

FDZ-Methodenreport

04/2013

DE

Methodische Aspekte zu Arbeitsmarktdaten

Programmierbeispiele zur Datenaufbereitung der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) in Stata

Generierung von Querschnittsdaten und biografischen
Variablen

August 2013 (2. aktualisierte Version)

Johanna Eberle,
Alexandra Schmucker,
Stefan Seth



Programmierbeispiele zur Datenaufbereitung der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) in Stata

Generierung von Querschnittsdaten und biografischen Variablen

Johanna Eberle (Institut für Arbeitsmarkt- und Berufsforschung)

Alexandra Schmucker (Institut für Arbeitsmarkt- und Berufsforschung)

Stefan Seth (Institut für Arbeitsmarkt- und Berufsforschung)

Die FDZ-Methodenberichte befassen sich mit den methodischen Aspekten der Daten des FDZ und helfen somit Nutzerinnen und Nutzern bei der Analyse der Daten. Nutzerinnen und Nutzer können hierzu in dieser Reihe zitationsfähig publizieren und stellen sich der öffentlichen Diskussion.

FDZ-Methodenberichte (FDZ method reports) deal with the methodical aspects of FDZ data and thus help users in the analysis of data. In addition, through this series users can publicise their results in a manner which is citable thus presenting them for public discussion.

Inhaltsverzeichnis

Zusammenfassung	4
Abstract	4
1 Einleitung	5
2 Kurzbeschreibung der Stata-Programme	6
3 Beschreibung der generierten biografischen Merkmale	7
3.1 Eintrittsdatum Erwerbsleben (ein_ew)	7
3.2 Dauer Erwerbsleben (tage_ew)	7
3.3 Eintrittsdatum Betrieb (ein_bet)	8
3.4 Dauer Betrieb (tage_bet)	8
3.5 Eintrittsdatum Beschäftigung (ein_job)	9
3.6 Dauer Beschäftigung (tage_job)	10
3.7 Anzahl bisheriger Leistungsbezüge (anz_ist)	10
3.8 Dauer bisheriger Leistungsbezüge (tage_ist)	11
4 Beschreibung der generierten Merkmale zu parallelen Status am Stichtag	12
4.1 Art der Nebenbeschäftigung (nb)	12
4.2 Nebenbeschäftigung im Betrieb der Hauptbeschäftigung (nb_betr)	12
4.3 Parallele Beobachtung aus LeH (leh)	13
4.4 Parallele Beobachtung aus (X)ASU (asu)	13
4.5 Parallele Beobachtung aus (X)LHG (lhg)	13
4.6 Gesamteinkommen aus allen Quellen (gtentgelt)	14
4.7 Stichtag des Querschnitts (stichtag)	14
Literatur	15
Anhang	16

Zusammenfassung

Der vorliegende FDZ-Methodenreport (einschließlich der Programmierbeispiele für Stata) beschreibt die Erstellung von Querschnittsdaten zu frei wählbaren Stichtagen und die Generierung von biografischen Merkmalen.

Abstract

The present FDZ-Methodenreport (including programming examples for Stata) describes the construction of cross-sectional data for freely selectable cutoff dates and the generation of biographic variables.

Keywords: Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB), Datenaufbereitung, Stichtagdaten, Datenstruktur

1 Einleitung

Der vorliegende FDZ-Methodenreport (einschließlich einiger Programmierbeispiele für Stata) soll Möglichkeiten zur Aufbereitung der SIAB-Daten aufzeigen und somit den Einstieg in die Arbeit mit diesen Daten erleichtern. Der Fokus der dargestellten Datenaufbereitungen und Variablengenerierungen liegt dabei auf der Vereinfachung der Datenstruktur.

Hauptziel der Datenaufbereitung ist hierbei die Erstellung von Querschnitten zu frei wählbaren Stichtagen. Zusätzlich sollen die Daten so reduziert werden, dass pro Person und Stichtag am Ende nur noch ein Haupterwerbsstatus vorliegt. Somit entsteht ein deutlich einfacher strukturierter Datensatz. Der Nachteil dieses Vorgehens besteht in dem Verlust an Informationen. Diese Einschränkungen können jedoch verringert werden, indem zuvor biografische Merkmale (wie beispielsweise Dauern oder Datumsangaben besonderer Ereignisse) generiert werden, die wichtige Informationen auch im späteren Querschnitt verfügbar machen. Darüber hinaus werden Variablen erzeugt, die ggf. parallele Erwerbszustände neben dem Hauptstatus anzeigen.

Die mitgelieferten Stata-Dofiles stellen lediglich einige Beispiele dar, die von den Nutzerinnen und Nutzern nach eigenen Vorstellungen geändert und erweitert werden können¹. Als Datenbasis wird die Stichprobe der Integrierten Arbeitsmarktbiografien 1975-2010 - SIAB 7510 (vom Berge et al. 2013) verwendet. Da die Aufbereitung der neueren administrativen Personendaten, die das FDZ anbietet, standardisiert ist, können die Programme nach marginalen Anpassungen auch für andere Datenprodukte wie beispielsweise ALWA-ADIAB (ALWA-Befragungsdaten verknüpft mit administrativen Daten des IAB) verwendet werden.

Das vorgestellte Vorgehen vereinfacht nur die Datenstruktur. Es werden keine Verfahren zur Verbesserung der Datenqualität bzw. zur Ergänzung fehlender Angaben verwendet. In diesem Bereich gibt es bereits FDZ-Methodenreporte beispielsweise zur Imputation von Entgelten über der Beitragsbemessungsgrenze (Gartner 2005) oder zur Qualitätsverbesserung der Bildungsvariable (Fitzenberger et al. 2005 und Drews 2006).

¹ So gibt es beispielsweise zahlreiche Definitionen von Arbeitslosigkeit, die sich auf die Berechnung der Arbeitslosigkeitsdauer auswirken (Kruppe et al. 2007).

2 Kurzbeschreibung der Stata-Programme

Die angehängten Stata-Programme ermöglichen es, unterschiedliche biografische Merkmale sowie beliebig viele Querschnitte zu frei wählbaren Stichtagen aus den Längsschnittdaten zu erstellen. Die Programme sind wie folgt aufgebaut:

master.do: Hier werden Makros für Pfadangaben und Dateinamen sowie für die gewünschten Stichtage und Jahre vergeben. Die entsprechenden Angaben müssen von den Nutzern und Nutzerinnen selbst ergänzt werden. Anschließend werden aus diesem Do-File heraus die Programme zur Generierung der biografischen Merkmale und der Querschnitte gestartet. Am Ende wird der nicht mehr benötigte Zwischendatensatz gelöscht.

SIAB_bio.do: Hier werden die biografischen Merkmale (siehe Kapitel 3) auf Basis der Längsschnittdaten erzeugt. Dabei werden Dauern immer bis zum Endedatum der jeweiligen Beobachtung berechnet. Der abschließend gespeicherte Zwischendatensatz enthält neben den neuen Variablen noch alle Informationen des Ausgangsdatensatzes.

SIAB_quer.do: Hier werden auf Basis des mit SIAB_bio.do erstellten Datensatzes für jeden ausgewählten Stichtag zunächst Merkmale generiert, die parallele Erwerbszustände neben dem Hauptstatus anzeigen (siehe Kapitel 4). Anschließend wird für jede Person nur die Beobachtung mit dem Hauptstatus behalten und die zuvor berechneten Dauern bis zum jeweiligen Stichtag gekürzt. Am Ende wird für jeden Stichtag ein eigener Datensatz abgespeichert. Anhand des Personenidentifikators (persnr) und der Stichtagsvariable (stichtag) können anschließend die einzelnen Datensätze ggf. zu einem Paneldatensatz zusammengeführt werden.

Die angehängten Stata-Programme (Stata-Version 11) sind für die SIAB 7510 V1 erstellt und getestet. Für die Erstellung von Querschnitten auf Basis anderer Versionen oder Datensätze müssen ggf. Änderungen von den Nutzerinnen und Nutzern selbst vorgenommen werden. Dabei sollte auch immer die unterschiedlichen Versionen der Integrierten Erwerbsbiografien des IABs, aus denen die Prozessdaten des FDZ generiert werden, berücksichtigt werden. Diese und weitere relevanten Informationen können aus den datensatzspezifischen FDZ-Datenreporten entnommen werden. Das FDZ übernimmt für die inhaltliche Korrektheit der Variablen keine Garantie. Hier ist an den Nutzerinnen und Nutzern zu prüfen, ob die generierten Merkmale richtig und für die eigne Forschungsfrage geeignet sind.

Zum besseren Verständnis der Programme können Forscherinnen und Forscher, die mit der Verarbeitung der Längsschnittdaten des IAB keine oder wenig Erfahrung haben, den FDZ-Methodenreport 6/2007 (Drews et al. 2007) heranziehen. Als allgemeine Einführung zur Datenanalyse mit Stata empfehlen wir Kohler und Kreuter (2012a und 2012b).

3 Beschreibung der generierten biografischen Merkmale

3.1 Eintrittsdatum Erwerbsleben (ein_ew)

Variablenlabel	Eintrittsdatum Erwerbsleben
Variablenname	ein_ew
Kategorie	Generierte biografische Merkmale
Herkunft	Generiert aus BeH
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal enthält das Beginndatum der ersten Beschäftigung einer Person in der SIAB. Nicht berücksichtigt sind Ausbildungszeiten (Stellung im Beruf = 0). Personen, die in der SIAB eine Ausbildungszeit aber keine sozialversicherungspflichtige oder geringfügige Beschäftigung aufweisen, haben immer ein Missing. Das Merkmal ist für den kompletten Zeitraum, der in den Daten für eine Person enthalten ist, konstant.</p> <p>Der Eintritt in das Erwerbsleben (ein_ew) kann lange nach dem Eintritt in den Betrieb (ein_bet) oder dem Eintritt in die Beschäftigung (ein_job) erfolgen, da bei letzteren Ausbildungszeiten einbezogen sind.</p>
Qualitätshinweis	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.

3.2 Dauer Erwerbsleben (tage_ew)

Variablenlabel	Dauer Erwerbsleben
Variablenname	tage_ew
Kategorie	Generierte biografische Merkmale
Herkunft	Generiert aus BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal gibt die Anzahl der Tage an, die eine Person bis zum Enddatum der gegenwärtigen Episode bzw. nach Erstellung der Querschnittsdaten bis zum jeweiligen Stichtag des Querschnittes beschäftigt war. Nicht berücksichtigt sind Ausbildungszeiten (Stellung im Beruf = 0). War eine Person nur in Ausbildung, hat das Merkmal den Wert 0.
Qualitätshinweis	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.
--	--

3.3 Eintrittsdatum Betrieb (ein_bet)

Variablenlabel	Eintrittsdatum Betrieb
Variablenname	ein_bet
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus BeH
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal enthält das Beginndatum der ersten Beschäftigungsmeldung einer Person in der SIAB im aktuellen Betrieb. Dabei werden auch Ausbildungszeiten berücksichtigt. Eine Unterbrechung der Beschäftigung im Betrieb ändert das Eintrittsdatum nicht, d.h. es ist eindeutig für jede Kombination von Person und Betriebsnummer. Bei fehlenden oder ungültigen Betriebsnummern erhält das Merkmal ein Missing.</p> <p>Der Eintritt in das Erwerbsleben (ein_erw) kann lange nach dem Eintritt in den Betrieb (ein_bet) oder dem Eintritt in die Beschäftigung (ein_job) erfolgen, da bei letzteren Ausbildungszeiten einbezogen sind.</p>
Qualitätshinweis	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.

3.4 Dauer Betrieb (tage_bet)

Variablenlabel	Dauer Betrieb
Variablenname	tage_bet
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal enthält die Anzahl der Tage, die eine Person im betreffenden Betrieb bis zum Endedatum der gegenwärtigen Episode bzw. nach Erstellung der Querschnittsdaten bis zum jeweiligen Stichtag des Querschnittes beschäftigt war. Dabei werden auch Ausbildungszeiten berücksichtigt, Beschäftigungslücken jedoch nicht.</p> <p>Berechnet man die Dauer im Betrieb alternativ mit dem Merkmal Ein-</p>

	<p>tritt Betrieb (ein_bet)</p> $\text{gen tage_bet_neu} = \text{endepi} - \text{ein_bet} + 1$ <p>erhält man Werte, die größer oder gleich tage_bet sind, da tage_bet Unterbrechungen der Beschäftigung nicht enthält.</p>
Qualitätshinweis	<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>

3.5 Eintrittsdatum Beschäftigung (ein_job)

Variablenlabel	Eintrittsdatum Beschäftigung
Variablenname	ein_job
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal enthält das Beginndatum der ersten Beschäftigungsmeldung einer Person im aktuellen Beschäftigungsverhältnis.</p> <p>Ausbildungszeiten (Stellung im Beruf = 0) im gleichen Betrieb werden als separate Beschäftigungsverhältnisse gezählt, auch wenn sie unmittelbar vor oder nach sonstigen Beschäftigungen anschließen.</p> <p>Beschäftigungen nach Unterbrechungen im selben Betrieb zählen als neues Beschäftigungsverhältnis, wenn</p> <ul style="list-style-type: none"> - der Meldegrund der letzten Meldung vor der Unterbrechung auf ein Ende des Beschäftigungsverhältnisses deutet (Abgabegrund der Meldung = 30, 34, 40 oder 49) und die Unterbrechung länger als 92 Tage dauert oder - der Meldegrund der letzten Meldung vor der Unterbrechung nicht auf ein Ende des vorhergehenden Beschäftigungsverhältnisses deutet (alle anderen Abgabegründe der Meldung) und die Unterbrechung länger als 366 Tage dauert. <p>Der Eintritt in das Beschäftigungsverhältnis (ein_job) kann nicht vor dem Eintritt in den Betrieb (ein_bet) liegen, wohl aber vor dem Eintritt in das Erwerbsleben (ein_erw).</p>
Qualitätshinweis	<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>

3.6 Dauer Beschäftigung (tage_job)

Variablenlabel	Dauer Beschäftigung
Variablenname	tage_job
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal enthält die Anzahl der Tage, die eine Person bis zum Enddatum der gegenwärtigen Episode bzw. nach Erstellung der Querschnittsdaten bis zum jeweiligen Stichtag des Querschnittes im betreffenden Beschäftigungsverhältnis beschäftigt war.</p> <p>Ausbildungszeiten (stib = 0) im gleichen Betrieb werden als separate Beschäftigungsverhältnisse gezählt, auch wenn sie unmittelbar vor oder nach sonstigen Beschäftigungen anschließen.</p> <p>Beschäftigungen nach Unterbrechungen im selben Betrieb zählen als neues Beschäftigungsverhältnis, wenn</p> <ul style="list-style-type: none"> - der Meldegrund der letzten Meldung vor der Unterbrechung auf ein Ende des Beschäftigungsverhältnisses deutet (Abgabegrund der Meldung = 30, 34, 40 oder 49) und die Unterbrechung länger als 92 Tage dauert oder - der Meldegrund der letzten Meldung vor der Unterbrechung nicht auf ein Ende des vorhergehenden Beschäftigungsverhältnisses deutet (alle anderen Abgabegründe der Meldung) und die Unterbrechung länger als 366 Tage dauert. <p>Berechnet man die Dauer Beschäftigung alternativ mit dem Merkmal Eintritt Beschäftigung (ein_job)</p> $\text{gen tage_job_neu} = \text{endepi} - \text{ein_job} + 1$ <p>erhält man Werte, die größer oder gleich tage_job sind, da tage_job Unterbrechungen der Beschäftigung nicht enthält.</p>
Qualitätshinweis	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.

3.7 Anzahl bisheriger Leistungsbezüge (anz_ist)

Variablenlabel	Anzahl bisheriger Leistungsbezüge
Variablenname	anz_ist
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus LeH/LHG/XLHG

Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal enthält die Anzahl der Leistungsbezüge einer Person einschließlich der gegenwärtigen Episode.</p> <p>Leistungsbezüge nach SGB II und SGB III gehen gleichermaßen ein. 2005 ändert sich also die Bedeutung des Merkmals.</p> <p>Bei Lücken von weniger als 10 Tagen zwischen zwei Leistungsbezugszeiten wird von weiter laufendem Leistungsbezug ausgegangen, also nicht hochgezählt. Wechselt die Leistungsart wird das ebenfalls nicht als neuer Leistungsbezug gezählt.</p>
Qualitätshinweis	Das Merkmal ist linkszensiert, für Westdeutsche am 1.1.1975. Für Ostdeutsche ist die Zensierungsgrenze nicht eindeutig. Die SIAB enthält keine Informationen zum Leistungsbezug vor 1990 und noch unvollständige für 1990 und 1991.

3.8 Dauer bisheriger Leistungsbezüge (tage_1st)

Variablenlabel	Dauer bisheriger Leistungsbezüge
Variablenname	tage_1st
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus LeH/LHG/XLHG
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal enthält die Dauer der Leistungsbezüge einer Person bis Enddatum der gegenwärtigen Episode bzw. nach Erstellung der Querschnittsdaten bis zum jeweiligen Stichtag des Querschnittes.</p> <p>Leistungsbezüge nach SGB II und SGB III gehen gleichermaßen ein. 2005 ändert sich also die Bedeutung des Merkmals.</p> <p>Lücken werden nicht aufaddiert.</p> <p>Aus verschiedenen Gründen ist es möglich, dass jemand einer (sozialversicherungspflichtigen oder geringfügigen) Beschäftigung nachgeht und gleichzeitig Leistungen erhält. In diesem Fall werden die Leistungsbezugszeiten in tage_1st mitgezählt.</p>
Qualitätshinweis	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.

4 Beschreibung der generierten Merkmale zu parallelen Status am Stichtag

4.1 Art der Nebenbeschäftigung (nb)

Variablenlabel	Art der Nebenbeschäftigung
Variablenname	nb
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal weist aus, ob und in welcher Form eine Nebenbeschäftigung zum jeweiligen Stichtag vorliegt. Dabei wird nur die 1. Nebenbeschäftigung berücksichtigt, Informationen zu weiteren parallelen Beschäftigungen gehen verloren. Dabei wird in erster Linie zwischen Voll- und Teilzeitbeschäftigungen unterschieden. Ab 1999 werden auch geringfüge Beschäftigungen erfasst und abgebildet. Nebenbeschäftigungen, die entweder in der Variable „Erwerbsstatus“ oder in der Variable „Stellung im Beruf“ keine gültige Angaben haben und daher nicht als Voll-, Teilzeit oder geringfügige Beschäftigung klassifiziert werden können, werden als „Nebenjob ohne nähere Angaben“ aufgeführt. Für Personen, die zum Stichtag keine Nebenbeschäftigung haben, enthält das Merkmal Missings.</p> <p>Ausprägungen und Labels:</p> <ol style="list-style-type: none"> 1 Vollzeit-Job 2 Teilzeit-Job 3 geringfügige Beschäftigung 4 Nebenjob ohne nähere Angaben
Qualitätshinweis	-

4.2 Nebenbeschäftigung im Betrieb der Hauptbeschäftigung (nb_betr)

Variablenlabel	Nebenbeschäftigung im Betrieb der Hauptbeschäftigung
Variablenname	nb_betr
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal weist aus, ob die Nebenbeschäftigung zum jeweiligen Stichtag vom gleichen Betrieb wie die Hauptbeschäftigung gemeldet wurde. Sollte keine gültige Betriebsnummer in der Haupt- oder der Nebenbeschäftigung vorliegen, enthält die Variable Missings.</p> <p>Ausprägungen und Labels:</p>

	0 anderer Betrieb 1 gleicher Betrieb
Qualitätshinweis	-

4.3 Parallele Beobachtung aus LeH (leh)

Variablenlabel	Parallele Beobachtung aus LeH
Variablenname	leh
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus LeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal weist aus, ob ob zum jeweiligen Stichtag zusätzlich zur Hauptbeobachtung eine Beobachtung aus der LeH vorliegt.
Qualitätshinweis	-

4.4 Parallele Beobachtung aus (X)ASU (asu)

Variablenlabel	Parallele Beobachtung aus (X)ASU
Variablenname	asu
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus ASU/XASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal weist aus, ob zum jeweiligen Stichtag zusätzlich zur Hauptbeobachtung eine Beobachtung aus der Arbeitssuchenden-Historik (ASU) oder aus XSozial-BA-SGB II (XASU) vorliegt.
Qualitätshinweis	-

4.5 Parallele Beobachtung aus (X)LHG (lhg)

Variablenlabel	Parallele Beobachtung aus (X)LHG
Variablenname	asu
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus LHG/XLHG
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal weist aus, ob zum jeweiligen Stichtag zusätzlich zur Hauptbeobachtung eine Beobachtung aus den Leistungs-Historiken Grundsicherung aus A2LL(LHG) oder aus XSozial-BA-SGB II

	(XLHG):vorliegt.
Qualitätshinweis	-

4.6 Gesamteinkommen aus allen Quellen (gtentgelt)

Variablenlabel	Gesamteinkommen aus allen Quellen
Variablenname	gtentgelt
Kategorie	Generierte biographische Merkmale
Herkunft	Generiert aus BeH/LeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal enthält die Summe aller Einkommen aus Beschäftigungsmeldungen und Leistungsbezugsätzen zum jeweiligen Stichtag.
Qualitätshinweis	-

4.7 Stichtag des Querschnitts (stichtag)

Variablenlabel	Stichtag des Querschnitts
Variablenname	stichtag
Kategorie	Generierte technische Merkmale
Herkunft	Generiert
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	Das Merkmal enthält das Datum des jeweiligen Stichtags, für den der Querschnitt erstellt wurde.
Qualitätshinweis	-

Literatur

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

Drews, Nils; Groll, Dominik; Jacobebbinghaus, Peter (2007): Programmierbeispiele zur Aufbereitung von FDZ Personendaten in STATA. FDZ Methodenreport, 06/2007

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

Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert (2005): Imputation rules to improve the education variable in the IAB employment subsample. FDZ Methodenreport, 03/2005

Gartner, Hermann (2005): The imputation of wages above the contribution limit with the German IAB employment sample. FDZ Methodenreport, 02/2005

Kohler, Ulrich; Kreuter, Frauke (2012a): Datenanalyse mit Stata * allgemeine Konzepte der Datenanalyse und ihre praktische Anwendung. 4. Auflage. München: Oldenbourg Wissenschaftsverlag

Kohler, Ulrich; Kreuter, Frauke (2012b): Data Analysis Using Stata. Third Edition. Stata Press

Kruppe, Thomas; Müller, Eva; Wichert, Laura; Wilke, Ralf A. (2007): On the definition of unemployment and its implementation in register data * the case of Germany. FDZ Methodenreport, 03/2007

Anhang

Download der Stata-Programme:

- Deutsche Version
http://doku.iab.de/fdz/reporte/2013/MR_04-13_Programme.ZIP
- Englische Version:
http://doku.iab.de/fdz/reporte/2013/MR_04-13_EN_programs.ZIP

Impressum

FDZ-Methodenreport 04/2013

Herausgeber

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

Redaktion

Stefan Bender, Heiner Frank

Technische Herstellung

Heiner Frank

Rechte

Nachdruck - auch auszugsweise - nur mit
Genehmigung des FDZ gestattet

Bezugsmöglichkeit

http://doku.iab.de/fdz/reporte/2013/MR_04-13.pdf

Internet

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

Rückfragen zum Inhalt an:

Alexandra Schmucker
Institut für Arbeitsmarkt und
Berufsforschung (IAB)
Regensburger Str. 104
90478 Nürnberg
Phone: 0911 / 179-1762
<mailto:Alexandra.Schmucker@iab.de>