriebe	Auswahlwahrscheinlichkeit	antwortende Betriebe	Rücklaufquote
1-4	1,072	.0011	625	.67
5-9	431	.0015	250	.64
10-19	466	.0030	299	.71
20-49	862	.0089	542	.70
50-99	535	.0153	350	.72
100-199	543	.0304	376	.77
200-499	923	.0862	615	.74
500-999	479	.1504	304	.71
1.000-4.999	1,497	.8765	924	.72
5.000 u.m.	115	.9127	71	.73
Insgesamt	6,923	.0043	4,356	.71

Quelle: IAB-Betriebspanel 1993, Kölling (2000), S. 294.<sup>3</sup>

An den Ziehungswahrscheinlichkeiten und Rücklaufquoten ändert sich im Verlauf der Jahre relativ wenig. Eine Ausnahme bilden hier die großen Betriebe. Panelmortalität kann in den oberen Größenklassen kaum ausgeglichen werden. Die Anzahl der verknüpfbaren Personen bleibt daher trotz steigender Anzahl von Betrieben aufgrund der Panelmortalität der Großbetriebe nahezu unverändert (zur Entwicklung der Fallzahlen siehe Abschnitt 5.3).

<sup>3</sup> Kölling, A. (2000): European Datawatch: *The IAB-Establishment-Panel*, Schmollers Jahrbuch, 120 Jg., 291 - 300.

## 4.2 Personendaten

Die Personendaten stammen aus der Beschäftigten-Leistungsempfänger-Historik-Datei (BLH) des IAB die bis ins Jahr 1975 zurückreicht. Angaben zu Beschäftigten stammen aus dem Meldeverfahren zur Sozialversicherung, wo Arbeitgeber mindestens einmal jährlich für jeden sozialversicherungspflichtig Beschäftigten eine Meldung abgeben. Diese Meldungen enthalten u.a.

- tagesgenaues Beginn- und Endedatum jeder Beschäftigtenmeldung
- Geburtsdatum, Geschlecht und Staatsangehörigkeit
- ausgeübter Beruf
- Lohn zensiert an den oberen und unteren Beitragsbemessungsgrenzen
- Bildungsniveau in Verbindung mit dem beruflichen Qualifikationsniveau
- Wirtschaftszweig
- Arbeits- und Wohnort
- Betriebsidentifikator

Die BLH enthält neben Angaben zu Beschäftigungszeiten auch Angaben zum Leistungsempfang aus Geschäftsstatistiken der Bundesagentur für Arbeit. Arbeitslosigkeit wird also nur dann erfasst, wenn die Person Leistungen der Bundesagentur für Arbeit bezieht. Gemeldete Arbeitslosigkeit ohne Anspruch auf Bezüge (z.B. nach einem Studium, bei Hilfe zum Lebensunterhalt oder nach Selbstständigkeit) ist somit nicht enthalten. Für Leistungsempfangszeiten sind u.a. folgende Informationen enthalten:

- tagesgenaues Beginn und Ende jeder Leistungsmeldung
- Alter, Geschlecht, Nationalität
- Beruf
- Art und Höhe der Leistung
- Bewilligungsart
- Abgabegrund der Meldung
- Beendigungsgrund
- Ort des zahlenden Arbeitsamtes

Leistungsmeldungen erfolgen nur bei leistungsrechtlich relevanten Veränderungen. Eine Leistungsmeldung kann im Gegensatz zu Beschäftigtenmeldungen einen längeren Zeitraum als ein Jahr beschreiben.

---

Die Beschäftigtenmeldungen gehen über den Umweg Kranken- und Rentenversicherung zeitverzögert in der BA ein. 6 Monate nach Jahresende liegen etwa 95% der Meldungen vor, nach 18 Monaten über 99%. Im letzten aufbereiteten Jahr können also noch Meldungen fehlen, wenn diese Daten auf einer 6-Monats-Datei basieren. Da Leistungsempfangsdaten direkt aus den Geschäftsstatistiken der BA generiert werden, ist hier der Füllgrad auch am aktuellen Rand nahezu vollständig.

---

## 5 Datenaufbereitung

Die Betriebsdaten stimmen bis auf die Selektion der Betriebe mit dem IAB-Betriebspanel überein. Dokumentationen für das IAB-Betriebspanel gelten also für den LIAB gleichermaßen. Bei den Personendaten streben wir ebenfalls eine Vereinheitlichung mit IABS, BAP und IEBS an, derzeit unterscheiden sich aber noch Dinge wie Variablennamen und Kodierung der Fehlenden Werte. Eine Gegenüberstellung gleicher Merkmale in unterschiedlichen Datensätzen bietet der FDZ-Datenreport 4/2006.

### 5.1 Grundsätzliche Verknüpfungsmöglichkeiten

Abbildung 1 soll die grundsätzlichen Verknüpfungsmöglichkeiten der tagesgenauen Spelldaten aus der BLH mit den jährlichen Stichtagsangaben des IAB-Betriebspanels veranschaulichen.<sup>4</sup>

---

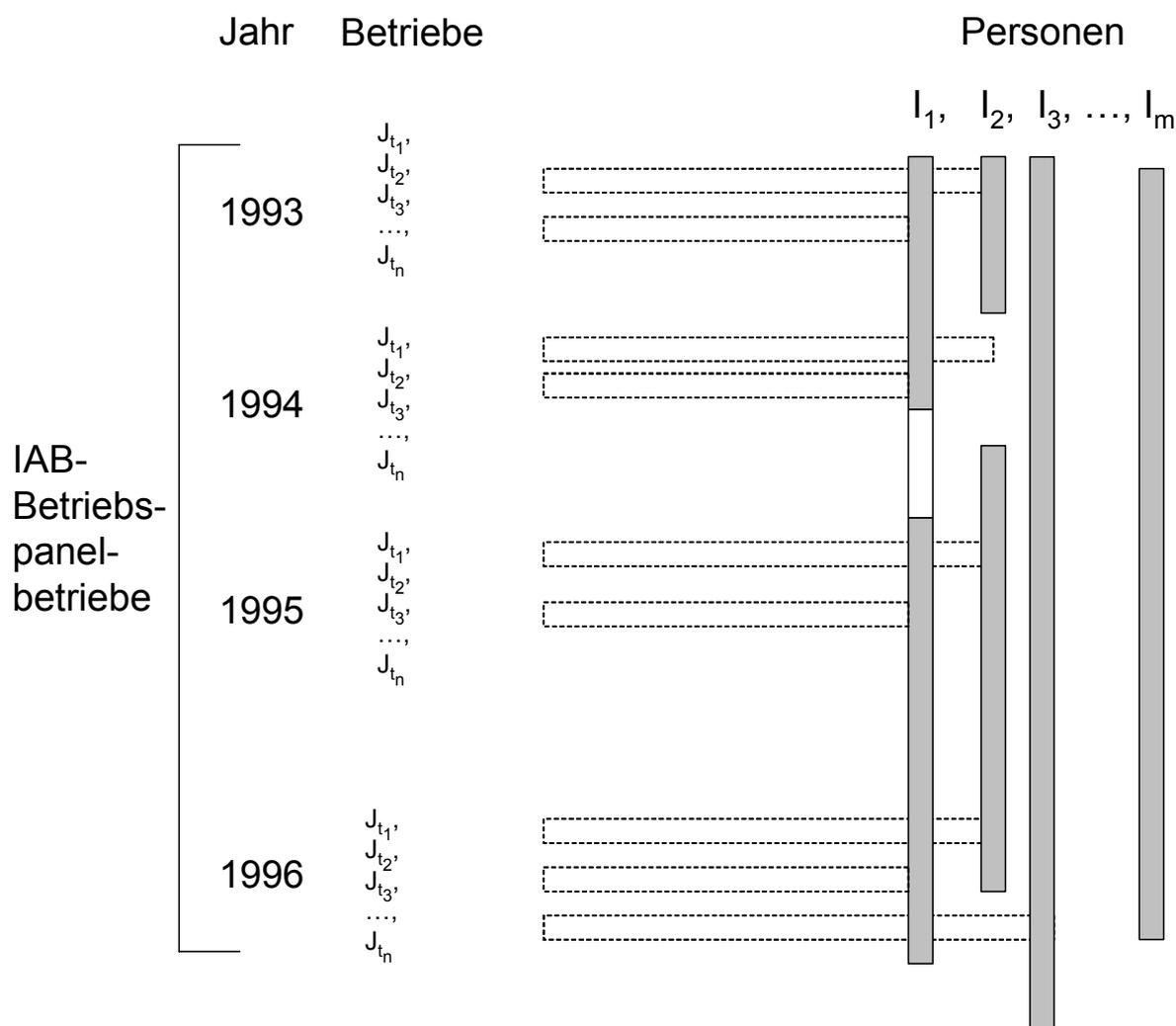
<sup>4</sup> Abowd/Kramarz (1999) geben einen Überblick über die grundsätzlichen Verknüpfungsmöglichkeiten von Betriebs- und Beschäftigtendaten. Sie unterscheiden u.a. zwischen Verknüpfungen ausschließlich administrativer Daten, ausschließlich von Befragungsdaten und diversen Mischformen. Für die hier beschriebenen LIAB-Daten geben Alda/Bender/Gartner (2005) einen umfassenderen Überblick über mögliche Anwendungsgebiete und -beispiele.

Abowd, J./Kramarz, F. (1999): *The analysis of labor markets using matched employer-employee data*. In: Ashenfelter, O./Card, D. (Hrsg.): *Handbook of Labor Economics*, Vol. 3B, Elsevier, Amsterdam, Chapter 40, pp. 2567 - 2627.

Alda, H./Bender, St./Gartner, H. (2005): *Der Linked-Employer-Employee-Datensatz aus IAB-Betriebspanel und den prozessproduzierten Daten des IAB (LIAB)*. IAB-Discussion-Paper No.6/2005

---

Abbildung 1: Betriebe und Beschäftigte in LIAB-Daten



Die horizontalen Linien stellen IAB-Betriebspanel-Betriebe dar, die vertikalen Linien Erwerbsverläufe von Personen in der BLH. Grau hinterlegte Balkenabschnitte kennzeichnen Zeiten der Beschäftigung und hell hinterlegte gemeldete Arbeitslosigkeit. Ein unterbrochener Balken bedeutet, dass die Person weder ein sozialversicherungspflichtiges Beschäftigungsverhältnis ausübt, noch Leistungen der BA in Anspruch nimmt. Vertikale Balken zu Beschäftigungsverhältnissen und horizontale Balken zu Betriebsbefragungen enthalten Betriebsnummern. Auf Basis dieser Betriebsnummern können die Personen- und Betriebsdaten verknüpft werden.<sup>5</sup>

Beim IAB-Betriebspanel handelt es sich um eine Stichprobe, bei der BLH um eine Vollerhebung. Jeder IAB-Betriebspanel-Betrieb lässt sich somit grundsätzlich mit Informationen zu jedem einzelnen sozialversicherungspflichtig Beschäftigten verknüp-

<sup>5</sup> Zu Einschränkungen bei der der Verknüpfung siehe Abschnitt 8 zur Datenqualität.

fen. In der anderen Richtung liegen aber nicht für jeden Beschäftigten Angaben aus der Betriebsbefragung vor. Abgesehen davon geht die Übertragung in beide Richtungen: Betriebsinformationen können auf die Personenebene übertragen werden und (auf Betriebsebene aggregierte) Personeninformationen können auf die Betriebsebene übertragen werden.

## 5.2 LIAB-Daten am FDZ

Für eine Verknüpfung kommen alle Personen in Betracht, die mindestens einmal in einem IAB-Betriebspanel-Betrieb gearbeitet haben. Das FDZ kann jedoch einen Datensatz mit all diesen Personendaten nicht bereitstellen, da dieser zu groß wäre.<sup>6</sup> Um die Dateigröße noch handhabbar und die Analysemöglichkeiten möglichst groß zu halten, wurden am IAB zwei Typen von Datensätzen entwickelt, die wir als Datenmodelle bezeichnen. Das eine Datenmodell enthält alle Betriebsinformationen des IAB-Betriebspanels, aber nur Personeninformationen zu den jeweiligen Stichtagen am 30. Juni der Befragungsjahre. Aus den Personendaten wird die Querschnittsinformation zu diesen Stichtagen für alle in einem der IAB-Betriebspanel-Betriebe Beschäftigten gezogen. Dieses Datenmodell bezeichnen wir als LIAB-Querschnittmodell, da es nur beschränkte Längsschnittinformationen für die Personen enthält. In Abbildung 1 ist das LIAB-Querschnittmodell durch alle Schnittpunkte der horizontalen und vertikalen Linien gekennzeichnet. Aus den Verlaufsdaten der Personen geht nur die Information an den Stichtagen ein.

Das zweite Datenmodell heißt LIAB-Längsschnittmodell und enthält alle Spells der Personen über einen bestimmten Zeitraum, so dass auch Informationen vor und nach Betriebseintritt vorliegen. Um die Dateigröße in Grenzen halten zu können diese Längsschnittinformationen jedoch nicht für alle IAB-Betriebspanel-Betriebe bereitgestellt werden, weshalb die Versionen des LIAB-Längsschnittmodells nur eine Teilstichprobe der Betriebsdaten enthalten. In Abbildung 1 ist das LIAB-Längsschnittmodell durch eine Auswahl der horizontalen Linien und längere Abschnitte der vertikalen Linien gekennzeichnet.

Um eine gute Datenqualität sicherzustellen und gute Analysemöglichkeiten zu erzielen, ist die Ziehung der Teilstichprobe der Betriebe für LIAB\_Längsschnittmodelle

---

<sup>6</sup> Die Dateigröße der BLH beträgt derzeit 277 Gigabyte.

nicht zufällig. Zunächst werden nur Betriebe ausgewählt die eine gute Verknüpfungsqualität zu den Personendaten aufweisen (vgl. Abschnitt 6: Datenqualität). Dann werden nur Betriebe ausgewählt, die an mehreren, aufeinander folgenden Jahren teilnehmen. Von diesen Betrieben wird dann eine Stichprobe gezogen.

Die beiden LIAB-Datenmodelle beschreiben zwei mögliche Entwürfe von Datensätzen, mit den LIAB-Versionen werden sie konkretisiert. LIAB-Versionen sind also reale Datensätze, die sich hinsichtlich des abgedeckten Zeitraums, des Variablenkanons und der ausgewählten Betriebe unterscheiden. Im Folgenden werden die vorhandenen LIAB-Versionen beschrieben.

### **5.2.1 LIAB-Querschnittmodell, Version 1**

In Personendaten sind die Daten der Beschäftigten in Betrieben abgelegt, die im jeweiligen Jahr an der Befragung des IAB-Betriebspanels teilgenommen haben. Es handelt sich dabei um die aktuelle Meldung der beschäftigten Personen am 30. Juni eines jeden Befragungsjahres. Es sind aber auch alle weiteren Informationen der BLH zu diesen Personen zu den Stichtagen enthalten. Dabei handelt es sich zum einen um parallele Beschäftigungen in Betrieben, die im entsprechenden Jahr nicht an der Befragung teilgenommen haben und zum anderen um Informationen zu parallelem Leistungsbezug.

Da in den einzelnen Querschnitten Identifikatoren für Personen und Betriebe vorliegen, können die Querschnitte miteinander verbunden werden. Somit ist es möglich, Verläufe von Personen in IAB-Betriebspanel-Betrieben zu verfolgen. Sowohl Betriebswechsel in einen nicht durch das Betriebspanel befragten Betrieb, als auch die Nichtteilnahme des Betriebes an der Befragung im Folgejahr führen dazu, dass eine Person nicht mehr in dieser Version des LIAB enthalten ist. Erweiterungen des LIAB-Querschnittmodells entstehen durch die Bereitstellung weiterer Erwerbsverlaufsinformationen, die aus der BLH generiert werden können, wie z.B. dem Ersteintritt in das Erwerbsleben oder den Betrieb oder die Anzahl von Leistungsbezügen bis zum Stichtag.

### **5.2.2 LIAB-Längsschnittmodell, Version 1**

Zunächst werden Betriebe gezogen, die 1999 bis 2001 an der Befragung teilgenommen haben. Betriebe mit schlechter Verknüpfungsqualität werden gelöscht. Von den

---

verbleibenden Betrieben werden etwa 2100 Betriebe in Ost- und Westdeutschland ausgewählt.

Die Personendaten enthalten alle Spells zu Beschäftigung und Leistungsempfang von 1990 bis 2006 der Personen, die zwischen 1996 und 2002 mindestens einen Tag in einem der etwa 4200 ausgewählten Betriebe beschäftigt waren.

**5.2.3 LIAB-Längsschnittmodell, Version 2**

Zunächst werden Betriebe gezogen, die 2000 bis 2002 an der Befragung teilgenommen haben. Betriebe mit schlechter Verknüpfungsqualität werden nicht gelöscht. zusätzlich sind alle Betriebe enthalten, die schon in Version 1 enthalten waren. Der Datensatz umfasst 4360 Betriebe in Ost- und 5293 Betriebe in Westdeutschland.

Die Personendaten enthalten alle Spells zu Beschäftigung und Leistungsempfang von 1993 bis 2006 der Personen, die zwischen 1997 und 2003 mindestens einen Tag in einem der ausgewählten Betriebe beschäftigt waren.

**5.2.4 LIAB-Längsschnittmodell, Version 3**

Zunächst werden Betriebe gezogen, die 1996 bis 2005 an der Befragung teilgenommen haben. Betriebe mit schlechter Verknüpfungsqualität werden nicht gelöscht.

Die Personendaten enthalten alle Spells zu Beschäftigung und Leistungsempfang von 1993 bis 2006 der Personen, die zwischen 1993 und 2006 mindestens einen Tag in einem der ausgewählten Betriebe beschäftigt waren.

**5.2.5 Vor- und Nachteile von Quer- und Längsschnittmodellen**

Vorteile Querschnittmodell	Vorteile Längsschnittmodell
Gesamtbestand der IAB-Betriebspanel-Betriebe kann analysiert werden schneller und relativ unkomplizierter Aufbau von Paneldatensätzen leichte und unkomplizierte Aggregation der Personendaten auf die Betriebsebene kürzere Rechenzeiten	Vor- und Nachgeschichte der Personen ist nachvollziehbar bei Panelmortalität liegen die Daten für die Betriebe in der Stichprobe auf der Betriebs- und Personenebene für einen gewissen Zeitraum weiterhin vor Informationen über Personen, die den Betrieb wechseln Verknüpfungsqualität von Betriebs- und Personendaten bereits geprüft

### 5.3 Mengengerüste

#### Fallzahlen im LIAB-Querschnittmodell, Version 1

Jahr	Personen	Betriebe mit Befragung	Betriebe mit Befragung und Verknüpfung	Betriebe einschließlich derer ohne Befragung
1993	2.585.830	4.265	4.189	10.805
1994	2.224.836	4.154	4.041	9.284
1995	1.885.357	4.114	3.934	9.011
1996	2.473.552	8.604	8.292	13.577
1997	2.122.084	8.917	8.385	12.946
1998	2.081.565	9.334	8.788	13.683
1999	2.020.784	9.915	9.433	54.000
2000	2.345.932	14.083	13.462	69.321
2001	2.538.832	15.783	14.982	73.294
2002	2.398.800	15.682	14.700	66.901
2003	2.181.089	16.165	15.030	72.307
2004	2.384.071	16.063	14.821	97.423
2005	2.352.338	16.280	14.675	95.035

Anmerkungen: Die Zahl der Datenzeilen in den Personendaten ist höher als die hier ausgewiesene Zahl der Personen, da es parallele Beschäftigung und parallelen Leistungsbezug gibt, die jeweils in eigene Datenzeile erhalten. Betriebe mit Befragung sind alle Betriebe des IAB-Betriebspanels mit den Wellenkennzeichen A, B, C, D, E und G. Nicht für alle diese Betriebe lässt sich eine Verknüpfung zu den Personendaten herstellen (siehe Abschnitt 6). Die letzte Spalte bezieht sich auf die Anzahl unterschiedlicher Betriebsnummern in den Personendaten, die aus paralleler Beschäftigung in Nicht-IAB-Betriebspanel-Betrieben resultiert.

Es fällt auf, dass sich die Anzahl der Betriebe über die Zeit vervierfacht hat, während die Zahl der dort Beschäftigten Personen etwa konstant geblieben ist. Ursache ist der Rückgang des Anteils von Großbetrieben im IAB-Betriebspanel.

## Fallzahlen im LIAB-Längsschnittmodell, Version 1

Jahr	Personen	Betriebe mit Befragung	Betriebe mit Befragung und Verknüpfung	Betriebe einschließlich derer ohne Befragung
1990	662.626	-	-	95.346
1991	949.762	-	-	129.095
1992	1.263.357	-	-	151.336
1993	1.295.660	905	882	152.819
1994	1.315.607	993	971	152.386
1995	1.353.374	1.099	1.083	143.602
1996	1.378.157	2.449	2.428	134.440
1997	1.409.216	2.723	2.695	134.649
1998	1.434.997	3.322	3.297	137.800
1999	1.483.394	4.264	4.247	187.303
2000	1.498.162	4.264	4.258	191.736
2001	1.492.905	4.264	4.201	189.410
2002	1.461.541	3.656	3.525	194.389
2003	1.422.898	3.256	3.096	207.312
2004	1.395.693	2.889	2.708	226.367
2005	1.200.211	2.586	2.357	220.502

Anmerkungen: Die Auszählungen beziehen sich nur auf Spells, die den 30.6. des Jahres einschließen. Die letzte Spalte enthält auch Betriebe, die sich nur in den Personendaten- oder nur in den Betriebsdaten finden.

Entsprechende Auszählungen für die weiteren LIAB-Längsschnittmodelle folgen nach deren Aktualisierung.

## 6 Datenqualität: Verknüpfung von Betrieben und Personen

Informationen zur Qualität der einzelnen Merkmale finden sich bei den Variablenbeschreibungen. In diesem Kapitel werden ausschließlich die Probleme bei der Verknüpfung von Betriebs- und Personendaten angesprochen.

In einem nicht zu vernachlässigenden Teil der Fälle ist die Verknüpfung der Betriebs- und Personendaten über die Betriebsnummer falsch. D.h. die Angaben aus dem IAB-Betriebspanel beziehen sich auf einen anderen Betrieb als die Angaben zu den Personen aus der BLH, obwohl die Betriebsnummer übereinstimmt. Die falschen Verknüpfungen werden offensichtlich, wenn man Informationen vergleicht, die sowohl auf der Personen- als auch auf der Betriebsseite vorliegen, wie z.B. die Anzahl der sozialversicherungspflichtig Beschäftigten zum Befragungszeitpunkt.

Die folgende Tabelle gibt einen Eindruck der Häufigkeit mangelhafter Verknüpfungen. Zunächst wurden die Anzahl der sozialversicherungspflichtig Beschäftigten pro Betrieb anhand der Personendaten berechnet. Anschließend wurden diese Zahlen mit den Angaben laut Fragebogen aus dem Betriebspanel verglichen. Die Zahl der Beschäftigten bezieht sich jeweils auf den Befragungszeitpunkt. In der ersten Zeile sehen wir, dass für eine steigende Anzahl an Betrieben keine Verknüpfung mit den Personendaten möglich ist.<sup>7</sup> Zeile 2 enthält die Anzahl der Betriebe, für die eine Verknüpfung zu den Personendaten möglich ist, es aber laut Fragebogen zum Befragungszeitpunkt keine sozialversicherungspflichtig Beschäftigten gibt. Für die übrigen Betriebe geben die Zeilen 3 bis 9 an, ob bzw. wie stark die Beschäftigtenzahlen nach Personendaten von den Angaben im Fragebogen des Betriebspanels abweichen. Relativ zur Gesamtzahl der Betriebe nehmen die geringen Abweichungen ab und die großen Abweichungen zu.

---

<sup>7</sup> Eine Ursache dürfte in der Dauer der Befragung liegen. Je länger ein Betrieb in der Befragung ist, desto höher ist die Wahrscheinlichkeit, dass sich zwischenzeitlich von uns ungesehen die Betriebsnummer ändert. Aber auch bei Neubefragten Betrieben beobachten wir eine steigende Zahl fehlender Verknüpfungen. Der hohe Anteil fehlender Verknüpfungen in 2005 liegt auch daran, dass zum Zeitpunkt der Erstellung noch nicht alle Meldungen der Betriebe vorliegen.

---

## Verknüpfungsqualität

	1993	1994	1995	1996	1997	1998	1999
keine Verknüpfung	76	113	180	312	532	546	482
keine soz-pfl. laut F	2	4	5	9	18	25	83
keine Abweichung	848	762	605	1,823	2,308	2,338	1,847
Abweichung >0 - 5%	1,345	1,072	951	2,245	1,926	2,163	2,159
Abweichung 5% - 10%	778	718	647	1,329	1,118	1,277	1,323
Abweichung 10% - 20%	538	633	662	1,238	1,192	1,182	1,269
Abweichung 20% - 50%	439	553	626	1,045	1,110	1,105	1,466
Abweichung 50% -100%	185	233	349	470	545	541	762
Abweichung >100%	54	66	89	133	168	157	524
<b>Total</b>	<b>4,265</b>	<b>4,154</b>	<b>4,114</b>	<b>8,604</b>	<b>8,917</b>	<b>9,334</b>	<b>9,915</b>
	2000	2001	2002	2003	2004	2005	2006
keine Verknüpfung	621	801	982	1,101	1,242	1,423	1,733
keine soz-pfl. laut F	94	124	139	188	228	247	261
keine Abweichung	2,591	2,813	2,764	2,911	2,708	2,649	2,496
Abweichung >0 - 5%	2,840	3,159	2,900	2,751	2,787	2,753	2,541
Abweichung 5% - 10%	1,798	1,931	1,878	1,827	1,678	1,748	1,674
Abweichung 10% - 20%	1,862	2,094	2,052	2,043	1,947	1,969	1,818
Abweichung 20% - 50%	2,180	2,474	2,517	2,518	2,516	2,472	2,402
Abweichung 50% -100%	1,127	1,325	1,373	1,521	1,529	1,518	1,567
Abweichung >100%	970	1,062	1,077	1,305	1,428	1,501	1,484
<b>Total</b>	<b>14,083</b>	<b>15,783</b>	<b>15,682</b>	<b>16,165</b>	<b>16,063</b>	<b>16,280</b>	<b>15,976</b>

### Auszählung auf Basis des LIAB-Querschnittmodellels, Version 1

Die Gründe für die falschen Verknüpfungen liegen zum einen im Verfahren der Vergabe der Betriebsnummern und zum anderen in der Zielrichtung des IAB-Betriebspanels. Die zu befragenden IAB-Betriebspanel-Betriebe werden geschichtet nach Branchen und Betriebsgröße aus der Beschäftigtenstatistik der BA gezogen. Bei der Erstbefragung gleicht der Interviewer neben anderen Merkmalen auch die Betriebsnummer ab, um sicher zu stellen, dass er den richtigen Betrieb befragt. Bei Wiederholungsbefragungen erfolgt der Abgleich über die Betriebsnummer nicht mehr, es wird über andere Merkmale sichergestellt, dass der gleiche Betrieb befragt wird wie im Vorjahr.<sup>8</sup> Wenn nun die Betriebsnummer z.B. durch einen neuen Eigentümer gewechselt hat, wird dies beim Interview nicht erfasst. Für das IAB-Betriebspanel ist dies i.d.R. unerheblich, sofern es sich um dieselbe Betriebseinheit handelt. Bei den LIAB-Daten führt dies jedoch dazu, dass ab dieser Welle entweder keine oder die falschen Personendaten verknüpft werden. Ein weiteres Problem liegt darin, dass eine Betriebsnummer nicht immer eine sinnvolle wirtschaftliche Einheit darstellt. So können mehrere Betriebe an einem Ort unter derselben Betriebsnummer

<sup>8</sup> Der Versuch den Abgleich der Betriebsnummer auch bei Wiederholungsbefragungen einzuführen, hat zu hohem Antwortausfall geführt, weshalb im Weiteren darauf verzichtet wurde.

geführt werden, z.B. eine Hauptgeschäftsstelle und mehrere Filialen. Das Interview wird dann mit einer Betriebseinheit geführt, im Idealfall mit der Hauptgeschäftsstelle. Alle Angaben im IAB-Betriebspanel beziehen sich folglich auf diese Einheit. Die verknüpften Personendaten beziehen sich in diesem Fall aber auf alle Betriebseinheiten mit dieser Betriebsnummer. Vorschläge zur Überprüfung der Verknüpfungsqualität finden sich in Abschnitt 9.2.2.

---

## 7 Variablenbeschreibung zu den Personendaten

### 7.1 Allgemeiner Teil

Die Variablen in den Personendaten der verschiedenen LIAB-Versionen unterscheiden sich nicht und werden daher gemeinsam beschrieben. Für die Datendokumentationen der Variablen des IAB-Betriebspanels wird auf das entsprechende Internet-Angebot des FDZ verwiesen.

Es gibt die folgende Kodierung für fehlende Werte:

-9	Keine Zuordnung möglich
-7	Keine Angabe
-6	Nicht zutreffend
-5	Quelle nicht gefüllt

Systemmissings gibt es ausschließlich bei Datumsangaben. In Ausnahmefällen werden unbekannte oder fehlende Werte in einzelnen Variablen als eine eigene Ausprägung explizit ausgewiesen.

Es gibt Veränderungen der Bedeutung von Variablen bzw. einzelnen Ausprägungen im Zeitablauf, die in den Abschnitten der Berechnungen der einzelnen Variablen erläutert werden. Viele Änderungen resultieren aus der Umstellung des Meldeverfahrens 1999. Begrenzte Gültigkeitsdauern haben insbesondere die Angaben zu Wirtschaftszweigen und Orten.

#### 7.1.1 Merkmalsliste

Die folgende Übersicht zeigt die Liste der Merkmale in den LIAB-Personendaten. Einige der Merkmale sind nur für Beobachtungen gefüllt, die aus den Beschäftigtendaten (BeH) oder den Leistungsempfängerdaten (LeH) stammen. So gibt es die Betriebsnummer nur für Beobachtungen zu Beschäftigung, die Leistungsart nur für Beobachtungen zum Leistungsbezug. Ein grau hinterlegtes Feld bedeutet, dass das Merkmal in der Datenquelle nicht gefüllt ist (Kodierung -5). Schließlich ist eingetragen, wenn Merkmale nur für einen bestimmten Zeitraum (konsistent) verfügbar sind.

### Merkmale in den LIAB-Personendaten

	Variablenname	BeH	LeH
<b>Identifikatoren</b>			
Personen-ID	id		
Betriebs-ID nach Betriebspanel	idnum		
Betriebs-ID nach BA-BNR	betnr		
<b>Technische Merkmale</b>			
Spellbeginn original	begorig		
Spellende original	endorig		
Spellbeginn Episode	begepi		
Spellende Episode	endepi		
Quelle	quelle		
Spellzähler der Person	spell_nr		
Spellzähler pro Episode und Quelle	level1		
Status der Betriebsnummer	betr_st		
Datumsbereinigung	dat_ber		
<b>Merkmale zur Person</b>			
Geschlecht	geschl		
Geburtsjahr	gebjahr		
Staatsangehörigkeit, gruppiert	nation		
Staatsangehörigkeit, original (sensibel)	nationor		
Ausbildung	ausbild		
Familienstand	fam		
Kinder	kind		
<b>Beschäftigung und Leistungsbezug</b>			
Personengruppe	pers_gr	ab 99	
Abgabegrund der Meldung	abg_gr	ab 99	
Stellung im Beruf	berstell		
Beruf	beruf		
Tagesentgelt/Leistungssatz	tag_entg		
Leistungsart	leistart		
Bewilligungsart	bew_art		
Abmeldegrund Leistungsbezug	abm_gr		
Wirtschaftszweig WZ73, 3-steller (sensibel)	wz73	93 - 02	
Wirtschaftszweig WZ93, 5-steller (sensibel)	wz93	99 - 03	
Wirtschaftszweig WZ03, 5-steller (sensibel)	wz03	ab 03	
<b>Generierte biographische Merkmale</b>			
Eintrittsdatum Erwerbsleben	ein_erw		
Eintrittsdatum Betrieb	ein_bet		
Dauer Betrieb	tage_bet		
Anzahl bisheriger Leistungsbezüge	anz_1st		
Dauer bisheriger Leistungsbezüge	tage_1st		
Erwerbsstatus 5 Jahre nach Stichtag (nicht in LM)	estat5jn	bis 00	
Erwerbsstatus 8 Tage vor Eintritt in den Betrieb	estat8tv		
Lohn 5 Jahre nach Stichtag (nicht in LM)	lohn5jn	bis 00	

	Variablenname	BeH	LeH
Lohn 8 Tage vor Eintritt in den Betrieb	lohn8tv		
<b>Ortsangaben</b>			
Arbeitsort Arbeitsagentur (sensibel)	arb_aa		
Arbeitsort Kreis (sensibel)	arb_kr		
Wohnort Arbeitsagentur (sensibel)	wohn_aa	ab 99	
Wohnort Kreis (sensibel)	wohn_kr	ab 99	
Ost-West-Kennzeichen, BeH: Arbeitsort, LeH: Wohnort	west		

### 7.1.2 Sensible Merkmale

Merkmale, die das Identifikationsrisiko von Personen oder Betrieben im Datensatz in besonderem Maße erhöhen, bezeichnen wir als sensible Merkmale. Beim LIAB sind dies folgende Merkmale, die alle aus den Personendaten stammen:

- Staatsangehörigkeit, original
- Wirtschaftszweig WZ73, 3-steller
- Wirtschaftszweig WZ93, 5-steller
- Wirtschaftszweig WZ03, 5-steller
- Arbeitsort Arbeitsagentur
- Arbeitsort Kreis
- Wohnort Arbeitsagentur
- Wohnort Kreis

Die sensiblen Merkmale stehen im Gastaufenthalt für Auswertungen nur dann zur Verfügung, wenn deren Erforderlichkeit schon im Antrag gesondert begründet und im Nutzervertrag festgehalten wird.

## 7.2 Beschreibung der einzelnen Variablen

### 7.2.1 Identifikatoren

#### 7.2.1.1 Personen-ID (id)

Herkunft	generiert aus der Sozialversicherungsnummer der BLH
Gefüllt für	BeH, LeH
Zeitbezug	personenbezogen
Detailbeschreibung	Der systemfreie Personenidentifikator wird aus der Sozialversicherungsnummer generiert und identifiziert Personen.
Besonderheiten	Die Personen-ID wird für jede Version und bei Updates neu erzeugt, so dass Personen aus verschiedenen LIAB-Versionen oder Updates einander nicht zuzuordnen sind (vgl. Abschnitt 8).

### 7.2.1.2 Betriebs-ID nach Betriebspanel (idnum)

Herkunft	IAB-Betriebspanel und generiert aus BLH-Betriebsnummer
Gefüllt für	BeH
Zeitbezug	Originalspell
Detailbeschreibung	Die Betriebs-ID nach Betriebspanel wird - sofern vorhanden - aus dem IAB-Betriebspanel übernommen. Für Betriebe, die keine idnum haben, da sie nie am IAB-Betriebspanel teilgenommen haben, werden zufällige Betriebsnummern vergeben. Diese zufälligen idnums sind größer als 300.000. Das Merkmal Status der Betriebsnummer (betr_st) gibt an, ob es sich um einen Betriebspanel-Betrieb handelt oder nicht.
Besonderheiten	Die Betriebs-ID für nicht-Betriebspanel-Betriebe wird für jede Version und bei Updates neu erzeugt, so dass diese Betriebe aus verschiedenen LIAB-Versionen oder Updates einander nicht zuzuordnen sind (vgl. Abschnitt 8).

### 7.2.1.3 Betriebs-ID nach BA-BNR (betnr)

Herkunft	generiert aus BLH-Betriebsnummer
Gefüllt für	BeH
Zeitbezug	Originalspell
Detailbeschreibung	In den ersten Jahren des Betriebspanels wurden Betriebe, die einmal teilgenommen hatten, dann aber nicht mehr, bei Ziehungen von Betrieben für Aufstockungen nicht berücksichtigt. Dies führte dazu, dass in manchen Zellen (Betriebsgrößen-/Branchenkombinationen) schließlich keine Betriebe für Aufstockungen mehr blieben. Daher wanderten "gebrauchte" Betriebsnummern nach einer Karenzzeit wieder in den Topf und die Betriebe wurden erneut angeschrieben. Nehmen diese Betriebe (erneut) an der Befragung teil, erhalten sie eine neue idnum. In der Regel wird es sich aber um den gleichen Betrieb handeln, der schon vor Jahren teilgenommen hat. Da wir dies nicht eindeutig klären können, stellen wir neben der idnum die Betriebs-ID nach BA-BNR bereit, die kennzeichnet, welche Betriebe die gleiche BA-Betriebsnummer haben, obwohl die idnum im Zeitablauf wechselt. Haben zwei Betriebspanel-Betriebe (2 idnums) die gleiche betnr, muss der Forscher durch einen Vergleich der Merkmale im Zeitablauf beurteilen, ob es sich um den gleichen Betrieb handelt wie zuvor.
Besonderheiten	Betriebs-ID nach BA-BNR (betnr) wird für jede Version und bei Updates neu erzeugt, so dass diese Betriebe aus verschiedenen LIAB-Versionen oder Updates einander nicht zuzuordnen sind (vgl. Abschnitt 8).

## 7.2.2 Technische Merkmale

### 7.2.2.1 Spellbeginn original (begorig)

Siehe Abschnitt 7.2.2.4.

### 7.2.2.2 Spellende original (endorig)

Siehe Abschnitt 7.2.2.4.

### 7.2.2.3 Spellbeginn Episode (begepi)

Siehe Abschnitt 7.2.2.4.

## 7.2.2.4 Spellende Episode (endepi)

Herkunft	BLH
Gefüllt für	BeH und LeH
Zeitbezug	begorig, endorig: Originalspell begepi, endepi: Episode
Detailbeschreibung	begorig und begepi geben tagesgenau den Beginn und das Ende des Originalspells an, begepi und endepi den Beginn und das Ende der überschneidungsfreien, gesplitteten Episoden. Beginn- und Endedatum der Spells sind wie alle Datumsangaben ganzzahlig gespeichert, wobei der 31.12.1959 den Wert -1 annimmt, der 1.1.1960 den Wert 0, der 2.1.1960 den Wert 1 u.s.w.
Besonderheiten	Wenn eine Meldung aufgrund von Überlappungen durch das Episodensplitting auf mehrere Episoden aufgeteilt wurde, dann stimmt der Beginn des Originalspells mit dem Beginn der <u>ersten</u> gesplitteten Episode überein und das Ende des Originalspells entspricht dem Ende der <u>letzten</u> gesplitteten Episode. Alle anderen Beginn- und Endedaten der Originalspells und der gesplitteten Episoden weichen voneinander ab. Sollen sich Auswertungen auf die ungesplitteten Spells beziehen, lässt sich ein Filter beispielsweise durch die Bedingung begorig = begepi bilden. Achtung: Manche Merkmale beziehen sich auf die Zeitangaben der Originalspells, nicht für die durch das Episodensplitting hinzugefügten Spells (abg_gr, abm_gr). Durch das Episodensplitting ergeben sich Episoden, die schon vor dem Beobachtungszeitraum einer Quelle beendet sind oder erst nach dem Beobachtungszeitraum einer Quelle beginnen.

## 7.2.2.5 Quelle (quelle)

Herkunft	BLH
Gefüllt für	BeH und LeH
Zeitbezug	Originalspell
Detailbeschreibung	Die Variable gibt an, ob es sich um eine Beschäftigungs- oder Leistungsmeldung handelt.

1	Beschäftigungsmeldung
2	Leistungsmeldung

## 7.2.2.6 Spellzähler der Person (spell\_nr)

Herkunft	BLH
Gefüllt für	BeH und LeH
Zeitbezug	Episode
Detailbeschreibung	Der Spellzähler zählt die Spells pro Person in der BLH, beginnend mit 1. Im LIAB-Querschnittmodell sind nur die Spells enthalten, die den 30.6. überlappen, daher ist die Nummerierung im Gegensatz zum LIAB-Längsschnittmodell nicht laufend.

## 7.2.2.7 Spellzähler pro Episode und Quelle (level1)

Herkunft	BLH
Gefüllt für	BeH und LeH
Zeitbezug	Episode
Detailbeschreibung	Der Spellzähler pro Episode und Quelle nummeriert parallele Spells einer

	<p>Episode von Null bis n.</p> <p>Dabei gilt folgende Sortierung: Innerhalb der Quelle BeH sind die sozialversicherungspflichtigen Beschäftigungsverhältnisse immer vor den geringfügigen einsortiert und innerhalb der gleichen Art von Beschäftigung hat stets diejenige mit dem höheren Tagesentgelt den Vorrang. Innerhalb der Quelle LeH haben die Haupt-Leistungsarten ALG, ALHI und UHG Vorrang vor den restlichen Ausprägungen.</p>
Besonderheiten	Die Variable wird u.a. dazu genutzt, um die Haupt- und Nebenbeschäftigung zu identifizieren. Die Hauptbeschäftigung hat immer den Wert Null, alle anderen einen Wert größer Null

#### 7.2.2.8 Status der Betriebsnummer (betr\_st)

Herkunft	generiert
Gefüllt für	BeH
Zeitbezug	Originalspell
Detailbeschreibung	<p>Die Variable zeigt an,</p> <ul style="list-style-type: none"> <li>– ob es sich die Beschäftigtenmeldung auf einen IAB-Betriebspanel-Betrieb bezieht,</li> <li>– ob dieser Betrieb in dieser LIAB-Version enthalten ist und</li> <li>– ob dieser Betrieb in diesem Jahr enthalten ist oder nicht.</li> </ul> <p>Im Querschnittmodell ergeben sich Beobachtungen in nicht-IAB-Betriebspanel-Betrieben durch parallele Beschäftigungen. Im Längsschnittmodell sind es zusätzlich frühere und spätere Beschäftigungsverhältnisse.</p>
Besonderheiten	Ausprägung 3 kommt im Querschnittmodell nicht vor, da im Gegensatz zum Längsschnittmodell alle IAB-Betriebspanel-Betriebe enthalten sind.

1	IAB-Betriebspanel-Betrieb in dieser LIAB-Version in diesem Jahr
2	IAB Betriebspanel-Betrieb in dieser LIAB-Version in anderen Jahr
3	IAB-Betriebspanel-Betrieb nicht in dieser LIAB-Version
4	Kein IAB-Betriebspanel-Betrieb

#### 7.2.2.9 Datumsbereinigung (dat\_ber)

Herkunft	BLH
Gefüllt für	LeH
Zeitbezug	Episode
Detailbeschreibung	<p>Bei der Bereinigung des Enddatums der Leistungsempfangsmeldungen wird dieses Kennzeichen gesetzt, das dazu dient, die betroffenen Datensätze und die jeweils durchgeführte Art der Bereinigung identifizieren zu können.</p> <p>Das Feld wurde als Bitmuster angelegt, so dass auch Kombinationen von Bereinigungen abgebildet werden können. Wenn eine Bereinigung durchgeführt wird, wird der entsprechende Wert der nachfolgenden Tabelle aufaddiert. 2006 wurden 7 neue, binäre Ausprägungen hinzugefügt.</p> <p>Beispiel: Der Wert 5 bedeutet, dass sowohl die Dauer von -1 als auch das Enddatum der Meldung durch die nachfolgende Meldung korrigiert wurde.</p>

0	keine Bereinigung
1	Dauer -1 korrigiert
2	Meldung liegt vollständig innerhalb des Zeitraums der vorhergehenden Meldung
4	Enddatum der Meldung durch nachfolgende Meldung korrigiert
8	Enddatum mit Missingwert wird durch nachfolgende Meldung korrigiert
16	Verkürzung durch Nachfolger

32	Spells mit gleicher Leistungsart und gleicher Leistungshöhe ohne Lücke wurden zu einem Spell zusammengefügt
64	leeres Enddatum im Anspruch wurde durch höchstes Enddatum der Zahlung ersetzt
128	LeH Datensatz identisch mit DWH Datensatz und wird deshalb gelöscht
256	Folgesatz in aktuellem Satz enthalten
512	aktueller Satz in Folgesatz enthalten
1024	Korrektur am 01.07.2004
2048	LeH Satz parallel zu DWH Satz

### 7.2.3 Merkmale zur Person

#### 7.2.3.1 Geschlecht (geschl)

Herkunft	BLH (generiert aus der Sozialversicherungsnummer)
Gefüllt für	BeH und LeH
Zeitbezug	personenbezogen
Detaillbeschreibung	Es gibt Frauen und Männer.

1	Männlich
2	Weiblich

#### 7.2.3.2 Geburtsjahr (gebjahr)

Herkunft	BLH (generiert aus der Sozialversicherungsnummer)
Gefüllt für	BeH und LeH
Zeitbezug	personenbezogen
Detaillbeschreibung	Das Geburtsjahr wird der Versicherungsnummer entnommen und bei Bedarf nach formalen Kriterien bereinigt. Das Jahrhundert wird logisch zum zeitlich ersten Satz eines jeden BLH-Kontos gebildet: Zunächst wird '19' als Jahrhundertangabe angenommen, sollte das Beschäftigungsjahr dadurch jedoch vor dem Geburtsjahr liegen, wird vom Geburtsjahr 100 subtrahiert. Anschließend wird das so ermittelte Geburtsjahr auf das gesamte Versichertenkonto übertragen, so dass das Geburtsjahr Konten übergreifend konstant ist.

#### 7.2.3.3 Staatsangehörigkeit, gruppiert (nation)

Siehe Abschnitt 7.2.3.4.

#### 7.2.3.4 Staatsangehörigkeit, original (nationor)

Herkunft	BLH
Gefüllt für	BeH und LeH
Zeitbezug	Originalspell
Detaillbeschreibung	Der Staatenschlüssel wird vom Statistischen Bundesamt gepflegt. Wie bieten das Merkmal in zwei Aggregationsgraden an: nation ist leicht aggregiert, während nationor die volle Information der BLH enthält. nationor ist ein sensibles Merkmal und muss daher im Antrag auf Datenzugang gesondert begründet werden.
Besonderheiten	Da die Staatsangehörigkeit nicht immer eingetragen ist, werden Lücken nach folgendem Algorithmus gefüllt: Fehlt die Staatsangehörigkeit in einem Spell, so wird innerhalb dieses Kontos ein Vorgänger und ein Nachfolger gesucht.

	Haben Vorgänger und Nachfolger dieselbe Ausprägung, so wird die Lücke mit diesem Eintrag gefüllt.
--	---

### 7.2.3.5 Ausbildung (ausbild)

Herkunft	BLH
Gefüllt für	BeH
Zeitbezug	Originalspell
Detailbeschreibung	<p>Bei dem Signierschlüssel 'Ausbildung' handelt es sich um einen kombinierten Schlüssel, durch den sowohl die erreichte Schulbildung als auch die abgeschlossene Berufsausbildung (jeweils höchster Abschluss) der/des Beschäftigten erhoben werden soll. Dabei ist grundsätzlich zunächst die Schulbildung festzustellen, danach die Berufsausbildung. Lediglich bei den Schlüsselpositionen 5 (Abschluss einer Fachhochschule) und 6 (Hochschul-/Universitätsabschluss) wird auf die Feststellung einer eventuell sonstigen Berufsausbildung verzichtet.</p> <p>Nicht als Schul- oder Berufsausbildung zählen berufliche Fortbildungen, z.B. Kurse an Sprach- und Dolmetscherschulen sowie Fachkenntnisse, die durch praktische Tätigkeiten erworben wurden.</p>
Besonderheiten	<p>„Für bestimmte Teilgruppen gibt es einen hohen Anteil fehlender Angaben, weil das Merkmal versicherungsrechtlich keine besondere Bedeutung hat [das betrifft z.B. die geringfügig entlohnten Beschäftigten; H.A.] (...). Zusammen mit einem Betriebswechsel treten manchmal 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 manchmal erst von einem neuen Beschäftigungsbetrieb erfasst“ (Meinken/Koch, 2004).</p> <p>Auffällig ist, dass die Ausprägung „Ausbildung unbekannt/fehlt“ zum aktuellen Rand hin deutlich zunimmt. Einen möglichen Umgang mit dem Problem inkonsistenter Werte beschreiben Osikominu/Fitzenberger/Völter (2004).</p>

1	bis mittlere Reife ohne Berufsausbildung
2	bis mittlere Reife mit Berufsausbildung
3	(Fach-)Hochschulreife ohne Berufsausbildung
4	(Fach-)Hochschulreife mit Berufsausbildung
5	Fachhochschulabschluss
6	Hochschulabschluss
7	Ausbildung unbekannt

### 7.2.3.6 Familienstand (fam)

Herkunft	BLH
Gefüllt für	LeH
Zeitbezug	Originalspell
Detailbeschreibung	Zeigt für LeH-Beobachtungen an, ob die Person verheiratet ist oder nicht.

0	nicht verheiratet
1	verheiratet

## 7.2.3.7 Kinder (kind)

Herkunft	BLH
Gefüllt für	LeH
Zeitbezug	Originalspell
Detailbeschreibung	Zeigt für LeH-Beobachtungen an, ob die Person eins oder mehrere Kinder hat oder nicht.
Besonderheiten	keine

0	Nein
1	Ja

## 7.2.4 Beschäftigung und Leistungsbezug

## 7.2.4.1 Personengruppe (pers\_gr)

Herkunft	BLH
Gefüllt für	BeH (konsistent ab 1999)
Zeitbezug	Originalspell
Detailbeschreibung	<p>Der Personengruppe ist seit Einführung des neuen Meldeverfahrens 1999 in den Beschäftigtenmeldungen enthalten. Bei Nachmeldungen nach dem neuen DEÜV-Verfahren, die sich auf vorhergehende Jahre beziehen, ist er ebenfalls enthalten. Für Meldungen aus dem alten DEVO/DÜVO, die vor 1999 eingegangen sind, wird versucht, aus den Merkmalen Tätigkeitsschlüssel und Betriebsnummer eine Zuordnung zu den Personengruppen durchzuführen. In vielen Fällen ist dies jedoch nicht eindeutig möglich.</p> <p>Zu beachten ist: Während in der BeH bis zum Jahr 1998 aus Vollständigkeitsgründen alle Personengruppen enthalten sind, ab 1999 hingegen nur alle mit dem JZM gelieferten Gruppen, sind in der BLH für alle Jahre durchgängig nur die JZM-Personengruppen 101, 102, 103, 105, 106, 109, 112, 118, 119, 120, 140, 141, 142, 143, 201, 203, 209 enthalten, sowie für die Jahre bis 1998 alle Meldungen ohne Angabe zur Personengruppe, da diese hier relativ häufig sind und nicht unterschlagen werden dürfen. Bei der Betrachtung von Zeitreihen oder beim Vergleich von BeH mit BLH müssen gegebenenfalls entsprechende Personengruppen gefiltert werden.</p>
Besonderheiten	Die Variable kann ab 1999 dazu verwendet werden, über die Werte 109 und 209 geringfügig Beschäftigte zu identifizieren.

101	Sozialvers. Besch. ohne bes. Merkmale
102	Auszubildende
103	Beschäftigte in Altersteilzeit
104	Hausgewerbetreibende
105	Praktikanten
106	Werkstudenten
107	Behinderte Menschen in Werkstätten
108	Bezieher von Vorruhestandsgeld
109	Geringfügig entlohnte Beschäftigte
110	Kurzfristig Beschäftigte
111	Pers. in Einrichtungen für behind. Menschen
112	Mitarbeit. Familienang. in der Landwirtschaft
113	Nebenerwerbslandwirte

114	Nebenerwerbslandwirte - saisonal beschäftigt
116	Ausgleichsgeldempfänger nach dem FELEG
118	Unständig Beschäftigte
119	Versicherungsfreie Altersvollrentner und Versorgungsbezieher wegen Alters
120	Personen, bei denen eine Beschäftigung vermutet wird
140	Seeleute
141	Auszubildende in der Seefahrt
142	Seeleute in Altersteilzeit
143	Seelotsen
201	Mit Haushaltsscheck-Verfahren gemeldete versicherungspflichtig Beschäftigte
202	Kurzfristig Beschäftigte
203	Versicherungspflichtige Künstler und Publizisten
204	Teilnehmer an Leistungen zur Teilhabe am Arbeitsleben
205	Unständig Beschäftigte
207	Pflegepersonen i.S. von § 19 SGB XI ohne Beihilfeberechtigung des Pflegebedürftigen
208	Pflegepersonen i.S. von § 19 SGB XI mit Beihilfeberechtigung des Pflegebedürftigen
209	Mit Haushaltsscheck-Verfahren gemeldete geringfügig entlohnte Beschäftigte
210	Mit Haushaltsscheck-Verfahren gemeldete kurzfristig Beschäftigte
301	Grundwehrdienstleistende
302	Wehrübungsleistende
303	Zivildienstleistende
304	Ableistende eines freiwilligen sozialen bzw. ökologischen Jahres anstelle des Zivildienstes

#### 7.2.4.2 Abgabegrund der Meldung (abg\_gr)

Herkunft	BLH
Gefüllt für	BeH, konsistent ab 1999
Zeitbezug	Originalspell
Detailbeschreibung	<p>In BeH-Meldungen kennzeichnet der Abgabegrund den Anlass, aus dem der Arbeitgeber den Sozialversicherungsträgern die betreffende Beschäftigungsmeldung erstattet hat. Die BLH enthält nur bei 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.</p> <p>Die Abgabegründe für Beschäftigungsmeldungen sind nach den Regeln des seit 1.1.1999 geltenden Meldeverfahrens (gemäß DEÜV) verschlüsselt. Im Vergleich zum alten Meldeverfahren (nach den DEVO-/DÜVO-Regeln) werden die Abgabegründe nach dem neuen Recht differenzierter erfasst. Alte Gründe sind auf die neuen umgeschlüsselt.</p> <p>Für LeH-Meldungen zeigt das Merkmal „Abgabegrund“ den Grund für das Ende des Bezugs von Arbeitslosengeld, -hilfe oder Unterhaltsgeld an.</p>
Besonderheiten	<p>Im LIAB-Querschnittmodell wird nicht empfohlen, den Abgabegrund „Ende der Beschäftigung“ als Austritt aus dem Betrieb zu werten, denn in einigen Fällen erfolgt direkt im Anschluss an die Abmeldung eine Wiederanmeldung bei dem gleichen Arbeitgeber. Das lässt sich im LIAB-Längsschnittmodell nachweisen, im LIAB-Querschnittmodell nicht. LEH-Meldungen: Für viele Analysezwecke wird bei Leistungsbezugsmeldungen die Verwendung der Variable abm_gr ausreichend sein.</p>

30	Abmeldung wegen Ende einer Beschäftigung
31	Abmeldung wegen Krankenkassenwechsel
32	Abmeldung wegen Beitragsgruppenwechsel
33	Abmeldung wegen sonstiger Gründe
34	Abmeldung wegen Ende einer soz.vers.pfl. Beschäftigung nach einer Unterbrechung von > 1 Monat
35	Abmeldung wegen Arbeitskampf von länger als 1 Monat
36	Abmeldung wegen Wechsel des Entgeltabrechnungssyst./Währungsumstellung während eines Kalenderjahres
40	Gleichzeitige An- und Abmeldung wegen Ende der Beschäftigung
49	Abmeldung wegen Tod
50	Jahresmeldung
51	Unterbrechungsmeldung wegen Bezug von bzw. Anspruch auf Entgeltersatzleistungen
52	Unterbr.-meld. wg. Erziehungsurlaub
53	Unterbr.-meld. wg. gesetzlicher Dienstpflicht
54	Meldung eines einmalig gezahlten Arbeitsentgelts
55	Meldung von nicht vereinbarungsgemäss verwendeten Wertguthaben (Störfall)
56	Meldung des Unterschiedsbeitrags bei Entgeltersatzleistungen während Altersteilzeit
59	Entgeltmeldung für unständig Beschäftigte
60	Änderung des Namens
61	Änderung der Anschrift
62	Änderung des Aktenzeichens/der Personalnummer des Beschäftigten
63	Änderung der Staatsangehörigkeit
70	Jahresmeldung für freigestellte Arbeitnehmer
71	Meldung des Vortages der Insolvenz/Freistellung
72	Entgeltmeldung zum rechtlichen Ende der Beschäftigung
89	Rückmeldung geringfügig Beschäftigter
90	Anforderung eines Sozialversicherungsausweises
1100	Eingliederungsgeld Wohnort-AA
1110	Arbeitsaufnahme
1130	Ausreise ohne E 303
1140	Mutterschaftsgeld
1150	Sonstige Gründe
1160	Bew. Abschnitt abgelaufen
1170	Sperrzeit (kurze Sperrzeit)
1180	Abbruch der Maßnahme
1117	Erlöschen des Anspruchs
1120	Leistungsempfänger verstorben
1121	Ausländische Sozialleistung
1122	Erwerbsminderungsrente (<15Std)
1123	Schulbesuch
1131	Teilarbeitslosengeld
1132	Fehlende Mitwirkung
1134	Fehlende Arbeitsbereitschaft
1135	Rentenbeantragung nicht erfolgt
1136	Beantragung Rehabilitation nicht erfolgt
1138	Erwerbsminderungsrente (15-30Std)
1199	Agenturzusammenlegung
1460	Meldung nicht erneuert

1470	Eingliederungsgeld, ab 2003 Erwinrente
1200	Anspruch erschöpft
1210	Bewilligung / Leistungsänd.-Diff-Nachz
1480	Keine Zahlung und keine KK-Abmeldung
1240	Einstellung
1241	keine Zahlung (programminterne Einstellung)
1490	Bestand (von ALG4)
1260	Gesetzl. Dienstpflicht
1270	Ortsabwesenheit
1280	UHG Wohnort-Arbeitsamt
1290	65. Lebensjahr
1300	Berufausbildungsbeihilfe
1310	Ende Leistungsfortz. / Anspruch Krankeng.
1320	Säumniszeit bis 2004. COLIBRI ab 2005
1330	Nichterscheinen Meldung
1340	bis 1998 Eingliederungsg, ab 2003 Zivild.
1350	Übergangsgeld
1360	Fortfall der Bund-ALHI, ab 2003 Wehrübung
1370	Ausbildung / Eingliederungsg. / Erwinrente
1380	Umzug
1390	Altersrente
1400	Unterhaltsgeld
1410	Kuraufenthalt
1420	Ausreise mit E 303
1430	Sperrzeit 12/6 Wochen
1440	Ablauf der Maßnahme
1450	Eigene Abmeldung
1171	Ablauf Bewilligungsabschnitt UHG
1211	Ablauf Bewilligungsabschnitt UHG
1500	Ende Leistungsfortz. / Anspruch Krankeng.
1510	and. Leistungssatz

#### 7.2.4.3 Stellung im Beruf (berstell)

Herkunft	BLH
Gefüllt für	BeH
Zeitbezug	Originalspell
Detailbeschreibung	<p>Die aktuelle berufliche Stellung des Beschäftigten wird vom Arbeitgeber im Rahmen der „Angaben zur Tätigkeit“ übermittelt. Das Merkmal „Berufsstellung“ unterscheidet zwischen Vollzeit- und Teilzeitbeschäftigten; maßgeblich ist das Verhältnis zwischen der arbeitsvertraglich vereinbarten und der betriebsüblichen Arbeitszeit. Nur für Vollzeitbeschäftigte liefert die Variable tatsächlich Angaben zur beruflichen Stellung, während für Teilzeitbeschäftigte nur qualitativ erfasst wird, ob sie mehr als die Hälfte der üblichen regelmäßigen Arbeitsstunden eines Vollzeitbeschäftigten geleistet haben oder nicht. Diese Grenze lag bis 1978 bei 20 Wochenstunden, zwischen 1979 und 1987 bei 15 Wochenstunden und seit 1988 bei 18 Wochenstunden.</p> <p>Die Unterscheidung zwischen vollzeitbeschäftigten Arbeitern (1,2) und Angestellten richtet sich ausschließlich nach der Art der Rentenversicherungspflicht: „Unterliegen Arbeitnehmer/innen der Versicherungspflicht in der bisherigen Arbeiterrentenversicherung, so sind sie nach der »Stellung im Beruf« als Arbeiter/innen (Schlüsselzahl 2) auszuweisen. Gehören die Arbeitnehmer/innen zum</p>

	<p>Personenkreis der bisherigen angestelltenversicherungspflichtigen Beschäftigten, so sind sie nach der »Stellung im Beruf« als Angestellte (Schlüsselzahl 4) zu kennzeichnen. Bei Beschäftigten, die von der Rentenversicherungspflicht befreit sind, ist zu entscheiden, zu welchem bisherigen Rentenversicherungszweig Versicherungspflicht bestünde, wenn ein Befreiungstatbestand nicht vorliegen würde“ (Bundesagentur für Arbeit, 2004).</p> <p>Falls mehrerer 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 (ebd.).</p>
--	---

0	in Ausbildung: 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 zu der Kategorie Beschäftigte in beruflicher Ausbildung.
1	nicht formal qualifiziert
2	Facharbeiter
3	Meister, Poliere (gleichgültig ob Arbeiter oder Angestellte)
4	Angestellter (aber nicht Meister im Angestelltenverhältnis)
7	Heimarbeiter, Hausgewerbetreibende
8	Teilzeit (bis zur Hälfte der Vollarbeitszeit, seit 1988: bis 18 Stunden)
9	Teilzeit (mehr als die Hälfte der Vollarbeitszeit, seit 1988: über 18 Stunden)

#### 7.2.4.4 Beruf (beruf)

Herkunft	BLH
Gefüllt für	BeH
Zeitbezug	Originalspell
Detailbeschreibung	<p>Die Berufsbezeichnung der aktuell vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Der Arbeitgeber verschlüsselt die Tätigkeit des Beschäftigten gemäß der „Klassifizierung der Berufe. Systematisches und alphabetisches Verzeichnis der Berufsbenennungen“, in dem ca. 25.000 Berufsbezeichnungen zu finden sind. Die Berufsordnung besteht aus einer 3-stelligen Kennziffer und umfasst rund 330 Berufe.</p> <p>Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die <i>überwiegend</i> ausgeübte Tätigkeit kennzeichnet.</p>

#### 7.2.4.5 Tagesentgelt/Leistungssatz (tag\_entg)

Herkunft	BLH
Gefüllt für	BeH (Tagesentgelt) und LeH (Leistungssatz)
Zeitbezug	Originalspell
Detailbeschreibung	<p>In BeH-Sätzen weist das Merkmal jeweils das Brutto-Tagesentgelt des Beschäftigten aus. Es wird berechnet aus dem vom Arbeitgeber gemeldeten Zeitragementgelt und der Dauer des Zeitraums der Originalmeldung in Kalendertagen und auf drei Nachkommastellen gerundet.</p> <p>In den Jahren 1975-1998 meldeten Arbeitgeber grundsätzlich nur das sozialversicherungspflichtige Arbeitsentgelt, also Entgelt oberhalb der Geringfügigkeitsgrenze und bis zur Beitragsbemessungsgrenze in der gesetzlichen Rentenversicherung. Mit der Einbeziehung der geringfügig Beschäftigten in das Meldeverfahren ab 1.1.1999 werden auch Entgelte unterhalb der Geringfügig-</p>

	<p>keitsgrenze ausgewiesen; die Beitragsbemessungsgrenze als obere Kapungsgrenze gilt weiterhin. Manchmal übersteigen die Entgeltangaben die Beitragsbemessungsgrenze dennoch. Dies dürfte in der Regel auf Jahressonderzahlungen zurückzuführen sein, die der Arbeitgeber bei den Jahres-, Unterbrechungs- oder Abmeldungen zum normalen Entgelt hinzurechnet. Dabei ist es 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 Beschäftigungszeit oder Entgelthöhe handeln, was aber aufgrund der Versicherungsrelevanz dieser Informationen unwahrscheinlich ist.</p> <p>Die Geringfügigkeits- und Beitragsbemessungsgrenzen unterscheiden sich von Jahr zu Jahr und in West- und Ostdeutschland (maßgeblich ist jeweils der Standort des Betriebs).</p> <p>Für LeH-Sätze gibt das Merkmal den täglichen Leistungssatz an.</p> <p>Alle Werte sind in Euro.</p>
Besonderheiten	<p>Tagesentgelte von Null entstehen für Zeiten, in denen ein Beschäftigungsverhältnis ohne Bezahlung fortgeführt wird, z.B. bei Erziehungszeiten, Krankheiten &gt;42 Tage oder Sabbaticals.</p> <p>Geringe Tagesentgelte bei Teilzeitbeschäftigten sind plausibel, bei Vollzeitbeschäftigten jedoch nicht, außer bei Auszubildenden oder Praktikanten. In den Daten des Meldeverfahrens haben Entgeltangaben eine hohe Zuverlässigkeit, die rein zu statistischen Zwecken erhobenen Informationen, wie die Stellung im Beruf, sind weniger zuverlässig. Bei Vollzeitbeschäftigung mit geringem Tagesentgelt dürften also entweder die Information Vollzeit oder die Datumsangaben der Meldung falsch sein.</p> <p>Im FDZ-Methodenreport Nr. 2 wird ein Verfahren zur Imputation von Entgelten oberhalb der vorgestellt. Es wird empfohlen, die auf der nachfolgenden Seite abgedruckten Grenzen um weitere zwei bis drei DM (Euro) nach unten (obere Grenze) bzw. oben (untere Grenze) zu korrigieren, weil sich bereits kurz vor der Beitragsbemessungsgrenze in den IAB-Personendaten Häufungspunkte befinden.</p>

#### 7.2.4.6 Leistungsart (leistart)

Herkunft	BLH
Gefüllt für	LeH
Zeitbezug	Originalspell
Detailbeschreibung	Die Variable fasst die vielfältigen Leistungsarten der Bundesagentur für Arbeit in vier Hauptkategorien zusammen.
Besonderheiten	Die Ausprägung 5 sind beispielsweise Zuschüsse zur Rentenversicherung, d.h. dort gibt es relativ häufig Überschneidungen mit Beschäftigungsmeldungen.

1	Arbeitslosengeld (ALG)
2	Arbeitslosenhilfe (ALH)
3	Unterhaltsgeld (UHG)
5	Zuschuss § 166b AFG/§ 207a SGB III

## 7.2.4.7 Bewilligungsart (bew\_art)

Herkunft	BLH
Gefüllt für	LeH
Zeitbezug	Originalspell
Detailbeschreibung	Das Merkmal enthält die Bewilligungsgründe für den Leistungsbezug.
Besonderheiten	Es gibt eine relativ hohe Anzahl der Ausprägung „sonstige Sätze“, aber die Variable ist zuverlässig gefüllt.

1	Erst-/Neubewilligung ALG (Abgleich mit Anspruchstabelle)
2	Weiterbewilligung ALG ohne Unterbrechung
3	Weiterbewilligung ALG nach Unterbrechung
4	Neubewilligung ALG mit Aufstockung (ein neu erworbener Anspruch wird durch einen bestehenden Restanspruch aufgestockt)
5	Weiterbewilligung ALG mit Kürzung (z.B. durch Sperrzeit)
6	ALHI-Satz, der zeitlich direkt an ALG-Satz anschließt
7	ALHI-Satz, der zeitlich direkt an vorhergehenden ALHI-Satz anschließt
8	ALHI-Satz mit Lücke zum vorhergehenden Satz
9	alle anderen Sätze
10	max.Anspr. aufg. d. Alters

## 7.2.4.8 Abmeldegrund Leistungsbezug (abm\_gr)

Herkunft	BLH
Gefüllt für	LeH
Zeitbezug	Originalspell
Detailbeschreibung	Für LeH-Meldungen fasst die Variable „Abmeldegrund“ die oft spezifischen Einzelgründe der Variable abg_gr zusammen. In den meisten Fällen ist diese Variable für Analysezwecke ausreichend.
Besonderheiten	

1	Arbeitsaufnahme
2	Anspruch auf andere Leistung
3	sonstige Gründe
4	steht dem Arbeitsmarkt nicht zur Verfügung
5	Ablauf der Maßnahme
6	Sperr-, Säumniszeiten
7	Anspruch erschöpft
8	Bewilligungsabschnitt abgelaufen
9	Sonstige

## 7.2.4.9 Wirtschaftszweig WZ73, 3-steller (wz73)

Siehe Abschnitt 7.2.4.11.

## 7.2.4.10 Wirtschaftszweig WZ93, 5-steller (wz93)

Siehe Abschnitt 7.2.4.11.

## 7.2.4.11 Wirtschaftszweig WZ03, 5-steller (wz03)

Herkunft	BLH
Gefüllt für	BeH wz73: bis 2002 wz93: 1999 - 2003 wz03: ab 2003
Zeitbezug	Originalspell
Detailbeschreibung	wz73 ist ein 3-steller, wz93 und wz03 sind 5-steller. In den oberen Hierarchieebenen entsprechen die Ausprägungen von wz93 und wz03 denen der NACE-Klassifikation. Das FDZ bietet derzeit keine Überführung der einen Klassifikation in die andere an.
Besonderheiten	Die Wirtschaftszweige sind sensible Merkmale, deren Nutzung muss im Antrag gesondert begründet werden (vgl. Abschnitt 7.1.2)

## 7.2.5 Generierte biographische Merkmale

## 7.2.5.1 Eintrittsdatum Erwerbsleben (ein\_erw)

Herkunft	generiert aus BLH
Gefüllt für	BeH und LeH
Zeitbezug	personenbezogen
Detailbeschreibung	Das Merkmal enthält das Beginndatum der ersten sozialversicherungspflichtigen Beschäftigung einer Person in der BLH. Nicht berücksichtigt sind Ausbildungszeiten (Stellung im Beruf = 0). Personen, die in der BLH eine Ausbildungszeit aber keine sozialversicherungspflichtige Beschäftigung aufweisen, haben Missing.
Besonderheiten	Der Eintritt in das Erwerbsleben kann lange nach dem Eintritt in den Betrieb (Abschnitt 7.2.5.2) eintreten, da bei letzterem Ausbildungszeiten einbezogen sind. Da es sich um eine Datumsvariable handelt, sind fehlende Werte Systemmissings und nicht weiter differenziert. Das Merkmal ist linkszensiert, für Westdeutsche am 1.1.1975. Für Ostdeutsche ist die Zensierungsgrenze nicht eindeutig. Eintritte am 1.1.1990 sind in jedem Fall zensiert, aber oftmals auch solche am 1.1.1991 und 1.1.1992, da für 1990 und 1991 nicht alle Beschäftigtenmeldungen vorliegen.

## 7.2.5.2 Eintrittsdatum Betrieb (ein\_bet)

Herkunft	generiert aus BLH
Gefüllt für	BeH
Zeitbezug	eindeutig für jede Kombination von Person und Betriebsnummer
Detailbeschreibung	Das Merkmal enthält das Beginndatum der ersten Beschäftigungsmeldung einer Person in der BLH im aktuellen Betrieb. Dabei werden auch Ausbildungszeiten berücksichtigt. Eine Unterbrechung der Beschäftigung im Betrieb ändert das Eintrittsdatum nicht.
Besonderheiten	Der Eintritt in das Erwerbsleben kann lange nach dem Eintritt in den Betrieb

	<p>(Abschnitt 7.2.5.2) eintreten, da bei letzterem Ausbildungszeiten einbezogen sind.</p> <p>Da es sich um eine Datumsvariable handelt, sind fehlende Werte Systemmissings und nicht weiter differenziert.</p> <p>Das Merkmal ist linkszensiert, für Westdeutsche am 1.1.1975. Für Ostdeutsche ist die Zensierungsgrenze nicht eindeutig. Eintritte am 1.1.1990 sind in jedem Fall zensiert, aber oftmals auch solche am 1.1.1991 und 1.1.1992, da für 1990 und 1991 nicht alle Beschäftigtenmeldungen vorliegen.</p>
--	---

### 7.2.5.3 Dauer Betrieb (tage\_bet)

Herkunft	generiert aus BLH
Gefüllt für	BeH
Zeitbezug	bezogen auf den 30.6. des Jahres für jede Kombination von Person und Betriebsnummer
Detailbeschreibung	<p>Das Merkmal enthält die Anzahl der Tage, die eine Person am 30.6. des Jahres im betreffenden Betrieb beschäftigt war.</p> <p>Dabei werden auch Ausbildungszeiten berücksichtigt, Beschäftigungslücken jedoch nicht.</p>
Besonderheiten	<p>Berechnet man die Dauer im Betrieb alternativ mit dem Merkmal Eintritt in den Betrieb (Abschnitt 7.2.5.2)</p> $\text{gen tage\_bet\_neu} = \text{d}(30.6.2005) - \text{ein\_bet} + 1$ <p>erhält man Werte die größer oder gleich tage_bet sind, da tage_bet Unterbrechungen der Beschäftigung nicht enthält.</p> <p>Das Merkmal ist linkszensiert, für Westdeutsche am 1.1.1975. Für Ostdeutsche ist die Zensierungsgrenze nicht eindeutig. Eintritte am 1.1.1990 sind in jedem Fall zensiert, aber oftmals auch solche am 1.1.1991 und 1.1.1992, da für 1990 und 1991 nicht alle Beschäftigtenmeldungen vorliegen.</p>

### 7.2.5.4 Anzahl bisheriger Leistungsbezüge (anz\_ist)

Herkunft	generiert aus BLH
Gefüllt für	BeH und LeH
Zeitbezug	bezogen auf den 30.6. des Jahres für jede Person
Detailbeschreibung	<p>Das Merkmal enthält die Anzahl der Leistungsbezüge, die für eine Person bis zu 30.6. des Jahres vorliegen.</p> <p>Bei Lücken von weniger als 10 Tagen zwischen zwei Leistungsbezugsspielen wird von weiter laufendem Leistungsbezug ausgegangen, also nicht hochgezählt. Wechsel der Leistungsart werden ebenfalls nicht als neuer Leistungsbezug gezählt.</p>
Besonderheiten	<p>Das Merkmal ist linkszensiert, für Westdeutsche am 1.1.1975. Für Ostdeutsche ist die Zensierungsgrenze nicht eindeutig. Die BLH enthält keine Informationen zum Leistungsbezug vor 1990 und noch unvollständige für 1990 und 1991.</p>

### 7.2.5.5 Dauer bisheriger Leistungsbezüge (tage\_ist)

Herkunft	generiert aus BLH
Gefüllt für	BeH und LeH
Zeitbezug	bezogen auf den 30.6. des Jahres für jede Person
Detailbeschreibung	<p>Das Merkmal enthält die Dauer der Leistungsbezüge, die für eine Person bis zu 30.6. des Jahres vorliegen.</p> <p>Bei Lücken werden nicht aufaddiert.</p>
Besonderheiten	<p>Das Merkmal ist linkszensiert, für Westdeutsche am 1.1.1975. Für Ostdeutsche ist die Zensierungsgrenze nicht eindeutig. Die BLH enthält keine Informationen</p>

	zum Leistungsbezug vor 1990 und noch unvollständige für 1990 und 1991.
--	--

### 7.2.5.6 Erwerbsstatus 5 Jahre nach Stichtag (estat5jn)

Das Merkmal ist nur im Querschnittmodell enthalten, nicht in den Versionen des Längsschnittmodells.

Siehe Abschnitt 7.2.5.7.

### 7.2.5.7 Erwerbsstatus 8 Tage vor Eintritt in den Betrieb (estat8tv)

Herkunft	generiert aus BLH
Gefüllt für	BeH estat5jn: bis 2000 (nur im Querschnittmodell enthalten) estat8tv: alle Jahre
Zeitbezug	estat5jn: bezogen auf den 30.6. des Jahres für jede Person estat8tv: bezogen auf 8 Tage vor dem Ersteintritt in den Betrieb (ein_bet) für jede Kombination von Person und Betriebsnummer
Detailbeschreibung	<p>Das Merkmal enthält Information zum Erwerbsstatus der Person zu zwei anderen Zeitpunkten als dem Befragungszeitpunkt des Betriebspanels. Zum einen wurde der Zeitpunkt 8 Tage vor dem Ersteintritt in den Betrieb gewählt. Damit steht neben Anzahl und Dauer bisheriger Leistungsbezüge eine weitere Information zur Erwerbshistorie vor dem aktuellen Beschäftigungsverhältnis zur Verfügung. Zum anderen wurde der Zeitpunkt 5 Jahre nach dem Referenztag 30.6. in jedem Jahr gewählt.</p> <p>Der Erwerbsstatus wurde generiert aus den Merkmalen: Stellung im Beruf (StiB): Vollzeit, Teilzeit, Azubi, Missing Personengruppe: Normal, geringfügig Beschäftigte, Sonstige Betriebsnummer: gleiche oder andere als zum Referenzzeitpunkt Leistungsart: ALG, ALH, UHG, Zuschuss § 166b AFG/§ 207a SGB III</p> <p>Liegen zum Zeitpunkt 8 Tage davor oder 5 Jahre danach mehrere parallele Einträge in der BLH vor, z.B. eine Beschäftigung und ein Leistungsbezug, wird genau ein Eintrag ausgewählt. Dabei hat Beschäftigung Vorrang vor Leistungsbezug, bei mehreren Beschäftigungen erhält die reguläre Beschäftigung vor der geringfügigen, bei mehreren Regulären die mit dem höchsten Tagesentgelt.</p> <p>Leistungsbezug wird also nur dann angezeigt, wenn es keine parallele Beschäftigung gibt, geringfügige Beschäftigung nur dann, wenn es keine höherwertige gibt. Bei estat8tv muss die vorangegangene Beschäftigung definitionsgemäß immer in einem anderen Betrieb sein.</p>
Besonderheiten	estat8tv ist linkszensiert, für Westdeutsche am 1.1.1975. Für Ostdeutsche ist die Zensierungsgrenze nicht eindeutig. Die BLH enthält keine Informationen für die neuen Bundesländer vor 1990 und noch unvollständige für 1990 und 1991.

0	kein Eintrag in BLH, es gibt weder BeH noch LeH Spell zu dem Zeitpunkt		
11	Info aus BeH: Betriebsnummer	Vollzeit - normal	- gleiche
12	Info aus BeH: Betriebsnummer	Teilzeit - normal	- gleiche

13	Info aus BeH: Betriebsnummer	Ausbildung - normal	- gleiche
15	Info aus BeH: Betriebsnummer	ohne StiB - normal	- gleiche
16	Info aus BeH: Betriebsnummer	Vollzeit - normal	- andere
17	Info aus BeH: Betriebsnummer	Teilzeit - normal	- andere
18	Info aus BeH: Betriebsnummer	Ausbildung - normal	- andere
20	Info aus BeH: Betriebsnummer	ohne StiB - normal	- andere
21	Info aus BeH: Betriebsnummer	Vollzeit - geringfügig	- gleiche
22	Info aus BeH: Betriebsnummer	Teilzeit - geringfügig	- gleiche
23	Info aus BeH: Betriebsnummer	Ausbildung - geringfügig	- gleiche
25	Info aus BeH: Betriebsnummer	ohne StiB - geringfügig	- gleiche
26	Info aus BeH: Betriebsnummer	Vollzeit - geringfügig	- andere
27	Info aus BeH: Betriebsnummer	Teilzeit - geringfügig	- andere
28	Info aus BeH: Betriebsnummer	Ausbildung - geringfügig	- andere
30	Info aus BeH: Betriebsnummer	ohne StiB - geringfügig	- andere
31	Info aus BeH: Betriebsnummer	Vollzeit - Ausbildung	- gleiche
32	Info aus BeH: Betriebsnummer	Teilzeit - Ausbildung	- gleiche
33	Info aus BeH: Betriebsnummer	Ausbildung - Ausbildung	- gleiche
36	Info aus BeH: Betriebsnummer	Vollzeit - Ausbildung	- andere
37	Info aus BeH: Betriebsnummer	Teilzeit - Ausbildung	- andere
38	Info aus BeH: Betriebsnummer	Ausbildung - Ausbildung	- andere
40	Info aus BeH: Betriebsnummer	ohne StiB - Ausbildung	- andere
41	Info aus BeH: Betriebsnummer	Vollzeit - Sonstige	- gleiche
42	Info aus BeH: Betriebsnummer	Teilzeit - Sonstige	- gleiche
43	Info aus BeH: Betriebsnummer	Ausbildung - Sonstige	- gleiche
45	Info aus BeH: Betriebsnummer	ohne StiB - Sonstige	- gleiche
46	Info aus BeH: Betriebsnummer	Vollzeit - Sonstige	- andere
47	Info aus BeH: Betriebsnummer	Teilzeit - Sonstige	- andere

48	Info aus BeH: Betriebsnummer	Ausbildung - Sonstige	- andere
50	Info aus BeH: Betriebsnummer	ohne StiB - Sonstige	- andere
91	Info aus LeH:	Arbeitslosengeld	
92	Info aus LeH:	Arbeitslosenhilfe	
93	Info aus LeH:	Unterhaltsgeld	
94	Info aus LeH: Zuschuss § 166b AFG/§ 207a SGB III		

#### 7.2.5.8 Lohn 5 Jahre nach Stichtag (lohn5jn)

Das Merkmal ist nur im Querschnittmodell enthalten, nicht in den Versionen des Längsschnittmodells.

Siehe Abschnitt 7.2.5.9.

#### 7.2.5.9 Lohn 8 Tage vor Eintritt in den Betrieb (lohn8tv)

Herkunft	generiert aus BLH
Gefüllt für	BeH lohn5jn: bis 2000 (nur im Querschnittmodell enthalten) lohn8tv: alle Jahre
Zeitbezug	lohn5jn: bezogen auf den 30.6. des Jahres für jede Person lohn8tv: bezogen auf 8 Tage vor dem Ersteintritt in den Betrieb (ein_bet) für jede Kombination von Person und Betriebsnummer
Detailbeschreibung	Das Merkmal enthält Information zum Tagesentgelt der Person zu zwei anderen Zeitpunkten als dem Befragungszeitpunkt des Betriebspanels. Analog zum Erwerbstatus wurden die Zeitpunkte 8 Tage vor dem Ersteintritt in den Betrieb und 5 Jahre nach dem Referenztag 30.6. in jedem Jahr gewählt.  Der Lohn zu diesen Zeitpunkten stammt aus dem Merkmal Tagesentgelt der BLH, d.h. es gelten die Hinweise in Abschnitt 7.2.4.5.  Liegen zum Zeitpunkt 8 Tage davor oder 5 Jahre danach mehrere parallele Beschäftigungen in der BLH vor, wird genau ein Eintrag ausgewählt. Dabei hat die reguläre Beschäftigung Vorrang vor der geringfügigen, bei mehreren Regulären die mit dem höchsten Tagesentgelt.
Besonderheiten	lohn8tv ist linkszensiert, für Westdeutsche am 1.1.1975. Für Ostdeutsche ist die Zensierungsgrenze nicht eindeutig. Die BLH enthält keine Informationen für die neuen Bundesländer vor 1990 und noch unvollständige für 1990 und 1991.

### 7.2.6 Ortsangaben

#### 7.2.6.1 Arbeitsort Arbeitsagentur (arb\_aa)

Herkunft	BLH
Gefüllt für	BeH
Zeitbezug	Originalspell
Detailbeschreibung	Angegeben ist jeweils das Arbeitsamt, in dessen Bezirk der Betrieb des Beschäftigten angesiedelt ist. Da sich die Grenzen der Arbeitsamtsbezirke im Zeitverlauf ändern, würden

	<p>ohne Gebietsstandsberichtigung Fälle auftreten, bei denen der Arbeitsamtsbezirk eines Betriebes wechselt, ohne dass dieser Betrieb seinen Standort verlagert hat. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, sind die Arbeitsamtsangaben daher derzeit auf den Gebietsstand vom 31.12.2005 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Betriebsortes zu einem Arbeitsamtsbezirk nach den Grenzen, welche die Arbeitsamtsbezirke zum 31.12.2005 hatten.</p> <p>Ein Problemfall stellt hierbei Berlin dar: Die Berliner Arbeitsämter wurden im Laufe der Jahre immer wieder neu zugeschnitten, was auch durch die Gebietsstandsberichtigung nicht vollständig bereinigt wird. Nähere Informationen zur Gebietsstandsberichtigung befinden sich in der Dokumentation der IAB-Regionalstichprobe. Die dortigen Informationen können auf die LIAB-Personendaten übertragen werden.</p>
Besonderheiten	arb_aa ist ein sensibles Merkmal und muss daher im Antrag auf Datenzugang gesondert begründet werden.

#### 7.2.6.2 Arbeitsort Kreis (arb\_kr)

Herkunft	BLH
Gefüllt für	BeH
Zeitbezug	Originalspell
Detailbeschreibung	<p>Die Kreiskennzifferinformation stammt aus der BeH. Angegeben ist jeweils der Kreis (kreisfreie Stadt oder Landkreis), in dem der Betrieb des Beschäftigten angesiedelt ist. Der 5-stellige Kreisschlüssel enthält in den ersten beiden Stellen den Schlüssel des Bundeslandes, in der ersten bis dritten Stelle den Regierungsbezirk und in der ersten bis fünften Stelle den Kreis. Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null. Angegeben ist jeweils der Kreis, in dem der Beschäftigungsbetrieb angesiedelt ist.</p> <p>Die Angaben sind gebietsstandsberichtigt zum Stand 31.12.2005.</p>
Besonderheiten	arb_kr ist ein sensibles Merkmal und muss daher im Antrag auf Datenzugang gesondert begründet werden.

#### 7.2.6.3 Wohnort Arbeitsagentur (wohn\_aa)

Herkunft	BLH
Gefüllt für	BeH (ab 1999) und LeH
Zeitbezug	Originalspell
Detailbeschreibung	<p>Für BeH ist die Arbeitsagentur enthalten, die für den Wohnort des Beschäftigten zuständig ist.</p> <p>Für die LeH ist die leistende Arbeitsagentur enthalten. Es wird angenommen, dass sich die Leistungsbezieher bei der zuständigen Geschäftsstelle bzw. Arbeitsagentur des Wohnortes melden müssen und damit betrifft die Angabe auch dem tatsächlichen Wohnort.</p> <p>Die Angaben der BeH sind gebietsstandsberichtigt zum Stand 31.12.2005, bei den Angaben der LeH ist keine Bereinigung möglich.</p>
Besonderheiten	wohn_aa ist ein sensibles Merkmal und muss daher im Antrag auf Datenzugang gesondert begründet werden.

#### 7.2.6.4 Wohnort Kreis (wohn\_kr)

Herkunft	BLH
Gefüllt für	BeH (ab 1999)
Zeitbezug	Originalspell

Detailbeschreibung	Für BeH ist der Kreis des Wohnortes des Beschäftigten enthalten. Die Angaben der BeH sind gebietsstandsbereinigt zum Stand 31.12.2005.
Besonderheiten	wohn_kr ist ein sensibles Merkmal und muss daher im Antrag auf Datenzugang gesondert begründet werden.

#### 7.2.6.5 Ost-West-Kennzeichen (west)

Herkunft	BLH
Gefüllt für	BeH und LeH
Zeitbezug	Originalspell
Detailbeschreibung	Die Angaben der BeH beziehen sich auf den Arbeitsort, die in der LeH auf den Wohnort.

0	Ostdeutschland
1	Westdeutschland

## 8 Datenanonymisierung

Die LIAB-Daten liegen derzeit ausschließlich in schwach anonymisierter Form vor. Schwach anonymisiert bedeutet hier, dass Betriebs- und Personenidentifikatoren (id, idnum und betnr) systemfrei ersetzt sind, alle weiteren Informationen aber unverändert vorliegen. id und betnr unterscheiden sich zwischen den LIAB-Versionen und bei Datenaktualisierungen, während die idnum gleich bleibt. Eine Verknüpfung von Dateien unterschiedlicher LIAB-Versionen oder Aktualisierungsständen ist also nur über die idnum möglich. Hintergrund für die Neuvergabe der Identifikatoren ist die Korrektur bzw. Stilllegung von Versicherungsnummern infolge von Kontenklärungen durch die Rentenversicherung. Dies macht eine zeitkonsistente Vergabe der systemfreien Identifikatoren aufwändig. Ein weiteres Argument ist der Datenschutz.

Bestimmte Merkmale, die die Deanonymisierung von Betrieben und Personen erleichtern, werden nur dann weitergegeben, wenn es für das Erreichen des Forschungsziels unerlässlich ist (-> siehe Abschnitt 7.1.2: sensible Merkmale).

## 9 Datennutzung

### 9.1 Datenzugang

LIAB-Daten können per Gastaufenthalt und anschließend per Datenfernverarbeitung bearbeitet werden. Näheres finden Sie auf unseren Internetseiten <http://fdz.iab.de> unter Datenzugang.

### 9.2 Hinweise für die Datenanalyse

#### 9.2.1 Testdaten

Testdaten dienen der Vorbereitung von Auswertungsprogrammen im Vorfeld der Datenfernverarbeitung und von Gastaufenthalten. Sie sind allgemein zugänglich und müssen daher absolut anonymisiert sein. Sie sollen aber die Datenstruktur möglichst genau wiedergeben, damit Auswertungsprogramme für möglichst viele unterschiedliche Auswertungen getestet werden können. Universal einsetzbare Testdaten sind bei beschränkter Dateigröße jedoch nicht zu erstellen. Daher empfehlen wir die Testdaten für die eigenen Zwecke anzupassen: sollen z.B. die Spanier untersucht werden, sollte man zuvor hinreichend viele Staatsangehörigkeiten auf Spanien umsetzen.

Testdaten sind Zufallszahlen: **Mit den Testdaten sind keine inhaltlichen Auswertungen möglich.**

Im Einzelnen gilt für die aktuellen Testdaten::

- alle Variablen der Originaldaten sind enthalten, haben die gleichen Variablennamen und Variablen- und Wertelabel,
  - sensible Merkmale sind in den Testdaten enthalten, in den Originaldaten nur auf Antrag,
  - relative Häufigkeiten von Ausprägungen entsprechen ungefähr den Verteilungen in den Originaldaten, ist eine Variable in einem Jahr oder einer Quelle nicht gefüllt, gilt das auch für die Testdaten,
  - seltene Ausprägungen wie einzelne Staaten bei der Staatsangehörigkeit können in den Testdaten fehlen,
-

- Korrelationen zwischen den Variablen sind nur teilweise erhalten, Variablen sind also häufig nicht konsistent und Schätzungen konvergieren nicht immer,
- wie in den Originaldaten gibt es Zu- und Abgänge bei Personen und Betrieben,
- wie in den Originaldaten stimmt die Zahl der sozialversicherungspflichtig Beschäftigten laut Fragebogen des IAB-Betriebspanels häufig nicht mit der Anzahl der Personen in den Personendaten überein.

### 9.2.2 Überprüfung der Verknüpfungsqualität

Da es keine für alle Fälle gleichsam gültige Antwort auf die Frage gibt, welche der Verknüpfungen von Personen- und Betriebsdaten zuverlässig sind, wird diese Einschätzung Datennutzer überlassen. Zunächst sollte anhand der Wellenkennzeichen (well`jahr') im IAB-Betriebspanel geprüft werden, ob es sich weiterhin um den erstbefragten Betrieb mit bei der Erstbefragung abgeglichener Betriebsnummer handelt oder nicht. Anschließend sollte die Übereinstimmung der Daten aus IAB-Betriebspanel und LIAB-Personendaten anhand von Informationen überprüft werden, die in beiden Quellen vorliegt:

- Sozialversicherungspflichtig Beschäftigte zum Befragungszeitpunkt oder Vorjahr
- Auszubildende zum Befragungszeitpunkt oder Vorjahr
- Ost-West-Kennzeichen
- Lohnsumme

Zu beachten ist, dass Befragungsangaben Rundungs- und Erinnerungsfehler enthalten können. Die Lohnsumme in den Betriebsdaten ist zudem nicht direkt vergleichbar mit dem Aggregat der Löhne aus den Personendaten, da Löhne oberhalb der Sozialversicherungsgrenze enthalten sind.

Als grobe Daumenregel könnten Verknüpfungen verworfen werden, deren Beschäftigtenzahl um mehr als 30% abweicht. Bei der Selektion der Betriebe in den LIAB-Längsschnittmodellen wird etwas differenzierter vorgegangen, damit Abweichungen um wenige Personen bei kleineren Betrieben weniger stark ins Gewicht fallen:

- Betriebe mit bis 5 Beschäftigten: 40 Prozent
- Betriebe mit 5 bis 19 Beschäftigten: 30 Prozent
- Betriebe mit 20 bis 100 Beschäftigten: 20 Prozent

- Betriebe mit mehr als 100 Beschäftigten: 10 Prozent

### 9.2.3 Weitere Tipps zur Datenaufbereitung

Durch die Verknüpfung der "langen" Personendaten (viele Beobachtungen) mit den "breiten" Betriebsdaten (viele Merkmale) können sich beim LIAB sehr große Dateien und ein entsprechend hoher Speicherbedarf und lange Rechenzeiten ergeben. Folgende Tipps helfen dies zu vermeiden:

- Merkmale und Beobachtungen, die nicht benötigt werden, sollten möglichst früh gelöscht bzw. gar nicht erst eingelesen werden.
  - Betriebs- und Personendaten sollten möglichst getrennt aufbereitet und erst nach der Selektion der benötigten Merkmale und Beobachtungen zusammengespielt werden.
  - Programme zur Aufbereitung der IAB-Betriebspaneldaten können bereits vor dem Gastaufenthalt per Datenfernverarbeitung erstellt werden.
  - Der Datenreport 2/2006 "[Aufbereitung eines Paneldatensatzes aus den Querschnittsdaten des IAB-Betriebspanels](#)" beschreibt die Aufbereitung eines Betriebspanel-Datensatzes mit über die Zeit konsistenten Codierungen der Merkmale. Die dort beschriebenen Programme können auch für den LIAB verwendet werden.
-

## 10 Abkürzungsverzeichnis

ALG	Arbeitslosengeld
ALH	Arbeitslosenhilfe
BA	Bundesagentur für Arbeit
BeH	Beschäftigten-Historik-Datei
BLH	Beschäftigten-Leistungsempfänger-Historik-Datei
DEÜV	Datenerfassungs- und -übermittlungsverordnung, in Kraft getreten am 1. Januar 1999
DEVO	Datenerfassungsverordnung, am 1. Januar 1999 durch die DEÜV abgelöst
DÜVO	Datenübermittlungsverordnung, am 1. Januar 1999 durch die DEÜV abgelöst
FDZ	Forschungsdatenzentrum
IAB	Institut für Arbeitsmarkt- und Berufsforschung
JZM	Jahreszeitraummaterial
LeH	Leistungsempfänger-Historik-Datei
LIAB	Linked Employer-Employee Daten des IAB
SGB	Sozialgesetzbuch
WZ	Wirtschaftszweigklassifikation

## 11 Literatur

Eine aktualisierte Literaturliste mit LIAB-Anwendungen und Dokumentationen finden Sie in unserer Literaturdatenbank im Internet unter <http://iab.de/asp/info/thema.asp>.

Vor dem Hintergrund der Vereinheitlichung der FDZ-Datenreporte und deren Übersetzung ins Englische wird mit dieser Neuauflage des LIAB-Datenhandbuchs versucht, die relevanten Informationen zu den LIAB-Versionen zusammenzuführen. Anlagen zum Handbuch sind eine Datei mit Merkmalsausprägungen für die Merkmale mit vielen der Ausprägungen und eine unformatierte log-Datei mit Auszählungen zu allen Merkmalen. Die log-Datei ist umfassend und bei späteren Updates leicht zu aktualisieren. Die bisherigen Dokumentationen werden nicht weiter gepflegt, bleiben aber dem deutschsprachigen Leser als reichhaltige Nachschlagewerke erhalten.

- FDZ-Methodenreport 1/2005: Die Verknüpfungsqualität der LIAB-Daten
  - FDZ-Datenreport 1/2005: Betriebe und Beschäftigte in den Linked Employer-Employee Daten des IAB
  - FDZ-Datenreport 2/2005: Datenbeschreibung der Version 1.1 des LIAB-Querschnittmodells
  - FDZ-Datenreport 3/2005: Datenbeschreibung der Version 1.1 des LIAB-Längsschnittmodells
  - FDZ-Datenreport 7/2005: LIAB-Datenhandbuch, Version 1.0
  - FDZ-Datenreport 8/2005: Datenbeschreibung der Version 2.0 des LIAB-Längsschnittmodells
  - FDZ-Datenreport 2/2007: LIAB-Datenhandbuch, Version 2.0
-

## Imprint

### FDZ *Datenreport*

No. 3/2007

#### **Publisher**

The Research Data Centre (FDZ)  
of the Federal Employment Service  
in the Institute for Employment Research  
Regensburger Str. 104  
D-90478 Nuremberg

#### **Editorial staff**

Stefan Bender, Dagmar Herrlinger

#### **Technical production**

Dagmar Herrlinger

#### **Copyright**

Reproduction – also in parts – only with permission of  
the FDZ

#### **Download**

[http://doku.iab.de/fdz/reporte/2008/DR\\_03-08.pdf](http://doku.iab.de/fdz/reporte/2008/DR_03-08.pdf)

#### **Internet**

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

#### **Corresponding author**

Peter Jacobebbinghaus, Institute for Employment  
Research, Regensburger Str. 104,  
D-90478 Nuremberg  
Tel.: +49 (0) 911/179-1765  
E-Mail: [Peter.Jacobebbinghaus@iab.de](mailto:Peter.Jacobebbinghaus@iab.de)