

## **Datenbeschreibung der Version 2.0 des LIAB-Längsschnittmodells**

*Holger Alda*

## Inhaltsverzeichnis

1. Allgemeines.....	3
2. Abdeckungsgrad über Betriebe und Beschäftigte .....	4
3. Variablen im Datensatz .....	6
4. Statistiken.....	8
Anhang.....	12
<i>Anhang A1: Die Variablen der Personendatenbanken der Version 2 des     LIAB-Längsschnittmodells .....</i>	<i>12</i>
<i>Anhangstabelle A2: Zusammenfassung von Wirtschaftszweigen.....</i>	<i>22</i>

## Datenverfügbarkeit

Der in diesem Beitrag beschriebene Datensatz ist für die Fachöffentlichkeit zugänglich. Nähere Informationen dazu erhalten Sie auf der Internetseite: <http://fdz.iab.de/> unter der Rubrik „Integrierte Betriebs- und Personendaten“.

## Danksagung

Das ITM-Team Datenmanagement des Instituts für Arbeitsmarkt- und Berufsforschung, IAB, hat die für die Erstellung der LIAB-Personendatenbanken erforderliche Dateninfrastruktur des IAB bereitgestellt. Zudem wird für die in diesem Rahmen erfolgte datentechnische Beratung und Unterstützung gedankt. Dem Forschungsbereich 6 des IAB „Betriebe und Beschäftigung“ wird für die Bereitstellung der IAB-Betriebspaneldaten und die Unterstützung bei inhaltlichen Fragen gedankt. Anmerkungen, Fragen und Kritik zu diesem Datenreport werden bevorzugt vom Autor entgegengenommen.

## 1. Allgemeines

Die vorliegende Datenbeschreibung für die zweite Version des LIAB-Längsschnittmodells bezieht sich auf je eine ost- und westdeutsche Personendatenbank, in der alle Beschäftigungs- und Leistungsmeldungen von Personen abgelegt sind, die mindestens einen Tag zwischen 1997 und 2002 in Betrieben des IAB-Betriebspanels gearbeitet haben, die im Zeitraum 2000 bis 2002 in jedem Jahr an der Befragung teilgenommen haben. Aufgrund des Umfangs der Personendatenbanken ist eine Zusammenspielung der beiden Landesdatensätze ohne eine vorherige Datenaufbereitung und Reduktion der Anzahl der Meldungen/Variablen nicht möglich.

Die Kenntnis der FDZ-Daten- und Methodenreports Nr. 1 erleichtern das Verständnis dieses Beitrags. Die Verknüpfung der Betriebs- und Personendaten ist eine von Anwendern eigenständig zu erbringende Leistung. Die Personendatenbanken liegen im SAS 8.0. und STATA 8.2. SE Format vor. Auch andere Softwareanalysetools können diese Formate lesen<sup>1</sup>. Nach der im Forschungsdatenzentrum jederzeit möglichen Konvertierung sind weitere Datenformate verfügbar (z.B. ASCII). Testdaten sind über das Internet zugänglich. Weitere Erläuterungen hierzu befinden sich im Internet unter [http://doku.iab.de/fdz/liab/Testdaten\\_lang.pdf](http://doku.iab.de/fdz/liab/Testdaten_lang.pdf).

---

<sup>1</sup> SPSS kann beispielsweise das SAS-Datenformat lesen.

## 2. Abdeckungsgrad über Betriebe und Beschäftigte

Ein Vorteil des LIAB-Längsschnittmodells ist die Analyse und Nutzung der Informationen der erwerbsbiographischen Ereignisse von Personen. In der zweiten Version des LIAB-Längsschnittmodells liegen auf der Personenebene lückenlose Beobachtungszeiträume und -zustände vom Jahr 1993 bis zum Jahr 2002 vor. Der Variablenkanon der Personendaten wird in Abschnitt 3 beschrieben.

Die Bedingung für die Aufnahme eines Personenkontos in dieser zweiten Längsschnittversion ist, mindestens einen Tag in einem der Längsschnittbetriebe von 2000 bis 2002 im Zeitraum 1997 bis 2002 gearbeitet zu haben. Für solche Personen sind Variablen der Beschäftigten-Leistungsempfänger-Historik (vgl. FDZ-Datenreport Nr. 1) für den Zeitraum 1993 bis 2002 in der jeweiligen Personendatenbank des Landesteils verfügbar. Im Anhang befinden sich Beschreibungen der einzelnen Variablen und Informationen über ihren Wertebereich, die im LIAB-Datenhandbuch (FDZ-Datenreport Nr. 7) vertieft werden.

Die für die zweite Version des LIAB-Längsschnittmodells ausgewählten Betriebe haben in den Jahren 2000 bis 2002 gültige Interviews im IAB-Betriebspanel. Damit kann für die entsprechenden Betriebe wenigstens in diesem Zeitraum das betriebliche Handeln und mögliche Veränderungen lückenlos untersucht werden. Vor dem Jahr 2000 werden die Betriebe „ihrem Schicksal überlassen“, d.h. ihr Überleben oder Sterben und die Panelmortalität ist *dann* zufällig verteilt. Auf der Personenebene sind alle Betriebe bis mindestens 1997 beobachtbar<sup>2</sup>. Die Entwicklung der im Betrieb Verbleibenden kann grundsätzlich bis zum Jahr 1993 zurückverfolgt werden. In dieser Version 2 des LIAB-Längsschnittmodells erfolgt – im Unterschied zur ersten Version - die Auswahl der Betriebe ohne Rücksicht auf die Toleranzintervalle, die im FDZ-Methodenreport Nr. 1 beschrieben werden. Die dort vorgenommene Ausfallanalyse zeigt, welche Betriebe besonders häufig außerhalb der Toleranzintervalle liegen.

Der Abdeckungsgrad über alle Betriebe und (sozialversicherungspflichtig) Beschäftigte des IAB-Betriebspanels wird wie im FDZ-Datenreport Nr. 2 bestimmt. Bezogen auf die Querschnittsfälle des IAB-Betriebspanels ergeben sich für die ausgewählten Längsschnittbetriebe der zweiten Version des LIAB-Längsschnittmodells die Werte der Tabelle 1:

---

<sup>2</sup> Viele Betriebscharakteristika können durch Aggregation der Personendaten gebildet werden, selbst wenn der Betrieb zwischen 1997 und 1999 kein gültiges Interview hat.

**Tabelle 1: Abdeckungsgrad über Betriebe und Beschäftigte des IAB-Betriebspanels (Stichtag 30. Juni) der Version 2 des LIAB-Längsschnittmodells ; Angabe in Prozent (Bezugsbasis: sozialversicherungspflichtig Beschäftigte in bzw. gültige(n) Querschnittsfälle(n) des IAB-Betriebspanels)**

a) ungewichtet	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Betriebe	30,9	34,5	38,5	42,9	48,7	57,1	67,7	69,3	62,1	58,5
Beschäftigte	30,7	34,1	39,7	46,6	53,1	58,9	68,9	74,9	68,7	66,0
b) gewichtet	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Betriebe	31,0	32,9	38,3	42,8	47,9	55,1	64,1	70,4	59,8	59,9
Beschäftigte	31,5	35,3	39,5	46,2	51,3	57,8	65,9	73,7	66,9	62,7

Quelle: IAB-Betriebspanel 1993 bis 2001; die Hochrechnung erfolgt mit dem Querschnittsgewichtungsfaktor des jeweiligen Jahres

Wie in der Datenbeschreibung der ersten Version des LIAB-Längsschnittmodells beschrieben liegt der Abdeckungsgrad über Betriebe und Beschäftigte aufgrund der Auswahl nach Längsschnittkriterien im Querschnitt je nach Beobachtungsjahr zwischen knapp 31 bis zu 75 Prozent. Die zweite Version des LIAB-Längsschnittmodells hat im Unterschied zur ersten den Vorteil, dass für die Jahre 2001 und 2002 ein eigener Längsschnitthochrechnungsfaktor zur Verfügung steht, der den beschriebenen geringeren Abdeckungsgrad über die Querschnittsfälle ausgleicht, wie Tabelle 2 zeigt.

**Tabelle 2: Vergleich der Anzahl Betriebe, der Gesamtanzahl sozialversicherungspflichtig Beschäftigter und aller Erwerbstätiger in der zweiten Version des LIAB-Längsschnittmodells und des jeweiligen Jahresquerschnitts im Zeitraum 2000 bis 2002 – hochgerechnete Werte, Angabe in Tsd.**

Anzahl ...	2001			2002		
	LS V2 <sup>+</sup>	QS <sup>++</sup>	Differenz <sup>+++</sup> in Prozent	LS V2 <sup>+</sup>	QS <sup>++</sup>	Differenz <sup>+++</sup> in Prozent
<b>Betriebe</b>	<b>2.145</b>	<b>2.145</b>	<b>± 0,00</b>	<b>2.131</b>	<b>2.128</b>	<b>+ 0,00</b>
<b>SvB*</b>	<b>28.123</b>	<b>27.840</b>	<b>+ 1,01</b>	<b>27.725</b>	<b>27.541</b>	<b>+ 0,67</b>
<b>Erwerbst.**</b>	<b>34.518</b>	<b>34.080</b>	<b>+ 1,29</b>	<b>33.971</b>	<b>33.902</b>	<b>+ 0,20</b>

\* sozialversicherungspflichtig Beschäftigte

\*\* alle Erwerbstätigen

+ LIAB-Längsschnittmodell, Version 2, gewichtet mit dem Längsschnitthochrechnungsfaktor 2000 auf 2001 bzw. 2002

++ Querschnitt des jeweiligen Jahres, gewichtet mit dem Querschnittshochrechnungsfaktor

+++ Wert (LSV2-QS) dividiert durch Wert QS mal einhundert

### 3. Variablen im Datensatz

Der Variablenkanon des IAB-Betriebspanels bleibt von der vorgenommenen Auswahl unberührt. Nur die Zahl der Betriebe wurde wie in Abschnitt 2 beschrieben und im nächsten Abschnitt dokumentiert reduziert. In den beiden Personendatenbanken sind Variablen aus der Beschäftigten-Leistungsempfänger-Historik-Datei abgelegt, die nachfolgend aufgelistet und im Anhang A2 ausführlich erläutert werden. Es wird zwischen drei Kategorien unterschieden. In der ersten Kategorie a) sind die Organisationsvariablen aufgelistet, die für die Arbeit mit den Daten unerlässlich sind. In der Kategorie b) befinden sich Variablen, die mehrheitlich bei Beschäftigungs- und Leistungsmeldungen gefüllt sind und in der Kategorie c) Variablen, die sich ausschließlich auf den Leistungsbezug durch die Bundesagentur für Arbeit beziehen.

#### Übersicht 1: Variablen der Personendatenbanken in der Version 1 des LIAB-Längsschnittmodells

Kurzbezeichnung
<b>a) Organisationsvariablen</b>
eindeutiger Personenidentifikator
eindeutiger Betriebsidentifikator
Spellzähler personenbezogen
Spellzähler personen- und quellenbezogen
Status der Betriebsnummer
<b>b) Merkmale überwiegend aus Beschäftigungsmeldungen</b>
Mann/Frau
Berufskennziffer (3-stellig + Null)
Arbeitsamt Wohnort
Kreiskennziffer Wohnort
Tagesentgelt
Beginndatum der Episode
Enddatum der Episode
Abgabegrund der Meldung
Personengruppe
Stellung im Beruf
Bildungsniveau in Verbindung mit beruflicher Qualifikation
Arbeitsamtsbezirk des Arbeitsortes
Kreiskennziffer des Arbeitsortes
Wirtschaftszweig nach der BA-Klassifikation 1973
Wirtschaftszweig nach der BA-Klassifikation 1993

(Fortsetzung nächste Seite)

<b>Kurzbezeichnung</b>
Dauer der Meldung in Tagen
Kennzeichen der Datumsbereinigung
Alter der Person
Währungskennzeichen
Ost/West-Kennzeichen der Meldung
Staatsangehörigkeit
<b>c) Merkmale, die sich auf Leistungen durch die Bundesagentur für Arbeit beziehen</b>
Beschäftigungs- oder Leistungsmeldung
gruppierte Leistungsart
Kennzeichen Bewilligungsart der Leistung
zusammengefasste Abmeldegründe bei Leistung
Zusammenfassung Familienstand
Zusammenfassung Kinder der Person

Die Merkmale aus der Kategorie a) sind in allen Meldungen gefüllt (Ausnahme Betriebsidentifikator bei Leistungsmeldungen). Die originalen Identifikatoren für Betriebe und Personen (Betriebs- und Sozialversicherungsnummern) sind durch systemfreie ersetzt. Es existieren zwei Datensätze mit Personendaten für die je etwa 4000 ost- und west-deutschen Betriebe. Ihre Auswahl wird im folgenden Abschnitt beschrieben.

#### 4. Statistiken

Im Vordergrund steht die Auswahl der Betriebe, da sich an Ihnen die Ziehung der Personenkonten orientiert (vgl. Abschnitt 2). Nach dieser Darstellung werden Kennzahlen der beiden Personendatenbanken ausgewiesen. Abschnitt 2 erläuterte die Vorgehensweise bezüglich der Auswahl der IAB-Betriebspanelbetriebe. Die Längsschnittbetriebe des IAB-Betriebspanels im Zeitraum 2000 bis 2002 haben die folgende Verteilung nach Wirtschaftszweigen und Betriebsgrößenklassen.

**Tabelle 3: ost- und westdeutsche Längsschnittsfälle des IAB-Betriebspanel im Zeitraum 2000 bis 2002: Betriebe für die zweite Version des LIAB-Längsschnittmodells (nach 20 Wirtschaftszweigen\* und vier Betriebsgrößenklassen)**

	1 bis 19 Beschäftigte		20 bis 99 Beschäftigte		100 bis 499 Beschäftigte		500 und mehr Beschäftigte	
	West	Ost	West	Ost	West	Ost	West	Ost
Land- und Forstwirtschaft	47	62	15	66	5	13	-	/
Bergbau /Energie/ Wasser	7	5	19	24	14	22	16	14
Nahrungs-/ Genussmittelherstellung	54	45	36	26	43	31	14	-
Verbrauchsgüterindustrie	73	62	67	28	74	15	16	/
Produktionsgüterindustrie	79	240	105	114	124	60	83	7
Investitionsgüterindustrie	114	266	151	187	161	79	142	23
Baugewerbe	265	291	113	129	46	48	/	5
Handel/Reparaturgewerbe	429	298	213	110	83	33	26	5
Verkehrs-/ Nachrichtengewerbe	76	42	56	25	28	21	23	15
Kredit- und Versicherungsgewerbe	46	16	43	10	46	21	50	9
Gastgewerbe	118	80	29	9	11	4	-	-
Erziehung und Unterricht	80	33	44	67	35	68	11	19
Gesundheits- und Sozialwesen	176	129	121	76	111	86	79	35
Datenverarbeitung	22	10	14	4	4	/	/	-
Forschung und Entwicklung	7	/	5	5	10	8	4	/

(Fortsetzung nächste Seite)



	1 bis 19 Beschäftigte		20 bis 99 Beschäftigte		100 bis 499 Beschäftigte		500 und mehr Beschäftigte	
	West	Ost	West	Ost	West	Ost	West	Ost
Rechtsberatung und Werbung	103	55	32	/	8	-	/	-
Grundstücks- und Wohnungsgewerbe	17	24	15	25	/	9	/	/
sonstige Dienstleistungen überwiegend für Unternehmen	91	81	57	48	44	19	11	5
sonstige Dienstleistungen	103	71	31	52	33	38	8	5
öffentliche Verwaltung u. ä.	111	105	157	150	160	199	65	78
insgesamt **	2018	1916	1323	1155	1040	775	556	229

\* Die Klassifizierung der Wirtschaftszweige erfolgt nach der Hochrechnungsmatrix des IAB-Betriebspanels ab dem Jahr 2000, die zehn Betriebsgrößenklassen werden zu vier Ausprägungen gruppiert

\*\* mit Ausnahme der Spalten drei, fünf und sechs inklusive nicht ausgewiesener Werte

- : keine Beobachtungen in der Zelle

/ : Fallzahl zu gering

Insgesamt stehen knapp 5.000 westdeutsche und etwas über 4000 ostdeutsche Betriebe für Auswertungen zur Verfügung. Da es – je nach der sektoralen Struktur der Wirtschaft in Ost- und Westdeutschland – einige Zellen mit geringen Fallzahlen gibt, wird empfohlen, neben der bereits in Tabelle 3 erfolgten Zusammenfassung für die Betriebsgrößenklassen auch die Wirtschaftszweige weiter zu verdichten.

Abschließend werden einige Kennzahlen und Kennziffern für die aus der Auswahl der Betriebe resultierende Ziehung der Personendatenbanken nach den Kriterien des Abschnitts zwei und basierend auf den Betrieben der Tabelle 3 beschrieben.

**Tabelle 4: Kennzahlen und Kennziffern der west- und ostdeutschen Personendatenbank der ersten Version des LIAB-Längsschnittmodells**

Kennzahl	Datensatz	
	West	Ost
Anzahl Betriebsnummern der Version 1 (Grundlage)	4.937	4.075
Anzahl Personen im Datensatz	2.358.450	1.100.954
Meldungen insgesamt	28.469.903	15.816.927
Meldungen von Leistungsbezug	3.180.677	4.462.824
Meldungen von Personen, die in beiden Datensätzen enthalten sind	211.368	211.368
Meldungen aus dem jeweils anderen Landesteil	871.801	1.295.233
Anzahl Personen, die mindestens einen Tag zeitgleich mehr als ein Beschäftigungsverhältnis haben	473.379	177.010
Gesamtanzahl der Betriebsnummern	899.971	469.153
Anzahl weiterer IAB-Betriebspanelnummern**	29.242	17.164
<b>ausgewählte Auszählungen von Meldungen</b>		
Meldungen von Frauen	10.532.429	7.758.912
Meldungen von Personen mit der deutschen Staatsbürgerschaft	25.224.753	15.648.658
Meldungen geringfügige Beschäftigung	1.132.096	625.774
Meldungen von Auszubildenden	1.753.255	785.627
Meldungen Teilzeit < 15 h bzw. 18 h	2.348.775	902.360
Meldungen Teilzeit > 15 h bzw. 18 h	2.603.070	1.522.514
Beschäftigungsmeldungen: Stellung im Beruf unbekannt	74.242	96.149
Beschäftigungsmeldungen: Ausbildung unbekannt	3.168.839	1.660.185
Beschäftigungsmeldungen: Fehlkodierung Ausbildung	23.930	33.288
Datensatzgröße in MB*	1.879	1.044
Anzahl Variablen	31	31

\* Größe der entsprechenden STATA® 8.2. SE-Datensätze;

\*\* zum Teil erst im Jahr 2003 erstmals befragt

In dem Längsschnittdatensatz sind die Ebenen Betriebe, Personen und Meldungen zu unterscheiden. Letztere liegen für einzelne Personen vor und bilden die Erwerbsgeschichte einer Person ab. Grundlage für die LIAB-Version sind 4937 Betriebe (Westdeutschland) bzw. 4075 Betriebe (Ostdeutschland). Diese jeweils etwa vier- bzw. fünftausend Betriebe haben Meldungen von etwa 2,3 Millionen (West) bzw. 1,1 Mio. sozialversicherungspflichtig Beschäftigten. Die Anzahl der Meldungen beträgt für Westdeutschland über 28 Mio. und für Ostdeutschland knapp 16 Mio. Meldungen. Von den Meldungen sind über 3,1 Mio. in Westdeutschland Leistungsbezugsmeldungen, im Os-

Osten sind es etwa 4,5 Millionen. 211 Tsd. Meldungen sind sowohl im Datensatz für Westdeutschland als auch im Datensatz für Ostdeutschland enthalten. Der Grund hierfür ist, dass etwa 11 Tsd. Personen nicht nur zwischen Betriebspanelbetrieben, sondern auch von West- nach Ostdeutschland (bzw. von Ost nach West) gewechselt sind. Bei einer Analyse für Deutschland muss dies berücksichtigt werden und diese Meldungen aus einem der beiden Datensätze gelöscht werden. Da es auch Migration außerhalb der Betriebspanelbetriebe gibt, sind in beiden Datensätze jeweils Meldungen aus dem anderen Geltungsbereich enthalten (West: ca. 870 Tsd. ; Ost: ca. 1,3 Mio.).

Da Personen auch mehr als ein Beschäftigungsverhältnis haben können, wird in Tabelle 4 auch die Anzahl Personen ausgewiesen, die mindestens einen Tag zeitgleich mehr als ein Beschäftigungsverhältnis haben (West: ca. 473.000; Ost: ca. 177.000).

Personen wechseln ihre Beschäftigungsverhältnisse, und deshalb sind in dem Datensatz nicht nur die knapp 5000 bzw. 4000 Betriebspanelbetriebe enthalten, die zu der Auswahl geführt haben, sondern insgesamt fast 900 Tsd. (West) bzw. 470 Tsd. (Ost) Betriebsnummern. Da Personen auch in Betriebspanelbetriebe wechseln können, die nicht für die Auswahl der Stichprobe herangezogen wurden, sind diese ebenfalls in Tabelle 4 ausgewiesen (West: 29.000; Ost: 17.000). Hier muss der Anwender darauf achten, dass eine Aggregation dieser Betriebsnummern zu unsinnigen Ergebnissen führt und sich wirklich nur auf die für die Stichprobe relevanten jeweils ca. fünf- oder viertausend Betriebe beschränken. Schließlich werden in der Tabelle einige Auszählungen von Meldungen vorgenommen, die den Anwendern eine Abschätzung der zu erwartenden Fallzahlen bei ausgewählten Meldegruppen erlaubt.

## Anhang

### Anhang A1: Die Variablen der Personendatenbanken der Version 2 des LIAB-Längsschnittmodells

Nachfolgend werden die Variablen in den LIAB-Personendatenbanken beschrieben. Für kategorisierte Variablen werden die einzelnen Ausprägungen benannt. Bei sehr vielen möglichen Ausprägungen werden manchmal Beispiele vorgestellt. Es wird gesagt, ob, wie und wo der vollständige Wertebereich für diese Variablen eingesehen werden kann. Manchmal werden auch kontinuierliche Wertebereichen erläutert. Eine genauere Übersicht über alle Variablen, ihre Bedeutung und ihren einzelnen Ausprägungen befinden sich im vierten Kapitel des LIAB-Datenhandbuchs.

#### a) Organisationsvariablen der Personendatenbanken

##### eindeutiger Personenidentifikator

###### *Beschreibung:*

Die ursprüngliche Sozialversicherungsnummer wird in den LIAB-Daten in *versionsspezifische* eindeutige systemfreie Personenidentifikatoren umgeschlüsselt. Versionspezifisch bedeutet, dass Personen aus verschiedenen LIAB-Versionen nicht eindeutig zugeordnet sind. Die Identifikationsnummer 1 beispielsweise bezeichnet in der Version 1 des LIAB-Quer- und Längsschnittmodells *verschiedene* Personen. In allen einzelnen Jahresscheiben bezeichnet der Identifikator auf der Personenebene immer die gleiche Person.

##### eindeutiger Betriebsidentifikator

###### *Beschreibung:*

Die ursprüngliche Betriebsnummer wird in den LIAB-Daten in eindeutige Betriebsidentifikatoren überführt. Diese Identifikationsnummer wird auf der Ebene des IAB-Betriebspanels und aller Versionen der LIAB-Quer- und Längsschnittmodelle beibehalten und bezeichnet immer exakt den gleichen Betrieb.

##### Spellzähler personenbezogen

###### *Beschreibung:*

Die Variable bezeichnet die Nummer der konkreten Meldung im Rahmen des Gesamtauszugs aus dem jeweiligen *Personenkonto*, d.h. sie zählt die Meldungen der entsprechenden Person, seitdem sie zum ersten Mal in der Beschäftigten-Leistungsempfänger-Historik aufgetaucht ist.

###### *Wertebereich:*

Die Variable kann Werte von 1 bis n annehmen, wobei n eine natürliche Zahl ist.

## Spellzähler der Episode pro Quelle

### *Beschreibung:*

Mit der Variable sind Meldungen *innerhalb* von Beschäftigungs- oder Leistungsmeldungen gekennzeichnet, die sich zeitlich überschneiden. Wichtig ist vor allem der jeweilige Wert bei Beschäftigungsmeldungen. Die Hauptbeschäftigung bekommt den Wert Null zugewiesen, danach wird pro Quelle wie bei dem personenbezogenen Spellzähler fortlaufend nummeriert. Eine Hauptbeschäftigung ist dabei wie folgt definiert.

1. Versicherungsnummer
2. Beginndatum
3. Geringfügigkeitskennzeichen
4. Tagesentgelt (absteigend)

Das höchste Tagesentgelt hat also i.d.R. den Wert Null und kennzeichnet die Hauptbeschäftigung.

### *Wertebereich:*

Die Reihenfolge der Nennung ist eine Priorisierung der Einstufung, d.h. die Variable hat nur dann größere Werte als Null, wenn sowohl die Versicherungsnummer wie auch das Beginndatum und das Geringfügigkeitskennzeichen gleiche Werte beinhalten. In diesem Fall hat die Variable den Wert Null für die Meldung, in der Personen das höchste Tagesentgelt erzielen (gleich Hauptbeschäftigung). Mit absteigendem Entgelt wird dann aufsummiert.

## Status der Betriebsnummer

### *Beschreibung:*

Die Variable dokumentiert, ob es sich um einen Betrieb handelt, der sich in der ersten Version des LIAB-Längsschnittmodells befindet, ob es sich um einen anderen IAB-Betriebspanel-betrieb handelt oder ob es eine Betriebsnummer ist, mit der keine Verbindung zum IAB-Betriebspanel hergestellt werden kann.

### *Wertebereich (Ausprägungen) :*

Betrieb der Version 1 des LIAB-Längsschnittmodells oder  
sonstiger IAB-Betriebspanelbetrieb oder  
sonstige Betriebsnummern

## **b) Merkmale überwiegend aus Beschäftigtenmeldungen**

### **Geschlecht der Person**

*Beschreibung:*

Aus der Sozialversicherungsnummer kann an bestimmten Satzstellen abgelesen werden, ob es sich bei der Person um einen Mann oder eine Frau handelt.

*Wertebereich (Ausprägungen):*

Mann            oder            Frau

### **Berufskennziffer (3-stellig + Null)**

*Beschreibung:*

Für die Beschäftigungsmeldungen wird von den Betrieben eine dreistellige Berufskennziffer eingetragen, hinter die eine Null zugefügt wird. Eingetragen werden soll von dem Betrieb die tatsächlich ausgeübte und nicht – falls abweichend – die erlernte Tätigkeit. Bei Leistungsmeldungen ist die Berufskennziffer auf der Vierstellerebene eingetragen.

*Wertebereich:*

Auf den Internet-Seiten des Forschungsdatenzentrums können die einzelnen Ausprägungen in einem PDF-Dokument eingesehen werden.

### **Arbeitsamtsbezirk am Wohnort**

*Beschreibung:*

Diese Variable ist ab dem Jahr 1999 gefüllt und enthält die Arbeitsamtsbezirksnummer des Wohnorts von Personen.

*Wertebereich:*

Die einzelnen Arbeitsamtsbezirksnummern sind im FDZ einsehbar. Auf den Internet-Seiten des Forschungsdatenzentrums befinden sich die einzelnen Ausprägungen in einem PDF-Dokument.

### **Kreiskennziffer des Wohnortes**

*Beschreibung:*

Die Variable ist ab dem Jahr 1999 gefüllt und enthält die Kreiskennziffer des Wohnorts von Personen.

*Wertebereich:*

Die einzelnen Kreiskennziffern sind im FDZ einsehbar. Auf den Internet-Seiten des Forschungsdatenzentrums befinden sich die einzelnen Ausprägungen in einem PDF-Dokument.

**sozialversicherungspflichtiges Tagesentgelt***Beschreibung:*

Eingetragen ist das sozialversicherungspflichtige Brutto-Tagesentgelt für eine Beschäftigtenmeldung bzw. der tägliche Leistungssatz.

*Wertebereich:*

Der Wertebereich ist links und rechts auf die untere und obere Beitragsbemessungsgrenze begrenzt. Diese Grenzen unterscheiden sich von Jahr zu Jahr und in West- und Ostdeutschland (maßgeblich ist der Standort des Betriebs). Eine Übersicht über diese Grenzen sind auf den Internet-Seiten des FDZ verfügbar. Dort befinden sich auch einige nützliche Erläuterungen, wie das Tagesentgelt berechnet wird.

**Beginndatum der Episode***Beschreibung:*

Eingetragen ist das Beginndatum der Meldung.

*Wertebereich:*

Die Werte sind im SAS-, bzw. STATA-Datumsformat kodiert. Die numerische Variable zählt die seit dem 01. 01. 1960 vergangenen Tage. Die frühesten Meldungen in den Personendatenbanken sind die des Jahres 1990 (in diesem Jahr nur Leistungsmeldungen). Zu beachten sind die Schaltjahre 1992, 1996 und 2000. Nachfolgend wird jeweils der erste Januar eines jeden Jahres von 1991 bis 2002 ausgewiesen.

<b>Jahr</b>	<b>Datumswert</b>	<b>Jahr</b>	<b>Datumswert</b>
1991	11323	1997	13515
1992	11688	1998	13880
1993	12054	1999	14245
1994	12419	2000	14610
1995	12784	2001	14976
1996	13149	2002	15341

### **Enddatum der Episode**

*Beschreibung:*

Eingetragen ist das Enddatum der Meldung.

*Wertebereich:*

Wie beim Beginndatum der Episode. Mit den beiden Variablen können Jahresmeldungen von Unterbrechungs-, Veränderungs- und Beendigungsmeldungen unterschieden werden (vgl. Variable Dauer der Meldung in Tagen auf der übernächsten Seite).

### **Abgabegrund der Meldung**

*Beschreibung:*

Hier werden Gründe für die Abgabe von Beschäftigungsmeldungen eingetragen. Meistens ist der Grund die Jahresmeldung, aber es kann sich auch um Beginn- oder Beendigungsmeldungen von Beschäftigungsverhältnissen handeln. Die einzelnen Abgabegründe sind umfangreich und vielschichtig. In den Testdaten zur ersten Version des LIAB-Querschnittmodells sind einige häufiger vorkommende Gründe mit Variablenlabels versehen.

*Wertebereich:*

Auf den Internet-Seiten des Forschungsdatenzentrums können die einzelnen Ausprägungen in einem PDF-Dokument eingesehen werden.

### **Personengruppe**

*Beschreibung:*

Die Variable kann u.a. dazu verwendet werden, um geringfügig Beschäftigte zu identifizieren. Eingetragen ist eine dreistellige numerische Zahl, die eine sehr detaillierte Unterscheidung von Beschäftigten erlaubt.

*Wertebereich:*

Auf den Internet-Seiten des Forschungsdatenzentrums können die einzelnen Ausprägungen in einem PDF-Dokument eingesehen werden.

### **Stellung im Beruf**

*Beschreibung:*

Das ist der so genannte B2-Schlüssel der Beschäftigtenmeldungen von den Betrieben. Es wird empfohlen, die Anmerkungen zur Stellung im Beruf auf den Internet-Seiten des FDZ zu beachten.



*Wertebereich (Ausprägungen):*

Auszubildende

nicht formal Qualifizierte

Facharbeiter

Meister, Poliere u.ä.

Angestellte

Personen mit diesen Merkmalen sind immer vollzeitbeschäftigt. Daneben gibt es noch folgende Ausprägungen:

Heimarbeiter

Teilzeitbeschäftigte ohne eigene Beitragszahlungen aus ihrem Beschäftigungsverhältnis (i.d.R. unter 18 Stunden/ Woche)

Teilzeitbeschäftigte mit eigenen Beitragszahlungen aus ihrem Beschäftigungsverhältnis (i.d.R. über 18 Stunden/ Woche)

### **Bildungsniveau in Verbindung mit der beruflichen Qualifikation**

*Beschreibung:*

Das ist der so genannte B1-Schlüssel der Beschäftigtenmeldungen von Betrieben. Es wird empfohlen, die Anmerkungen zur Stellung im Beruf auf den Internet-Seiten des FDZ zu beachten.

*Wertebereich (Ausprägungen):*

ohne abgeschlossene Berufsausbildung

bis mittlere Reife mit Berufsausbildung

(Fach)Hochschulreife ohne abgeschlossene Berufsausbildung

(Fach)Hochschulreife mit Berufsausbildung

Fachhochschulabschluss

Hochschulabschluss

unbekannt, Angabe nicht möglich

### **Arbeitsamtbezirk des Arbeitsortes**

*Beschreibung:*

Diese Variable ist ab dem Jahr 1999 gefüllt und enthält die Arbeitsamtsbezirksnummer des Arbeitsortes von Personen (Betrieb).

*Wertebereich:*

Auf den Internet-Seiten des Forschungsdatenzentrums können die einzelnen Ausprägungen in einem PDF-Dokument eingesehen werden.

### **Kreiskennziffer des Arbeitsortes**

*Beschreibung:*

Die Variable ist ab dem Jahr 1999 gefüllt und enthält die Kreiskennziffer des Arbeitsortes von Personen (Betrieb).

*Wertebereich:*

Auf den Internet-Seiten des Forschungsdatenzentrums können die einzelnen Ausprägungen in einem PDF-Dokument eingesehen werden.

### **Wirtschaftszweig nach der BA-Klassifikation 1973**

*Beschreibung:*

Die Variable enthält auf der Dreisteller-Ebene den Wirtschaftszweig des Beschäftigungsbetriebs.

*Wertebereich:*

Es wird auf die hiesigen Dokumentationen der Bundesagentur für Arbeit verwiesen.

### **Wirtschaftszweig nach der BA-Klassifikation 1993**

*Beschreibung:*

Die Variable enthält auf der Fünfsteller-Ebene den Wirtschaftszweig des Beschäftigungsbetriebs (ab 1999 gefüllt).

*Wertebereich:*

Es wird auf die hiesigen Dokumentationen der Bundesagentur für Arbeit verwiesen.

### **Dauer der Meldung in Tagen**

*Beschreibung:*

Die Variable wird generiert durch Enddatum minus Beginndatum und enthält die Dauer der Meldung in Tagen.

*Wertebereich:*

Bei Beschäftigungsmeldungen wird mindestens einmal jährlich eine Meldung abgegeben, deswegen ist dort das Wertebereich [1; 366]. Leistungsmeldungen können unter Umständen (i.d.R. nur bei Arbeitslosenhilfe) auch länger andauern.

### **Alter der Person**

*Beschreibung:*

Die Werte aus der Sozialversicherungsnummer generiert, die u.a. das Geburtsjahr der Personen enthält.

*Wertebereich:*

Meist ist das Wertebereich [16;65], in Ausnahmefällen auch darüber.

### **Währungskennzeichen**

*Beschreibung:*

Die Variable gibt an, ob das sozialversicherungspflichtige Brutto-Tagesentgelt in DM oder Euro angegeben ist. In der Regel sind bis 1999 DM Beträge angegeben, danach Euro-Beträge (der amtliche Umrechnungskurs beträgt 1.95883 DM = 1 Euro).

*Wertebereich (Ausprägungen)*

Euro und DM

### **Ost/Westkennzeichen der Meldung**

*Beschreibung:*

Es ist angegeben, aus welchem Landesteil die Meldung erfolgte.

*Wertebereich (Ausprägungen):*

Westdeutschland und Ostdeutschland

### **Staatenschlüssel für Personen**

*Beschreibung:*

Die Variable gibt die Staatsangehörigkeit der Person an und ist bei Beschäftigungs- und Leistungsmeldungen gefüllt.

*Wertebereich:*

Der Wertebereich der bis zu dreistelligen Variable kann auf den Internet-Seiten des Forschungsdatenzentrums bzw. die einzelnen Ausprägungen in einem PDF-Dokument eingesehen werden.

**c) Merkmale, die sich mehrheitlich auf Leistungsbezug durch die Bundesagentur für Arbeit beziehen****Beschäftigung oder Leistung***Beschreibung:*

Mit der Variable können Beschäftigungs- und Leistungsmeldungen unterschieden werden. Leistungsmeldungen sind nur verfügbar, wenn die Person Leistungen durch die Bundesagentur für Arbeit erhält. Gemeldete Arbeitslosigkeit ohne Anspruch auf Bezüge ist *nicht* enthalten. Die nachfolgenden Variablen sind alle nur bei dem Wert zwei für diese Variable gefüllt.

*Wertebereich (Ausprägungen):*

Beschäftigungsmeldung oder Leistungsmeldung

**gruppierte Leistungsart***Beschreibung:*

In der Variable werden die vielfältigen BA-Leistungsarten zusammengefasst.

*Wertebereich (Ausprägungen):*

Arbeitslosengeld

Arbeitslosenhilfe

Fort- und Weiterbildung (Unterhaltsgeld)

Beitrag § 166b AFG/§ 207a SGB III (z.B. Zuschüsse zur Rentenversicherung)

Fehlkodierung

**Kennzeichen Art der Bewilligung***Beschreibung:*

In der Variable werden die Bewilligungsgründe für den Leistungsbezug zusammengefasst.

*Wertebereich (Ausprägungen):*

Erst-/Neubewilligung Arbeitslosengeld

Weiterbewilligung Arbeitslosengeld ohne Unterbrechung

Weiterbewilligung Arbeitslosengeld nach Unterbrechung

Neubewilligung Arbeitslosengeld mit Aufstockung (ein neu erworbener Anspruch wird durch einen bestehenden Restanspruch aufgestockt)

Weiterbewilligung Arbeitslosengeld mit Kürzung (z.B. durch Sperrzeit)

Arbeitslosenhilfe-Satz, der zeitlich direkt an Arbeitslosengeld-Satz anschließt

Arbeitslosenhilfe-Satz, der zeitlich direkt an vorhergehenden Arbeitslosenhilfe-Satz anschließt

Arbeitslosenhilfe-Satz mit Lücke zum vorhergehenden Satz

alle anderen Sätze

### **Zusammengefasste Abmeldegründe**

*Beschreibung:*

Hier wird angegeben, warum eine Leistungsbezugsmeldung endet.

*Wertebereich (Ausprägungen):*

Arbeitsaufnahme

Anspruch auf andere Leistung

sonstige Gründe

steht dem Arbeitsmarkt nicht zur Verfügung

Ablauf der Maßnahme

Sperr-, Säumniszeiten

Anspruch erschöpft

Bewilligungsabschnitt abgelaufen

Fehlkodierung/fehlt

### **Zusammenfassung Familienstand**

*Beschreibung:*

Hier ist bei Leistungsempfang ausgewiesen, ob die Person ledig oder verheiratet ist. Die Variable ist nicht sonderlich zuverlässig gefüllt.

*Wertebereich (Ausprägungen):*

nicht verheiratet      oder      verheiratet

### **Kinder der Person**

*Beschreibung:*

Die Variable enthält eine qualitative Information über Kinder der Person bei Leistungsbezug. Die Variable ist nicht sonderlich zuverlässig gefüllt.

*Wertebereich (Ausprägungen):*

keine Kinder      oder      Kinder vorhanden

Anhangstabelle A2: Zusammenfassung von Wirtschaftszweigen

Nr.	Wirtschaftszweig IAB-Betriebspanel	Wirtschaftszweig Version 1	Nr.
1	Land- und Forstwirtschaft, Fischereiwesen	<b>Verarbeitendes Gewerbe I (inkl. primärem Sektor)</b>	<b>1</b>
2	Bergbau, Energie- und Wasserversorgung		
3	Nahrungs- und Genussmittelherstellung		
4	Textil- und Bekleidungsindustrie		
5	Papier- und Druckgewerbe, Verlage		
6	Holzgewerbe (ohne Möbelherstellung)		
7	Chemische Industrie, Mineralölverarbeitung		
8	Herstellung von Gummi- und Kunststoffwaren		
9	Glasgewerbe, Keramik, Verarb. von Steinen und Erden		
10	Metallerzeugung und -bearbeitung	<b>wie Nr. 2</b>	<b>2</b>
11	Recycling	<b>wie Nr. 1</b>	<b>1</b>
12	Herstellung von Metallerzeugnissen, Stahlbau	<b>Metall- und Elektrowerkstoffe</b>	<b>2</b>
13	Maschinenbau		
14	Herstellung von Kraftwagen und Kraftwagenteilen		
15	Sonstiger Fahrzeugbau		
16	Elektrotechnik, Herstellung von Büromaschinen		
17	Feinmechanik und Optik		
18	Herst. von Möbeln, Schmuck, Musikinstrumenten etc.	<b>wie Nr. 1</b>	<b>1</b>
19	Bauhauptgewerbe	<b>Bauhauptgewerbe</b>	<b>3</b>
20	Baunebengewerbe		
21	Kraftfahrzeughandel und -reparatur	<b>Handel/ Verkehr/ Nachrichten</b>	<b>4</b>
22	Großhandel und Handelsvermittlung		
23	Einzelhandel (ohne Kfz), Rep. von Gebrauchsgütern		
24	Verkehr		
25	Nachrichtenübermittlung	<b>Finanzdienstleistungen</b>	<b>5</b>
26	Kreditinstitute		
27	Versicherungsgewerbe		
28	Datenverarbeitung und Datenbanken	<b>unternehmensnahe Dienstleistungen</b>	<b>6</b>
29	Forschung und Entwicklung		
30	Rechts-, Steuer- und Unternehmensberatungen		
31	Grundstücks- und Wohnungswesen		
32	Vermietung bewegl. Sachen, sonstige Dienstleistungen		
33	Gaststätten, Beherbergungsgewerbe	<b>andere Dienstleistungen</b>	<b>7</b>
34	Erziehung und Unterricht		
35	Gesundheits-, Veterinär- und Sozialwesen		

(Fortsetzung nächste Seite)

Nr.	Wirtschaftszweig IAB-Betriebspanel	Wirtschaftszweig Version 1	Nr.
36	Entsorgung, Abwasser- und Abfallbeseitigung		
37	Kultur, Sport, Unterhaltung		
38	Andere Dienstleistungen (Wäscherei, Reinigung etc.)		
39	Interessenvertretungen, Verbände, Kirchen u.a.		
40	Private Haushalte	<b>öffentlicher Sektor</b>	<b>8</b>
41	Öffentliche Verwaltung, Sozialversicherung		

**Imprint**

**FDZ *Datenreport***

No. 8/2005

**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/2005/DR\\_8.pdf](http://doku.iab.de/fdz/reporte/2005/DR_8.pdf)

**Internet**

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

**Corresponding author**

Dr. Holger Alda, Tel.: +49 (0) 911/179-3138  
E-Mail: [holger.alda@iab.de](mailto:holger.alda@iab.de)