

FDZ-Datenreport

Dokumentation zu Arbeitsmarktdaten

06/2010

DE

Codebuch und Dokumentation des 'Panel Arbeitsmarkt und soziale Sicherung' (PASS)

Band 1: Datenreport Welle 3

Marco Berg,
Ralph Cramer,
Christian Dickmann,
Reiner Gilberg,
Birgit Jesske,
Karen Marwinski,
Daniel Gebhardt,
Claudia Wenzig,
Martin Wetzel



Codebuch und Dokumentation des 'Panel Arbeitsmarkt und soziale Sicherung' (PASS)

Datenreport Welle 3

Marco Berg, Ralph Cramer, Christian Dickmann, Reiner Gilberg, Birgit Jesske, Karen Marwinski, (alle infas Institut für angewandte Sozialwissenschaft GmbH), Daniel Gebhardt, Claudia Wenzig, Martin Wetzel (alle Institut für Arbeitsmarkt- und Berufsforschung (IAB))

Die FDZ-Datenreporte beschreiben die Daten des FDZ im Detail. Diese Reihe hat somit eine doppelte Funktion: Zum einen stellen Nutzerinnen und Nutzer fest, ob die angebotenen Daten für das Forschungsvorhaben geeignet sind, zum anderen dienen sie zur Vorbereitung der Auswertungen. Dieser Datenreport dokumentiert die Aufbereitung der 3. Welle von PASS. Das Dokument basiert auf dem Datenreport der 2. Welle: Gebhardt, Daniel; Müller, Gerrit; Bethmann, Arne; Trappmann, Mark; Christoph, Bernhard; Gayer, Christine; Müller, Bettina; Tisch, Anita; Siflinger, Bettina; Kiesel, Hans; Huyer-May, Bernadette; Achatz, Juliane; Wenzig, Claudia; Rudolph, Helmut; Graf, Tobias; Biedermann, Anika (2009): Codebuch und Dokumentation des 'Panel Arbeitsmarkt und soziale Sicherung' (PASS) Welle 2 (2007/2008), (FDZ Datenreport, 06/2009 (de), Nürnberg, 1097 Seiten. Passagen, an denen die Prozesse unverändert geblieben sind, wurden unverändert übernommen (das betrifft die Kapitel 1.1, 1.2). Andere Passagen wurden modifiziert (1.3, 2, 3, 4, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 6). Zudem gibt es auch völlig neu entwickelte Kapitel (5.7, 5.8).

FDZ-Datenreporte (FDZ data reports) describe FDZ data in detail. As a result, this series of reports has a dual function: on the one hand, those using the reports can ascertain whether the data offered is suitable for their research task; on the other, the data can be used to prepare evaluations. This data report documents the data preparation of the third PASS wave and is based upon the second wave's data report: Gebhardt, Daniel; Müller, Gerrit; Bethmann, Arne; Trappmann, Mark; Christoph, Bernhard; Gayer, Christine; Müller, Bettina; Tisch, Anita; Siflinger, Bettina; Kiesel, Hans; Huyer-May, Bernadette; Achatz, Juliane; Wenzig, Claudia; Rudolph, Helmut; Graf, Tobias; Biedermann, Anika (2009): Codebuch und Dokumentation des 'Panel Arbeitsmarkt und soziale Sicherung' (PASS) Welle 2 (2007/2008), (FDZ Datenreport, 06/2009 (de), Nuremberg, 1097 pages. Sections whose procedures remain the same were adopted without any alterations (this applies to the chapters 1.1, 1.2). Other sections were modified (1.3, 2, 3, 4, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 6). In addition, there are also completely new chapters (5.7, 5.8).

Inhaltsverzeichnis

1	Einleitung	10
1.1	Ziele und Fragestellungen des ‚Panel Arbeitsmarkt und soziale Sicherung‘	10
1.2	Instrumente und Befragungsprogramm	11
1.3	Besonderheiten und Neuerungen der 3. Welle	12
2	Kennwerte	17
2.1	Stichprobenumfang	17
2.2	Response-Raten	20
2.3	Panelbereitschaft, Zuspielbereitschaft, Verknüpfung mit Prozessdaten	22
2.4	Abgespaltene Haushalte	23
3	Datensatzstruktur	24
4	Generierte Variablen	26
4.1	Vercodung offener Angaben	26
4.2	Harmonisierung	29
4.3	Dependent Interviewing	32
4.4	Einfache Generierungen	35
4.5	Konstruktvariablen	61
5	Datenaufbereitung	86
5.1	Strukturprüfungen und entfernte Interviews	87
5.2	Filterbereinigung	91
5.3	Plausibilitätsprüfungen	92
5.4	Rückwirkende Veränderungen der 1. Welle und 2. Welle	94
5.5	Anonymisierung	106
5.6	Arbeitslosengeld-II-Bezüge	111
5.7	Erwerbsbiografien	115
5.8	Maßnahmeteilnahmen	121
6	Gewichtung Welle 3	123
6.1	Designgewichte für die Bestandshaushalte der 3. Welle	123
6.2	Designgewichte für die Auffrischungsstichprobe der 3. Welle	124
6.3	Wiederteilnahme-Propensity Haushalte	124
6.4	Teilnahme-Propensity erstbefragte Split-Haushalte	129
6.5	Ausfallgewichtung für Haushalte aus Auffrischungsstichprobe der 3. Welle	132
6.6	Wiederteilnahme-Propensity Personen	136
6.7	Integration der Gewichte zum Gesamtgewicht vor Kalibrierung	142
6.8	Integration temporärer Ausfälle (Haushalte)	142
6.9	Kalibrierung zum Haushaltsgewicht 3. Welle Querschnitt	144
6.10	Kalibrierung zum Personengewicht 3. Welle Querschnitt	162
6.11	Schätzung der BA-Querschnittgewichte für Haushalte und Personen ohne Arbeitslosengeld-II-Bezug	185
7	Anhang: Kurzbeschreibung des Datensatzes	186

Tabellen- und Abbildungsverzeichnis

Tabelle 1:	Panelstichprobenumfang auf Haushaltsebene nach Wellen und Teilstichproben	18
Tabelle 2:	Panelstichprobenumfang auf Personenebene nach Wellen und Teilstichproben	18
Tabelle 3:	Panelstichprobenumfang fremdsprachige Interviews nach Wellen.....	19
Abbildung 1:	Realisierte Panelstichprobe Haushalte und Personen nach Erhebungswellen	19
Tabelle 4:	Response-Rate der 3. Welle auf Haushaltsebene nach Teilstichproben	20
Tabelle 5:	Durchschnittliche Realisierungsrate innerhalb der befragten Haushalte nach Wellen und Teilstichproben	21
Tabelle 6:	Anteil Personeninterviews in Welle 2 und Welle 3 mit panelbereiten Befragten der Vorwelle nach Teilstichproben.....	21
Tabelle 7:	Panelbereitschaft bei erstbefragten Haushalten nach Wellen	22
Tabelle 8:	Zustimmung zur Zuspiegelung von Prozessdaten in Personeninterviews (15 bis unter 65 Jahre), in denen in der jeweiligen Welle die Zuspieffrage gestellt wurde, nach Wellen.....	23
Abbildung 2:	Datensatzstruktur des PASS in Welle 3	26
Tabelle 9:	Vercodung offener Angabe auf Haushaltsebene in Welle 3	27
Tabelle 10:	Vercodung offener Angaben auf Personenebene in Welle 3	28
Tabelle 11:	Harmonisierte Variablen im Personendatensatz (<i>PENDDAT</i>)	30
Tabelle 12:	Wellenübergreifend generierte, aber nicht vollständig harmonisierte Variablen im Personendatensatz (<i>PENDDAT</i>)	31
Tabelle 13:	Für alle Wellen generierte, aber nicht im Längsschnitt auswertbare Variablen im Personendatensatz (<i>PENDDAT</i>)	32
Tabelle 14:	Fortgeschriebene Vorwelleninformationen in Welle 3, wiederholt befr. Haushalte	33
Tabelle 15:	Fortgeschriebene Vorwelleninformationen in Welle 3, erstmals befr. Haushalte	34
Tabelle 16:	Fortgeschriebene Vorwelleinformationen in Welle 3, Personenfragebogen	34
Tabelle 17:	Typen der einfachen generierten Variablen in den Querschnittdatensätzen (<i>HHENDDAT</i> ; <i>PENDDAT</i>) bei Haushalten bzw. Personen, die schon in einer früheren Welle Angaben zum jeweiligen Thema gemacht haben	36
Tabelle 18:	Einfache Generierungen für Welle 3 im Haushaltsdatensatz (<i>HHENDDAT</i>) (alphabetische Reihenfolge)	37
Tabelle 19:	Einfache Generierungen für Welle 3 im Personendatensatz (<i>PENDDAT</i>) (alphabetische Reihenfolge)	38
Tabelle 20:	Einfache Generierungen für Welle 3 im Arbeitslosengeld-II-Spellldatensatz (<i>alg2_spells</i>) (Reihenfolge analog zur Reihenfolge im Datensatz)	48
Tabelle 21:	Einfache Generierungen für Welle 3 im Erwerbstätigkeitsspellldatensatz (<i>et_spells</i>) (Reihenfolge analog zur Reihenfolge im Datensatz).....	51

Tabelle 22:	Einfache Generierungen für Welle 3 im Arbeitslosigkeitsspelldatensatz (<i>al_spells</i>) (Reihenfolge analog zur Reihenfolge im Datensatz).....	53
Tabelle 23:	Einfache Generierungen für Welle 3 im Lückenspelldatensatz (<i>lu_spells</i>) (Reihenfolge analog zur Reihenfolge im Datensatz)	55
Tabelle 24:	Einfache Generierungen für Welle 3 im Maßnahmespelldatensatz (<i>mn_spells</i>) (Reihenfolge analog zur Reihenfolge im Datensatz)	57
Tabelle 25:	Einfache Generierungen für Welle 3 im Personenregisterdatensatz (<i>p_register</i>) (alphabetische Reihenfolge)	58
Abbildung 3:	Übersicht über generierte Variablen auf Personenebene in Welle 3	60
Abbildung 3:	Übersicht über generierte Variablen auf Personenebene in Welle 3 (Fortsetzung)	61
Tabelle 26:	Übersicht über die Schritte der Datenaufbereitung in der 3. Welle des PASS.....	87
Tabelle 27:	Übersicht über die vergebenen Missing Codes.....	92
Tabelle 28:	Übersicht über rückwirkende Veränderungen im Haushaltsdatensatz (<i>HHENDDAT</i>)	95
Tabelle 29:	Übersicht über rückwirkende Veränderungen im Personendatensatz (<i>PENDDAT</i>)	97
Tabelle 30:	Übersicht über rückwirkende Veränderungen in den Spelldaten auf Haushaltsebene (<i>alg2_spells</i>).....	104
Tabelle 31:	Übersicht über rückwirkende Veränderungen in den Spelldaten auf Personenebene (<i>et_spells</i> ; <i>al_spells</i> ; <i>lu_spells</i> ; <i>mn_spells</i>)	105
Tabelle 32:	Übersicht über rückwirkende Veränderungen in Registerdatensätzen (<i>hh_register</i> ; <i>p_register</i>).....	105
Tabelle 33:	Übersicht über rückwirkende Veränderungen in den Gewichtungsdatsätzen (<i>hweights</i> ; <i>pweights</i>)	106
Tabelle 34:	Übersicht über die in Welle 3 anonymisierten Variablen im Personendatensatz (<i>PENDDAT</i>).....	107
Tabelle 35:	Übersicht über die in Welle 3 anonymisierten Variablen im Erwerbstätigkeitsspelldatensatz (<i>et_spells</i>).....	111
Tabelle 36:	Querschnittvariablen im ALG2-Spelldatensatz (<i>alg2_spells</i>).....	113
Tabelle 37:	Querschnittvariablen im ET-Spelldatensatz (<i>et_spells</i>)	117
Tabelle 38:	Querschnittvariablen im AL-Spelldatensatz (<i>al_spells</i>).....	117
Tabelle 39:	Überblick über die Informationen zum Enddatum im integrierten MN-Spelldatensatz der Welle 2 und 3 (<i>mn_spells</i>)	122
Tabelle 40:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholerhaushalte.....	125
Tabelle 40:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholerhaushalte (Fortsetzung 1).....	126
Tabelle 40:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholerhaushalte (Fortsetzung 2).....	127
Tabelle 41:	Logit-Modelle zur Wiederteilnahme für Panelbereitschaft, Erreichbarkeit und Teilnahme	127
Tabelle 41:	Logit-Modelle zur Wiederteilnahme für Panelbereitschaft, Erreichbarkeit und Teilnahme (Fortsetzung 1)	128
Tabelle 41:	Logit-Modelle zur Wiederteilnahme für Panelbereitschaft, Erreichbarkeit und Teilnahme (Fortsetzung 2)	129

Tabelle 42:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der erstbefragten Split-Haushalte	130
Tabelle 43:	Logit-Modelle zur Erstteilnahme von Split-Haushalten für Erreichbarkeit und Teilnahme	131
Tabelle 44:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Auffrischungsstichprobe Welle 3.....	132
Tabelle 44:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Auffrischungsstichprobe Welle 3 (Fortsetzung 1).....	133
Tabelle 45:	Logit-Modelle zur Erstteilnahme für Erreichbarkeit und Teilnahme	134
Tabelle 45:	Logit-Modelle zur Erstteilnahme für Erreichbarkeit und Teilnahme (Fortsetzung)	135
Tabelle 46:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholer	136
Tabelle 46:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholer (Fortsetzung 1)	137
Tabelle 46:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholer (Fortsetzung 2)	138
Tabelle 46:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholer (Fortsetzung 3)	139
Tabelle 47:	Logit-Modelle zur Wiederteilnahme für Erreichbarkeit und Teilnahme.....	139
Tabelle 47:	Logit-Modelle zur Wiederteilnahme für Erreichbarkeit und Teilnahme (Forts. 1)	140
Tabelle 47:	Logit-Modelle zur Wiederteilnahme für Erreichbarkeit und Teilnahme (Forts. 2)	141
Tabelle 48:	Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der temporären Ausfälle	143
Tabelle 49:	Logit-Modelle zur Wiederteilnahme in Welle 3 gegeben Nichtteilnahme in Welle 2 für Erreichbarkeit und Teilnahme	144
Tabelle 50:	Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Haushalte)	146
Tabelle 50:	Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Haushalte) (Fortsetzung 1)	147
Tabelle 50:	Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Haushalte) (Fortsetzung 2)	148
Tabelle 51:	Kenngößen der Verteilung der Gewichte	148
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte)	150
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 1)	151
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 1)	151
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 2)	152
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 2)	152
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 3)	153

Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 3)	153
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 4)	154
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 4)	154
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 5)	155
Tabelle 52:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 5)	155
Tabelle 53:	Kenngößen der Verteilung der Gewichte	155
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte)	157
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 1)	158
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 1)	158
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 2)	159
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 2)	159
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 3)	160
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 3)	160
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 4)	161
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 4)	161
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 5)	162
Tabelle 54:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 5)	162
Tabelle 55:	Kenngößen der Verteilung der Gewichte	162
Tabelle 56:	Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Personen)	164
Tabelle 56:	Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Personen) (Fortsetzung)	165
Tabelle 57:	Kenngößen der Verteilung der Gewichte	166
Tabelle 58:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen)	167
Tabelle 59:	Kenngößen der Verteilung der Gewichte	175
Tabelle 60:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen)	177
Tabelle 60:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 1)	178
Tabelle 60:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 2)	179

Tabelle 60:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 3)	180
Tabelle 60:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 4)	181
Tabelle 60:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 5)	182
Tabelle 60:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 6)	183
Tabelle 60:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 7)	184
Tabelle 60:	Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 8)	185
Tabelle 61:	Kenngößen der Verteilung der Gewichte	185

Datenverfügbarkeit

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

1 Einleitung

1.1 Ziele und Fragestellungen des ‚Panel Arbeitsmarkt und soziale Sicherung‘

Mit dem Panel ‚Arbeitsmarkt und soziale Sicherung‘ (PASS) baut das Institut für Arbeitsmarkt- und Berufsforschung (IAB) einen neuen Datensatz für die Arbeitsmarkt-, Sozialstaats- und Armutsforschung in Deutschland auf, der eine neue empirische Grundlage für Wissenschaft und Politikberatung schafft.

Die Studie ist im Rahmen der SGB II-Forschung¹ am IAB angesiedelt, deren gesetzlicher Auftrag es ist, die Wirkungen der Leistungen zur Eingliederung und der Leistungen zur Sicherung des Lebensunterhalts des SGB II zu untersuchen. Sie erlaubt jedoch aufgrund des komplexen Stichprobendesigns auch die Beantwortung von Fragen, die über diesen Rahmen weit hinausgehen. Bei der Entwicklung der neuen Studie waren vor allem fünf Kernfragestellungen leitend, die in Achatz et al. (2007) ausführlich erläutert werden:

1. Welche Wege führen aus dem Bezug von Arbeitslosengeld II?
2. Wie verändert sich die soziale Lage der betroffenen Haushalte im Leistungsbezug?
3. Wie entwickeln sich Bewältigungsprozesse und Handlungsorientierungen betroffener Personen?
4. Wie gestalten sich Kontakte zu den Trägern der Grundsicherung, und wie sieht die institutionelle Handlungspraxis aus?
5. Welche Erwerbsverlaufsmuster oder Haushaltsdynamiken führen in den Bezug von Arbeitslosengeld II?

Dieser Datenreport gibt einen Überblick über die dritte Befragungswelle, für die zwischen Dezember 2008 und August 2009 13.439 Personen in 9.535 Haushalten² befragt wurden. Mit 11.300 Personen und 8.207 Haushalten wurde bereits zum wiederholten Mal ein Interview im Rahmen von PASS geführt.

Mit dem vorliegenden Datenreport der 3. Welle wurde erstmalig der Report in zwei Bestandteile aufgeteilt. Für die Dokumentation der 3. Welle sind also relevant:

1. der wellenspezifische Datenreport (inkl. Codebuch) und
2. der wellenübergreifende User Guide³.

Der wellenübergreifende User Guide wird in der Verantwortung des PASS-Projektteams beim IAB erstellt. Die Dokumentation des wellenspezifischen Datenreports der Welle 3 wurde von infas erstellt. Grundlage dafür bildete der Datenreport der Welle 2.

¹ Sozialgesetzbuch (SGB) Zweites Buch (II) – Grundsicherung für Arbeitsuchende.

² Die Zahlen umfassen nur auswertbare Interviews. Für wiederholungsbefragte Haushalte wurden auch solche berücksichtigt, für die nur ein Haushaltsinterview ohne Personen- bzw. Senioreninterview abgeschlossen werden konnte.

³ Bis zum Erscheinen des User Guides kann auf den Datenreport der Welle 2 zurückgegriffen werden, der auch wellenübergreifende Informationen enthält, bspw. zum Umgang mit den Daten.

Der vorliegende wellenspezifische Datenreport soll die wellenbezogenen Aspekte der Studie dokumentieren⁴. Nach einem Kurzüberblick zu den Neuerungen und Besonderheiten der 3. Welle (Kapitel 1.3) werden die zentralen Kennzahlen zu Stichprobe und Ausschöpfungsquoten der 3. Welle berichtet (Kapitel 2). Weiterhin werden die Schritte der Datenaufbereitung und die dabei getroffenen Entscheidungen erläutert (Kapitel 5) sowie ein Überblick über die generierten Variablen gegeben (Kapitel 4). Zusätzlich wird das Gewichtungsverfahren vorgestellt (Kapitel 6). In den gesonderten Tabellenbänden werden die Häufigkeiten aller im Scientific Use File enthaltenen Variablen, unterteilt nach den jeweiligen Datensätzen, aufgelistet, die in Welle 3 erhoben wurden (Band II bis Band V).

1.2 Instrumente und Befragungsprogramm

Im PASS werden sowohl auf der Ebene der Haushalte als auch auf der Ebene der Personen Informationen mittels jeweils eigener Fragebögen erhoben. Zunächst wird mit jedem Haushalt ein Haushaltsinterview geführt, in dem Informationen, die den gesamten Haushalt betreffen, erhoben werden. Die Zielperson für dieses Haushaltsinterview⁵ wird bereits in der Kontaktphase ausgewählt, welche den eigentlichen Interviews vorausgeht. An das Haushaltsinterview schließen sich Personeninterviews mit den einzelnen Haushaltsmitgliedern an. Ziel ist es, mit allen im Haushalt lebenden Personen ab 15 Jahre ein Personeninterview zu führen – Haushaltsmitglieder ab 65 Jahre erhalten dabei eine Kurzfassung des Personenfragebogens (Seniorenbogen), in dem für diese Altersgruppe irrelevante Fragen nicht enthalten sind.

Erhebungsinstrumente und Befragungsprogramm der 3. Welle basieren auf der 2. Welle vom PASS, wobei jedoch Einzelfragen und –module überarbeitet oder neu entwickelt wurden (siehe wellenübergreifender User Guide sowie für einen Überblick Kapitel 1.3).

⁴ Der wellenübergreifende User Guide soll im Gegensatz dazu der Dokumentation der Studie insgesamt dienen. Hier werden ausführlich die Ziele und das Design vom PASS beschrieben wie auch die Inhalte und Instrumente der Befragung vorgestellt. Des Weiteren werden die Struktur des Scientific Use Files und das Konzept der Variablentypen und -benennung erläutert. Abschließend wird anhand von Beispielen die Verwendung der verschiedenen Datensätze erläutert.

⁵ Die Zielperson für das Haushaltsinterview sollte sich möglichst gut mit allgemeinen Fragen zum Haushalt auskennen. In wiederbefragten Haushalten war das die Person, die schon in der Vorwelle den Haushaltsfragebogen beantwortet hatte. War diese Person während der gesamten Feldzeit nicht verfügbar oder kein Haushaltsmitglied mehr, dann wurde eine andere erwachsene Person ausgewählt, die sich gut mit dem Haushalt auskennt. In der Auffrischungsstichprobe, die aus den Daten der BA gezogen wurde, sollte der bei der BA registrierte ALG2-Antragsteller die Fragen zum Haushalt beantworten. Bei Splithaushalten sollte ein ehemaliges, mindestens 15-jähriges Mitglied des Ursprungshaushalts als Zielperson ausgewählt werden. Lebte die Person, mit der in der Vorwelle das Haushaltsinterview im Ursprungshaushalt geführt wurde, nun im abgespaltenen Haushaltsteil, dann sollte diese Person als Zielperson des Haushaltsinterviews im Splithaushalt ausgewählt werden. Immer dann, wenn eine bestimmte, bereits namentlich bekannte Zielperson während der Feldzeit nicht erreichbar war, wurde versucht, das Haushaltsinterview mit einer Person ab 15 Jahre zu führen, die sich möglichst gut mit allgemeinen Fragen zum Haushalt auskennt.

Auch in der 3. Welle erlauben die Instrumente neben der erstmaligen Befragung⁶ die Wiederbefragung von Haushalten und Personen, die bereits an einer der Vorwellen teilgenommen haben. Um bei der erneuten Befragung Seam-Effekte⁷ zu vermeiden und die Datenqualität zu erhöhen, wurden bereits seit der 2. Welle bei bestimmten Fragen per dependent interviewing Informationen, die der Befragte beim letzten Interview gegeben hatte, fortgeschrieben. Außerdem wurde weiterhin auf die erneute Erhebung unveränderlicher Merkmale weitgehend verzichtet. Auf Haushaltsebene existiert vergleichbar zur 2. Welle aufgrund der komplexen Fortschreibung der Haushaltsstruktur ein eigener Fragebogen für wiederbefragte Haushalte (HHalt) und für erstmals befragte Haushalte (HHneu).

Die einzelnen Instrumente und das Befragungsprogramm werden ausführlich im wellenübergreifenden User Guide dargestellt. Der nachfolgende Abschnitt liefert einen Überblick zu den Besonderheiten und Neuerungen der 3. Welle.

1.3 Besonderheiten und Neuerungen der 3. Welle

Nutzern, die bereits mit den Daten der ersten beiden Panelwellen gearbeitet haben, möchten wir an dieser Stelle einen schnellen Überblick über Besonderheiten der 3. PASS-Welle geben.

Die Besonderheiten und Neuerungen in Welle 3 betreffen das Frageprogramm⁸ (Fortschreibung der in Welle 2 erstmals erhobenen erwerbsbiografischen Informationen, Einsatz von Schwerpunktmodulen in den Bereichen „Netzwerke“, „Gesundheit“ und „Altersvorsorge“ sowie Aussetzen von bestehenden Modulen), die Stichprobe, die Aufbereitung der Daten und die Dokumentation.

Im Rahmen der 3. Erhebungswelle werden zum ersten Mal mit Hilfe des sog. dependent interviewing die in Welle 2 erhobenen erwerbsbiografischen Informationen fortgeschrieben. Neben den Erwerbstätigkeitsepisoden⁹ wird auch der weitere Verlauf von den in der Vorwelle andauernden Arbeitslosigkeiten¹⁰ und der Bezug von Arbeitslosengeld I¹¹ erhoben. Andere

⁶ Zu den erstmals befragten Haushalten gehören in der 3. Welle: (1) Haushalte der Auffrischungstichprobe der 3. Welle sowie (2) Haushalte, die sich von bereits in der 1. oder 2. Welle befragten Haushalten abgespalten haben (Splithaushalte). Weiterhin werden zwei Typen von Personen erstmals befragt: (1) Personen, die in der 3. Welle erstmals Mitglied eines PASS-Haushalts sind sowie (2) Personen, die bereits in der 1. oder 2. Welle Mitglied eines PASS-Haushalts waren, für die aber kein Interview aus einer der Vorwellen vorliegt.

⁷ Die Anzahl der Veränderungen an der Schnittstelle (Seam) von einem Interview eines Panels auf das nächste ist häufig deutlich höher als die innerhalb eines Interviews beobachteten Veränderungen (vgl. Jäckle 2008).

⁸ Kleinere Änderungen im Frageprogramm (Neuaufnahme, Umformulierung bzw. Löschen von Einzelfragen) werden in diesem Rahmen nicht aufgeführt.

⁹ Fragen E 38_X – E63_X im Personenfragebogen.

¹⁰ Fragen A106 – A111, A117 im Personenfragebogen.

¹¹ Fragen A112a – A116 im Personenfragebogen.

Aktivitäten (z.B. berufliche Ausbildung, Hausfrau/Hausmann, Rentner), die zum Zeitpunkt der letzten Befragung andauert haben, werden nicht explizit fortgeschrieben¹².

Wiederholt Befragte, die zum Zeitpunkt des letzten Interviews angegeben hatten, erwerbstätig (mit einem Einkommen von über 400 Euro) gewesen zu sein, werden nun gefragt, ob sie diese damals genannte Tätigkeit immer noch ausüben bzw. bis wann sie diese Tätigkeit ausgeübt haben. Für diese fortgeschriebene Tätigkeit werden die folgenden Informationen erneut erhoben: (1) die berufliche Stellung¹³ und (2) die Arbeitszeit¹⁴ sowie, (3) ob vormals befristete Tätigkeiten in unbefristete umgewandelt wurden und zusätzlich, (4) wie das Arbeitsverhältnis beendet wurde (nur bei Tätigkeiten, die vor dem Interviewdatum der 3. Welle beendet wurden).

Nach einer vergleichbaren Logik werden auch die zum letzten Interviewzeitpunkt andauernde Arbeitslosigkeitsepisode und der Bezug von Arbeitslosengeld I fortgeschrieben. So wird mit Hilfe von dependent interviewing für wiederholt Befragte in Welle 3 erhoben, bis wann die damals angegebene Arbeitslosigkeitsmeldung andauerte sowie ggf. die Beendigungsgründe. Für den Zeitraum seit der letzten Befragung wird in diesen Fällen weiterhin erhoben, ob und wenn ja, wie lange Arbeitslosengeld I bezogen wurde. Bei Befragten, die zum Interviewzeitpunkt der dritten Welle Arbeitslosengeld I beziehen, wird zusätzlich die Höhe des Betrags erfragt.

Nach der Fortschreibung der in Welle 2 genannten Erwerbstätigkeiten bzw. Arbeitslosigkeitsepisoden werden dann ggf. weitere Erwerbstätigkeiten (über 400 Euro), Arbeitslosmeldungen und alle weiteren Aktivitäten innerhalb von sog. Lücken in der Erwerbsbiografie (eine Lücke ist definiert als eine mehr als dreimonatige Zeitdauer, in der weder eine Erwerbstätigkeit noch eine Arbeitslosigkeit berichtet wird) seit dem letzten Interviewdatum erhoben sowie der aktuelle (Nicht-)Erwerbsstatus zum Interviewzeitpunkt der 3. Welle¹⁵. Die Erhebung der weiteren Erwerbstätigkeiten, Arbeitslosmeldungen, Bezugsepisoden von Arbeitslosengeld I und sonstigen Aktivitäten erfolgt mit dem bereits bekannten Frageprogramm aus Welle 2.

Die erwerbsbiografischen Informationen von Neubefragten Personen werden ebenfalls analog zu Welle 2 erhoben, lediglich der Anknüpfungzeitpunkt unterscheidet sich. Für diese Personen wird die Erwerbsbiografie seit Januar 2006 erhoben (in Welle 2 war der Anknüpfungspunkt für Neubefragte Januar 2005).

¹² Das Lückenmodul, in dem diese Aktivitäten erfasst werden, dient in Welle 2 und 3 ausschließlich dazu, Aktivitäten in Zeiträumen von mehr als drei Monaten zu erheben, für die keine Erwerbstätigkeit oder Arbeitslosigkeit vorliegt.

¹³ Bei abgeschlossenen fortgeschriebenen Tätigkeiten wird die berufliche Stellung am Ende der Tätigkeit erfragt und bei andauernden fortgeschriebenen Tätigkeiten wird die derzeitige berufliche Stellung erhoben.

¹⁴ Bei abgeschlossenen fortgeschriebenen Tätigkeiten wird die Arbeitszeit am Ende der Tätigkeit und bei andauernden fortgeschriebenen Tätigkeiten die derzeitige Arbeitszeit erhoben.

¹⁵ Fragen P126 – P132 im Personenfragebogen.

Der Scientific Use File enthält innerhalb der relevanten Spelldatensätze¹⁶ sowohl die in Welle 2 als auch die in Welle 3 erfassten Erwerbstätigkeits-, Arbeitslosigkeits- und Nichterwerbstätigkeitsepisoden. Die Integration aller Episoden in die jeweiligen Spelldatensätze erfolgt nach bestimmten Regeln (siehe dazu Kapitel 5.6, 5.7, 5.8). Werden Episoden über mehrere Wellen hinweg fortgeschrieben, kann ein Spell auch mehrere wellenspezifische Informationen enthalten (z.B. Arbeitsstunden zum Interviewzeitpunkt der Welle 2 und 3). Sie werden in wellenspezifischen Variablen abgelegt. Wellenspezifische Variablen bezogen auf die Welle 2 enden auf der Ziffer „0“, bezogen auf die Welle 3 enden diese auf der Ziffer „1“ usw. (siehe dazu wellenübergreifender User Guide).

Die Teilnahme an arbeitsmarktpolitischen Maßnahmen wird in der 3. Welle mit dem bekannten Konzept der 2. Welle erhoben. Eine Fortschreibung von genannten Maßnahmen findet allerdings nicht statt, da für jede Maßnahme bereits das tatsächliche bzw. das geplante Ende oder die Dauer aus Welle 2 vorliegt. Lediglich der Anknüpfungzeitpunkt unterscheidet sich zwischen wiederholt und Neubefragten Personen. Bei Neubefragten Personen sind alle Maßnahmen seit Januar 2007 für die Befragung relevant (in Welle 2 war es Januar 2006), wiederholt Befragte sollen alle Maßnahmen und finanzierten Programme nennen, an denen sie seit dem letzten Interviewdatum teilgenommen haben. Der Maßnahmenspelldatensatz (*mn_spells*) des Scientific Use Files enthält alle genannten Maßnahmen und finanzierten Programme, die im Rahmen der 2. und der 3. Welle erhoben wurden.

Des Weiteren werden in der 3. Welle zu drei Bereichen vertiefende Fragen gestellt. Schwerpunktthemen sind die Bereiche „Netzwerke“, „Gesundheit“ und „Altersvorsorge“.

Zusätzlich zum Standardfrageprogramm des Moduls „Netzwerke“ werden in der 3. Welle Detailangaben zu den drei wichtigsten Freunden erhoben (Geschlecht, Schulabschluss, Erwerbsstatus, Art des Verhältnisses, Klima der Freundschaft). Außerdem wird abgefragt, inwieweit die Befragten über soziale Ressourcen verfügen. Hierzu werden zum einen zehn mögliche Situationen genannt, in denen Menschen häufig andere Personen um Unterstützung bitten, und zum anderen der private Kontakt zu bestimmten Personengruppen (z.B. Unternehmer, Straftäter) abgefragt.¹⁷

Auch zur Erfassung der Gesundheit werden in der 3. Welle zum bereits vorhandenen Frageprogramm Zusatzfragen an alle Befragten ab 15 Jahre gestellt.¹⁸ Erstmals erhoben werden somit im PASS die gesundheitsbezogene Lebensqualität mit Hilfe des sog. SF-12v2¹⁹, die Einschränkungen der Erwerbstätigkeit, das Gesundheitsverhalten (Sport, Alkohol- und Tabakkonsum) sowie das Übergewicht.

Es ist geplant, beide Schwerpunktthemen auch in einer der zukünftigen Wellen einzusetzen, um langfristige Veränderungen beobachten zu können.

¹⁶ Erwerbstätigkeitsspell: *et_spells*; Arbeitslosigkeitsspell: *al_spells*; Spells zu Zeiten der Nichterwerbstätigkeit: *lu_spells*.

¹⁷ Fragen N1 – N17 im Personen- und Seniorenfragebogen.

¹⁸ Fragen G1 – G 18 im Personen- und Seniorenfragebogen..

¹⁹ Im PASS wurde nicht die Originalversion (Ware et al., 2002), sondern mit Erlaubnis des DIW die für das SOEP entwickelte Version eingesetzt (siehe dazu Andersen et al., 2007a).

Das Schwerpunktmodul „Altersvorsorge“ wird sowohl im Haushaltsbogen als auch im Personenbogen (für alle Personen zwischen 40 und 64 Jahre sowie ihre Partner im Haushalt, und zwar unabhängig vom Alter) erhoben.²⁰ Im Rahmen des Haushaltsinterviews wird der Haushaltsvorstand gebeten, Detailfragen zum selbst genutzten Wohneigentum (Art, Größe, Wert) zu beantworten. Im Personeninterview sollen die Befragten Angaben zu ihrer gesetzlichen und privaten Altersvorsorge (auch Zusatzversorgungen des öffentlichen Dienstes, betriebliche Altersvorsorge, sog. Riesterrente, Lebensversicherung) machen. So soll jeweils die Dauer der bisherigen Einzahlung, die Art der Auszahlung, die geschätzte Höhe der späteren Rente(n) und ggf. das Alter, ab dem man über die Rente verfügen kann, angegeben werden. Des Weiteren werden sie zur (vorzeitigen) Kündigung bzw. Beitragsfreiheit der privaten Altersvorsorge gefragt (Datum der Kündigung, Kündigungsgründe, Höhe der Auszahlung). Abschließend wird die Zufriedenheit mit dem Lebensstandard (aktuell und zum Zeitpunkt des Rentenalters bzw. vor dem Rentenalter) erhoben. Im Rahmen des Schwerpunktmoduls werden die Befragten auch gebeten, ihre Erlaubnis für die Zuspierung von Daten der Deutschen Rentenversicherung zu geben. Die Fragen zur Altersvorsorge werden in jeweils einem separaten Datensatz auf Haushalts- und Personenebene (*HAVDAT*, *PAVDAT*) abgelegt und nicht in den regulären Haushalts- oder Personendatensatz integriert.

Daneben werden in der 3. Welle drei Fragemodule auf Personenebene nicht erhoben. Dies betrifft die Einstellungen zu Familie und Beruf, Einstellungen zu Partnerschaft und Rollenbeziehungen sowie den Themenbereich Religion. Dabei handelt es sich nur um ein kurzfristiges Aussetzen der Themenbereiche, die Module werden in späteren Wellen gemäß einem Rotationsplan wieder aufgenommen.

Des Weiteren wurde in der 3. Welle ebenfalls eine sog. Auffrischungstichprobe für die BA-Teilstichprobe²¹ gezogen. Ziel dabei ist es, die Repräsentativität der BA-Stichprobe im Querschnitt sicherzustellen und im Zeitverlauf ausreichend neue Übergänge in den Arbeitslosengeld II-Bezug beobachten zu können. Für die Auffrischung werden Bedarfsgemeinschaften gezogen, die zum 1. Juli 2008 Arbeitslosengeld II bezogen haben, nicht jedoch zum Zeitpunkt der Stichprobenziehung der 1. Welle bzw. 2. Welle (vgl. Kapitel 2.1) und zum Auffrischungskonzept Trappmann et al 2009: 11 ff.). Diese in der 3. Welle neubefragten Haushalte können wiederum über den Stichprobenindikator (*sample*) identifiziert werden.

Die Aufbereitung der Daten der 3. Welle wird erstmalig vom Erhebungsinstitut infas Institut für angewandte Sozialwissenschaft in Bonn durchgeführt. Die ersten beiden Wellen wurden vom IAB selbst bereinigt und aufbereitet. Um eine zu den Vorwellen möglichst analoge Aufbereitung zu gewährleisten, wurden die durchzuführenden Arbeitsschritte vorab vom IAB definiert und die dazu benötigten Materialien und Aufbereitungs-Do-Files an infas übergeben. Die eigentliche Datenaufbereitung erfolgte in enger Abstimmung mit dem IAB. Grundsätzliche Vorgehensweisen z.B. zur Fortschreibung der Datensätze und zur Bereinigung von

²⁰ Fragen im Haushaltsbogen HV 1 – HV 9 und Fragen im Personenbogen V1 – V99.

²¹ Die 1. Welle des PASS besteht aus zwei Teilstichproben: (1) eine aus den Prozessdaten der BA gezogene Stichprobe von Arbeitslosengeld II-Empfängerhaushalten und (2) eine aus einer Datenbank des kommerziellen Anbieters MICROM gezogene, statusgeschichtete allg. Bevölkerungstichprobe.

Problemen in den Haushaltsstrukturen wurden im Aufbereitungsprozess diskutiert und durch das IAB entschieden.

Eine weitere Neuerung in der 3. Welle betrifft die Dokumentation und die Arbeitshilfen, die Nutzern zum Einstieg in die Arbeit mit PASS zur Verfügung stehen. Während die Fragebögen sowie der Methodenbericht, in dem die Feldarbeit des Erhebungsinstituts beschrieben wird, auch weiterhin unverändert zur Verfügung stehen, wurde das Konzept des Datenreports überarbeitet.

In den Wellen 1 und 2 enthielt der Datenreport sowohl *wellenspezifische* als auch *wellenübergreifende* Informationen mit eher allgemeinem Charakter. Anstatt diese Informationen auch weiterhin in einem Dokument zusammenzufassen, werden sie aufgeteilt. Auch in Zukunft wird mit jeder Welle ein Datenreport erscheinen. Dieser wird allerdings einen deutlich klareren Schwerpunkt haben und sich auf *wellenspezifische* Informationen konzentrieren. Der Datenreport enthält also Informationen zu den Inhalten einer Welle, deren Datenaufbereitung und Auszählungen der in der Welle erhobenen Variablen in den verschiedenen Datensätzen. Darüber hinaus informiert der Datenreport auch über relevante Kennzahlen und das Vorgehen bei der Erstellung der Gewichte der jeweiligen Welle.

Neben den Datenreport tritt ab Welle 3 ein weiteres Dokument: der User Guide. Während sich der Datenreport auf wellenspezifische Informationen konzentriert und als Dokument für jede veröffentlichte Welle vorliegt, enthält der User Guide *wellenübergreifende, allgemeine* Informationen zu PASS und dem Scientific Use File. Dieser neue Teil der Dokumentation beantwortet allgemeine Fragen zum PASS, bspw. zum Studien- und Stichprobendesign, zum Frageprogramm über die Wellen hinweg, zur Datenstruktur und zum Gewichtungskonzept. Darüber hinaus enthält der User Guide Beispiele zur Nutzung der Datensätze, darunter zum Schlüsseln zwischen den Datensätzen und zur Verwendung der Gewichte. Auch der User Guide wird in der Zukunft angepasst und überarbeitet werden. Hier wird allerdings immer das neue Dokument die vorangegangene Version ersetzen. Der Datenreport einer Welle wird dagegen nicht durch den Datenreport der folgenden Welle ersetzt, da sich die Dokumente jeweils auf andere Wellen beziehen.

In der 3. Welle ergibt sich durch die Konzeptänderung eine Übergangsphase, in der zwar der Datenreport nach dem neuen Konzept vorliegt, der User Guide aber noch nicht fertig gestellt ist. Hier kann vorerst auf den Datenreport der Welle 2 zurückgegriffen werden, der nach dem alten Konzept auch noch allgemeine Informationen zum Erhebungs- und Stichprobendesign, zu den Befragungsinstrumenten, zum Variablenkonzept sowie praktische Beispiele für den Umgang mit dem PASS enthält.

2 Kennwerte

Im folgenden Kapitel wird ein kurzer Überblick über zentrale Kennziffern der Studie, wie Stichprobengröße brutto und netto oder Ausschöpfungsquoten, gegeben. Diese werden für die Panelstichprobe über den Verlauf der bisherigen drei Wellen dargestellt und sowohl für die beiden ursprünglichen Teilstichproben und die Zugangsstichproben getrennt als auch für die Gesamtstudie ausgewiesen.

- Teilstichprobe 1 (BA-Stichprobe) bezeichnet dabei im Folgenden die Leistungsempfängerstichprobe aus den Prozessdaten der Bundesagentur für Arbeit.
- Teilstichprobe 2 (MICROM-Stichprobe) verweist auf die geschichtete Bevölkerungsstichprobe.
- Zugangsstichprobe 1 (BA-Stichprobe) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 1 und Welle 2.
- Zugangsstichprobe 2 (BA-Stichprobe) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 2 und Welle 3.

2.1 Stichprobenumfang

Ausgangsbasis für den Stichprobenumfang in einem Panel bilden die befragten Haushalte der Erstbefragung. Im PASS beinhaltet die Bruttopanelstichprobe die befragten Haushalte aus der 1. Welle, aber auch die erstbefragten Haushalte aus den Zugangsstichproben in den Wellen 2 und 3. Dabei muss berücksichtigt werden, dass nur solche erstbefragten Haushalte auch bei den Wiederholungsbefragungen zur Verfügung stehen, für welche die Panelbereitschaft²² vorliegt. Die Panelbereitschaft wird nur bei der Erstbefragung erhoben. In den Folgewellen ist eine erneute Bereitschaftserklärung für diese Haushalte nicht notwendig. Der Zugang zum Panel wird neben der Panelbereitschaftserklärung bereits in der ersten Befragung durch die grundsätzliche Teilnahmebereitschaft, d.h. durch die Realisierung eines Interviews bedingt. Maßnahmen, die im Rahmen vom PASS einen möglichst selektionsfreien Zugang zum Panel sicherstellen sollten, werden ausführlich im Methoden- und Feldbericht der Wellen 1 bis 3 beschrieben²³.

PASS ist in der 1. Welle mit 12.794 realisierten Haushaltsinterviews gestartet, von denen 12.000 Haushalte ihre Panelbereitschaft erklärt haben. Die panelbereiten Haushalte der 1. Welle bilden den Stichprobenumfang für den Start der ersten Wiederholungsbefragung.

Das Panelkonzept im PASS sieht vor, dass durch Auszug von Personen aus Panelhaushalten neue Haushalte bzw. abgespaltene Haushalte (Splithaushalte) entstehen, die als eigene Haushalte gezählt werden, sobald dort ein Haushaltsinterview realisiert werden konnte. Damit nimmt die Anzahl der Haushalte gegenüber der Ausgangsstichprobe zu. Das Vorgehen im Panelkonzept von PASS wird im Abschnitt „Abspaltene Haushalte“ näher beschrieben.

²² Die Panelbereitschaft wird nur im ersten Interview mit dem Haushaltsbevollmächtigten erhoben und zählt damit für alle Haushaltsmitglieder. Panelbereite Haushalte haben der Speicherung ihrer Adresse zum Zwecke von Wiederholungsbefragungen im Rahmen der Studie zugestimmt.

²³ Vgl. Hartmann et al. (2008); Büngeler et al. (2009); Büngeler et al. (im Erscheinen).

Neben der Erweiterung des Panels kann es aufgrund von Panelmortalität auch zum Verlust von Haushalten kommen. Haushalte, in denen alle Befragungspersonen verstorben oder ins Ausland verzogen sind, werden in den Folgewellen nicht mehr zum Panelbrutto gezählt. Panelverluste entstehen ebenfalls dadurch, dass für einen Haushalt über zwei Wellen in Folge kein Haushaltsinterview realisiert werden konnte. Dieser Fall kann erstmalig am Ende der 3. Welle eintreten und wirkt sich dann auf das Panelbrutto in der 4. Welle aus.

Die Fallzahlen für den Bruttostichprobenumfang der jeweiligen Erhebungswellen und Teilstichproben wird in der nachfolgenden Tabelle ausgewiesen. In der 3. Welle konnte in 8.349 Haushalten der 11.982 Panelhaushalte mindestens ein Interview geführt werden. Dazu kommen 1.186 befragte Haushalte aus der BA-Auffrischungstichprobe, von denen 1.145 Haushalte ihre Panelbereitschaft erklärt haben.

Tabelle 1: Panelstichprobenumfang auf Haushaltsebene nach Wellen und Teilstichproben

n	Stichprobe				Gesamt
	BA	Microm	BA-Zugang 1	BA-Zugang 2	
Welle 1 HH-Interview realisiert	6.804	5.990			12.794
	davon: HH panelbereit	6.452	5.548		
Welle 2 Panel-HH brutto	6.520	5.611			12.131
	HH-Interview realisiert	3.491	3.897	1.041	
davon: HH panelbereit	3.360	3.766	1.003		8.129
Welle 3 Panel-HH brutto	5.833	5.141	1.008		11.982
	HH-Interview realisiert	3.754	3.901	694	1.186
davon: HH panelbereit	3.576	3.777	669	1.145	9.167

Quelle: HH-Register und PENDDAT; Scientific Use File IAB

Den 9.535 realisierten Haushaltsinterviews in der 3. Welle entsprechen 13.439 realisierte Personeninterviews. Die Verteilung der befragten Personen über die Teilstichproben und die jeweiligen Erhebungswellen ist der nachfolgenden Tabelle zu entnehmen.

Tabelle 2: Panelstichprobenumfang auf Personenebene nach Wellen und Teilstichproben

Personeninterview realisiert		Stichprobe				Gesamt
		BA	Microm	BA-Zugang 1	BA-Zugang 2	
Welle 1	abs.	9.386	9.568			18.954
Welle 2	abs.	4.753	6.392	1.342		12.487
Welle 3	abs.	4.913	6.207	898	1.421	13.439

Quelle: P_Register; Scientific Use File IAB

Für Befragte ohne ausreichende Deutschkenntnisse gab es die Möglichkeit, sich in einer anderen Sprache als Deutsch befragen zu lassen. Als alternative Interviewsprachen wurden dabei Türkisch und Russisch angeboten. Tabelle 3 gibt an, wie viele Haushalte bzw. Personen in den beiden Erhebungssprachen befragt wurden.

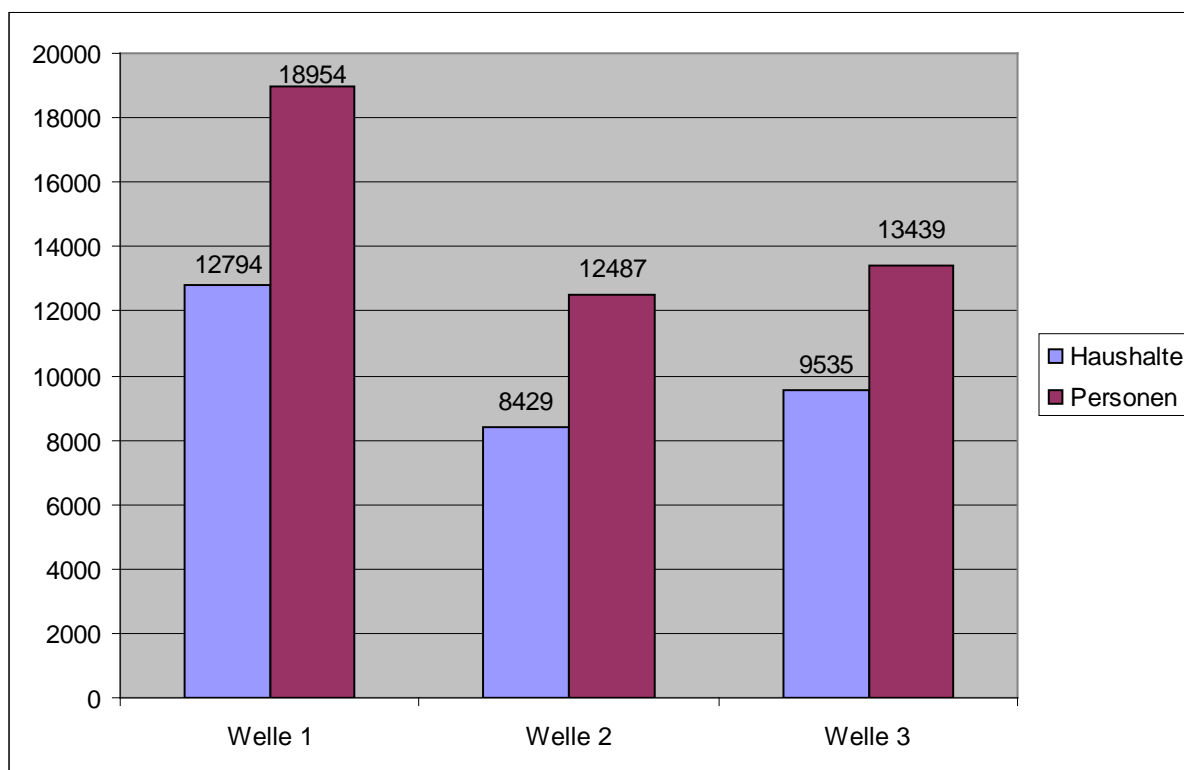
Tabelle 3: Panelstichprobenumfang fremdsprachige Interviews nach Wellen

		Russisch abs.	Türkisch abs.
Welle 1	Haushalte	275	163
	Personen	432	305
Welle 2	Haushalte	156	39
	Personen	219	31
Welle 3	Haushalte	210	69
	Personen	330	109

Quelle: PENDDAT; Scientific Use File IAB

Damit ergibt sich für den Gesamtbestand der realisierten Panelstichprobe folgendes Bild im Überblick für Haushalte und Personen über die drei Erhebungswellen.

Abbildung 1: Realisierte Panelstichprobe Haushalte und Personen nach Erhebungswellen



2.2 Response-Raten

Die Berechnung der Response-Rate erfolgt nach AAPOR-Standards (AAPOR 2006). Berichtet wird die Ausschöpfungsquote RR1, in der auch alle Fälle ungeklärter Eligibilität in den Nenner eingehen und die somit den geringsten Wert aller Ausschöpfungsquoten annimmt²⁴. Als Ausschöpfungsquote auf der Haushaltsebene wird der Anteil der auswertbaren Haushaltsinterviews an der Summe aus auswertbaren Haushaltsinterviews und nicht neutralen Ausfällen berechnet. Als neutrale Ausfälle gelten nur Haushalte, in denen alle Personen verstorben und Haushalte, die dauerhaft ins Ausland verzogen sind. Haushalte gelten als auswertbar, wenn zumindest ein vollständiges Haushaltsinterview vorliegt. Neuhaushalte gelten nur dann als auswertbar, wenn neben dem Haushaltsinterview auch mindestens ein vollständiges Personeninterview vorliegt.

Auf der Haushaltsebene ergaben sich für die 3. Welle folgende Ausschöpfungsquoten:

Tabelle 4: Response-Rate der 3. Welle auf Haushaltsebene nach Teilstichproben

Welle 3		Stichprobe				Gesamt
		BA	Microm	BA-Zugang 1	BA-Zugang 2	
HH brutto	abs.	5.833	5.141	1.008	3.801	15.783
	%	100,0	100,0	100,0	100,0	100,0
neutrale Ausfälle	abs.	16	37	2	12	67
	%	0,3	0,7	0,2	0,3	0,4
HH brutto bereinigt*	%	5.817	5.104	1.006	3.789	15.716
HH-Interview realisiert	abs.	3.754	3.901	694	1.186	9.535
	%	64,5	76,4	69,0	31,3	60,7
davon: HH panelbereit	abs.				1.145	
	%				30,2	

* HH brutto - neutrale Ausfälle

Quelle: HH-Register; Scientific Use File IAB - für BA-Zugang 2: Bruttodatensatz Welle 3 IAB

In einer Haushaltserhebung kann zwischen der Response-Rate auf Haushaltsebene und der Response-Rate innerhalb von Haushalten unterschieden werden.

Mit Response-Rate innerhalb von Haushalten bezeichnen wir den durchschnittlichen Anteil aller Personen ab 15 Jahre innerhalb der Haushalte mit einem auswertbaren Haushaltsinterview, für die ein vollständiges Personeninterview vorliegt.

²⁴ Dies wird in Deutschland sehr uneinheitlich gehandhabt. Häufig werden sehr viele nicht befragte Personen oder Haushalte als „ineligible“ verbucht und bei der Berechnung der Ausschöpfung aus dem Nenner gestrichen. Im Falle einer Ziehung aus Registern darf aber weder ein Haushalt, der nicht unter der erwarteten Adresse wohnt, noch ein Haushalt, der behauptet, nicht zur Zielgruppe zu gehören, als neutraler Ausfall verbucht werden. Zudem beschränkt sich die Population von PASS nicht auf deutschsprachige Befragte und auf befragungsfähige Personen, so dass auch die Ausfallgründe „spricht kein Deutsch“ oder „Person ist krank/nicht befragungsfähig“ nicht als neutrale Ausfälle gelten können.

Im Durchschnitt ergeben sich innerhalb der befragten Haushalte folgende Realisierungsraten:

Tabelle 5: Durchschnittliche Realisierungsrate innerhalb der befragten Haushalte nach Wellen und Teilstichproben

		Stichprobe				Gesamt
		BA	Microm	BA-Zugang 1	BA-Zugang 2	
Welle 1	%	85,6	84,2			84,9
Welle 2	%	85,5	85,1	86,2		85,4
Welle 3	%	83,1	83,6	84,3	84,2	83,5

Quelle: P_Register; Scientific Use File IAB

Neben den Response-Raten auf der Haushaltsebene und innerhalb der Haushalte sind im Folgenden die Wiederbefragungsquoten auf der Personenebene angegeben. Dabei handelt es sich um den Anteil von panelbereiten Personen, mit denen in der Folgewelle ein Interview realisiert werden konnte.

Tabelle 6: Anteil Personeninterviews in Welle 2 und Welle 3 mit panelbereiten Befragten der Vorwelle nach Teilstichproben

		Stichprobe			Gesamt
		BA	Microm	BA-Zugang 2	
Welle 2	panelbereite Befragungspersonen W1	abs.	8.925	8.938	17.863
	in W2 wiederholt befragte Personen	abs.	4.274	5.829	10.103
	<i>Anteil</i>	%	47,9	65,2	56,6
Welle 3	panelbereite Befragungspersonen W2	abs.	4.686	6.292	12.276
	in W3 wiederholt befragte Personen	abs.	3.365	4.956	9.141
	<i>Anteil</i>	%	71,8	78,8	74,5

Quelle: PENDDAT; Scientific Use File IAB

2.3 Panelbereitschaft, Zuspielbereitschaft, Verknüpfung mit Prozessdaten

Zur Speicherung der Adressen zum Zwecke der Wiederbefragung in der Folgewelle und zum Zusammenspielen der Erhebungsdaten mit Prozessdaten der Bundesagentur für Arbeit ist jeweils die Einverständniserklärung der Befragten notwendig.

Die Panelbereitschaft wurde schon im Kapitel 2.1 beim Stichprobenumfang näher erläutert. Die Panelbereitschaft für in einer Welle erstmals befragte Haushalte²⁵ im PASS sieht wie folgt aus:

Tabelle 7: Panelbereitschaft bei erstbefragten Haushalten nach Wellen

	Realisierte HH-Interviews mit erstmalig befragten HH	Realisierte HH-Interviews mit erstmalig befragten HH mit Panelbereitschaft	Anteil mit Panelbereitschaft
	abs.	abs.	%
Welle 1	12.794	12.000	93,8
Welle 2	1.086	1.048	96,5
Welle 3	1.327	1.285	96,8

*erstmalig befragte HH aus Auffrischung und Split

Quelle: PENDDAT; Scientific Use File IAB

Die Panelbereitschaft wurde bei erstmalig befragten Haushalten in Welle 3 im Anschluss an das erste Personeninterview erfasst. Die Angabe dieser Person, wurde dann auf den Haushalt übertragen. War die Person panelbereit, wurde der Haushalt als panelbereit gewertet. War die Person nicht panelbereit, wurde der Haushalt als nicht panelbereit gewertet²⁶.

Im Gegensatz zur Panelbereitschaft wurde die Erlaubnis zur Zuspielung von Prozessdaten der Bundesagentur für Arbeit an die Befragungsdaten für jeden Befragten erhoben, der mit dem Personenfragebogen interviewt wurde. Für Personen ab 65 Jahre entfällt die Nachfra-

²⁵ Bei allen Haushalten der Welle 1 handelt es sich um erstmalig befragte Haushalte. Ab Welle 2 handelt es sich nur bei Haushalten der jeweiligen Auffrischungstichproben um erstmalig befragte Haushalte sowie bei den erstmalig teilnehmenden Split-Haushalten. Ab Welle 2 sind erstmalig befragte Haushalte daher in der Minderheit – bei einem Großteil der in diesen Wellen realisierten Haushaltsinterviews handelt es sich um Interviews mit Haushalten, die bereits zu einem früheren Zeitpunkt befragt wurden

²⁶ Die Angabe zur Panelbereitschaft wird also von einer Person für den gesamten Haushalt gemacht. Die auf Haushaltsebene vorliegende Information wurde im Zuge der Datenaufbereitung in den Personendatensatz (*PENDDAT*) integriert. Die einzelnen, in einem Haushalt befragten Personen übernahmen den entsprechenden Wert, der für den Haushalt vorlag. Analog wurde auch in Welle 2 vorgegangen. In Welle 1 wurde die Panelbereitschaft dagegen nach jedem Personen- und Senioreninterview personenspezifisch erhoben – hier können also unterschiedliche Werte innerhalb eines Haushalts vorkommen. Ein Haushalt, in dem mindestens eine panelbereite Person lebte, wurde dabei als panelbereit gewertet.

Bei wiederbefragten Haushalten wurde im Rahmen der Adressaktualisierung nach dem ersten Personeninterview angekündigt, dass sich im kommenden Jahr wieder ein Interviewer melden wird. Widerspruch der Befragte dieser Ankündigung nicht explizit, wurde der Haushalt als weiterhin panelbereit gewertet und die Variable *panel* im Personendatensatz (*PENDDAT*) entsprechend fortgeschrieben.

ge, da sie im Seniorenfragebogen nicht enthalten ist. Die Zuspielbereitschaft wird dabei nicht in jeder Welle neu erhoben²⁷.

Tabelle 8 gibt einen Überblick über die Zuspielbereitschaft in den einzelnen Wellen. Aufgeführt sind nur diejenigen Interviews, in denen die Frage zur Zuspielbereitschaft in der betreffenden Welle im Rahmen des Personenfragebogens gestellt wurde.

Tabelle 8: Zustimmung zur Zuspielung von Prozessdaten in Personeninterviews (15 bis unter 65 Jahre), in denen in der jeweiligen Welle die Zuspielfrage gestellt wurde, nach Wellen

	Realisierte Personen-Interviews der Welle, in der die Zuspielfrage gestellt wurde	Realisierte Personen-Interviews der Welle, in der die Zuspielbereitschaft erteilt wurde	Anteil mit Zuspielbereitschaft
	abs.	abs.	%
Welle 1	17.249	13.766	79,8
Welle 2	3.358	2.560	76,2
Welle 3	2.656	2.128	80,1

Basis: Personen 15 bis 64 Jahre

Quelle: PENDDAT; Scientific Use File IAB

Von den 2.128 in Tabelle 8 dargestellten Personen, die in Welle 3 ihre Zuspielerlaubnis gegeben hatten, konnte für 1.769 (83.1 Prozent) eine Prozessdatenverknüpfung hergestellt werden.

Die Zuspielerlaubnis gilt auch für die Zukunft, sofern sie vom Befragten nicht widerrufen wird²⁸. Berücksichtigt man die Zustimmungen, die bereits in Welle 1 und Welle 2 gegeben und seitdem nicht widerrufen wurden, liegen für 11.098 der insgesamt 12.104 in Welle 3 realisierten Personeninterviews (15 bis unter 65 Jahre) eine Zuspielerlaubnis vor, was einem Anteil von 91.7 Prozent entspricht. 10.436 (94,0 Prozent) der Welle 3-Personeninterviews mit Zuspielerlaubnis seit Welle 1 konnten tatsächlich mit den Prozessdaten verknüpft werden. Insgesamt sind damit 86,2 Prozent der Personeninterviews (15 bis unter 65 Jahre) der Welle 3 mit den Prozessdaten der Bundesagentur für Arbeit verknüpft.

2.4 Abgespaltene Haushalte

PASS ist als dynamisches Panel angelegt. Personen, die in Befragungshaushalte einziehen oder dort geboren werden, werden ebenfalls befragt, sofern sie 15 Jahre oder älter sind. Aber auch Personen, die aus den Befragungshaushalten ausziehen bzw. ein Jahr oder länger nicht im Haushalt leben, sollen weiterhin befragt werden. Ihre neuen Haushalte werden

²⁷ Auf Grund von Umstellungen bei der Filterung konnte es vorkommen, dass in Welle 2 und 3 die Frage nach der Zuspielbereitschaft erneut gestellt wurde, sofern die befragte Person ihre Zuspielbereitschaft in einer der Vorwellen noch nicht erklärt hatte.

²⁸ Befragte, die in der Vergangenheit ihre Zustimmung zur Verknüpfung der Befragungsdaten mit Prozessdaten der Bundesagentur für Arbeit gegeben haben, können diese natürlich jederzeit widerrufen.

als Abspaltungen des ursprünglichen Befragungshaushalts verstanden. Diese abgespaltenen Haushaltsteile (oder Splithaushalte) werden zu Befragungshaushalten des PASS. Alle in ihnen lebenden Personen ab 15 Jahre werden zu Zielpersonen für Personeninterviews. Sollte sich in einer Folgewelle wiederum ein Haushaltsteil aus dem Splithaushalt abspalten, wird auch dieser zu einem Befragungshaushalt im PASS, unabhängig davon, ob darin noch Personen aus einer der ursprünglichen Stichproben leben („infinite degree contagion model“, Rendtel und Harms 2009, 267). Ins Ausland verzogene Personen werden dagegen nicht weiter befragt, da sie nicht mehr zur Population gehören und auch inhaltlich die SGB II-spezifischen Fragestellungen nicht mehr zutreffen. Personen, die für weniger als ein Jahr nicht im Haushalt leben, werden weiterhin als Haushaltsmitglieder gezählt und bilden ebenfalls keinen neuen PASS-Haushalt.

Zwischen den Befragungszeitpunkten der 1. und 2. Welle haben sich von den bereits in der 1. Welle befragten Haushalten insgesamt 344 Haushalte abgespalten, von denen 46 während der Feldzeit der 2. Welle befragt werden konnten. Die nicht befragten Splithaushalte werden in der 3. Welle erneut kontaktiert, sofern sie die Teilnahme nicht endgültig verweigert haben. In Welle 3 gibt es 358 Splithaushalte, von denen insgesamt 142 befragt werden konnten.

Die befragten Splithaushalte können in den Datensätzen durch den Vergleich der aktuellen (*hnr*) und der ursprünglichen Haushaltsnummer (*uhnr*) identifiziert werden, die sich in diesen Fällen unterscheidet. Die Ursprungshaushaltsnummer (*uhnr*) enthält dabei die Haushaltsnummer des Panelhaushalts, von dem sich der neue Haushalt abgespalten hat. Splithaushalte übernehmen das Stichprobenkennzeichen (*sample*), die Informationen zum Jahr der Stichprobenziehung (*jahrsamp*), die Primary Sampling Unit (*psu*) und deren Stratifizierung (*strpsu*) von ihrem Ursprungshaushalt.

3 Datensatzstruktur

Die übliche Struktur für die Aufbereitung eines Paneldatensatzes, wie sie z.B. in Erhebungen wie dem Sozio-oekonomischen Panel (SOEP) oder dem British Household Panel Survey (BHPS) praktiziert wird, ist eine Ablage der Informationen zu Personen und Haushalten in jährlichen Einzeldatensätzen. Diese können ggf. um spezifische, auch wellenübergreifende Datensätze ergänzt werden, wie z.B. für Register- oder Spelldaten.

Diese Datenstruktur erlaubt es, die Informationen platzsparend zu speichern. Welche Variablen in welchen Jahren erhoben wurden, ist nach dem Blick in die Datensätze unmittelbar zu erkennen. Auch die Zuschlüsselung von Informationen – über Schlüsselvariablen wie Haushalts- oder Personennummern – ist vergleichsweise einfach. Allerdings ist diese für Paneldaten übliche Struktur auch mit Nachteilen verbunden, welche die Arbeit mit diesen Datensätzen deutlich erschweren. Sollen Analysen nicht nur im Quer- sondern auch im Längsschnitt durchgeführt werden, müssen zunächst alle relevanten Variablen aus den Einzeldatensätzen der jeweiligen Wellen in einen gemeinsamen Datensatz integriert werden, wobei darauf geachtet werden muss, dass es sich inhaltlich tatsächlich um dieselben Kon-

strukture handelt. Für typische Längsschnittanalysen muss der so erstellte, wellenübergreifende Datensatz dann in das so genannte Long-Format umstrukturiert werden. Im Gegensatz zum Wide-Format, in dem in der Datenmatrix für jede Beobachtungseinheit (z.B. ein Haushalt oder eine Person) genau eine Zeile vorhanden ist und dann für jede Erhebungswelle mehrere Datensätze existieren, werden im Long-Format alle einer Beobachtungseinheit zugeordneten Erhebungswellen untereinander angeordnet. Statt die Informationen in wellenspezifischen Variablen in derselben Zeile anzuordnen, werden im Long-Format die Informationen in wellenspezifischen Zeilen der Beobachtungseinheiten der jeweils gleichen Variable zugeordnet.

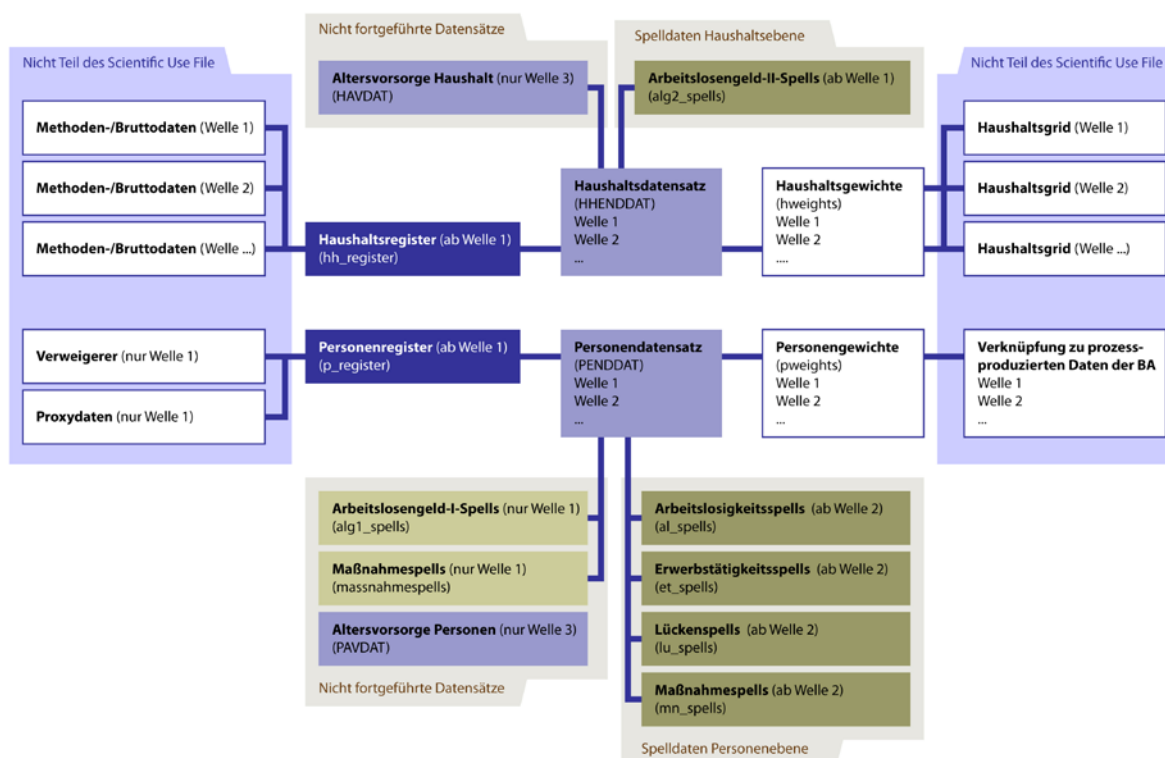
Auch die Aufbereitung im Long-Format ist mit Vor- und Nachteilen verbunden. Der entscheidende Vorteil dieser Variante ist, dass die Daten bereits in der für viele Längsschnittanalysen (wie z.B. Verweildaueranalysen) benötigten Struktur vorliegen. Der zusätzliche Aufwand für die Erstellung eines wellenübergreifenden Files entfällt. Der Umstieg vom Long- ins Wide-Format ist zudem vergleichsweise einfach zu realisieren. STATA bietet dazu bspw. mit dem „reshape“-Befehl eine Möglichkeit, ohne großen Aufwand zwischen den beiden Formaten zu wechseln. Das bis vor wenigen Jahren noch zentrale Argument gegen diese Art der Datensatzstruktur ist der deutlich höhere Bedarf an Speicherplatz, der sich insbesondere daraus ergibt, dass auch Variablen, die nur in einer oder in wenigen Befragungswellen erhoben wurden, immer eine komplette, alle Wellen übergreifende Spalte im Datensatz benötigen. Zudem werden die Long-Files mit zunehmender Laufzeit des Panels auch allein durch das wiederholte Aneinanderreihen der Einzelwellen relativ groß, wodurch der Speicherbedarf und die Dauer einzelner Operationen mit den Daten deutlich steigen. Die breite Verfügbarkeit schneller Prozessoren und großer Speicherkapazitäten selbst auf einfachen Desktop-PCs, lässt dieses Gegenargument jedoch mittlerweile als vernachlässigbar erscheinen. Einen weiteren Nachteil stellt das Zuschlüsseln von Informationen dar. Im Gegensatz zu jenen im Wide-Format aufbereiteten Datensätzen benötigt man jetzt eine zusätzliche Schlüsselvariable, mit der eine Beobachtung eindeutig identifiziert werden kann. Dies kann ein Wellenindikator in den Haushalts- oder Personendatensätzen sein oder aber die Spellnummer in den Spelldatensätzen, die ebenfalls im Long-Format vorliegen. Weiterhin ist nicht auf den ersten Blick ersichtlich, welche Variablen für welche Wellen erhoben wurden, da alle jemals erfragten Variablen im Datensatz vorhanden sind. Diese Variablen sind für die Wellen, in denen sie nicht erhoben wurden, mit einem Sondercode (-9) gefüllt.

In der Abwägung der Vor- und Nachteile des Long-Formats für die Nutzer, überwiegen u.E. deutlich die Vorteile. Entsprechend wurden der Haushalts- und der Personendatensatz des PASS (*HHENDDAT*; *PENDDAT*) sowie die entsprechenden Gewichtungsdaten (*hweights*; *pweights*) im Long-Format aufbereitet. Die im Rahmen des Schwerpunktmoduls zur Altersvorsorge auf Haushalts- und Personenebene erhobenen Informationen wurden jeweils in separate Datensätze ausgelagert (*HAVDAT*; *PAVDAT*), die über die entsprechenden Schlüsselvariablen an den Haushalts- bzw. Personendatensatz angespielt werden können.²⁹

²⁹ Die Datensätze enthalten jeweils alle in Welle 3 befragten Haushalte bzw. Personen. Wurden für einen Fall keine Altersvorsorgeinformationen erhoben, so sind die inhaltlichen Variablen mit dem Code ‚-3‘ (trifft nicht zu, Filter) gefüllt.

Weiterhin enthält der Scientific Use File auf Haushaltsebene die in Spellform aufbereiteten Daten zum Arbeitslosengeld II-Bezug des Haushalts (*alg2_spells*). Auf Personenebene existieren vier Spelldatensätze, und zwar (1.) Spelldaten zu Erwerbsepisoden (*et_spells*), (2.) Zeiten der Arbeitslosigkeit (*al_spells*) und (3.) der Nichterwerbstätigkeit (*lu_spells*) jeweils seit Januar 2005 sowie (4.) zur Teilnahme an arbeitsmarktpolitischen Maßnahmen (*mn_spells*) seit Januar 2006. Das Haushalts- und das Personenregister (*hh_register*; *p_register*) liegen im Wide-Format vor.

Abbildung 2: Datensatzstruktur des PASS in Welle 3



4 Generierte Variablen

4.1 Vercodung offener Angaben

Einige Items der Befragung wurden als geschlossene Items mit offener Restkategorie bzw. als offene Items erhoben. In diesen Fällen wurden i.d.R. zusätzliche Variablen generiert³⁰, die sich von der ursprünglichen Version der Variable lediglich dahingehend unterscheiden, dass die offenen Angaben soweit als möglich den entsprechenden Kategorien zugecodet wurden. In einigen Fällen wurden zudem auch neue Kategorien auf Basis der offenen Angaben gebildet. Die Bezeichnung dieser Variablen unterscheidet sich von der ursprünglichen Variable häufig lediglich dadurch, dass die „0“ in der letzten Ziffer durch eine „1“ ersetzt wur-

³⁰ Andere offene Angaben wurden dagegen nicht vercodet, bspw. der Name des Trägers der Grundsicherung (P138).

de. Die Items zu Geburtsland, Staatsangehörigkeit und den Zuzugsländern von Eltern/Großeltern wurden zudem anonymisiert und tragen sprechende Variablennamen³¹. Tabelle 9 geben einen Überblick über die offenen Angaben, die in der 3. Welle vercodet wurden³².

Tabelle 9: Vercodung offener Angabe auf Haushaltsebene in Welle 3

Fragebogennummer wieder befragte HH	neu befragte HH	Zugecodet zu Variable	Datensatz	Bezeichnung
HH64	n. in Fbvers.	HW0881a-j	HHENDDAT	anderer, nicht in Liste enthaltener Auszugsgrund
HH85	HH42	HD0601	HHENDDAT	HH-Sprache: andere, nicht in Liste enthaltene Sprache
HH87	HH44	HD0801	HHENDDAT	HH-Sprache, nach Nachfrage zu weiterer Sprache: andere, nicht in Liste enthaltene Sprache
HH88	HH45	HD0901	HHENDDAT	HH-Sprache, gleiche Nutzung zweier Sprachen: erste Sprache ist andere, nicht in Liste enthaltene Sprache
HH89	HH46	HD1001	HHENDDAT	HH-Sprache, gleiche Nutzung zweier Sprachen: zweite Sprache ist andere, nicht in Liste enthaltene Sprache
HH99	HH56	AL21301a-f AL21401a-f AL21501a-f AL21601a-f AL21701a-f AL21801a-f AL21851a-f AL21901 a-f AL22001a-f AL22101a-f AL22102a-f AL22103a-f	alg2_spells	anderer, nicht in Liste enthaltener Grund für Kürzung
Z1	Z1	AL22201a-f	alg2_spells	anderer, nicht in Liste enthaltener Grund, warum kein ALG2 mehr bezogen wurde

³¹ ogebland (Geburtsland); ostaatant (Staatsangehörigkeit); ozulanda bis ozulandf (Zuzugsländer Eltern/Großeltern)

³² Variablen, für die zwar in der 1. Welle offene Angaben erhoben und vercodet wurden, jedoch nicht in den Folgewellen, sind (mit Ausnahme des ALG2-Spelldatensatzes) nicht aufgeführt. Diese Variablen sind für die Beobachtungen ab der 2. Welle mit -9 gefüllt (Item in Welle nicht erhoben) und im Datenreport der 1. Welle dokumentiert.

Tabelle 10: Vercodung offener Angaben auf Personenebene in Welle 3

Fragebogennummer Personen	Senioren	Zugecodet zu Variable	Datensatz	Bezeichnung
P8_6	n. in Fbvers	PB0231	PENDDAT	anderer, nicht in Liste enthaltener deutscher Schulabschluss (Fortschreibung)
P8_7	n. in Fbvers	PB0231	PENDDAT	anderer, nicht in Liste enthaltener ausländischer Schulabschluss (Fortschreibung)
P10_9	P5_9	PB0401	PENDDAT	anderer, nicht in Liste enthaltener deutscher Schulabschluss (Ersterhebung o. keine Angabe in VW)
P10_10	P5_10	PB0401	PENDDAT	anderer, nicht in Liste enthaltener ausländischer Schulabschluss (Ersterhebung o. keine Angabe in VW)
P11	n. in Fbvers	PB1001	PENDDAT	anderer, nicht in Liste enthaltener ausländischer Schulabschluss (Ersterhebung o. keine Angabe in VW)
P26_9	P7_9	PB1301a-j	PENDDAT	anderer, nicht in Liste enthaltener deutscher Berufsschluss (Fortschreibung o. Ersterhebung)
P26_10	P7_10	PB1301a-j	PENDDAT	anderer, nicht in Liste enthaltener ausländischer Berufsabschluss (Fortschreibung o. Ersterhebung)
P28	n. in Fbvers	PB1601	PENDDAT	anderer, nicht in Liste enthaltener Abschluss dem der ausländische Abschluss entspricht
P111	n. in Fbvers	AL0601	al_spells	anderer, nicht in Liste enthaltener Grund, weshalb nicht mehr arbeitslos gemeldet
P129	n. in Fbvers	LU0101	lu_spells	anderer, nicht in Liste enthaltener Lückenstatus
P143	n. in Fbvers	PTK0321a-g	PENDDAT	anderer, nicht in Liste enthaltener Grund, weshalb keine Arbeit gesucht werden muss
P162	n. in Fbvers	MN0201a-h MN0202h	mn_spells	anderer, nicht in Liste enthaltener Teil, aus dem das Programm besteht
P167	n. in Fbvers	MN1001a-e	mn_spells	anderer, nicht in Liste enthaltener Grund, warum Programm beendet wurde
Z2	n. in Fbvers	ET2401	et_spells	andere, nicht in Liste enthaltene Art, von der Tätigkeit zu erfahren
P184	n. in Fbvers	PAS0901a-i	PENDDAT	andere, nicht in Liste enthaltene Stellen bei denen sich ZP über offene Stellen informiert hat
P219	P51	PG0901a-g	PENDDAT	andere, nicht in Liste enthaltene gesundheitl. Einschränkung
P223	P54	PG1301	PENDDAT	andere, nicht in Liste enthaltene Krankenversicherung
P264	P73	ogebland	PENDDAT	anderes, nicht in Liste enthaltene Geburtsland
P267	P76	ostaatan	PENDDAT	andere, nicht in Liste enthaltene Staatsangehörigkeit

Tabelle 10: Vercodung offener Angaben auf Personenebene in Welle 3 (Fortsetzung 1)

Fragebogennummer Personen	Senioren	Zugecodet zu Variable	Datensatz	Bezeichnung
P274	P80	<i>ozulanda-f</i>	<i>PENDDAT</i>	anderes, nicht in der Liste enthaltenes Zuzugsland des Elternteils/ Großeltern- teils
P275	P81	<i>PMI1111</i>	<i>PENDDAT</i>	Sprache im Freundeskreis: andere, nicht in Liste enthaltene Sprache
P276	P82	<i>PMI1121</i>	<i>PENDDAT</i>	Sprache im Freundeskreis, gleiche Nut- zung zweier Sprachen: erste Sprache ist andere, nicht in Liste enthaltene Spra- che
P277	P83	<i>PMI1131</i>	<i>PENDDAT</i>	Sprache im Freundeskreis, gleiche Nut- zung zweier Sprachen: zweite Sprache ist andere, nicht in Liste enthaltene Sprache
P278_9	n. in Fbvers	<i>PSH0201</i>	<i>PENDDAT</i>	anderer, nicht in Liste enthaltener deut- scher Schulabschluss der Mutter
P278_10	n. in Fbvers	<i>PSH0201</i>	<i>PENDDAT</i>	anderer, nicht in Liste enthaltener aus- ländischer Schulabschluss der Mutter
P279_7	n. in Fbvers	<i>PSH0301a-i</i>	<i>PENDDAT</i>	anderer, nicht in Liste enthaltener deut- scher Berufsabschluss der Mutter
P279_8	n. in Fbvers	<i>PSH0301a-i</i>	<i>PENDDAT</i>	anderer, nicht in Liste enthaltener aus- ländischer Berufsabschluss der Mutter
P289_9	n. in Fbvers	<i>PSH0501</i>	<i>PENDDAT</i>	anderer, nicht in Liste enthaltener deut- scher Schulabschluss des Vaters
P289_10	n. in Fbvers	<i>PSH0501</i>	<i>PENDDAT</i>	anderer, nicht in Liste enthaltener aus- ländischer Schulabschluss des Vaters
P290_7	n. in Fbvers	<i>PSH0601a-i</i>	<i>PENDDAT</i>	anderer, nicht in Liste enthaltener deut- scher Berufsabschluss des Vaters
P290_8	n. in Fbvers	<i>PSH0601a-i</i>	<i>PENDDAT</i>	anderer, nicht in Liste enthaltener aus- ländischer Berufsabschluss des Vaters

4.2 Harmonisierung

Bei einigen Variablen gab es über die Wellen hinweg Veränderungen an den Erhebungsinstrumenten. Vor allem die Integration des Erwerbsbiografiemoduls in Welle 2 führte dazu, dass zentrale Informationen zum Erwerbsstatus, der aktuellen Haupterwerbstätigkeit, dem Status der Nichterwerbstätigkeit und dem Bezug von Arbeitslosengeld I anders erhoben wurden als noch in der 1. Welle. Informationen werden seitdem nicht mehr nur bezogen auf den Befragungszeitpunkt erhoben, sondern in Episodenform für bestimmte Zeiträume erfasst.

Um in solchen Fällen wellenübergreifende Analysen zu vereinfachen, werden für wichtige Konstrukte über die Wellen hinweg harmonisierte Variablen generiert. Harmonisierungen stellen damit eine Sondergruppe der generierten Variablen dar (siehe Abschnitt 4.4), die dazu dienen, unterschiedlich erhobene Konstrukte im Nachhinein zu vereinheitlichen.

Veränderungen zwischen den Wellen können das gesamte Erhebungskonzept, Kategorien und die befragten Gruppen betreffen. Harmonisierte Variablen berücksichtigen deshalb unterschiedliche Quellvariablen, die sich aus veränderten Erhebungskonzepten ergeben, Veränderungen an Kategorien sowie an den befragten Gruppen und versuchen, diese soweit möglich über die Wellen hinweg zu vereinheitlichen, bevor auf ihrer Basis die Generierung durchgeführt werden.

Eine Harmonisierung wurde bisher für den Erwerbsstatus (*erwerb2*) und die einfache Einteilung der Stellung im Beruf (*stibkz*) vorgenommen. Allerdings ist damit zu rechnen, dass sich die Zahl der notwendigen Harmonisierungen mit zunehmender Laufzeit des Panels erhöht.

Tabelle 11: Harmonisierte Variablen im Personendatensatz (PENDDAT)

Variable	Thematischer Bereich	Bezeichnung
<i>erwerb2</i>	Erwerbstätigkeit	Erwerbsstatus, generiert (alle Wellen)
<i>stibkz</i>	Erwerbstätigkeit	Jetzige Stellung im Beruf, einfache Einteilung, harmonisiert (anonymisiert)

Während die ausdrücklich harmonisierten Variablen – neben Veränderungen am Erhebungskonzept – auch Veränderungen an Kategorien und den befragten Gruppen über die Wellen berücksichtigen, werden von einem zweiten Variablentyp Veränderungen an den befragten Gruppen nicht explizit berücksichtigt. Diese Variablen werden zwar für alle Welle generiert, können jedoch je nach Welle Informationen für verschiedene Befragtengruppen enthalten. Diese Unterschiede ergeben sich aus Überarbeitungen der Filterführung, die zwischen den Wellen vorgenommen wurden und sich auf die jeweiligen Quellvariablen einer generierten Variable auswirken.

Wellenübergreifende Variablen dieses Typs treten somit neben die eigentlichen Harmonisierungen und vereinheitlichen einzelne Aspekte zwischen den Wellen. Im Gegensatz zu den harmonisierten Variablen werden sie in jeder Welle jeweils für alle Gruppen generiert, für die in dieser Welle die jeweiligen Quellvariablen erhoben wurden. Für Auswertungen im Querschnitt einer spezifischen Welle sind sie somit einfach nutzbar. Im Längsschnitt müssen diese Unterschiede jedoch beachtet werden, bevor Aussagen über Veränderungen zwischen den Wellen getroffen werden können.

Aus diesem Grund sollte vor der Arbeit mit den wellenübergreifenden, jedoch nicht harmonisierten Variablen geprüft werden, ob Unterschiede in den befragten Gruppen problematisch für die jeweiligen Auswertungen sein könnten und ob ggf. eine Vereinheitlichung notwendig ist³³.

Insbesondere die folgenden, wellenübergreifenden Variablen weisen Unterschiede hinsichtlich der Gruppen auf, für die sie generiert sind:

³³ Bspw. wurden in Welle 1 andere Befragtengruppen nach einer Erwerbstätigkeit gefragt als in den Folgewellen. Entsprechend haben auch jeweils andere Gruppen Angaben zur StiB, zur beruflichen Tätigkeit, zur Arbeitszeit, zur Befristung etc. gemacht.

Tabelle 12: Wellenübergreifend generierte, aber nicht vollständig harmonisierte Variablen im Personendatensatz (PENDDAT)

Variable	Thematischer Bereich	Bezeichnung
<i>nichterw</i>	Erwerbstätigkeit	Erwerbsstatus, generiert (alle Wellen)
<i>nichtew2</i>	Erwerbstätigkeit	jetzige Stellung im Beruf, einfache Einteilung, harmonisiert (anonymisiert)
<i>isco88</i>	Erwerbstätigkeit	ISCO 88 (ZUMA-Verkodung), aktuelle Erwerbstätigkeit, generiert
<i>isco88it</i>	Erwerbstätigkeit	ISCO 88 (Infratest-Verkodung), aktuelle Erwerbstätigkeit, generiert
<i>kldb_it</i>	Erwerbstätigkeit	Klassifizierung der Berufe 1992 (Infratest-Verkodung), aktuelle Erwerbstätigkeit
<i>arbzeit</i>	Erwerbstätigkeit	wöchentliche Arbeitszeit inkl. Angaben bei unregelmäßiger Arbeitszeit, generiert
<i>befrist</i>	Erwerbstätigkeit	jetzige Tätigkeit: Befristung? Generiert (alle Wellen)
<i>mps</i>	Erwerbstätigkeit	Magnitude-Prestigeskala, aktuelle Erwerbstätigkeit, generiert
<i>siops</i>	Erwerbstätigkeit	Standard International Occupational Prestige Scale, aktuelle Erwerbstätigkeit, generiert
<i>isei</i>	Erwerbstätigkeit	International Socio-Economic Index, aktuelle Erwerbstätigkeit, generiert
<i>egp</i>	Erwerbstätigkeit	Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), aktueller Beruf, generiert
<i>esec</i>	Erwerbstätigkeit	European Socio-economic Classification (ESeC), aktueller Beruf, generiert
<i>stib</i>	Erwerbstätigkeit	Stellung im Beruf, Kennziffer, aktuelle Erwerbstätigkeit, generiert
<i>alg1abez</i>	Leistungsbezug	aktueller ALG1-Bezug, generiert
<i>aktmassn</i>	Maßnahmeteilnahme	aktuelle Teilnahme an einem vom Arbeitsamt finanzierten/geförderten Programm, generiert

Weiterhin sind Variablen im Datensatz enthalten, die zwar für alle Wellen generiert wurden, aber nicht im Längsschnitt auswertbar sind. Dabei handelt es sich um die generierten Einkommensvariablen auf Personenebene. Die Unterschiede in den Erhebungskonzepten zwischen Welle 1 und den Folgewellen wurden hier bei der Generierung nicht berücksichtigt. Für Welle 1 enthalten die betreffenden Variablen Einkommensangaben, die sich auf die Haupterwerbstätigkeit beziehen, wenn der Befragte für mindestens eine Stunde in der Woche erwerbstätig war. Ab Welle 2 beziehen sich die Angaben nicht mehr nur auf die Haupterwerbstätigkeit, sondern auf alle zum Befragungszeitpunkt aktuell andauernden Erwerbstätigkeiten mit einem Einkommen von über 400 Euro – die Variablen enthalten also hier kumulierte Angaben. Hintergrund ist auch die Einführung des Erwerbsbiografiemoduls ab Welle 2.

Eine Auswertung dieser Variablen im Längsschnitt würde also zu Fehlern führen, da die enthaltenen Angaben auf verschiedenen Erhebungskonzepten basieren, inhaltlich unterschiedliche Konstrukte enthalten und für verschiedene Befragtengruppen vorliegen.

Eine Überarbeitung und Vereinheitlichung in Form wellenübergreifender oder harmonisierter Variablen ist geplant und wird mit dem Scientific Use File einer späteren Welle ausgeliefert.

Tabelle 13: Für alle Wellen generierte, aber nicht im Längsschnitt auswertbare Variablen im Personendatensatz (PENDDAT)

Variable	Thematischer Bereich	Bezeichnung
<i>brutto</i>	Einkommen	Bruttoeinkommen inkl. kategorisierter Angaben, generiert
<i>bruttokat</i>	Einkommen	kategorisiertes Bruttoeinkommen, generiert
<i>netto</i>	Einkommen	Nettoeinkommen inkl. kategorisierter Angaben, generiert
<i>nettokat</i>	Einkommen	kategorisiertes Nettoeinkommen, generiert

4.3 Dependent Interviewing

An verschiedenen Stellen wurden sowohl in den Haushaltsinterviews als auch in den Personeninterviews Informationen per Dependent Interviewing, d.h. in Abhängigkeit von Angaben in der Vorwelle, erhoben. Dabei wurden Angaben aus der letzten Befragung zur Steuerung der Filterführung genutzt oder direkt als Teil des Fragetextes im aktuellen Interview einblendet.

Mit der Nutzung von Vorwelleninformationen wurden insbesondere zwei Ziele verfolgt. Erstens sollten an bestimmten Stellen nur Veränderungen seit der Vorwelle erfasst werden, teilweise in Abhängigkeit davon, ob schon in der Vorwelle Informationen zu einem bestimmten Fragekomplex vorlagen³⁴. An diesen Stellen wurden Vorwelleninformationen zur Filtersteuerung genutzt. Zweitens sollten dem Befragten inhaltliche Informationen gegeben werden. An den Stellen, an denen Veränderungen seit der Vorwelle erfasst werden sollten, wurde deshalb das Interviewdatum der Vorwelle in den Fragetext eingebunden, um den Berichtszeitraum deutlicher abzugrenzen³⁵. An anderen Stellen, insbesondere bei der Fortschreibung von Episodeninformationen³⁶, wurden auch Antworten des Befragten aus der Vorwelle in die Fragetexte integriert, die dazu dienten, den Befragten an die Angaben in der Vorwelle zu erinnern. Damit sollte verhindert werden, dass Statuswechsel berichtet werden, die in der Realität nicht aufgetreten sind, sondern ein Artefakt der offenen Erhebung und damit von Erinnerungsfehlern oder ungenauen Angaben sind.

Betrachtet man also die Informationen einer einzelnen Welle im Datensatz, so liegen in Folge des Dependent Interviewings für einige Befragte nur unvollständige Informationen vor, die lediglich die Veränderung zwischen zwei Befragungszeitpunkten widerspiegeln. Für Befragte, die erstmals zu einem Thema befragt wurden, können dagegen Informationen vorliegen, die bezogen auf diese Welle vollständig sind³⁷.

³⁴ Bspw. wurden Personen nur einmal nach ihrem höchsten Schulabschluss gefragt. Haben sie einmal eine entsprechende Angabe gemacht, werden in den folgenden Wellen nur noch neue Schulabschlüsse erfragt, die seit der jeweils letzten Befragung erworben wurden.

³⁵ Sollten bspw. nur neue Schulabschlüsse seit dem letzten Interview berichtet werden, wurde zunächst gefragt: „Haben Sie seit unserem letzten Interview am [Einblendung Interviewdatum der Vorwelle] einen allgemeinbildenden Schulabschluss gemacht?“

³⁶ Beispiele sind die Fortschreibung des Arbeitslosengeld-II-Bezugs aus der Vorwelle im Haushaltsinterview der jeweils aktuellen Welle, bzw. die Fortschreibung von Erwerbstätigkeiten oder Arbeitslosigkeiten im Personeninterview.

³⁷ Personen, die erstmals zu ihrem Schulabschluss befragt werden, berichten den jeweils höchsten Schulabschluss. Die Information zum höchsten Schulabschluss liegt in dieser Welle also vollständig in den erhobenen Variablen vor. In der Folgewelle wird nur noch nach neu erworbenen

Im Zuge der Datenaufbereitung werden die erhobenen Veränderungen mit Informationen aus der Vorwelle kombiniert, um auch Variablen bzw. Datensätze mit vollständigen Informationen zu erzeugen. Mit den neu erhobenen Episodeninformationen werden die Spells in den bestehenden Spelldatensätzen fortgeschrieben. In den Querschnittdatensätzen (*HHENDDAT*, *PENDDAT*) werden dagegen generierte Variablen erstellt, in denen die Informationen der Vorwelle mit den erhobenen Veränderungen kombiniert werden.

Im Folgenden geben die Tabelle 14, Tabelle 15 und Tabelle 16 einen kurzen Überblick über alle relevanten Stellen in den Fragebögen und zeigen, in welcher Variable die fortgeschriebene Information zu finden ist. Die Fälle, in denen die Fortschreibung in generierten Variablen erfolgt, sind zusätzlich in Kapitel 4.4 des vorliegenden Datenreports aufgeführt.

Tabelle 14: Fortgeschriebene Vorwelleninformationen in Welle 3, wiederholt befr. Haushalte

Haushaltsfragebogen für wiederholt befragte Haushalte (HHalt)			
Konstrukt	Fragenr.	Anmerkung	Fortschreibung in Var.
Wohnsituation		Wohnform, Art des Mietverhältnisses und die Art des Wohnheims von Infratest während des Interviews fortgeschrieben	HHENDDAT: <i>HW0200 bis HW0400</i>
Haushaltsstruktur	HH1-HH60	Haushaltsgröße von Infratest während des Interviews fortgeschrieben	HHENDDAT: <i>HA0100</i>
		Geschlecht der Personen im HH von Infratest während des Interviews ggf. korrigiert	HHENDDAT: <i>HD0100a bis HD0100o</i>
		Alter der Personen im HH von Infratest während des Interviews fortgeschrieben	HHENDDAT: <i>HD0200a bis HD0200o</i>
		Verwandtschaftsbeziehungen von Infratest während des Interviews fortgeschrieben	nicht im SUF ausgeliefert
Jahr des Einzugs in aktuelle Wohnung	HH66	in generierter Variable fortgeschrieben	HHENDDAT: <i>einzuwj</i>
Wohnfläche der Wohnung in qm Arbeitslosengeld-II-Bezug	HH65/HH69 HH91-HH104	in generierter Variable fortgeschrieben	HHENDDAT: <i>wohnfl</i>
		in Arbeitslosengeld-II-Spell-datensatz fortgeschrieben	alg2_spells: Variablen des Arbeitslosengeld-II-Spell-datensatzes
		Informationen zum aktuellen Arbeitslosengeld-II-Bezug des HH	HHENDDAT: <i>alg2abez</i> PENDDAT: <i>hhalg2</i>
		Information zum Arbeitslosengeld-II-Bezug der Bedarfsgemeinschaft	p_register: <i>bgbez3; bgbezb3</i>

Schulabschlüssen gefragt. Wurde bspw. ein Schulabschluss neu erhoben, so liegt diese Information zwar in den erhobenen Variablen vor, es ist allerdings unklar, ob es sich tatsächlich um den höchsten Schulabschluss handelt. Die Information in der Folgewelle ist in diesem Sinne in den berichteten Variablen unvollständig.

Tabelle 15: Fortgeschriebene Vorwelleninformationen in Welle 3, erstmals befr. Haushalte

Haushaltsfragebogen erstmals befragte Haushalte (HHneu)			
Konstrukt	Fragenr.	Anmerkung	Fortschreibung in Var.
Arbeitslosengeld-II-Bezug	HH48-HH60	in Arbeitslosengeld-II-Spell-datensatz fortgeschrieben Informationen zum aktuellen Arbeitslosengeld-II-Bezug des HH Information zum Arbeitslosengeld-II-Bezug der Bedarfsgemeinschaft	alg2_spells: <i>Variablen des Arbeitslosengeld-II-Spell-datensatzes</i> HHENDDAT: <i>alg2abez</i> PENDDAT: <i>hhalg2</i> p_register: <i>bgbez3; bgbezb3</i>

Tabelle 16: Fortgeschriebene Vorwelleinformationen in Welle 3, Personenfragebogen

Personenfragebogen			
Konstrukt	Fragenr.	Anmerkung	Fortschreibung in Var.
Höchster allgemeinbildender Schulabschluss	P7-P12	in generierter Variable fortgeschrieben	PENDDAT: <i>schul1</i> (ohne offene Ang.) <i>schul2</i> (mit offenen Ang.)
Jahr des höchsten Schulabschlusses	P13	in generierter Variable fortgeschrieben	PENDDAT: <i>schulabj</i>
Berufsabschluss	P23-P29	höchster Berufsabschluss in generierter Variable fortgeschrieben	PENDDAT: <i>beruf1</i> (ohne offene Ang.) <i>beruf2</i> (mit offenen Ang.) <i>berabj</i>
Jahr des Berufsabschlusses	P30	in generierter Variable fortgeschrieben	
Zeiten von Erwerbsätigkeiten mit einem Einkommen von über 400 Euro	E38-E63, P38-P63	in Spelldatensatz zu Erwerbstätigkeiten fortgeschrieben; Informationen zur aktuellen Erwerbstätigkeit in generierten Variablen fortgeschrieben Informationen zum aktuellen (Nicht-)Erwerbstätigkeitsstatus in generierten Variablen fortgeschrieben	et_spells: <i>Variablen des Spelldatensatzes zu Erwerbstätigkeiten</i> PENDDAT: <i>isco88; isco88_it; kldb_it;stib; stibkz; arbeitszeit; befrist; mps; siops; isej; egp; esec</i> PENDDAT: <i>erwerb2; nichterw; nichtew2</i>
Zeiten von gemeldeten Arbeitslosigkeiten, inkl. Arbeitslosengeld-I-Bezug	A106-A117 P106-P117	in Spelldatensatz zu Arbeitslosigkeiten fortgeschrieben	al_spells: <i>Variablen des Spelldatensatzes zu Arbeitslosigkeiten</i>

Tabelle 16: Fortgeschriebene Vorwelleinformationen in Welle 3, Personenfragebogen (Fortsetzung)

Personenfragebogen			
Konstrukt	Fragenr.	Anmerkung	Fortschreibung in Var.
		Informationen zum aktuellen (Nicht-)Erwerbstätigkeitsstatus in generierten Variablen fortgeschrieben	PENDDAT: <i>erwerb2; nichterw; nichtew2</i>
		Informationen zum aktuellen Arbeitslosengeld-I-Bezug	al_spells: <i>Variablen zum Arbeitslosengeld-I-Bezug im Spelldatensatz zu Arbeitslosigkeiten</i> PENDDAT: <i>alg1abez</i>

Von diesen Merkmalen, bei denen in der Vergangenheit erhobene Informationen mit Angaben zu Veränderungen zwischen den Befragungszeitpunkten fortgeschrieben bzw. aktualisiert werden, sind die so genannten „unveränderlichen Merkmale“ zu unterscheiden. Bei ihnen wird davon ausgegangen, dass sie sich im Zeitverlauf nicht verändern. Entsprechend werden sie im PASS nur einmal erhoben, auch wenn teilweise spätere Korrekturen möglich sind. Da Angaben zu diesen Merkmale in der Regel nur für den Zeitpunkt der ersten Befragung in den erhobenen Variablen verfügbar sind, werden sie danach in Form generierter Variablen zur Verfügung gestellt (siehe dazu Kapitel 4.4, den PASS User Guide bzw. bis zu dessen Erscheinen Kapitel 10.3 im PASS Datenreport der Welle 2).

4.4 Einfache Generierungen

Die einfachen Generierungen umfassen Variablen, bei denen mehrere, aus erhebungstechnischen Gründen getrennte Items eines Konstrukts zusammengefasst oder Informationen der aktuellen Welle mit Vorwelleinformation kombiniert (siehe Kapitel 4.3) (wie z. B. beim höchsten Bildungsabschluss) oder wichtige Informationen aus anderen Teildatensätzen zugespielt wurden (z.B. Indikatoren für den aktuellen Bezug von Arbeitslosengeld I oder Arbeitslosengeld II).

Bei Haushalten bzw. Personen, die erstmals zu einem Thema befragt werden, können die einfachen generierten Variablen immer auf Basis der in der aktuellen Welle erhobenen Informationen erzeugt werden. Bei Haushalten bzw. Personen, die bereits in einer früheren Welle Angaben zu einem Thema gemacht haben, können sie allerdings in den Querschnittdatensätzen (*HHENDDAT; PENDDAT*) hinsichtlich der Herkunft der zu ihrer Generierung jeweils notwendigen Variablen unterschieden werden. Die drei verschiedenen Typen der einfachen generierten Variablen sind in Tabelle 17 aufgeführt.

Tabelle 17: Typen der einfachen generierten Variablen in den Querschnitt Datensätzen (HHENDDAT; PENDDAT) bei Haushalten bzw. Personen, die schon in einer früheren Welle Angaben zum jeweiligen Thema gemacht haben

Typ	Generierung auf Basis von Quelldaten aus		Beschreibung
	Welle der erstmaligen Erfassung des Themas bei HH/Person	Aktuelle Welle	
<i>unveränderlich (uv)</i>	ja	nein	Die bei erstmaliger Erhebung erfassten Informationen werden in die Folgewelle grundsätzlich übernommen – außer in der aktuellen Welle wurden Eingabefehler korrigiert. <u>Bsp:</u> <i>zpsex</i> (Geschlecht)
<i>fortgeschrieben (fs)</i>	ja	ja	Die in der Vorwelle aktuellen Informationen werden mit Angaben der aktuellen Welle kombiniert und ggf. fortgeschrieben. <u>Bsp:</u> <i>schul1</i> (Höchster Schulabschluss)
<i>unabhängig neu (neu)</i>	nein	ja	Die Variable wird in jeder Welle unabhängig von Angaben aus der Vorwelle aus den Angaben der aktuellen Welle neu generiert. <u>Bsp:</u> <i>hhincome</i> (Haushaltsnettoeinkommen)

Die einfachen generierten Variablen sind in den datensatzspezifischen Tabelle 18 bis Tabelle 25 dargestellt. Die einzelnen Variablen werden dabei jeweils kurz beschrieben. Darüber hinaus werden die für die Generierung der Variable in Welle 3 notwendigen Quellvariablen aufgeführt³⁸. Für die Querschnitt Datensätze (HHENDDAT; PENDDAT) ist zudem jeweils angegeben, um welchen Typ der in Tabelle 16 dargestellten einfacherer Generierung es sich handelt (uv; fs; neu). Für die Spelldatensätze ist diese Unterteilung nicht sinnvoll, da es hier keine wellenspezifischen Beobachtungen gibt. Stattdessen werden die generierten Variablen jeweils auf Spellebene neu erzeugt, wenn der Spell in der aktuellen Welle neu angelegt wur-

³⁸ Wie die Variablen in den Querschnitt Datensätzen (HHENDDAT; PENDDAT) für Beobachtungen der Welle 1 bzw. Welle 2 generiert wurden, ist in den jeweiligen Datenreports dokumentiert. Ebenfalls in der Dokumentation der jeweiligen Welle beschrieben ist die Generierung der wellenspezifischen Variablen in den Registerdatensätzen. Die generierten Variablen in den Spelldatensätzen wurden immer in den bereits fortgeschriebenen Datensätzen erzeugt. Wurde ein Spell nicht fortgeschrieben, blieben die entsprechenden generierten Variablen unverändert (ggf. mit Ausnahme, dass im Zensierungsindikator ein Sondercode gesetzt wurde, falls der Spell aus technischen Gründen nicht fortgeführt werden konnte). Wurde ein Spell fortgeschrieben, wurde immer auf die aktuellsten Informationen zurückgeriffen, d.h. auf die mit den Informationen aus der aktuellen Welle gefüllten Variablen bzw. die für die aktuelle Welle relevanten Querschnittvariablen in den Spells.

de oder mit in der aktuellen Welle erhobenen Angaben fortgeschrieben wurde. Auch die Registerdatensätze folgen einer anderen Logik, so dass auch hier auf eine weitere Differenzierung verzichtet wurde.

Tabelle 18: Einfache Generierungen für Welle 3 im Haushaltsdatensatz (HHENDDAT) (alphabetische Reihenfolge)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>alg2abez</i>	<i>Aktueller ALG2-Bezug des HH, generiert</i> Indikator für aktuellen Arbeitslosengeld-II-Bezug des Haushalts (neu)	<i>zensiert; AL20300; AL20400; AL20500 (alg2_spells);</i> Information zu weiteren Arbeitslosengeld-II-Bezügen (HHalt: HH104; HHneu: HH61); <i>hintjahr (HHENDDAT)</i>
<i>bik</i>	<i>BIK-Regionsgrößenklassen (GKBIK10), generiert</i> Die Information zur Regionsgrößenklasse wurde von TNS Infratest generiert, indem von der in den Adressdaten vorliegenden Postleitzahl auf GKBIK10 umgeschlüsselt wurde (neu)	vom Erhebungsinstitut geliefert
<i>blneuallt</i>	<i>alte oder neue Bundesländer, generiert</i> Zusammenfassung der Bundesländer zu alten Bundesländern (ohne Berlin) und neuen Bundesländern (mit Berlin). Das Bundesland wurde von TNS Infratest auf Basis der in den Adressdaten vorliegenden Postleitzahlen bestimmt (neu)	vom Erhebungsinstitut generiert und gelieferte Information zum Bundesland, in dem Haushalt zum Befragungszeitpunkt lebt
<i>einzugj</i>	<i>Einzugsjahr in akt. Wohnung, generiert</i> Information, in welchem Jahr der Haushalt in die aktuelle Wohnung eingezogen ist. Bei wiederbefragten Haushalten wurde das Einzugsjahr ab der 2. Welle nur erhoben, wenn es sich um einen in einem Wohnheim lebenden Haushalt handelte oder wenn seit der Vorwelle ein Umzug stattgefunden hatte (fs)	<u>Bei Ersterfassung:</u> <i>HW0900 (HHENDDAT)</i> <u>Bei wiederholter Erfassung:</u> <i>einzugj</i> aus Vorwelle; <i>HW0900, HW0200; umzug (HHENDAT)</i>
<i>hhinckat</i>	<i>kategorisiertes Haushaltseinkommen pro Monat (in Euro), generiert</i> Zusammenfassung der mehrstufig erhobenen kategorialen Angaben zum Haushaltseinkommen zu einer Variablen (neu)	<i>HEK0700; HEK0800; HEK0900; HEK1000; HEK1100 (HHENDDAT)</i>
<i>hhincome</i>	<i>Haushaltseinkommen pro Monat (in Euro) inkl. kategorisierter Angaben, generiert</i> Generierung einer integrierten Variable aus kategorisierter und offener Angabe zum Haushaltsnettoeinkommen (neu)	<i>HEK0600; HEK0700; HEK0800; HEK0900; HEK1000; HEK1100 (HHENDDAT)</i>
<i>hintdat</i>	<i>Interviewdatum Haushaltsinterview</i> Generierte Variable mit dem Datum, an dem das Haushaltsinterview geführt wurde im Format JMMTT (neu)	<i>hintjahr, hintmon, hinttag (HHENDDAT)</i>
<i>kindu4</i>	<i>Steuervariable: Kind unter 4 Jahre im HH</i> Die Variable zeigt an, dass in der Welle mindestens eine Person im Haushalt unter vier Jahre alt ist. Da die Generierung nur auf den Altersangaben im Haushaltsdatensatz basiert, ist unerheblich, ob diese Person im Alter von vier Jahren auch tatsächlich Kind einer anderen im Haushalt lebenden Person ist (neu)	<i>HD0200a - HD0200o (HHENDDAT)</i>

Tabelle 18: Einfache Generierungen für Welle 3 im Haushaltsdatensatz (HHENDDAT) (alphabetische Reihenfolge) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>kindu13</i>	<i>Steuervariable: Kind unter 13 Jahre im HH</i> Die Variable zeigt an, dass in der Welle mindestens eine Person im Haushalt unter 13 Jahre alt ist. Da die Generierung nur auf den Altersangaben im Haushaltsdatensatz basiert, ist unerheblich, ob diese Person im Alter von 13 Jahren auch tatsächlich Kind einer anderen im Haushalt lebenden Person ist (neu)	<i>HD0200a - HD0200o</i> (HHENDDAT)
<i>kindu15</i>	<i>Steuervariable: Kind unter 15 Jahre im HH</i> Die Variable zeigt an, dass in der Welle mindestens eine Person im Haushalt unter 15 Jahre alt ist. Da die Generierung nur auf den Altersangaben im Haushaltsdatensatz basiert, ist unerheblich, ob diese Person im Alter von 15 Jahren auch tatsächlich Kind einer anderen im Haushalt lebenden Person ist. Fehlte die offene Altersangabe, wurde die kategoriale Nachfrage zu den Altersgruppen in die Generierung einbezogen (neu)	<i>HD0200a - HD0200o</i> ; kategoriale Nachfrage zu Altergruppe (bei KA in <i>HD0200</i>) (HHENDDAT)
<i>wohnfl</i>	<i>Wohnfläche in qm, generiert</i> Information zur Wohnfläche der aktuellen Wohnung des Haushalts. Bei wiederbefragten Haushalten, wurde die Wohnungsgröße ab der 2. Welle nur erhoben, wenn seit der Vorwelle ein Umzug oder eine Veränderung der Wohnung stattgefunden hatte (fs)	<u>Bei Ersterfassung:</u> <i>HW1000</i> (HHENDDAT) <u>Bei wiederholter Erfassung:</u> <i>wohnfl</i> aus Vorwelle; <i>HW1000</i> ; <i>HW0910</i> ; <i>HW0920</i> (HHENDDAT)

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>aktmassn</i>	<i>Akt. Teiln. an einem v. Arbeitsamt finanz./geförd. Programm, generiert</i> Indikator: Befragungsperson nimmt zum Interviewzeitpunkt an einer arbeitsmarktpolitischen Maßnahme teil (neu)	<i>MN0500</i> ; zensiert (<i>mn_spells</i>); <i>PA0711b-f</i> ; <i>PA0721a-f</i> (PENDDAT); Information aus Prüfnachfrage P 178_X_Prüf (Personenfragebogen)
<i>alg1abez</i>	<i>aktueller ALG1-Bezug, generiert</i> Indikator: Befragungsperson bezieht zum Interviewzeitpunkt Arbeitslosengeld I. In der 3. Welle wurden die Zeiten seit Januar 2006 erhoben, in denen der Befragte arbeitslos gemeldet war. Für jede Episode wurde zudem erfragt, ob und in welchem Zeitraum Arbeitslosengeld I bezogen wurde. Diese Information wurde mit einer Nachfrage für Befragte kombiniert, die 58 Jahre oder älter waren und deshalb Arbeitslosengeld I beziehen können, ohne arbeitslos gemeldet zu sein (neu)	<i>AL0700</i> ; <i>AL1000</i> ; <i>AL1100</i> ; <i>AL1200</i> ; <i>alg1bj</i> ; <i>alg1ej</i> (<i>al_spells</i>); <i>PA0405</i> (PENDDAT); Information ob eine weitere AL vorliegt (P117/A117)

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 1)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>alg1s05</i>	<p><i>Indikator: Erhalt ALG1 seit Jan. 2005? Generiert (alle Wellen)</i></p> <p>Indikator: Befragungsperson hat irgendwann seit Januar 2005 Arbeitslosengeld I bezogen. In der 3. Welle wurden die Zeiten seit Januar 2006 erhoben, in denen der Befragte arbeitslos gemeldet war. Für jede Episode wurde zudem erfragt, ob und in welchem Zeitraum Arbeitslosengeld I bezogen wurde. Diese Information wurde mit einer Nachfrage für Befragte kombiniert, die 58 Jahre oder älter waren und deshalb Arbeitslosengeld I beziehen können, ohne arbeitslos gemeldet zu sein (fs)</p>	<p><u>Bei Ersterfassung:</u> AL0700; AL1200; alg1bj; alg1ej (al_spells); PET0911; PA0405 (PENDDAT); Information ob eine weitere AL vorliegt (P117/A117)</p> <p><u>Bei wiederholter Erfassung:</u> AL0700; AL1200; alg1bj; alg1ej (al_spells); alg1s05 aus Vorwelle; PET0911; PA0405 (PENDDAT); Information ob eine weitere AL vorliegt (P117/A117)</p>
<i>apartner</i>	<p><i>Steuervariable: Unverh. Partner im HH</i></p> <p>Indikator: Befragungsperson hat einen unspezifischen oder unverheirateten Partner im HH (neu)</p>	Informationen zu Beziehungen im Haushalt (Haushaltsgrid); PD0500 - PD0900 (PENDDAT)
<i>arbeitszeit</i>	<p><i>wöchentliche Arbeitszeit inkl. Angaben bei unregelmäßiger Arbeitszeit, gen.</i></p> <p>Wöchentliche Arbeitszeit der Erwerbstätigkeit, die der Befragte zum Interviewzeitpunkt ausübt, generiert aus offenen Angaben zur Arbeitszeit und kategorialer Nachfrage bei unregelmäßigen Arbeitszeiten (neu)</p>	ET2101; ET2201 (et_spells); PET0510; PET0700 (PENDDAT)
<i>befrist</i>	<p><i>jetzige Tätigkeit: Befristung? Generiert (alle Wellen)</i></p> <p>Indikator: Die Erwerbstätigkeit, die der Befragte zum Interviewzeitpunkt ausübt, ist befristet (neu)</p>	PET2510a; PET2510b (PENDDAT)
<i>begjeewt</i>	<p><i>Anfangsjahr der ersten Erwerbstätigkeit, generiert</i></p> <p>Jahr, in dem der Befragte erstmals hauptberuflich tätig war. Für die Generierung wurden Informationen zur ersten Erwerbstätigkeit mit Informationen aus den Erwerbstätigkeitsepisoden kombiniert, wenn der Befragte die erste Erwerbstätigkeit bereits bei der Erhebung der Erwerbstätigkeitsepisoden seit Januar 2006 genannt hat (uv)</p>	<p><u>Bei Ersterfassung:</u> bjahr (et_spells); PET0150; PET0151; PET3200b (PENDDAT)</p> <p><u>Nach Ersterfassung:</u> begjeewt aus Vorwelle (PENDDAT)</p>
<i>begmeewt</i>	<p><i>Anfangsmonat der ersten Erwerbstätigkeit, generiert</i></p> <p>Monat, in dem der Befragte erstmals hauptberuflich tätig war (Generierung siehe begjeewt) (uv)</p>	<p><u>Bei Ersterfassung</u> bmonat (et_spells); PET0150; PET0151; PET3200a (PENDDAT)</p> <p><u>Nach Ersterfassung:</u> begmeewt aus Vorwelle (PENDDAT)</p>
<i>berabj</i>	<p><i>Jahr des höchsten Berufsabschlusses</i></p> <p>Jahr, in dem der Befragte seinen zum Interviewzeitpunkt höchsten Berufsabschluss erworben hat (fs)</p> <p><u>Anmerkung:</u> Die Abschlussjahre der in der 1. Welle berichteten Berufsabschlüsse wurden in der 2. Welle nacherhoben.</p>	<p><u>Bei Ersterfassung:</u> PB1300a-j; PB1310am-km; PB1310aj-kj (PENDDAT)</p> <p><u>Bei wiederholter Erfassung:</u> berabj aus Vorwelle; PB1300a-j; PB1310am-km; PB1310aj-kj (PENDDAT)</p>

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 2)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>beruf1</i>	<i>höchster beruflicher Abschluss, ohne ausl. Abschl. und off. Angaben, generiert</i> Identifikation des zum jeweiligen Interviewzeitpunkt höchsten Berufsabschlusses durch Hierarchisierung der von den Befragungspersonen genannten beruflichen Abschlüsse, ohne offene Angaben (fs)	<u>Bei Ersterfassung:</u> <i>PB0100; PB0200; PB0300; PB1200b; PB1200c; PB1300a-j; (PENDDAT)</i> <u>Bei wiederholter Erfassung:</u> <i>beruf1</i> aus Vorwelle; <i>PB0100; PB0200; PB1200a; PB1300a-j (PENDDAT)</i>
<i>beruf2</i>	<i>höchster beruflicher Abschluss, mit ausl. Abschl. und off. Angaben, generiert</i> Wie <i>beruf1</i> mit folgenden Unterschieden: 1. Einbeziehen von offenen Angaben; 2. Einbeziehen von Angaben zu ausländischen Abschlüssen; 3. Differenzierung der Studienabschlüsse nicht nach FH/Uni, sondern nach Abschlussniveau (Bachelor; Master/Diplom; Promotion). (fs)	<u>Bei Ersterfassung:</u> <i>PB0200; PB1301a-j; PB1500a; PB1500b; PB1500c; PB1601 (PENDDAT)</i> <u>Bei wiederholter Erfassung:</u> <i>beruf2</i> aus Vorwelle; <i>PB0200; PB1301a-j; PB1500a; PB1500b; PB1500c; PB1601 (PENDDAT)</i>
<i>brutto</i>	<i>Bruttoeinkommen inkl. kategorisierter Angaben, generiert</i> Generierung einer integrierten Variable aus kategorisierter und offener Angabe zum Bruttoeinkommen (neu)	<i>PEK0100b; PEK0200; PEK0300; PEK0400; PEK0500; PEK0600 (PENDDAT)</i>
	<u>Anmerkung:</u> Die Variable ist für alle Wellen generiert, aber aktuell nicht im Längsschnitt auswertbar (siehe dazu Kapitel 4.2).	
<i>bruttokat</i>	<i>kategorisiertes Bruttoeinkommen, generiert</i> Zusammenfassung der mehrstufig erhobenen kategorialen Bruttoeinkommensangaben zu einer Variable (neu)	<i>PEK0200; PEK0300; PEK0400; PEK0500; PEK0600 (PENDDAT)</i>
	<u>Anmerkung:</u> Die Variable ist für alle Wellen generiert, aber aktuell nicht im Längsschnitt auswertbar (siehe dazu Kapitel 4.2).	
<i>ejhrlewt</i>	<i>Endzeitpunkt letzte Erwerbstätigkeit (Jahr)</i> Jahr, in dem der Befragte zuletzt erwerbstätig war. Für die Generierung wurden Informationen aus den Erwerbstätigkeitsepisoden mit den Angaben zur letzten Erwerbstätigkeit kombiniert, wenn der Befragte seit Januar 2005 nicht erwerbstätig war (fs)	<u>Bei Ersterfassung:</u> <i>PET1200b (PENDDAT); ejahr; emonat (et_spells)</i> <u>Bei wiederholter Erfassung:</u> <i>ejhrlewt</i> aus Vorwelle (<i>PENDDAT</i>); <i>ejahr; emonat (et_spells)</i>
<i>ekin1517</i>	<i>Steuervariable: eigenes Kind zwischen 15 und 17 Jahre im HH</i> Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/Adoptivkind im Alter zwischen 15 und 17 Jahren im Haushalt hat (neu)	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 3)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>ekind</i>	<i>Steuervariable: eigenes Kind im HH</i> Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/Adoptivkind beliebigen Alters im Haushalt hat (neu)	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<i>ekin614</i>	<i>Steuervariable: eigenes Kind zwischen 6 und 14 Jahre im HH</i> Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/Adoptivkind im Alter zwischen 6 und 14 Jahren im Haushalt hat (neu)	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<i>ekinu15</i>	<i>Steuervariable: eigenes Kind unter 15 Jahre im HH</i> Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/Adoptivkind im Alter von unter 15 Jahren im Haushalt hat (neu)	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<i>ekinu18</i>	<i>Steuervariable: eigenes Kind unter 18 Jahre im HH</i> Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/Adoptivkind im Alter von unter 18 Jahren im Haushalt hat (neu)	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<i>emonlewt</i>	<i>Endzeitpunkt letzte Erwerbstätigkeit (Monat)</i> Monat, in dem der Befragte zuletzt erwerbstätig war (Generierung siehe <i>ejhrlewt</i>) (fs)	<u>Bei Ersterfassung:</u> <i>PET1200a</i> (<i>PENDDAT</i>); <i>emonat2</i> (<i>et_spells</i>) <u>Bei wiederholter Erfassung</u> <i>emonlewt</i> aus Vorwelle (<i>PENDDAT</i>); <i>emonat</i> (<i>et_spells</i>)
<i>epartner</i>	<i>Steuervariable: Ehepartner oder eingetr. Lebenspartner im HH</i> Indikator, dass Befragungsperson einen Ehepartner oder einen gleichgeschlechtlichen, eingetragenen Lebenspartner im Haushalt hat (neu)	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 4)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>erwerb2</i>	<p><i>Erwerbsstatus, generiert (alle Wellen)</i> Integrierte, für 1. Welle harmonisierte Erwerbsstatusvariable. Die in der 1. Welle erstellte Variable <i>erwerb</i> konnte aufgrund der Umstellung auf die Erwerbsbiografien ab der 2. Welle nicht fortgeschrieben werden. Daher wurde eine neue Statusvariable erstellt, die für die 1. Welle auf der bisherigen Erwerbsstatusvariable <i>erwerb</i> basiert und ab der 2. Welle auf Basis des Nichterwerbsstatus (mit offenen Angaben), dem Status als Schüler/Student/Azubi, der aktuellen Arbeitszeit und der Spellinformation zu aktuell andauernden Erwerbstätigkeiten generiert wurde.</p> <p>Grundlage für die Generierung ist die Information aus dem Episodendatensatz der jeweiligen Welle, ob eine aktuell andauernde Episode eines bestimmten Typs vorliegt. Im Falle einer aktuell andauernden Nichterwerbstätigkeit aus dem Lückendatensatz wird die Art dieser Nichterwerbstätigkeit über die Variable <i>LU0101</i> (d.h. inkl. offener Angaben) bestimmt. Im Rahmen der Harmonisierung werden die Welle 1 Kategorien 2 („arbeitslos“) und 3 („ABM, 1-Euro-Job u.ä.“) zu einer gemeinsamen Kategorie 2 („arbeitslos“) zusammengefasst. Auch die früheren Kategorien 8 („Lehre/ Ausbildung/ Weiterbildung/ Umschulung“) und 10 („Student“) werden zwecks Integration mit den Daten ab Welle 2 in eine Kategorie zusammengeführt (neu)</p>	<p><i>PB0100; arbeit; nichtew2</i> (PENDDAT) <i>ET0601 (et_spells)</i></p>
<i>famstand</i>	<p><i>Familienstand, generiert</i> Generierung einer integrierten Familienstandsvariable aus den Angaben im Personenbogen und der aus dem Haushaltsdatensatz generierten Steuerungsvariable <i>epartner</i> (neu)</p>	<p><i>epartner; PD0500; PD0700</i> (PENDDAT)</p>
<i>gebhalbj</i>	<p><i>Geburtshalbjahr, generiert</i> Indikator, ob Geburtstag im 1. oder 2. Halbjahr des Geburtsjahres liegt (neu)</p>	<p>Informationen zum Geburtsmonat</p>
<i>hhalg2</i>	<p><i>Steuervariable: Alg2-Bezug aktuell</i> Indikator, für den Arbeitslosengeld-II-Bezug des Haushalts zum Zeitpunkt des Haushaltsinterviews (neu)</p>	<p><i>HA0250b (HHENDDAT)</i> <i>AL20400; AL20500 (alg2_spells)</i></p>
<i>kindzges</i>	<p><i>Anzahl eigene Kinder gesamt (inner- und außerhalb HH), generiert</i> Gesamtzahl der Kinder der Befragungsperson als Summe der Kinder, die im Haushalt leben und der Anzahl der Kinder außerhalb des Haushalts (neu)</p>	<p>Informationen zu Beziehungen im Haushalt (Haushaltsgrid); <i>PD0900; PD1000; PD1100</i> (PENDDAT)</p>

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 5)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>kindzihh</i>	<i>Anzahl eigene Kinder innerhalb des HH, generiert</i> Auf Basis der Angaben im HH-Bogen generierte Variable für die Zahl der Kinder, die eine Person im HH hat (Summe aus der Zahl der Personen in der HH-(Halb-) Matrix, die als Kinder der Befragungsperson zählen und der Anzahl der Personen in der HH-(Halb-)Matrix, für die angegeben wird, dass es sich bei der HH-Person um einen Elternteil handelt) (neu)	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<i>mberuf1</i>	<i>höchster ber. Abschluss der Mutter, inkl. Mut. im HH, ohne of. Ang., gen.</i> Der Berufsabschluss der Mutter wurde in der 1. Welle nur dann erhoben, wenn diese nicht im Haushalt wohnt. Wohnte sie im Haushalt, wurden die Angabe zum Berufsabschluss aus ihrem Personeninterview ergänzt. Ab der 2. Welle wurde die Information zum Berufsabschluss der Mutter bei allen Neubefragten Personen erhoben, unabhängig davon, ob die Mutter im Haushalt lebte oder nicht. Ab der 2. Welle wiederholt Befragte wurden die Werte aus der generierten Variable <i>mberuf1</i> aus der Vorwelle übernommen (uv)	<u>Bei Ersterfassung:</u> <i>PSH0300a-i (PENDDAT)</i> <u>Nach Ersterfassung:</u> <i>mberuf1</i> aus Vorwelle (PENDDAT)
<i>mberuf2</i>	<i>höchster ber. Abschluss der Mutter, inkl. Mut. im HH, mit of. Ang., gen.</i> Wie <i>mberuf1</i> mit dem Unterschied, dass bei der Generierung von <i>mberuf2</i> auch die gemachten offenen Angaben berücksichtigt wurden (uv)	<u>Bei Ersterfassung:</u> <i>PSH0301a-i (PENDDAT)</i> <u>Nach Ersterfassung:</u> <i>mberuf2</i> aus Vorwelle (PENDDAT)
<i>mhh</i>	<i>Steuervariable: Mutter lebt im HH</i> Indikator, der anzeigt, dass die unspezifische, leibliche, Stief-/Adoptivmutter der Befragungsperson im Haushalt lebt (neu)	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 6)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>migration</i>	<p><i>Migrationshintergrund d. Befragten, generiert</i> Generierte Variable für eine vierstufige Unterscheidung des Migrationshintergrunds: ohne Migrationshintergrund; eigene Migration (1. Generation); Migration mind. eines Elternteils, aber nicht selbst migriert (2. Generation); Migration mind. eines Großelternteils, aber nicht selbst migriert und kein Elternteil migriert (3. Generation). (uv)</p> <p><u>Anmerkung:</u> Das Generierungskonzept der Variable wurde ab der 2. Welle überarbeitet. Statt nur die Information, ob die Befragungsperson in Deutschland geboren ist, und die Angaben, welche Vorfahren nach Deutschland zugezogen sind, zur Generierung zu verwenden, wird nun auch die Information, ob und ggf. welcher Vorfahr außerhalb von Deutschland geboren ist, einbezogen. Um eine konsistente Logik über die Wellen zu gewährleisten, wurde die Variable für 1. Welle ebenfalls neu generiert.</p>	<p><u>Bei Ersterfassung:</u> <i>PMI0100; PMI0700; PMI0800a-f; PMI0900a-f (PENDDAT)</i></p> <p><u>Nach Ersterfassung:</u> <i>migration</i> aus Vorwelle (PENDDAT)</p>
<i>mschul1</i>	<p><i>höchster allg. Schulabschluss d. Mutter, inkl. Mutter im HH, o. offene Ang., gen.</i> Der höchste Schulabschluss der Mutter wurde in der 1. Welle nur dann erhoben, wenn diese nicht im Haushalt wohnt. Wohnte sie im Haushalt, wurden die Angabe zum Schulabschluss aus ihrem Personeninterview ergänzt (uv)</p> <p>Ab der 2. Welle wurde die Information zum höchsten Schulabschluss der Mutter bei allen neubefragten Personen erhoben, unabhängig davon ob die Mutter im Haushalt lebte oder nicht.</p>	<p><u>Bei Ersterfassung:</u> <i>PSH0200 (PENDDAT)</i></p> <p><u>Nach Ersterfassung:</u> <i>mschul1</i> aus Vorwelle (PENDDAT)</p>
<i>mschul2</i>	<p><i>höchster allg. Schulabschluss d. Mutter, inkl. Mut. im HH, inkl. of. Ang., gen.</i> Wie <i>mschul1</i> mit dem Unterschied, dass bei der Generierung von <i>mberuf2</i> auch die gemachten offenen Angaben berücksichtigt wurden (uv)</p>	<p><u>Bei Ersterfassung:</u> <i>PSH0201 (PENDDAT)</i></p> <p><u>Nach Ersterfassung:</u> <i>mschul2</i> aus Vorwelle (PENDDAT)</p>
<i>mstib</i>	<p><i>Stellung im Beruf der Mutter, Kennziffer, generiert</i> Detaillierte Stellung der Mutter im Beruf, generiert aus den Einzelvariablen (uv)</p>	<p><u>Bei Ersterfassung:</u> <i>PSH0320; PSH0330; PSH0340; PSH0360; PSH0370; PSH0380 (PENDDAT)</i></p> <p><u>Nach Ersterfassung:</u> <i>mstib (PENDDAT)</i></p>

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 7)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>netto</i>	<i>Nettoeinkommen inkl. kategorisierter Angaben, generiert</i> Generierung einer integrierten Variable aus kategorisierter und offener Angabe zum Nettoeinkommen (neu)	<i>PEK0700b; PEK0800; PEK0900; PEK1000; PEK1100; PEK1200 (PENDDAT)</i>
	<u>Anmerkung:</u> Die Variable ist für alle Wellen generiert, aber aktuell nicht im Längsschnitt auswertbar (siehe dazu Kapitel 4.2).	
<i>nettokat</i>	<i>kategorisiertes Nettoeinkommen, generiert</i> Zusammenfassung der mehrstufig erhobenen kategorialen Nettoeinkommensangaben zu einer Variable (neu)	<i>PEK0800; PEK0900; PEK1000; PEK1100; PEK1200 (PENDDAT)</i>
	<u>Anmerkung:</u> Die Variable ist für alle Wellen generiert, aber aktuell nicht im Längsschnitt auswertbar (siehe dazu Kapitel 4.2).	
<i>nichterw</i>	<i>Status Nichterwerbstätigkeit, generiert (alle Wellen)</i> Integrierte Variable für den Status der Nichterwerbstätigkeit der Befragungsperson. Für Welle 1 generiert aus <i>PET0800</i> . Ab Welle 2 generiert aus Informationen zur Art der aktuellen Nichterwerbstätigkeit aus dem Lückenmodul (<i>LU0100</i> , d.h. ohne Berücksichtigung der offenen Angaben) bzw. aus dem Arbeitslosigkeitsmodul zu aktuell andauernder Arbeitslosigkeit (neu)	<i>LU0100 (lu_spells); zensiert (al_spells); PET0151; PET0911 (PENDDAT);</i> Indikator für Fälle, für die fälschlicherweise kein Lückenstatus erhoben wurde
<i>nichtew2</i>	<i>Status Nichterwerbstätigkeit, generiert, inkl. off. Ang. (alle Wellen)</i> Integrierte Variable für den Status der Nichterwerbstätigkeit der Befragungsperson. Bei der Generierung von <i>nichtew2</i> wurden auch die gemachten offenen Angaben berücksichtigt. Für Welle 1 generiert aus <i>PET0801</i> . Ab Welle 2 generiert aus Informationen zur Art der aktuellen Nichterwerbstätigkeit aus dem Lückenmodul (<i>LU0101</i> , d.h. inklusive der offenen Angaben) bzw. aus dem Arbeitslosigkeitsmodul zu aktuell andauernder Arbeitslosigkeit (neu)	<i>LU0101 (lu_spells); zensiert (al_spells); PET0151; PET0911 (PENDDAT);</i> Indikator für Fälle, für die fälschlicherweise kein Lückenstatus erhoben wurde
<i>palter</i>	<i>Alter (aus p1), generiert</i> Alter der Befragungsperson, generiert auf Basis der Angaben zum Geburtsdatum und des Datums des Personeninterviews in der aktuellen Welle (neu)	<i>p1; pintjahr, pintmon, pinttag (PENDDAT)</i>
<i>panel</i>	<i>Panelbereitschaft</i> (neu)	Vom Erhebungsinstitut gelieferte Information zur Panelbereitschaft der Haushalte

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 8)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>pintdat</i>	<i>Interviewdatum Personeninterview</i> Generierte Variable mit dem Datum, an dem das Personeninterview geführt wurde im Format JMMTT (neu)	<i>pintjahr, pintmon, pinttag</i> (PENDDAT)
<i>schul1</i>	<i>höchster Schulabschluss, ohne ausl. Abschl. und offene Angaben</i> Variable für den höchsten Schulabschluss; äquivalente west- und ostdeutsche Abschlüsse wurden zusammengefasst (z.B. EOS und Abitur); ohne offene Angaben (fs)	<u>Bei Ersterfassung:</u> <i>PB0200; PB0220; PB0230; PB0300; PB0400</i> (PENDDAT) <u>Bei wiederholter Erfassung:</u> <i>schul1</i> aus Vorwelle; <i>PB0200; PB0220; PB0230; PB0300; PB0400</i> (PENDDAT)
<i>schul2</i>	<i>höchster Schulabschluss, mit ausl. Abschl. und offenen Angaben</i> Wie <i>schul1</i> mit folgenden Unterschieden: 1. Einbeziehen von offenen Angaben; 2. Einbeziehen von Angaben zu ausländischen Abschlüssen (fs)	<u>Bei Ersterfassung:</u> <i>PB0200; PB0220; PB0231; PB0300; PB0401</i> (PENDDAT) <u>Bei wiederholter Erfassung:</u> <i>schul2</i> aus Vorwelle; <i>PB0200; PB0220; PB0231; PB0300; PB0401</i> (PENDDAT)
<i>schulabj</i>	<i>Jahr des höchsten Schulabschlusses</i> Jahr, in dem die Befragungsperson ihren höchsten Schulabschluss erworben hat (fs) <u>Anmerkung:</u> Wiederholt Befragte, für die bereits Informationen zum Schulabschluss aus einer Vorwelle vorlagen, wurden in der aktuellen Welle nicht zum Abschlussjahr gefragt, wenn sie seit der Vorwelle einen neuen Abschluss erworben haben. In diesem Fall wurde das Abschlussjahr in Abhängigkeit des Interviewjahrs und -monats geschätzt. Wurde das Interview in der 3. Welle vor Mai 2009 geführt, wurde davon ausgegangen, dass der Abschluss in 2008 erworben wurde, bei später geführten Interviews wurde 2009 als Abschlussjahr angenommen.	<u>Bei Ersterfassung:</u> <i>PB0220; PB0230; PB0400; PB0410; ;pintjahr; pintmon</i> (PENDDAT) <u>Bei wiederholter Erfassung:</u> <i>schulabj</i> aus Vorwelle; <i>PB0220; PB0230; PB0400; PB0410; pintjahr; pintmon</i> (PENDDAT)
<i>stib</i>	<i>Stellung im Beruf, Kennziffer, generiert</i> Generierung der detaillierten Kennziffer der Stellung im Beruf aus den Einzelvariablen. Generierung der Variable mit Informationen aus dem Erwerbstätigkeitsmodul (<i>ET0601-ET1201</i>). Bei mehreren aktuell andauernden Erwerbstätigkeits-episoden wurde die Episode mit dem höheren Stundenumfang gewählt. Bei mehreren aktuell andauernden Episoden mit exakt gleichem Stundenumfang, diejenige die eher begonnen hat (neu)	<i>ET0500; ET0601 ET0701; ET0801; ET0901; ET1001; ET1101; ET1201</i> (<i>et_spells</i>)

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 9)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>stibewt</i>	<i>Stellung im Beruf, erste Erwerbstätigkeit, Kennziffer, generiert</i> Detaillierte Kennziffer der Stellung im Beruf, in dem der Befragte erstmals hauptberuflich tätig war. Für die Generierung wurden Informationen zur ersten Erwerbstätigkeit mit Informationen aus den Erwerbstätigkeitsepisoden kombiniert, wenn der Befragte angab, die erste Erwerbstätigkeit bereits bei der Erhebung der Erwerbstätigkeitsepisoden seit Januar 2006 genannt zu haben (uv)	<u>Bei Ersterhebung:</u> PET3300b; PET3000; PET3400; PET3500; PET3600; PET3700; PET3800; PET3900 (PENDDAT) ET0601; ET0701; ET0801; ET0901; ET1001; ET1101; ET1201 (et_spells) <u>Nach Ersterhebung:</u> stibewt aus Vorwelle (PENDDAT)
<i>stibkz</i>	<i>jetzige Stellung im Beruf, einf. Eint., harmonisiert (anonymisiert)</i> Generierung der einfachen Kennziffer der Stellung im Beruf aus den Einzelvariablen (neu)	PET1510 (PENDDAT)
<i>stiblewt</i>	<i>Stellung im Beruf, letzte Erwerbstätigkeit, Kennziffer, generiert</i> Detaillierte Kennziffer der Stellung im Beruf, in dem der Befragte zuletzt erwerbstätig war. Für die Generierung wurden Informationen aus den Erwerbstätigkeitsepisoden mit den Angaben zur letzten Erwerbstätigkeit kombiniert, wenn der Befragte seit Januar 2006 nicht erwerbstätig war (fs)	<u>Bei Ersterfassung:</u> PET1210b; PET1210; PET1220; PET1230; PET1240; PET1250; PET1260; PET1270 (PENDDAT) ET0601; ET0701; ET0801; ET0901; ET1001; ET1101; ET1200 (et_spells) <u>Bei wiederholter Erfassung:</u> stiblewt aus Vorwelle (PENDDAT); ET0601; ET0701; ET0801; ET0901; ET1001; ET1101; ET1200 (et_spells)
<i>vberuf1</i>	<i>höchster ber. Abschluss des Vaters, inkl. Vater im HH, ohne of. Ang., gen.</i> Generierung für den höchsten Berufsabschluss des Vaters analog zu <i>mberuf1</i> (uv)	<u>Bei Ersterfassung:</u> PSH0600a-i (PENDDAT) <u>Nach Ersterfassung:</u> vberuf1 aus Vorwelle (PENDDAT)
<i>vberuf2</i>	<i>höchster ber. Abschluss des Vaters, inkl. Vater im HH, mit of. Ang., gen.</i> Generierung für den höchsten Berufsabschluss des Vaters (unter Berücksichtigung der offenen Angaben) analog zu <i>mberuf2</i> (uv)	<u>Bei Ersterfassung:</u> PSH0601a-i (PENDDAT) <u>Nach Ersterfassung:</u> vberuf2 aus Vorwelle (PENDDAT)
<i>vhh</i>	<i>Steuervariable: Vater lebt im HH</i> Indikator, der anzeigt, dass der unspezifische, leibliche, Stief-/Adoptivvater der Befragungsperson im Haushalt lebt (neu)	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<i>vschul1</i>	<i>höchster allg. Schulabschluss d. Vaters, inkl. Vater im HH, o. of. Ang., gen</i> Generierung für den höchsten Schulabschluss des Vaters analog zu <i>mschul1</i> (uv)	<u>Bei Ersterfassung:</u> PSH0500 (PENDDAT) <u>Nach Ersterfassung:</u> vschul1 aus Vorwelle (PENDDAT)

Tabelle 19: Einfache Generierungen für Welle 3 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge) (Fortsetzung 10)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>vschul2</i>	<i>höchster allg. Schulabschluss d. Vater, inkl. Vater im HH, inkl. of. Ang., gen</i> Generierung für den höchsten Schulabschluss des Vaters (unter Berücksichtigung der offenen Angaben) analog zu <i>mschul2</i> (uv)	<u>Bei Ersterfassung:</u> <i>PSH0501</i> (PENDDAT) <u>Nach Ersterfassung:</u> <i>vschu2</i> aus Vorwelle (PENDDAT)
<i>vstib</i>	<i>Stellung im Beruf des Vaters, Kennziffer, generiert</i> Detaillierte Stellung im Beruf des Vaters, generiert aus den Einzelvariablen (uv)	<u>Bei Ersterfassung:</u> <i>PSH0620; PSH0630; PSH0640; PSH0660; PSH0670; PSH0680</i> (PENDDAT) <u>Nach Ersterfassung:</u> <i>vstib</i> aus Vorwelle (PENDDAT)

Tabelle 20: Einfache Generierungen für Welle 3 im Arbeitslosengeld-II-Spelldatensatz (*alg2_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>bmonat</i>	<i>ALG2-Episode: Anfangsmonat, generiert</i> Beginnmonat der ALG2-Episode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden. <u>Anmerkung:</u> Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben des Befragten sind ab der 2. Wellen in den Quellvariablen erhalten. Die Jahreszeitangaben wurden nach folgenden Regeln in Monatswerte recodiert 21 Jahresanfang/Winter → Januar 24 Frühjahr/Ostern → April 27 Jahresmitte/Sommer → Juli 30 Herbst → Oktober 32 Jahresende → Dezember	<i>AL20100</i> (<i>alg2_spells</i>)
<i>bjahr</i>	<i>ALG2-Episode: Anfangsjahr, generiert</i> Endjahr der Arbeitslosengeld-II-Episode.	<i>AL20200</i> (<i>alg2_spells</i>)
<i>emonat</i>	<u>Anmerkung:</u> siehe <i>bmonat</i> <i>ALG2-Episode: Endmonat, generiert</i> Endmonat der Arbeitslosengeld-II-Episode in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden und bei rechtszensierten (d.h. zum Befragungszeitpunkt des Haushalts aktuell andauernden) Spells der Interviewmonat eingesetzt wurde. <u>Anmerkung:</u> siehe <i>bmonat</i>	<i>AL20300</i> (<i>alg2_spells</i>) <i>hintmon</i> (HHENDDAT)

Tabelle 20: Einfache Generierungen für Welle 3 im Arbeitslosengeld-II-Spell Datensatz (*alg2_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung 1)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>ejahr</i>	<p><i>ALG2-Episode: Endjahr, generiert</i></p> <p>Endjahr der Arbeitslosengeld-II-Episode, in der bei rechtszensierten (d.h. zum Befragungszeitpunkt des Haushalts aktuell andauernden) Spells das Interviewjahr eingesetzt wurde.</p>	<p><i>AL20400 (alg2_spells)</i></p> <p><i>hintjahr (HHENDDAT)</i></p>
<i>alg2kbma</i> - <i>alg2kbmf</i>	<p><u>Anmerkung:</u> siehe <i>bmonat</i></p> <p><i>ALG2: 1. Kürzung: Anfangsmonat, generiert</i></p> <p>Beginnmonat der Arbeitslosengeld-II-Kürzung, in der für die Generierung die Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden.</p> <p><u>Anmerkung:</u> Die Arbeitslosengeld-II-Kürzungen sind in die Arbeitslosengeld-II-Bezugsspell eingebettet. Über den Indikator am Ende der jeweiligen Variablen (a-f) können die Informationen zu den einzelnen Kürzungsepisoden unterschieden werden. Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben des Befragten sind ab der 2. Welle in den Quellvariablen enthalten.</p>	<p>1. Kürzung: <i>AL21000a (alg2_spells)</i></p> <p>bis</p> <p>6. Kürzung: <i>AL21000f (alg2_spells)</i></p>
<i>alg2kbja</i> – <i>alg2kbjf</i>	<p><i>ALG2: 1. Kürzung: Anfangsjahr, generiert</i></p> <p>Beginnjahr der Arbeitslosengeld-II-Kürzung.</p> <p><u>Anmerkung:</u> siehe <i>alg2kma - alg2kbmf</i></p>	<p>1. Kürzung: <i>AL21100a (alg2_spells)</i></p> <p>bis</p> <p>6. Kürzung: <i>AL21100f (alg2_spells)</i></p>
<i>alg2kema</i> - <i>alg2kemf</i>	<p><i>ALG2: 1. Kürzung: Endmonat, generiert</i></p> <p>Endmonat der Arbeitslosengeld-II-Kürzung, in der für die Generierung die Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden. Wurde vom Befragten eine Dauer der Kürzung berichtet, wurde aus dieser auf Basis des generierten Beginndatums der Kürzung das jeweilige Enddatum berechnet.</p>	<p>1. Kürzung: <i>alg2kbma; alg2kbmja; AL21200a; AL21201a; AL21202a (alg2_spells)</i></p> <p>bis</p> <p>6. Kürzung: <i>alg2kbmf; alg2kbmjf; AL21200f; AL21201f; AL21202f (alg2_spells)</i></p>
<i>alg2keja</i> – <i>alg2kejf</i>	<p><u>Anmerkung:</u> siehe <i>alg2kma - alg2kbmf</i></p> <p><i>ALG2: 1. Kürzung: Endjahr, generiert</i></p> <p>Endjahr der Arbeitslosengeld-II-Kürzung. Wurde vom Befragten eine Dauer der Kürzung berichtet, wurde aus dieser auf Basis des generierten Beginndatums der Kürzung das jeweilige Enddatum berechnet.</p> <p><u>Anmerkung:</u> siehe <i>alg2kma - alg2kbmf</i></p>	<p>1. Kürzung: <i>alg2kbma; alg2kbmja; AL21200a; AL21201a; AL21202a (alg2_spells)</i></p> <p>bis</p> <p>6. Kürzung: <i>alg2kbmf; alg2kbmjf; AL21200f; AL21201f; AL21202f (alg2_spells)</i></p>

Tabelle 20: Einfache Generierungen für Welle 3 im Arbeitslosengeld-II-Spell Datensatz (*alg2_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung 2)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>AL22150a bis AL22150f</i>	<i>ALG2: Kürzung: Welchem HH-Mitglied wurde gekürzt, generiert</i> In dieser Variable wurde codiert, für welche Haushaltsmitglieder eine Kürzung des Arbeitslosengelds II stattgefunden hat. Es handelt sich um eine Stringvariable mit 15 Stellen. Von links beginnend, steht dabei jede Stelle dieser Variable für die Position einer Person im Haushaltsgrid. Die erste Stelle der Variable zeigt so bspw. an, ob in der jeweiligen Kürzungsepisode für die erste Person im Haushalt das Arbeitslosengeld II gekürzt wurde, die zweite Stelle, ob für die zweite Person gekürzt wurde usw. Da die Quellinformationen für die Generierung erst ab der 2. Welle erhoben wurden, ist die Frage in allen in der ersten Welle berichteten Kürzungen an allen 15 Stellen mit „I“ (Item in Welle nicht erhoben) gefüllt (s. u.)	Information, für welches Haushaltsmitglied in der jeweiligen Kürzungsepisode gekürzt wurde (HH102 im Haushaltsfragebogen für wiederbefragte Haushalte; HH53 im Haushaltsfragebogen für abgespaltene und Neubefragte Haushalte).

Jede der 15 Stellen der Variable, die für eine der maximal 15 Personen in der Haushaltsstruktur steht, hat einen der folgenden Codes, der den Kürzungsstatus der jeweiligen Person angibt.

Codes:

- 1 - dem HH-Mitglied wurde das ALG2 gekürzt
- 2 - em HH-Mitglied wurde das ALG2 nicht gekürzt
- W - weiß nicht
- K - keine Angabe
- T - trifft nicht zu (Filter)
- F - Frage irrtümlich nicht gestellt
- U - unplausibler Wert
- I - Item in Welle nicht erhoben

Tabelle 20: Einfache Generierungen für Welle 3 im Arbeitslosengeld-II-Spelldatensatz (*alg2_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung 3)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>zensiert</i>	<p><i>ALG2-Episode: Spell dauert z. Zeitpkt. d. letzten HH-Int. an (Rechtszens.), gen.</i></p> <p>Der Zensierungsindikator gibt an, ob ein Spell zum Zeitpunkt des letzten Haushaltsinterviews noch andauerte.</p> <p><u>Anmerkung:</u> Ein Spell gilt als zensiert, wenn eine der folgenden Bedingungen erfüllt ist:</p> <p>(a) Es handelt sich um einen in den Vorwellen zensierten Spell eines Haushalts, der in den Folgewellen bis hin zur aktuellen Welle nicht wiederbefragt wurde.</p> <p>(b) Ein in Welle 3 befragter Haushalt gibt in H91/H93 (HHalt) / H48/H50 (HHneu) an, dass ein ALG2-Spell auch zum Interviewdatum der Welle 3 noch andauert. Oder in H91/H93 (HHalt) / H48/H50 (HHneu) wird ein Enddatum genannt, das identisch mit dem Interviewdatum der Welle 3 ist und in der Nachfrage in H94 (HHalt) / HH51 (HHneu) wird bestätigt, dass der Bezug aktuell noch andauert.</p> <p>Der Code -5 wurde vergeben, wenn die Haushaltsauskunftsperson der Vorwelle in Welle 3 nicht mehr im Haushalt lebt und in Welle 3 nicht befragt wurde.</p>	<i>AL20100; AL20500 (alg2_spells)</i>

Tabelle 21: Einfache Generierungen für Welle 3 im Erwerbstätigkeitsspelldatensatz (*et_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>bmonat</i>	<p><i>Erwerbstätigkeit: Anfangsmonat, generiert</i></p> <p>Beginnmonat der Erwerbstätigkeitsepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden.</p> <p><u>Anmerkung:</u> Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben des Befragten sind in den Quellvariablen erhalten.</p> <p>Die Jahreszeitangaben wurden nach folgenden Regeln in Monatswerte recodiert:</p> <p>21 Jahresanfang/Winter → Januar 24 Frühjahr/Ostern → April 27 Jahresmitte/Sommer → Juli 30 Herbst → Oktober 32 Jahresende → Dezember</p>	<i>ET0100 (et_spells)</i>
<i>bjahr</i>	<p><i>Erwerbstätigkeit: Anfangsjahr, generiert</i></p> <p>Beginnjahr der Erwerbstätigkeitsepisode.</p> <p><u>Anmerkung:</u> siehe <i>bmonat</i></p>	<i>ET0200 (et_spells)</i>

Tabelle 21: Einfache Generierungen für Welle 3 im Erwerbstätigkeitsspell Datensatz (et_spells) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>emonat</i>	<i>Erwerbstätigkeit: Endmonat, generiert</i> Endmonat der Erwerbstätigkeitsepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden und bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells der Interviewmonat eingesetzt wurde.	<i>ET0300; ET0500 (et_spells)</i> <i>pintmon (PENDDAT)</i>
<i>ejahr</i>	<u>Anmerkung:</u> siehe <i>bmonat</i> <i>Erwerbstätigkeit: Endjahr, generiert</i> Endjahr der Erwerbstätigkeitsepisode, in der bei rechtszensierten (d.h. zum Befragungszeitpunkt des Haushalts aktuell andauernden) Spells das Interviewjahr eingesetzt wurde.	<i>ET0400; ET0500 (et_spells)</i> <i>pintjahr (PENDDAT)</i>
<i>zensiert</i>	<u>Anmerkung:</u> siehe <i>bmonat</i> <i>Erwerbstätigkeit: ggf. noch andauernder Spell (Rechtszensierung)</i> Der Zensierungsindikator gibt an, ob ein Spell zum Zeitpunkt des Personeninterviews in der letzten Befragungswelle noch andauerte, d.h. es handelt sich um einen rechtszensierten Spell. <u>Anmerkung:</u> Ein Spell gilt als zensiert, wenn eine der beiden folgenden Bedingungen erfüllt ist: Die Person gibt bei Frage P42 bezüglich Enddatum des Erwerbstätigkeitsspells an, zum Zeitpunkt des Interviews die Tätigkeit noch auszuüben (P42 Ende=0). Oder es wird bei P42 ein Enddatum genannt, welches identisch mit dem Interviewdatum ist und in der Nachfrage P43 bestätigt, dass die Erwerbstätigkeit derzeit noch andauert. Im Lückenmodul berichtete zusätzliche Erwerbstätigkeitsepisoden bzw. Datumskorrekturen wurden vor der Generierung berücksichtigt.	<i>ET0300; ET0400; ET0500 (et_spells)</i>
<i>stib</i>	<i>Stellung im Beruf, Kennziffer, generiert</i> Generierung der detaillierten Kennziffer der Stellung im Beruf aus den Einzelvariablen.	<u>Erfassung von</u> <u>Spellinformationen in Welle 3</u> <i>ET0601; ET0701; ET0801;</i> <i>ET0901; ET1001; ET1101;</i> <i>ET1201 (et_spells)</i>
<i>arbeitszeit</i>	<i>wöchentliche Arbeitszeit inkl. Angaben bei unregelmäßiger Arbeitszeit, gen.</i> Integrierte Variable zur wöchentlichen Arbeitszeit der Erwerbstätigkeit, zusammengefasst aus der offenen Angabe zur Arbeitszeit und der kategorialen Nachfrage. Für die geschlossenen Kategorien der Nachfrage wurden die Kategoriemittelwerte eingesetzt, für die offene Kategorie (40h und mehr) der Median der offenen angegebenen Wochenstunden.	Ansonsten bleibt der Wert der Vorwelle bestehen <u>Erfassung von</u> <u>Spellinformationen in Welle 3</u> <i>ET2101; ET2201 (et_spells)</i> Ansonsten bleibt der Wert der Vorwelle bestehen

Tabelle 22: Einfache Generierungen für Welle 3 im Arbeitslosigkeitsspell Datensatz (*al_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>bmonat</i>	<p><i>AL-Meldung: Anfangsmonat, generiert</i></p> <p>Beginnmonat der Arbeitslosmeldungsepisode in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden.</p> <p><u>Anmerkung:</u> Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben des Befragten sind in den Quellvariablen erhalten.</p> <p>Die Jahreszeitangaben wurden nach folgenden Regeln in Monatswerte recodiert: 21 Jahresanfang/Winter → Januar 24 Frühjahr/Ostern → April 27 Jahresmitte/Sommer → Juli 30 Herbst → Oktober 32 Jahresende → Dezember</p>	<i>AL0100 (al_spells)</i>
<i>bjahr</i>	<p><i>AL-Meldung: Anfangsjahr, generiert</i></p> <p>Beginnjahr der Arbeitslosmeldungsepisode.</p>	<i>AL0200 (al_spells)</i>
<i>emonat</i>	<p><u>Anmerkung:</u> siehe <i>bmonat</i></p> <p><i>AL-Meldung: Endmonat, generiert</i></p> <p>Endmonat der Arbeitslosmeldungsepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden und bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells der Interviewmonat eingesetzt wurde.</p>	<i>AL0300; AL0500 (al_spells)</i>
<i>ejahr</i>	<p><u>Anmerkung:</u> siehe <i>bmonat</i></p> <p><i>AL-Meldung: Endjahr, generiert</i></p> <p>Endjahr der Arbeitslosmeldungsepisode, in der bei rechtszensierten (d.h. zum Befragungszeitpunkt des Haushalts aktuell andauernden) Spells das Interviewjahr eingesetzt wurde.</p>	<i>AL0400; AL0500 (al_spells)</i>
<i>alg1bm</i>	<p><u>Anmerkung:</u> siehe <i>bmonat</i></p> <p><i>ALG1-Bezug: Anfangsmonat, generiert</i></p> <p>Beginnmonat des Arbeitslosengeld-I-Bezugs, in der für die Generierung die Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden.</p> <p><u>Anmerkung:</u> Die Arbeitslosengeld-I-Bezüge sind in die Arbeitslosmeldungsspell eingebettet. Pro Arbeitslosmeldung ist maximal ein Arbeitslosengeld-I-Bezugszeitraum vorgesehen. Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben des Befragten sind in den Quellvariablen enthalten.</p> <p>Umrechnung der Monatsangaben siehe <i>bmonat</i>.</p>	<i>AL0800 (al_spells)</i>

Tabelle 22: Einfache Generierungen für Welle 3 im Arbeitslosigkeitsspell Datensatz (*al_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung 1)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>alg1bj</i>	<i>ALG1-Bezug: Anfangsjahr, generiert</i> Beginnjahr des Arbeitslosengeld-I-Bezugs.	<i>AL0900 (al_spells)</i>
<i>alg1em</i>	<u>Anmerkung:</u> siehe <i>alg1bm</i> <i>ALG1-Bezug: Endmonat, generiert</i> Endmonat des Arbeitslosengeld-I-Bezugs, in der für die Generierung die Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden und bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells das Interviewdatum eingesetzt wurde.	<i>AL1000; AL1200 (al_spells)</i> <i>pintmon (PENDDAT)</i>
<i>alg1ej</i>	<u>Anmerkung:</u> siehe <i>alg2kma - alg2kbme</i> <i>ALG1-Bezug: Endjahr, generiert</i> Endjahr des Arbeitslosengeld-I-Bezugs, in der bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells das Interviewdatum eingesetzt wurde.	<i>AL1100; AL1200 (al_spells)</i> <i>pintjahr (PENDDAT)</i>
<i>alg1akt</i>	<u>Anmerkung:</u> siehe <i>alg2kma - alg2kbme</i> <i>ALG1-Bezug: ggw. noch andauernder Spell (Rechtszensierung)</i> Der Zensierungsindikator gibt an, ob der ALG1-Bezug zum Zeitpunkt des Personeninterviews in der letzten Welle noch andauerte, d.h. es sich um einen rechtszensierten Spell handelt. <u>Anmerkung:</u> Ein Spell gilt als zensiert, wenn eine der beiden folgenden Bedingungen erfüllt ist: Die Person gibt bei Frage P114 bezüglich Enddatum des Arbeitslosengeld-I-Bezug-Spells an, zum Zeitpunkt des Interviews noch Arbeitslosengeld I zu erhalten (P114 Ende=0). Oder es wird bei P114 ein Enddatum genannt, welches identisch mit dem Interviewdatum ist und in der Nachfrage P115 bestätigt, dass der Bezug derzeit noch andauert. Die Generierung basiert auf den plausibilitätsgeprüften, generierten Datumsvariablen.	<i>emonat, ejahr, AL1000; AL1100;</i> <i>AL1200 (al_spells)</i>

Tabelle 22: Einfache Generierungen für Welle 3 im Arbeitslosigkeitsspell Datensatz (*al_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung 2)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>zensiert</i>	<p><i>AL-Meldung: ggw. noch andauernder Spell (Rechtszensurierung)</i></p> <p>Der Zensurierungsindikator gibt an, ob ein Spell zum Zeitpunkt des Personeninterviews in der letzten Befragungswelle noch andauerte, d.h. es handelt sich um einen rechtszensierten Spell.</p> <p><u>Anmerkung:</u> Ein Spell gilt als zensiert, wenn eine der beiden folgenden Bedingungen erfüllt ist: Die Person gibt bei Frage P109 bezüglich Enddatum der Arbeitslosenmeldung an, zum Zeitpunkt des Interviews noch arbeitslos gemeldet zu sein (P109 Ende=0). Oder es wird bei P109 ein Enddatum genannt, welches identisch mit dem Interviewdatum ist und in der Nachfrage P110 bestätigt, dass die Arbeitslosenmeldung derzeit noch andauert. Im Lückenmodul berichtete zusätzliche Erwerbstätigkeitsepisoden bzw. Datumskorrekturen wurden vor der Generierung berücksichtigt.</p>	<p>AL0300; AL0400; AL0500 (<i>al_spells</i>)</p>

Tabelle 23: Einfache Generierungen für Welle 3 im Lückenspell Datensatz (*lu_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>bmonat</i>	<p><i>Episode: Anfangsmonat, generiert</i></p> <p>Beginnmonat der Nichterwerbstätigkeitsepisode in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden.</p> <p><u>Anmerkung:</u> Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben des Befragten sind in den Quellvariablen erhalten. Die Jahreszeitangaben wurden nach folgenden Regeln in Monatswerte recodiert: 21 Jahresanfang/Winter → Januar 24 Frühjahr/Ostern → April 27 Jahresmitte/Sommer → Juli 30 Herbst → Oktober 32 Jahresende → Dezember</p>	<p>LU0200 (<i>lu_spells</i>)</p>
<i>bjahr</i>	<p><i>Episode: Anfangsjahr, generiert</i></p> <p>Beginnjahr der Nichterwerbstätigkeitsepisode.</p> <p><u>Anmerkung:</u> siehe <i>bmonat</i></p>	<p>LU0300 (<i>lu_spells</i>)</p>

Tabelle 23: Einfache Generierungen für Welle 3 im Lückenspell Datensatz (*lu_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>emonat</i>	<p><i>Episode: Endmonat, generiert</i></p> <p>Endmonat der Nichterwerbstätigkeitsepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden und bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells das Interviewdatum eingesetzt wurde.</p>	<p>LU0400; LU0600 (<i>lu_spells</i>)</p> <p><i>pintjahr (PENDDAT)</i></p>
<i>ejahr</i>	<p><u>Anmerkung:</u> siehe <i>bmonat</i></p> <p><i>Episode: Endjahr, generiert</i></p> <p>Endjahr der Nichterwerbstätigkeitsepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden und bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells das Interviewdatum eingesetzt wurde.</p>	<p>LU0500; LU0600 (<i>lu_spells</i>)</p> <p><i>pintjahr (PENDDAT)</i></p>
<i>zensiert</i>	<p><u>Anmerkung:</u> siehe <i>bmonat</i></p> <p><i>Episode: ggw. noch andauernder Spell (Rechtszensierung)</i></p> <p>Der Zensierungsindikator gibt an, ob ein Spell zum Zeitpunkt des Personeninterviews in der letzten Befragungswelle noch andauerte, d.h. es handelt sich um einen rechtszensierten Spell.</p> <p><u>Anmerkung:</u> Ein Spell gilt als zensiert, wenn eine der beiden folgenden Bedingungen erfüllt ist: Die Person gibt bei Frage P130 bezüglich des Enddatum an, zum Zeitpunkt des Interviews den Status der Nicht-Erwerbstätigkeit noch zu haben (P130 Ende=0). Oder es wird bei P130 ein Enddatum genannt, welches identisch mit dem Interviewdatum ist und in der Nachfrage P131 bestätigt, dass der Status der Nicht-Erwerbstätigkeit noch andauert.</p>	<p>LU0400; LU0500; LU0600 (<i>lu_spells</i>)</p>

Tabelle 24: Einfache Generierungen für Welle 3 im Maßnahmespell Datensatz (*mn_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>bmonat</i>	<p><i>Maßnahme: Anfangsmonat, generiert</i></p> <p>Beginnmonat der Maßnahmeepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden.</p> <p><u>Anmerkung:</u> Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben (mit Ausnahme der in den Wertebereichsprüfungen als unplausibel identifizierten Werte) des Befragten sind in den Quellvariablen erhalten.</p> <p>Die Jahreszeitangaben wurden nach folgenden Regeln in Monatswerte recodiert: 21 Jahresanfang/Winter → Januar 24 Frühjahr/Ostern → April 27 Jahresmitte/Sommer → Juli 30 Herbst → Oktober 32 Jahresende → Dezember</p>	<i>MN0300 (mn_spells)</i>
<i>bjahr</i>	<p><i>Maßnahme: Anfangsjahr, generiert</i></p> <p>Beginnjahr der Maßnahmeepisode.</p>	<i>MN0400 (mn_spells)</i>
<i>emonat</i>	<p><u>Anmerkung:</u> siehe <i>bmonat</i></p> <p><i>Maßnahme: Endmonat, generiert</i></p> <p>Endmonat der Maßnahmeepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden und bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells das Interviewdatum eingesetzt wurde. Wurde statt eines Enddatums die Maßnahmedauer berichtet, so wurde aus Beginndatum und Dauer das Enddatum berechnet.</p>	<i>MN0300; MN0400; MN0500; MN0600; MN0700; MN1100; MN1200 (mn_spells)</i> <i>pintjahr (PENDDAT)</i>
<i>ejahr</i>	<p><u>Anmerkung:</u> siehe <i>bmonat</i></p> <p><i>Maßnahme: Endjahr, generiert</i></p> <p>Endjahr der Maßnahmeepisode, in der bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells das Interviewdatum eingesetzt wurde. Wurde statt eines Enddatums die Maßnahmedauer berichtet, so wurde aus Beginndatum und Dauer das Enddatum berechnet.</p>	<i>MN0300; MN0400; MN0500; MN0600; MN0800; MN1100; MN1300 (mn_spells)</i> <i>pintjahr (PENDDAT)</i>
<i>zensiert</i>	<p><u>Anmerkung:</u> siehe <i>bmonat</i></p> <p><i>Maßnahme: ggw. noch andauernder Spell (Rechtszensierung)</i></p> <p>Der Zensierungsindikator gibt an, ob ein Spell zum Zeitpunkt des Personeninterviews in der letzten Befragungswelle noch andauerte, d.h. es handelt sich um einen rechtszensierten Spell.</p> <p><u>Anmerkung:</u> Ein Spell gilt als zensiert, wenn die Person bei Frage P164 angibt, derzeit noch an einem Programm teilzunehmen. (P164=1)</p>	<i>MN0500 (mn_spells)</i>

Tabelle 25: Einfache Generierungen für Welle 3 im Personenregisterdatensatz (*p_register*) (alphabetische Reihenfolge)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>alter3</i>	<p><i>Alter der Person in Welle 3 (2008/2009)</i> Variable enthält die „beste“ verfügbare Altersinformation einer Person. Dabei handelt es sich entweder (a) um das aus dem Geburtsdatum, das in Welle 3 berichte wurde, berechnete Alter oder (b), wenn kein Geburtsdatum aus der 3. Welle vorliegt, um die Altersangabe aus dem Haushaltsinterview. Die Information aus <i>alter3</i> wurde auch in den Haushaltsdatensatz übernommen und entspricht den Angaben in <i>HD0200a</i> bis <i>HD0200o</i>. Dieses Vorgehen ist konsistent mit dem Vorgehen von Infratest. Bereits während der Feldzeit wurde die Altersvariable in der Datenbank mit den jeweils „besten“ Informationen gefüllt. Dort wird eine Variable in der Datenbank zunächst mit der Altersangabe laut Haushaltsinterview belegt. Wird ein Personeninterview geführt, wird diese Variable in der Datenbank mit dem Alter überschrieben, das auf Basis der Angaben des Personeninterviews berechnet wurde (Geburtsdatum, Personeninterviewdatum). Sowohl die im Haushaltsdatensatz als auch die im Personendatensatz ausgelieferten Altersangaben basieren auf dieser Variable der Datenbank. Die für Welle 3 im Haushaltsdatensatz enthaltene „beste“ Altersinformation wurde in den Plausibilitätsprüfungen und bei der Generierung der BG- und Haushaltstypen berücksichtigt.</p>	<p><i>p1, pintjahr, pintmon, pinttag</i> (PENDDAT) <i>HD0200a</i> bis <i>HD0200o</i> (HHENDDAT)</p>
<i>korrsex</i>	<p><i>Angabe zum Geschlecht wurde zwischen Befragungswellen korrigiert</i> Variable gibt für Personen, die in mehr als einer Welle Mitglied eines Befragungshaushalts waren an, ob im Rahmen des Haushaltsinterviews das Geschlecht korrigiert wurde.</p>	<p><i>HD0100a</i> bis <i>HD0100o</i> aller Wellen (HHENDDAT)</p>
<i>lastint</i>	<p><i>Befragungswelle des letzten Interviews auf Personenebene</i> Variable gibt an, in welcher Welle das letzte Interview auf Personenebene (Personeninterview oder Senioreninterview) mit der Person geführt wurde.</p>	<p>Personeninterviews aller Wellen (PENDDAT)</p>
<i>neuj3</i>	<p><i>Jahr, seit dem P. neu im aktuellen HH ist, berichtet in Welle 3 (2008/2009)</i> Variable zeigt an, seit welchem Jahr die Person neu zum dem Haushalt gehört, dessen Haushaltsmitglied sie in der 3. Welle ist.</p>	<p>Informationen zum Datum, seit dem eine Person neu im Haushalt ist. Erhoben im Haushaltsfragebogen für wiederbefragte Haushalte (HH18, HH37)</p>
<p><u>Anmerkung:</u> Information zum Datum stammt aus dem Welle 3-Interview mit dem wiederbefragten Haushalt, in den die Person seit der Vorwelle zugezogen ist bzw. in dem sie geboren wurde.</p>		

Tabelle 25: Einfache Generierungen für Welle 3 im Personenregisterdatensatz (*p_register*) (alphabetische Reihenfolge) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 3
<i>neum3</i>	<p><i>Monat, seit dem P. neu im aktuellen HH ist, berichtet in Welle 3 (2008/2009)</i></p> <p>Variable zeigt an, seit welchem Monat die Person neu zum dem Haushalt gehört, dessen Haushaltsmitglied sie in der 3. Welle ist.</p>	Informationen zum Datum, seit dem eine Person neu im Haushalt ist. Erhoben im Haushaltsfragebogen für wiederbefragte Haushalte (HH18, HH37)
<i>wegj3</i>	<p><u>Anmerkung:</u> siehe <i>neuj3</i></p> <p><i>Jahr, seit dem P. nicht mehr im vorherigen HH ist, berichtet in Welle 3 (2008/2009)</i></p> <p>Variable zeigt an, seit welchem Jahr die Person nicht mehr zum Haushalt der Vorwelle gehört.</p> <p><u>Anmerkung:</u> Information zum Datum stammt aus dem Welle 3-Interview mit dem Haushalt, in dem die Person in der Vorwelle gelebt hat.</p>	Informationen zum Datum, seit dem eine Person nicht mehr im Haushalt ist. Erhoben im Haushaltsfragebogen für wiederbefragte Haushalte (HH8, HH28)
<i>wegm3</i>	<p><i>Monat, seit dem P. nicht mehr im vorherigen HH ist, berichtet in Welle 3 (2008/2009)</i></p> <p>Variable zeigt an, seit welchem Monat die Person nicht mehr zum Haushalt der Vorwelle gehört.</p>	Informationen zum Datum, seit dem eine Person nicht mehr im Haushalt ist. Erhoben im Haushaltsfragebogen für wiederbefragte Haushalte (HH8, HH28)
<i>zmhh3</i>	<p><u>Anmerkung:</u> siehe <i>wegj3</i></p> <p><i>Zeiger: Personnummer der Mutter der Zielperson im HH in Welle 3 (2008/2009)</i></p> <p>Identifiziert die Personnummer der Mutter, sofern sie im Haushalt lebt. Als Mutter gelten: leibliche Mütter, Stiefmütter, Adoptivmütter oder Pflegemütter sowie Mütter mit nicht weiter spezifiziertem Status.</p>	Informationen zu Beziehungen im Haushalt der 3. Welle (Haushaltsgrid)
<i>zparth3</i>	<p><i>Zeiger: Personnummer des Partners der Zielperson im HH in Welle 3 (2008/2009)</i></p> <p>Identifiziert die Personnummer eines im Haushalt lebenden Partners. Als Partner gelten: Ehepartner, eingetragene Lebenspartner, unverheiratete Partner und Partner mit nicht weiter spezifiziertem Status.</p>	Informationen zu Beziehungen im Haushalt der Welle 3 (Haushaltsgrid)
<i>zupanel</i>	<p><i>Befragungswelle des Zugangs der Person zum Panel</i></p> <p>Variable gibt an, in welcher Welle die Person erstmals Mitglied eines Befragungshaushaltes war.</p>	Informationen zu den im Haushalt lebenden Personen aller Wellen (Haushaltsgrid)
<i>zvhh3</i>	<p><i>Zeiger: Personnummer des Vaters der Zielperson im HH in Welle 3 (2008/2009)</i></p> <p>Identifiziert die Personnummer des Vaters, sofern er im Haushalt lebt. Als Vater gelten: leibliche Väter, Stiefväter, Adoptivväter oder Pflegeväter sowie Väter mit nicht weiter spezifiziertem Status.</p>	Informationen zu Beziehungen im Haushalt der Welle 3 (Haushaltsgrid)

In den Datensätzen auf Personenebene sind eine Vielzahl generierter Variablen und Konstruktvariablen enthalten. Darunter auch Variablen (z.B. für die Stellung im Beruf), die in mehreren Datensätzen zu finden sind. Abbildung 3 gibt einen Überblick über die auf Personenebene enthaltenen einfachen und komplexen generierten Variablen.

Abbildung 3: Übersicht über generierte Variablen auf Personenebene in Welle 3

	PENDDAT				ET-Spells	AL-Spells	MN-Spells	
	Aktueller Status	Erwerbshistorie		Soziale Herkunft		Erwerbsbiografie	Arbeitslosigkeitsbiografie	Maßnahmeteilnahme
		letzte ET	erste ET	Mutter	Vater			
Bildung	berabj			mberuf1	vberuf1			
	beruf1			mberuf2	vberuf2			
	beruf2							
	schulabj			mschul1	vschul1			
	schul1			mschul2	vschul2			
Bildungs-klassifikationen	casmin			mcasmin	vcasmin			
	isced97			misced97	visced97			
	blzeit			mblzeit	vblzeit			
Informationen zum aktuellen Status	aktmassn							
	erwerb2							
Sozio-ökonomische Position	egp	egplewt	egpsewt	megp	vegp	egp		
	esec	eseclewt	esecsewt	mesec	vesec	esec		
	isei	iseilewt	iseisewt	misei	visei	isei		
	mpps	mpsewt	mpseewt	mmps	vmps	mpps		
	siops	siopsewt	siopseewt	msiops	vsiops	siops		
Stellung im Beruf	stib	stiblewt	stibeewt	mstib	vstib	stib		
	stibkz							
Datierung der Erwerbstätigkeit			begmeewt			bmonat		bmonat
			beoleewt			bjahr		bjahr
		emonlewt				emonat		emonat
Informationen zur Erwerbstätigkeit	arbeit	eihrlwt				ejahr		ejahr
	befrist							
Berufliche Tätigkeit	isco88	iscolewt	iscoeewt	misco	visco	isco88		
	isco88it	iscolewt_it	iscoeewt_it	misco_it	visco_it	isco88it		
	kldb_it	kldblewt	kldbeewt	mkldb	vkldb	kldb_it		
Tätig in Branche	branche					branche		mnbranche

Abbildung 3: Übersicht über generierte Variablen auf Personenebene in Welle 3 (Fortsetzung)

PENDDAT		ET-Spells		AL-Spells	MN-Spells			
	Aktueller Status	Erwerbshistorie		Soziale Herkunft		Erwerbsbiografie	Arbeitslosigkeitsbiografie	Maßnahmeteilnahme
		letzte ET	erste ET	Mutter	Vater			
Einkommen	brutto							
	bruttokat							
	netto							
	nettokat							
Leistungsbezug	alg1abez						alg1akt	
	alg1s05							
	hhalg2							
	hhalg2s05							
	hhalg2s06							
Haushaltskontext und Familienstand	hhgr							
	famstand							
	vhh							
	mhh							
	apartner							
	epartner							
	ekind							
	ekin614							
	ekin15							
	ekin18							
	ekin1517							
	kindzges							
kindzihh								
Migrationshintergrund	ogebland							
	ostaatan							
	ozulanda							
	ozulandb							
	ozulandc							
	ozulandd							
	ozulande							
	ozulandf							
migraton								
Informationen zur Person	gebhalb							
	palter							
	zpalthh							
	zpsex							
Gesundheit	pcs							
	mcs							
Allgemein	altbefr							
	fb_vers							
	panel							
	pintdat							
	RegP0100							
	sample							

4.5 Konstruktvariablen

Unter Konstruktvariablen fallen diejenigen Variablen, für deren Generierungen umfangreichere Umcodierungen und/oder Vercodungsarbeiten notwendig sind. In den meisten Fällen handelt es sich um bereits an anderer Stelle erprobte, theoretisch untermauerte und zumindest teilweise standardisierte sozial- oder wirtschaftswissenschaftliche Konstrukte. Beispiele hierfür sind die European Socio-economic Classification (ESeC), die International Standard Classification of Education (ISCED) oder das bedarfsgewichtete Haushaltseinkommen. Auf den folgenden Seiten werden die im PASS bereitgestellten Konstruktvariablen detailliert dargestellt, inkl. eines kurzen Überblicks über den theoretischen Hintergrund des Konstrukts und der wichtigsten Literaturverweise.

4.5.1 Personenebene

Bildung in Jahren

<u>Variablenname</u>	<i>bilzeit</i>																										
<u>Variablenlabel</u>	Schul- und Berufsbildung in Jahren, generiert																										
<u>Ausgangsvariablen</u>	<i>schul2; beruf2</i>																										
<u>Typ / Datensatz</u>	Bildung / Personendaten																										
<u>Bearbeitung</u>	Bernhard Christoph																										
<u>Erläuterung</u>	<p>Für viele Analyseverfahren ist die Verwendung einer linearen Bildungsvariable deutlich geeigneter als die einer kategorialen. Recht einfach lässt sich eine solche Umrechnung bei den Schulabschlüssen realisieren. Hier kann einfach die reale Dauer bis zum Bildungsabschluss verwendet werden. Dabei ist lediglich zu beachten, dass äquivalente Abschlüsse immer mit der gleichen Zeitdauer veranschlagt werden. So sollte z. B. eine Hochschulreife immer mit der gleichen Zeitdauer in die Variable eingehen, unabhängig davon, ob sie nach zwölf oder 13 Jahren erreicht wurde. Hier wurden die Schulabschlüsse wie folgt veranschlagt:</p> <table border="0" style="width: 100%;"> <tr> <td>Hauptschule; POS 8. Klasse; and. Abschluss:</td> <td style="text-align: right;">9 Jahre</td> </tr> <tr> <td>Mittlere Reife; POS 10. Klasse:</td> <td style="text-align: right;">10 Jahre</td> </tr> <tr> <td>Fachhochschulreife:</td> <td style="text-align: right;">12 Jahre</td> </tr> <tr> <td>Allg. oder fachgeb. Hochschulreife (incl. EOS)</td> <td style="text-align: right;">13 Jahre</td> </tr> </table> <p>Anders stellt sich die Situation bei den beruflichen Abschlüssen dar. Aufgrund der Unterschiedlichkeit der Bildungswege in der beruflichen Bildung und der mit diesen verbundenen – trotz vergleichbarer Dauer – teilweise recht stark divergierenden Einkommenserträge, verbietet sich hier eine einfache Umrechnung der Ausbildungsdauer. Dieses Problem kann dadurch umgangen werden, dass versucht wird, den mit einem bestimmten Berufsabschluss verknüpften Zuwachs an Humankapital zu operationalisieren (vgl. z. B. Helberger 1988).</p> <p>Ein ähnlicher Ansatz wurde in der vorliegenden Studie verfolgt. Dabei wurde jeweils nur der höchste Berufsabschluss der Befragungsperson berücksichtigt, und die für den jeweiligen Abschluss veranschlagten Zuwächse an Bildungsjahren wurden zur Dauer der Schulzeit hinzuaddiert.</p> <table border="0" style="width: 100%;"> <tr> <td>Anlernausbildung:</td> <td style="text-align: right;">+1 Jahr</td> </tr> <tr> <td>Lehre, Berufsfachschule, Schule des Gesundheitswesens:</td> <td style="text-align: right;">+1,5 Jahre</td> </tr> <tr> <td>Meister:</td> <td style="text-align: right;">+3 Jahre</td> </tr> <tr> <td>Berufsakademie:</td> <td style="text-align: right;">+3 Jahre</td> </tr> <tr> <td>Fachhochschule/Bachelor:</td> <td style="text-align: right;">+3 Jahre</td> </tr> <tr> <td>Uni/Master:</td> <td style="text-align: right;">+5 Jahre</td> </tr> <tr> <td>Promotion:</td> <td style="text-align: right;">+8 Jahre</td> </tr> <tr> <td>Anderer Abschluss, D:</td> <td style="text-align: right;">+1,5 Jahre</td> </tr> <tr> <td>Anderer Abschluss, Ausl.:</td> <td style="text-align: right;">+1,5 Jahre</td> </tr> </table>	Hauptschule; POS 8. Klasse; and. Abschluss:	9 Jahre	Mittlere Reife; POS 10. Klasse:	10 Jahre	Fachhochschulreife:	12 Jahre	Allg. oder fachgeb. Hochschulreife (incl. EOS)	13 Jahre	Anlernausbildung:	+1 Jahr	Lehre, Berufsfachschule, Schule des Gesundheitswesens:	+1,5 Jahre	Meister:	+3 Jahre	Berufsakademie:	+3 Jahre	Fachhochschule/Bachelor:	+3 Jahre	Uni/Master:	+5 Jahre	Promotion:	+8 Jahre	Anderer Abschluss, D:	+1,5 Jahre	Anderer Abschluss, Ausl.:	+1,5 Jahre
Hauptschule; POS 8. Klasse; and. Abschluss:	9 Jahre																										
Mittlere Reife; POS 10. Klasse:	10 Jahre																										
Fachhochschulreife:	12 Jahre																										
Allg. oder fachgeb. Hochschulreife (incl. EOS)	13 Jahre																										
Anlernausbildung:	+1 Jahr																										
Lehre, Berufsfachschule, Schule des Gesundheitswesens:	+1,5 Jahre																										
Meister:	+3 Jahre																										
Berufsakademie:	+3 Jahre																										
Fachhochschule/Bachelor:	+3 Jahre																										
Uni/Master:	+5 Jahre																										
Promotion:	+8 Jahre																										
Anderer Abschluss, D:	+1,5 Jahre																										
Anderer Abschluss, Ausl.:	+1,5 Jahre																										
<u>Literatur</u>	Helberger (1988)																										

Bildung in Jahren, Mutter

<u>Variablenname</u>	<i>mbilzeit</i>
<u>Variablenlabel</u>	Schul- und Berufsbildung in Jahren, generiert
<u>Ausgangsvariablen</u>	<i>mschul2; mberuf2</i>
<u>Typ / Datensatz</u>	Bildung / Personendaten
<u>Bearbeitung</u>	Bernhard Christoph
<u>Erläuterung</u>	Generelle Darstellung: vgl. Bildung in Jahren Bei den Bildungsabschlüssen der Eltern unterscheiden sich die für die Berufsbildung hinzuaddierten Jahre von denen beim Befragten, da die Berufsbildung der Eltern weniger detailliert erhoben wurde (insbes. im Bereich der tertiären Bildung). Die verwendeten Ausprägungen sind wie folgt: Anlernausbildung: +1 Jahr Lehre, Berufsfachschule, Schule des Gesundheitswesens: +1,5 Jahre Meister: +3 Jahre Berufsakademie: +3 Jahre Fachhochschule: +3 Jahre Universität: +5 Jahre Anderer Abschluss, D: +1,5 Jahre Anderer Abschluss, Ausl.: +1,5 Jahre
<u>Literatur</u>	Helberger (1988)

Bildung in Jahren, Vater

<u>Variablenname</u>	<i>vbilzeit</i>
<u>Variablenlabel</u>	Schul- und Berufsbildung in Jahren, generiert
<u>Ausgangsvariablen</u>	<i>vschul2; vberuf2</i>
<u>Typ / Datensatz</u>	Bildung / Personendaten
<u>Bearbeitung</u>	Bernhard Christoph
<u>Erläuterung</u>	Generelle Darstellung: vgl. Bildung in Jahren Bei den Bildungsabschlüssen der Eltern unterscheiden sich die für die Berufsbildung hinzuaddierten Jahre von denen beim Befragten, da die Berufsbildung der Eltern weniger detailliert erhoben wurde (insbes. im Bereich der tertiären Bildung). Die verwendeten Ausprägungen sind wie folgt: Anlernausbildung: +1 Jahr Lehre, Berufsfachschule, Schule des Gesundheitswesens: +1,5 Jahre Meister: +3 Jahre Berufsakademie: +3 Jahre Fachhochschule: +3 Jahre Universität: +5 Jahre Anderer Abschluss, D: +1,5 Jahre Anderer Abschluss, Ausl.: +1,5 Jahre
<u>Literatur</u>	Helberger (1988)

CASMIN

Variablenname

Variablenlabel

Ausgangsvariablen

Typ / Datensatz

Bearbeitung

Erläuterung

casmin

Bildung klassifiziert nach CASMIN, aktual. Version, generiert

schul2; beruf2

Bildung / Personendaten

Bernhard Christoph

Die CASMIN-Bildungsklassifikation wurde im Rahmen des CASMIN-Projekts (Comparative Analysis of Social Mobility in Industrial Nations) als international vergleichende Skala zur Erfassung von schulischen und beruflichen Bildungsabschlüssen entwickelt (König et al. 1987). Sie liegt mittlerweile in einer aktualisierten Fassung vor (Brauns & Steinmann 1999).

Die hier durchgeführte Umcodierung der Bildungsabschlüsse in CASMIN orientiert sich dabei – insbesondere, was die Behandlung problematischer Fälle angeht – an den in Lechert et al. (2006) sowie Granato (2000) beschriebenen Vorgehensweisen (selbstverständlich unter Berücksichtigung der leicht unterschiedlichen Erfassung der Bildungsvariablen im vorliegenden Datensatz). Die Details können der folgenden tabellarischen Darstellung entnommen werden. Zellen mit Kombinationen, die gültige CASMIN-Werte ergeben, sind hellgrau, solche mit definierten missing values dunkelgrau unterlegt.

schul beruf	nicht erhob.	Schüler	n. gest.	TNZ	KA	WN	ohne Abschl.	Sonder- schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. aus. Abschl.
nicht erhob.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-
unplaus. Wert	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Schüler	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
nicht gest.	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
TNZ	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
KA	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
WN	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
ohne Abschl.	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Anlern- ausbild.	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Lehre	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Berufs- fachsch.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Schul. d. Gesewes.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Meister	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
BA	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
FH/ Bachelor	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
Uni/ Master	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
Dissert.	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
And. dt. Abschl.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
And. aus Abschl.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

Literatur

Brauns et al. (1999); Granato (2000); König et al. (1987); Lechert et al. (2006)

MCASMIN

Variablenname

Variablenlabel

Ausgangsvariablen

Typ / Datensatz

Bearbeitung

Erläuterung

mcasmin

Bildung der Mutter, klassifiziert nach CASMIN, aktual. Version, generiert

mschul2; mberuf2

Bildung / Personendaten

Bernhard Christoph

Generelle Darstellung: vgl. CASMIN

Aufgrund der unterschiedlichen Erhebung der Bildungsvariablen für Befragte und deren Eltern weicht das Vercodungsschema von *mcasmin* und *vcasmin* leicht von dem bei *casmin* verwendeten ab. Zu Details vgl. die folgende Abbildung.

Schul- Beruf	nicht erhob.	Plnt fehlt	Elternt. unbek.	nicht gest.	TNZ	KA	WN	ohne Abschl.	Sonder- Schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. au. Abschl.
nicht erhob.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
unplaus. Wert	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Plnt fehlt	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Elternt. unbek.	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
nicht gest.	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
TNZ	-	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
KA	-	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
WN	-	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
ohne Abschl.	-	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Anlern- ausbild.	-	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Lehre	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Meister	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
BA	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
FH	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
Uni	-	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
And. dt. Abschl.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
And. aus Abschl.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

Literatur

Brauns et al. (1999); Granato (2000); König et al. (1987); Lechert et al. (2006)

VCASMIN

Variablenname

Variablenlabel

Ausgangsvariablen

Typ / Datensatz

Bearbeitung

Erläuterung

vcasmin

Bildung des Vaters klassifiziert nach CASMIN, aktual. Version, generiert

vschul2; vberuf2

Bildung / Personendaten

Bernhard Christoph

Generelle Darstellung: vgl. CASMIN

Aufgrund der unterschiedlichen Erhebung der Bildungsvariablen für Befragte und deren Eltern weicht das Vercodungsschema von *mcasmin* und *vcasmin* leicht von dem bei *casmin* verwendeten ab. Zu Details vgl. die folgende Abbildung.

Schul- Beruf	nicht erhob.	Plnt fehlt	Elternt. unbek.	nicht gest.	TNZ	KA	WN	ohne Abschl.	Sonder- Schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. au. Abschl.
nicht erhob.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
unplaus. Wert	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Plnt fehlt	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Elternt. unbek.	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
nicht gest.	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
TNZ	-	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
KA	-	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
WN	-	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
ohne Abschl.	-	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Anlern- ausbild.	-	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Lehre	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Meister	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
BA	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
FH	-	-	-	-	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a	3a
Uni	-	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
And. dt. Abschl.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
And. aus Abschl.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

Literatur

Brauns et al. (1999); Granato (2000); König et al. (1987); Lechert et al. (2006)

ISCED 97

Variablenname

Variablenlabel

Ausgangsvariablen

Typ / Datensatz

Bearbeitung

Erläuterung

iscsed97

Bildung klassifiziert nach isced97, aktual. Version, generiert

schul2; beruf2

Bildung / Personendaten

Bernhard Christoph

Als Alternative zu CASMIN bietet sich die Bildungsklassifikation ISCED-97 (International Standard Classification of Education) der OECD an (OECD 1999, für eine Kurzdarstellung vgl. auch BMBF 2003).

Bei der Vercodung der ISCED-97 ist zu beachten, dass die Klassifikation auch Ausprägungen enthält, die sich anhand der vorhandenen Daten nicht sinnvoll zuordnen lassen. Die ISCED Stufen 0 (Vorschulerziehung/Kindergarten) und 1 (Grundschule) sind aufgrund des Alters der Stichprobenpersonen (mind. 15 Jahre) nicht adäquat. Stattdessen wurde eine eigene Gruppe für Personen mit einer niedrigeren Bildung als ISCED Stufe 2 (ISCED 2=Haupt- bzw. Realschulabschluss) generiert. Aus diesem Grund umfasst die hier vorgenommene Codierung lediglich die ISCED Stufen 2-6.

Die Details der Zuordnung können der folgenden Tabelle entnommen werden. Zeilen mit Kombinationen, die gültige ISCED-Werte ergeben, sind hellgrau, solche mit definierten missing values dunkelgrau unterlegt.

Schul- Beruf	nicht erhob.	Schüler	nicht gest.	TNZ	KA	WN	ohne Abschl.	Sonder- schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. aus. Abschl.
nicht erhob.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-
unplaus. Wert	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Schüler	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
nicht gest.	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
TNZ	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
KA	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
WN	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
ohne Abschl.	-	-	-	-3	-2	-1	1	1	2	2	3a	3a	2	2
Anlern- ausbild.	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
Lehre	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Berufs- fachs.	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Schul. d. Ges.wes. Meister	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
BA	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
FH/ Bachelor	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Univ/ Master	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Dissert.	-	-	-	6	6	6	6	6	6	6	6	6	6	6
And. dt. Abschl.	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
And. aus Abschl.	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2

Literatur

BMBF (2003); OECD (1999)

MISCED 97

Variablenname

Variablenlabel

Ausgangsvariablen

Typ / Datensatz

Bearbeitung

Erläuterung

misced97

Bildung der Mutter klassifiziert nach isced97, aktual. Version, generiert

mschul2; mberuf2

Bildung / Personendaten

Bernhard Christoph

Zum theoretischen Hintergrund und zur Generierung vgl. ISCED-97.

Im Gegensatz zur Umsetzung von ISCED-97 bei den Befragten selbst lässt sich ISCED Stufe 6 für die Eltern nicht generieren, da die entsprechenden Abschlüsse (Promotion oder Habilitation) nicht separat erfragt wurden. Aus diesem Grund umfasst die hier vorgenommene Codierung lediglich die ISCED Stufen 2-5. Zu Details der Vercodung vgl. die folgende Abbildung.

Schulberuf	nicht erhob.	PInt fehlt	Elternt. unbek.	nicht gest.	TNZ	KA	WN	ohne Abschl.	Sonder-Schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. au. Abschl.
nicht erhob.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
umplaus. Wert	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
PInt fehlt	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Elternt. unbek.	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
nicht gest.	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
TNZ	-	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
KA	-	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
WN	-	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
ohne Abschl.	-	-	-	-	-3	-2	-1	1	1	2	2	3a	3a	2	2
Anlernausbild.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
Lehre	-	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Meister	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
BA	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
FH	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Uni	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
And. dt. Abschl.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
And. aus Abschl.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2

Literatur

BMBF (2003); OECD (1999)

VISCED 97

Variablenname

Variablenlabel

Ausgangsvariablen

Typ / Datensatz

Bearbeitung

Erläuterung

visced97

Bildung des Vaters klassifiziert nach isced97, aktual. Version, generiert

vschul2; vberuf2

Bildung / Personendaten

Bernhard Christoph

Zum theoretischen Hintergrund und zur Generierung vgl. ISCED-97.

Im Gegensatz zur Umsetzung von ISCED-97 bei den Befragten selbst lässt sich ISCED Stufe 6 für die Eltern nicht generieren, da die entsprechenden Abschlüsse (Promotion oder Habilitation) nicht separat erfragt wurden. Aus diesem Grund umfasst die hier vorgenommene Codierung lediglich die ISCED Stufen 2-5. Zu Details der Vercodung vgl. die folgende Abbildung.

Schulberuf	nicht erhob.	PInt fehlt	Elternt. unbek.	nicht gest.	TNZ	KA	WN	ohne Abschl.	Sonder-Schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. au. Abschl.
nicht erhob.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
umplaus. Wert	-	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
PInt fehlt	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Elternt. unbek.	-	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
nicht gest.	-	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
TNZ	-	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
KA	-	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
WN	-	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
ohne Abschl.	-	-	-	-	-3	-2	-1	1	1	2	2	3a	3a	2	2
Anlernausbild.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
Lehre	-	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Meister	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
BA	-	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
FH	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
Uni	-	-	-	-	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a	5a
And. dt. Abschl.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
And. aus Abschl.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2

Literatur

BMBF (2003); OECD (1999)

International Standard Classification of Occupations 1988 (ISCO88); ZUMA-Vercodung

<u>Gebildet</u>	<u>Erwerbstätigkeit</u>	<u>Variablenname</u>	<u>Ausgangsvariablen</u>
	aktuelle	<i>isco88</i>	W1: P46; ab W2: P40_X
	Spelldaten (<i>et_spells</i>)	<i>isco88</i>	ab W2: P40_X
	erste	<i>iscoeewt</i>	ab W2: P40_X, P91, P100
	letzte	<i>iscolewt</i>	ab W2: P40_X, P91
	des Vaters	<i>visco</i>	ab W2: P299
	der Mutter	<i>misco</i>	ab W2: P288
<u>Variablenlabel</u>	aktuelle Ewt.: ISCO 88 (ZUMA-Vercodung), generiert		
	Spelldaten (<i>et_spells</i>): ISCO 88 (ZUMA-Vercodung), generiert		
	erste Ewt.: ISCO 88 (ZUMA-Vercodung), erste Erwerbstätigkeit, generiert		
	letzte Ewt.: ISCO 88 (ZUMA-Vercodung), letzte Erwerbstätigkeit, generiert		
	Vater: ISCO 88 (ZUMA-Vercodung) des Vaters, generiert		
	Mutter: ISCO 88 (ZUMA-Vercodung) der Mutter, generiert		
<u>Typ / Datensatz</u>	Beruf / Personendaten		
<u>Ansprechpartner</u>	Bernhard Christoph		
<u>Erläuterung</u>	<p>Die International Standard Classification of Occupations (ISCO) ist eine von der ILO entwickelte, international vergleichbare Berufsklassifikation. Besonderheit der ISCO-88 ist, speziell im Vergleich zur hier ebenfalls zur Verfügung gestellten Klassifizierung der Berufe des statistischen Bundesamts, dass neben der beruflichen Tätigkeit auch das zur Ausübung dieser Tätigkeit i.d.R. notwendige Bildungsniveau bei der Zuordnung zu einem bestimmten Berufscode berücksichtigt wird.</p> <p>Die Vercodung erfolgte durch das Leibniz-Institut für Sozialwissenschaften (GESIS, vormals ZUMA). Diese Vercodung von ISCO-88 stellt im Gegensatz zur Infratest-Vercodung eine originäre Vercodung von ISCO-88 dar. Sie bildet die Grundlage für die Generierung der ISCO-basierten Status- und Prestigemaße.</p>		
<u>Literatur</u>	ILO (1990)		

International Standard Classification of Occupations 1988 (ISCO88); Infratest Vercodung

<u>Gebildet</u>	<u>Erwerbstätigkeit</u>	<u>Variablenname</u>	<u>Ausgangsvariablen</u>
	aktuelle	<i>isco88it</i>	W1: P46; ab W2: P40_X
	Spelldaten (<i>et_spells</i>)	<i>isco88it</i>	ab W2: P40_X
	erste	<i>iscoeewt_it</i>	ab W2: P40_X, P91, P100
	letzte	<i>iscolewt_it</i>	ab W2: P40_X, P91
	des Vaters	<i>visco_it</i>	ab W2: P299
	der Mutter	<i>misco_it</i>	ab W2: P288
<u>Variablenlabel</u>	aktuelle Ewt.: ISCO 88 (Infratest -Vercodung), generiert		
	Spelldaten (<i>et_spells</i>): ISCO 88 (Infratest-Vercodung), generiert		
	erste Ewt.: ISCO 88 (Infratest -Vercodung), erste Erwerbstätigkeit, generiert		
	letzte Ewt.: ISCO 88 (Infratest -Vercodung), letzte Erwerbstätigkeit, generiert		
	Vater: ISCO 88 (Infratest -Vercodung) des Vaters, generiert		
	Mutter: ISCO 88 (Infratest -Vercodung) der Mutter, generiert		
<u>Typ / Datensatz</u>	Beruf / Personendaten		
<u>Ansprechpartner</u>	Bernhard Christoph		
<u>Erläuterung</u>	<p>Die ISCO ist eine von der ILO entwickelte, international vergleichbare Berufsklassifikation. Besonderheit der ISCO-88 ist, insbesondere im Vergleich zur hier ebenfalls zur Verfügung gestellten Klassifizierung der Berufe des statistischen Bundesamts, dass neben der beruflichen Tätigkeit auch das zur Ausübung dieser Tätigkeit i.d.R. notwendige Bildungsniveau bei der Zuordnung zu einem bestimmten Berufscode berücksichtigt wird.</p> <p>Die Vercodung erfolgte durch Infratest, das Feldinstitut des PASS für die Wellen 1-3, als Umschlüsselung aus der Klassifizierung der Berufe des Statistischen Bundesamts.</p>		
<u>Literatur</u>	ILO (1990)		

Klassifizierung der Berufe 1992 (KIdB92); Infratest Vercodung

<u>Gebildet</u>	<u>Erwerbstätigkeit</u>	<u>Variablenname</u>	<u>Ausgangsvariablen</u>
	aktuelle	<i>kldb_it</i>	W1: P46; ab W2: P40_X
	Spelldaten (<i>et_spells</i>)	<i>kldb_it</i>	ab W2: P40_X
	erste	<i>kldbeewt</i>	ab W2: P40_X, P91, P100
	letzte	<i>kldblewt</i>	ab W2: P40_X, P91
	des Vaters	<i>vkldb</i>	ab W2: P299
	der Mutter	<i>mkldb</i>	ab W2: P288
<u>Variablenlabel</u>	aktuelle Ewt.: KIdB 92 (Infratest -Vercodung), generiert		
	Spelldaten (<i>et_spells</i>): KIdB 92 (Infratest-Vercodung), generiert		
	erste Ewt.: KIdB 92 (Infratest -Vercodung), erste Erwerbstätigkeit, generiert		
	letzte Ewt.: KIdB 92 (Infratest -Vercodung), letzte Erwerbstätigkeit, generiert		
	Vater: KIdB 92 (Infratest -Vercodung) des Vaters, generiert		
	Mutter: KIdB 92 (Infratest -Vercodung) der Mutter, generiert		
<u>Typ / Datensatz</u>	Beruf / Personendaten		
<u>Ansprechpartner</u>	Bernhard Christoph		
<u>Erläuterung</u>	Die KIdB92 ist die aktuelle Version der Klassifizierung der Berufe in der Variante des Statistischen Bundesamts. Sie ist eine spezifisch auf Deutschland bezogene Berufsklassifikation. Die Klassifikation basiert ausschließlich auf der beruflichen Tätigkeit.		
	Die Vercodung erfolgte durch Infratest, das Feldinstitut des PASS für die Wellen 1-3.		
<u>Literatur</u>	StBA (1992).		

Klassenschema nach Erikson, Goldthorpe und Portocarrero (EGP)

<u>Gebildet</u>	<u>Erwerbstätigkeit</u>	<u>Variablenname</u>	<u>Ausgangsvariablen</u>
	aktuelle	<i>egp</i>	<i>isco88, stib</i>
	Spelldaten (<i>et_spells</i>)	<i>egp</i>	<i>isco88, stib</i>
	erste	<i>egpeewt</i>	<i>iscoeewt, stibeewt</i>
	letzte	<i>egplewt</i>	<i>iscolewt, stiblewt</i>
	des Vaters	<i>vegp</i>	<i>visco, vstib</i>
	der Mutter	<i>megp</i>	<i>misco, mstib</i>
<u>Variablenlabel</u>	<p>aktuelle Ewt.: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), akt. Beruf, gen.</p> <p>Spelldaten (<i>et_spells</i>): Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), gen.</p> <p>erste Ewt.: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), erste Erwerbst., gen.</p> <p>letzte Ewt.: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), letzte Erwerbst., gen.</p> <p>Vater: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), Beruf d. Vaters, gen.</p> <p>Mutter: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), Beruf d. Mutter, gen.</p>		
<u>Typ / Datensatz</u>	sozioökonomische Position / Personendaten		
<u>Bearbeitung</u>	Bernhard Christoph		
<u>Erläuterung</u>	<p>Das Klassenschema von Erikson, Goldthorpe und Portocarero (Erikson et al. 1979, 1982; Erikson & Goldthorpe 1992) ist eines der gebräuchlichsten Instrumente zur Operationalisierung der Klassenposition.</p> <p>Die Vercodung erfolgte hier ausschließlich auf Basis der ISCO-88-Berufsklassifikation und der Stellung im Beruf. Sie basiert auf einem früher erarbeiteten Vercodungsansatz von Christoph et al. (2005), wo sich auch eine detaillierte Darstellung des Verfahrens findet. Im Unterschied zum dort beschriebenen Vorgehen wurden hier allerdings – analog zur im Folgenden beschriebenen Vercodung der European Socio-Economic Classification (ESeC) – mithelfende Familienangehörige nicht als Selbständige, sondern als abhängig Erwerbstätige vercodet.</p> <p>Ein Unterschied zwischen den hier durchgeführten Codierungen von EGP und ESeC besteht darin, dass bei EGP solche Fälle auf „Missing“ (-7) gesetzt wurden, bei denen die berufliche Tätigkeit und die Stellung im Beruf nicht miteinander vereinbar schien (z. B. „Direktoren und Hauptgeschäftsführer“ [ISCO=1210], die angaben, „Angestellte mit einfacher Tätigkeit“ [StiB=51] zu sein). Im Falle von ESeC wurde aus Gründen der Kompatibilität mit der von uns übernommenen, stark standardisierten Vercodung dieses Instruments auf eine analoge Überprüfung verzichtet.</p>		
<u>Literatur</u>	Christoph et al. (2005); Erikson & Goldthorpe (1992); Erikson et al. (1982); Erikson et al. (1979):		

European Socio-economic Classification (ESeC)

<u>Gebildet</u>	<u>Erwerbstätigkeit</u>	<u>Variablenname</u>	<u>Ausgangsvariablen</u>
	aktuelle	<i>esec</i>	<i>isco88, stib, PET2000, PET2700</i>
	Spelldaten (<i>et_spells</i>)	<i>esec</i>	<i>isco88, stib, ET1100, ET1300</i>
	erste	<i>eseceewt</i>	<i>iscoeewt, stibeewt, PET1261,</i>
	letzte	<i>eseclewt</i>	<i>iscolewt, stiblewt, PET3801</i>
	des Vaters	<i>vesec</i>	<i>visco, vstib, PSH0670</i>
	der Mutter	<i>mesec</i>	<i>misco, mstib, PSH0370</i>
<u>Variablenlabel</u>	aktuelle Ewt.: European Socio-economic Classification (ESeC), akt. Beruf, gen. Spelldaten (<i>et_spells</i>): European Socio-economic Classification (ESeC), gen. erste Ewt.: European Socio-economic Classification (ESeC), erste Erwerbst., gen. letzte Ewt.: European Socio-economic Classification (ESeC), letzte Erwerbst., gen. Vater: European Socio-economic Classification (ESeC), Beruf d. Vaters, gen. Mutter: European Socio-economic Classification (ESeC), Beruf d. Mutter, gen.		
<u>Typ / Datensatz</u>	sozioökonomische Position / Personendaten		
<u>Bearbeitung</u>	Bernhard Christoph		
<u>Erläuterung</u>	Die European Socio-economic Classification orientiert sich von ihrer theoretischen Konzeption her stark am EGP-Klassenschema. Im Unterschied zu Letzterem wurde bei ESeC jedoch hoher Wert auf eine international vergleichbare Operationalisierung und eine ausführliche Validierung der Klassifikation gelegt (vgl. für eine generelle Darstellung: Rose & Harrison 2007; für Deutschland Müller et al. 2006, 2007). Der Stata do-file für die Generierung von ESeC wurde uns dankenswerterweise von Heike Wirth von GESIS-ZUMA zur Verfügung gestellt (Fischer & Wirth 2007) und von uns lediglich für die vorliegende Studie angepasst. Dieser do-file stellt eine Stata-Umsetzung der von Harrison & Rose (2006) geschriebenen Standard-SPSS-Syntax zur Generierung von ESeC dar.		
<u>Literatur</u>	Fischer & Wirth (2007); Harrison & Rose (2006); Müller et al. (2006, 2007); Rose & Harrison (2007)		

Magnitude-Prestigeskala (MPS)

<u>Gebildet</u>	<u>Erwerbstätigkeit</u>	<u>Variablenname</u>	<u>Ausgangsvariablen</u>
	aktuelle	<i>mps</i>	<i>isco88</i>
	Spelldaten (<i>et_spells</i>)	<i>mps</i>	<i>isco88</i>
	erste	<i>mpseewt</i>	<i>iscoeewt</i>
	letzte	<i>mpslewt</i>	<i>iscolewt</i>
	des Vaters	<i>vmps</i>	<i>visco</i>
	der Mutter	<i>mmmps</i>	<i>misco</i>
<u>Variablenlabel</u>	aktuelle Ewt.:Magnitude-Prestigeskala, akt. Beruf, gen. Spelldaten (<i>et_spells</i>): Magnitude-Prestigeskala, generiert erste Ewt.: Magnitude-Prestigeskala, erste Erwerbst., gen. letzte Ewt.: Magnitude-Prestigeskala, letzte Erwerbst., gen. Vater: Magnitude-Prestigeskala, Beruf d. Vaters, gen. Mutter: Magnitude-Prestigeskala, Beruf d. Mutter, gen.		
<u>Typ / Datensatz</u>	sozioökonomische Position / Personendaten		
<u>Ansprechpartner</u>	Bernhard Christoph		
<u>Erläuterung</u>	Die Magnitude-Prestigeskala [MPS] (Wegener 1985, 1988) ist bis heute die einzige spezifisch deutsche Prestigeskala auf Basis detaillierter Berufsangaben. Sie wurde ursprünglich für die ältere 1968er Version der International Standard Classification of Occupations (ISCO-68) entwickelt. Da die Berufsvercodung in der vorliegenden Studie lediglich anhand der aktuelleren ISCO-88-Klassifikation sowie der KIdB 1992 des Statistischen Bundesamts vorgenommen wurde, fand hier die auf ISCO-88 transferierte Variante der Skala (Christoph 2005) Verwendung. Die Zuspiegelung erfolgte im Rahmen der Berufsvercodung durch das Zentrum für Umfragen, Methoden und Analysen (GESIS-ZUMA).		
<u>Literatur</u>	Christoph (2005); Wegener (1985, 1988)		

Standard International Occupational Prestige Scale (SIOPS/Treiman-Skala)

<u>Gebildet</u>	<u>Erwerbstätigkeit</u>	<u>Variablenname</u>	<u>Ausgangsvariablen</u>
	aktuelle	<i>siops</i>	<i>isco88</i>
	Spelldaten (<i>et_spells</i>)	<i>siops</i>	<i>isco88</i>
	erste	<i>siopseewt</i>	<i>iscoeewt</i>
	letzte	<i>siopslewt</i>	<i>iscolewt</i>
	des Vaters	<i>vslops</i>	<i>visco</i>
	der Mutter	<i>msiops</i>	<i>misco</i>
<u>Variablenlabel</u>	aktuelle Ewt.: Standard International Occupational Prestige Scale, akt. Beruf, gen. Spelldaten (<i>et_spells</i>): Standard International Occupational Prestige Scale, generiert erste Ewt.: Standard International Occupational Prestige Scale, erste Erwerbst., gen. letzte Ewt.: Standard International Occupational Prestige Scale, letzte Erwerbst., gen. Vater: Standard International Occupational Prestige Scale, Beruf d. Vaters, gen. Mutter: Standard International Occupational Prestige Scale, Beruf d. Mutter, gen.		
<u>Typ / Datensatz</u>	sozioökonomische Position / Personendaten		
<u>Ansprechpartner</u>	Bernhard Christoph		
<u>Erläuterung</u>	Die ursprünglich von Treiman (1977) für ISCO-68 entwickelte Treiman-Prestigeskala ist die erste und bis heute einzige international vergleichende Berufsprestigeskala. Seit der Umsetzung der Skala auf ISCO-88 (Ganzeboom & Treiman 1996, 2003) bürgerte sich die dort von den Autoren verwendete Bezeichnung als „Standard International Occupational Prestige Scale“ ein. Die Zuspiegelung erfolgte im Rahmen der Berufsvercodung durch das Zentrum für Umfragen, Methoden und Analysen (GESIS-ZUMA).		
<u>Literatur</u>	Ganzeboom & Treiman (1996, 2003); Treiman (1977)		

International Socio-Economic Index (ISEI)

<u>Gebildet</u>	<u>Erwerbstätigkeit</u>	<u>Variablenname</u>	Ausgangsvariablen
	aktuelle	<i>isei</i>	<i>isco88</i>
	Spelldaten (<i>et_spells</i>)	<i>isei</i>	<i>isco88</i>
	erste	<i>iseieewt</i>	<i>iscoeewt</i>
	letzte	<i>iseilewt</i>	<i>iscolewt</i>
	des Vaters	<i>visei</i>	<i>visco</i>
	der Mutter	<i>misei</i>	<i>misco</i>
<u>Variablenlabel</u>	aktuelle Ewt.: International Socio-Economic Index, akt. Beruf, gen. Spelldaten (<i>et_spells</i>): International Socio-Economic Index, generiert erste Ewt.: International Socio-Economic Index, erste Erwerbst., gen. letzte Ewt.: International Socio-Economic Index, letzte Erwerbst., gen. Vater: International Socio-Economic Index, Beruf d. Vaters, gen. Mutter: International Socio-Economic Index, Beruf d. Mutter, gen.		
<u>Typ / Datensatz</u>	sozioökonomische Position / Personendaten		
<u>Ansprechpartner</u>	Bernhard Christoph		
<u>Erläuterung</u>	<p>Der International Socio-Economic Index ist sicherlich einer der gebräuchlichsten Indizes seiner Art. Dies verdankt er nicht zuletzt der Tatsache, dass ihm, im Gegensatz zu den meisten SEIs, ein eigenständiges theoretisches Konzept zugrunde liegt, in dem der Beruf bzw. dessen sozioökonomischer Status als „intervening variable“ zwischen Bildung und Einkommen begriffen wird.</p> <p>Der ISEI wurde zunächst für ISCO-68 entwickelt (Ganzeboom et al. 1992), später jedoch auch für ISCO-88 umgesetzt (Ganzeboom & Treiman 1996, 2003).</p> <p>Die Zuspiegelung erfolgte im Rahmen der Berufsvercodung durch das Zentrum für Umfragen, Methoden und Analysen (GESIS-ZUMA).</p>		
<u>Literatur</u>	Ganzeboom et al. (1992); Ganzeboom & Treiman (1996, 2003)		

Klassifikation der Wirtschaftszweige 2003 (WZ2003)

<u>Gebildet</u>	<u>Erwerbstätigkeit</u>	<u>Variablenname</u>	Ausgangsvariablen
	aktuelle	<i>branche</i>	P61_X
	Spelldaten (<i>et_spells</i>)	<i>branche</i>	P61_X
	Spelldaten (<i>mn_spells</i>)	<i>mnbranche</i>	P176_X
<u>Variablenlabel</u>	aktuelle Ewt.: Jetzige Tätigkeit: Wirtschaftszweig/Branche (WZ2003) Spelldaten (<i>et_spells</i>): Wirtschaftszweig/Branche (WZ 2003), generiert Spelldaten (<i>mn_spells</i>): Maßnahme: Wirtschaftszweig/Branche (WZ 2003)		
<u>Typ / Datensatz</u>	sozioökonomische Position / Personendaten		
<u>Ansprechpartner</u>	Bernhard Christoph		
<u>Erläuterung</u>	<p>Die Vercodung der offenen Branchenangabe erfolgte anhand des 2-Stellers der Klassifikation der Wirtschaftszweige des statistischen Bundesamts (WZ2003). Auf der Ebene des Zweistellers entspricht diese Klassifikation weitgehend der europäischen Nomenclature générale des Activités économiques dans les Communautés Européennes (NACE) in der Revision 1.1.</p> <p>Die Vercodung erfolgte durch Infratest, das Feldinstitut des PASS für die Wellen 1-3.</p>		
<u>Literatur</u>	StaBA (2002); EG (2002)		

Körperliche Skala des SF12v2 (SOEP-Version, NBS)

<u>Variablenname</u>	<i>pcs</i>
<u>Variablenlabel</u>	körperliche Skala des SF12v2 (SOEP-Version, NBS), generiert
<u>Ausgangsvariablen</u>	<i>PG1200; PG1205; PG1210; PG1215*</i>
<u>Typ / Datensatz</u>	Gesundheit / Personendaten
<u>Bearbeitung</u>	Christian Dickmann
<u>Erläuterung</u>	<p>Der SF12-Fragebogen ist ein aus dem SF36 hervorgegangener Kurzfragebogen zur Ermittlung der gesundheitsbezogenen Lebensqualität. Seit 2002 werden die international anerkannten und angewendeten SF12-Indikatoren (Version 2 – SF12v2) im SOEP erhoben. Die SOEP-Version weicht allerdings in Formulierung, Reihenfolge der Fragen und Layout teilweise vom Original-SF12v2 ab. Für PASS wurden die SF12-Indikatoren analog zum SOEP erhoben.</p> <p>Die Generierung von pcs in PASS basiert auf der bei Nübling et al. (2006); wiedergegebenen SPSS-Syntax.</p>
<u>Literatur</u>	Nübling et al. (2006); Andersen et al. (2007)

Psychische Skala des SF12v2 (SOEP-Version, NBS)

<u>Variablenname</u>	<i>mcs</i>
<u>Variablenlabel</u>	psychische Skala des SF12v2 (SOEP-Version, NBS), generiert
<u>Ausgangsvariablen</u>	<i>PG1200; PG1205; PG1210; PG1215*</i>
<u>Typ / Datensatz</u>	Gesundheit / Personendaten
<u>Bearbeitung</u>	Christian Dickmann
<u>Erläuterung</u>	<p>Der SF12-Fragebogen ist ein aus dem SF36 hervorgegangener Kurzfragebogen zur Ermittlung der gesundheitsbezogenen Lebensqualität. Seit 2002 werden die international anerkannten und angewendeten SF12-Indikatoren (Version 2 – SF12v2) im SOEP erhoben. Die SOEP-Version weicht allerdings in Formulierung, Reihenfolge der Fragen und Layout teilweise vom Original-SF12v2 ab. Für PASS wurden die SF12-Indikatoren analog zum SOEP erhoben.</p> <p>Die Generierung von mcs in PASS basiert auf der bei Nübling et al. (2006); wiedergegebenen SPSS-Syntax.</p>
<u>Literatur</u>	Nübling et al. (2006); Andersen et al. (2007)

4.5.2 Haushalts- oder Bedarfsgemeinschaftsebene

Bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung

<u>Variablenname</u>	<i>oecdinca</i>
<u>Variablenlabel</u>	bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung (gerundet)
<u>Ausgangsvariablen</u>	<i>HD0200a-HD0200o; HA0100; hhincome</i>
<u>Typ / Datensatz</u>	sozioökonomische Position / Haushaltsdaten
<u>Bearbeitung</u>	Bernhard Christoph
<u>Erläuterung</u>	<p>Mittels eines bedarfsgewichteten Haushaltseinkommens (auch Äquivalenzeinkommen genannt) wird versucht, die in Mehrpersonenhaushalten durch gemeinsames Wirtschaften im Vergleich zu Einpersonenhaushalten erreichbaren Einsparungen zu berücksichtigen. Dies geschieht dadurch, dass in Mehrpersonenhaushalten bei der Berechnung des Pro-Kopf-Einkommens nicht die reale Haushaltsgröße als Divisor verwendet wird, sondern eine über den unterstellten Bedarf der Personen berechnete, in der Regel niedrigere Zahl (die bedarfsgewichtete Haushaltsgröße).</p> <p>Bei der alten OECD-Skala wird dabei nur für die erste (mind. 15-jährige) Person des Haushalts ein Bedarfsgewicht von 1,0 unterstellt. Alle weiteren Personen ab 15 erhalten ein Bedarfsgewicht von 0,7; Personen bis einschl. 14 gehen mit einem Gewicht von 0,5 in die bedarfsgewichtete Haushaltsgröße ein.</p> <p>Zur alten OECD-Skala vgl. OECD (1982) für einen Überblick zum Thema vgl. Hauser (1996).</p>
<u>Literatur</u>	Hauser (1996); OECD (1982)

Bedarfsgewichtetes Haushaltseinkommen, neue OECD-Gewichtung

<u>Variablenname</u>	<i>oecdincn</i>
<u>Variablenlabel</u>	bedarfsgewichtetes Haushaltseinkommen, neue OECD-Gewichtung (gerundet)
<u>Ausgangsvariablen</u>	<i>HD0200a-HD0200o; HA0100; hhincome</i>
<u>Typ / Datensatz</u>	sozioökonomische Position / Haushaltsdaten
<u>Bearbeitung</u>	Bernhard Christoph
<u>Erläuterung</u>	<p><u>Generelle Darstellung:</u> vgl. Bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung.</p> <p>Bei der neuen OECD-Skala wird nur für die erste (mind. 15-jährige) Person des Haushalts ein Bedarfsgewicht von 1,0 unterstellt. Alle weiteren Personen ab 15 erhalten ein Bedarfsgewicht von 0,5; Personen bis einschl. 14 gehen mit einem Gewicht von 0,3 in die bedarfsgewichtete Haushaltsgröße ein.</p> <p>Zur neuen OECD-Skala vgl. Hagenaars et al. (1994).</p>
<u>Literatur</u>	Hagenaars et al. (1994)

Deprivationsindex, ungewichtet

<u>Variablenname</u>	<i>depindug</i>
<u>Variablenlabel</u>	Deprivationsindex, ungew. (n. aus fin. Gründ. fehl. Items; ungew. Itemssumme: 26)
<u>Ausgangsvariablen</u>	<i>HLS0100a-HLS2600a; HLS0100b-HLS2600b</i>
<u>Typ / Datensatz</u>	materielle Situation / Haushaltsdaten
<u>Bearbeitung</u>	Bernhard Christoph
<u>Erläuterung</u>	<p>In der Armutsforschung wird im Anschluss an Ringen (1988) i.d.R. zwischen direkter und indirekter Messung von Armut unterschieden. Bei der indirekten Messung konzentriert man sich dabei auf die zur Erreichung eines bestimmten Lebensstandards zur Verfügung stehenden Ressourcen, insbesondere das (bedarfsgewichtete Haushalts-) Einkommen. Aus diesem Grund spricht man hier auch vom Ressourcenansatz.</p> <p>Die direkte Messung versucht hingegen, die tatsächlich vorhandene Güterausstattung der Haushalte zu erfassen bzw. festzustellen, inwieweit die Haushalte aus finanziellen Gründen nicht in der Lage sind, sich bestimmte, als relevant erachtete Dinge zu leisten. Hier spricht man auch vom Deprivationsansatz (vgl.: hierzu z. B. Halleröd 1995).</p> <p>Als genereller Tenor der Forschung ist festzuhalten, dass der Ressourcen- und der Deprivationsansatz zum Teil unterschiedliche Personengruppen erfassen bzw. als arm klassifizieren. Für eine Abgrenzung der Armenpopulation im engeren Sinne wird daher auch oft vorgeschlagen, Maße für Einkommensarmut und Deprivation zu kombinieren und nur diejenigen, die von beiden Ansätzen als arm klassifiziert werden, als Armutspopulation im engeren Sinne zu betrachten (vgl. Halleröd 1995; Nolan & Whelan 1996; Andreß und Lipsmeier 2001).</p> <p>Dem Index liegt eine Liste von 26 Gütern oder Aktivitäten zugrunde, für die jeweils erhoben wird, ob der Haushalt der Befragten diese Güter hat bzw. an diesen Aktivitäten partizipiert. Der hier berechnete ungewichtete Index summiert lediglich die Anzahl der Items, bei denen dies nicht der Fall ist. Allerdings wird das Fehlen eines Items nur dann berücksichtigt, wenn es auf finanzielle Gründe zurückzuführen ist. Dies geschieht, um auszuschließen, dass bestimmte Konsumpräferenzen (z.B. der bewusste Verzicht auf ein Auto oder einen Fernseher) als Minderung des Lebensstandards fehlinterpretiert werden.</p> <p>Zudem gilt, dass ein Fehlen aus finanziellen Gründen nur dann angenommen wurde, wenn beide Fragen explizit in diesem Sinne beantwortet wurden. WN- bzw. KA-Angaben wurden also gewertet, als ob das entsprechende Gut vorhanden ist bzw. als ob es aus anderen als aus finanziellen Gründen fehlt. Diese Annahme ist sicherlich nicht in jedem Fall korrekt. Alternativ wäre es möglich gewesen, in allen Fällen, in denen die Frage nach (mindestens) einem Gut nicht beantwortet wurde, überhaupt keinen Indexwert zu berechnen ("istwise deletion"). Bei der Abfrage von 26 Gütern hätte dies allerdings sehr schnell zu einer hohen Zahl fehlender Indexwerte führen können, weshalb das oben beschriebene alternative Vorgehen gewählt wurde. Im Vergleich zur listwise deletion besteht hierbei allerdings die Gefahr, die Zahl der fehlenden Güter zu unterschätzen.</p>
<u>Literatur</u>	Andreß & Lipsmeier (2001); Halleröd (1995); Nolan & Whelan (1996); Ringen (1988)

Deprivationsindex, gewichtet

<u>Variablenname</u>	<i>depindg</i>
<u>Variablenlabel</u>	Deprivationsindex, gew. (n. aus fin. Gründ. fehl. Items; gew. Itemssumme: 12,8)
<u>Ausgangsvariablen</u>	<i>HLS0100a-HLS2600a; HLS0100b-HLS2600b; PLS0100-PLS2600</i>
<u>Typ / Datensatz</u>	materielle Situation / Haushaltsdaten (Gewichtung auf Personenebene)
<u>Bearbeitung</u>	Bernhard Christoph
<u>Erläuterung</u>	<p>Für eine generelle Darstellung: vgl. Deprivationsindex, ungewichtet</p> <p>Ein Problem, welches häufig an ungewichteten Indizes, wie dem oben dargestellten kritisiert wird, ist die identische Gewichtung aller in den Index eingehenden Items. Vergleicht man z.B. zwei Items wie die Frage nach dem Vorhandensein einer Toilette innerhalb der Wohnung oder nach einem Videorecorder bzw. DVD-Player, so wird unmittelbar deutlich, dass mit dem Fehlen dieser beiden Items jeweils sehr unterschiedliche Einschränkungen des Lebensstandards einhergehen. Eine Gewichtung der einzelnen Items scheint daher konzeptionell sinnvoll, selbst wenn empirische Untersuchungen gezeigt haben, dass sich gewichtete und ungewichtete Indexvarianten im Ergebnis i.d.R. nur geringfügig unterscheiden (vgl. Lipsmeier 1999).</p> <p>Wir haben uns hier entschieden, die Gewichtung der Items über den Anteil der Befragten, die ein Item als notwendig erachten, vorzunehmen. Dieses Vorgehen wurde nicht nur ausgewählt, weil es ein konzeptionell überzeugendes und übliches Verfahren ist (es findet sich z.B. bei Halleröd 1995), sondern auch, weil es sich vergleichsweise sparsam umsetzen ließ. Da für die im Personenbogen zu erhebenden Deprivationsgewichte eine hohe Stabilität über die Zeit unterstellt werden kann, müssen sie nur einmalig bzw. in vergleichbar langen Intervallen erhoben werden. Zudem war es aufgrund der großen Stichprobe des PASS möglich, die Befragten in mehrere zufällig ausgewählte Subsamples aufzuteilen, denen jeweils nur ein Teil der Items vorgelegt wurde.</p> <p>Alternative Arten der Gewichtung, wie z.B. eine Beschränkung des Indizes auf die von mind. einem bestimmten Anteil der Befragten als notwendig erachteten Items (z. B. Andreß & Lipsmeier 1995, Andreß et al. 1996) oder eine theoretische Beschränkung auf wenige, als grundlegend angesehene Items (z. B. Nolan & Whelan 1996) wurden hier nicht umgesetzt, lassen sich aber ggf. auf Basis der ausgelieferten Daten generieren. Für eine zusammenfassende Diskussion unterschiedlicher Vorgehensweisen bei der Indexgewichtung vgl. Andreß & Lipsmeier (2001, insbes. S. 28ff.)</p>
<u>Literatur</u>	Andreß & Lipsmeier (1995, 2001); Andreß et al. (1996); Halleröd (1995); Lipsmeier (1999); Nolan & Whelan (1996)

Haushaltstypologie

Variablenname

hhtyp

Variablenlabel

Haushaltstyp, generiert

Ausgangsvariablen

Haushaltsinformationen zu Alter und Verwandtschaftsbeziehungen

Typ / Datensatz

Haushaltsstruktur / Haushaltsdaten

Bearbeitung

Daniel Gebhardt

Erläuterung

Für die Typologisierung von Haushalten existiert eine Vielzahl von Varianten und Vorschlägen (vgl. z.B. Lengerer et al. 2005 für die Haushaltstypologie des Mikrozensus, Porst (1984) sowie Beckmann & Trometer 1991 für die Typologie des ALLBUS und Frick et al. (o.J.) für das SOEP). Die Haushaltstypologie des PASS orientiert sich an der letztgenannten Typologie. Die entscheidenden Differenzierungskriterien bilden dabei Partnerschaften sowie Anzahl und Alter der Kinder sowie das Vorhandensein von Generationenbeziehungen. Während im SOEP jedoch lediglich die Stellung aller Haushaltsmitglieder zum Haushaltsvorstand die Basis für die Generierung des Haushaltstyps bildet, werden im PASS die Angaben zu den Beziehungen zwischen allen Haushaltsmitgliedern untereinander für die Generierung verwendet. Daneben werden das Alter der Haushaltsmitglieder laut Haushaltsinterview und die Haushaltsgröße bei der Generierung mit einbezogen.

Definition der Beziehungen bei der Generierung des Haushaltstyps:

- Paare: Ehepaare; eingetragene Lebenspartnerschaften; unverheiratete Partnerschaften sowie Partnerschaften mit nicht weiter spezifiziertem Status (fehlender Wert in Nachfrage nach Art der Partnerschaft).
- Kind einer Person: leibliches Kind; Stief-, Adoptiv- oder Pflegekind; Kind mit nicht weiter spezifiziertem Status (fehlender Wert in Nachfrage nach Art der Beziehung zum Kind).
- Elternteil einer Person: leibliches Elternteil; Stief-, Adoptiv- und Pflegeelternteil; Eltern mit nicht weiter spezifiziertem Status (fehlender Wert in Nachfrage nach Art der Elternschaft).

Definition der Haushaltstypen:

- 1-Personen Haushalt: Haushalt mit nur einer Person im Haushalt.
- Paar ohne Kinder: Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen.
- 1-Eltern-Haushalt: Haushalt besteht ausschließlich aus einem Elternteil und dessen Kindern. Hinsichtlich des Alters der Kinder werden keine Einschränkungen vorgenommen.
- Paar mit Kindern < 16 Jahre: Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen und deren jeweiligen bzw. gemeinsamen Kindern. Alle Kinder sind unter 16 Jahre alt.
- Paar mit Kindern >= 16 Jahre: Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen und deren jeweiligen bzw. gemeinsamen Kindern. Alle Kinder sind 16 Jahre oder älter.
- Paar mit Kindern < 16 Jahre und >= 16 Jahre: Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen und deren jeweiligen bzw. gemeinsamen Kindern. Es leben sowohl Kinder unter 16 Jahre als auch im Alter von 16 Jahren oder älter im Haushalt.
- Mehrgenerationenhaushalt: Haushalt besteht aus Angehörigen mindestens dreier linear aufeinanderfolgender Generationen. Der Haushalt besitzt einen Mehrgenerationenkern, d.h. mindestens eine Person im Haushalt ist sowohl Kind als auch Elternteil einer anderen Personen im Haushalt. Ausgehend von den zu diesem Mehrgenerationenkern gehörenden Personen leben im Haushalt ausschließlich Eltern, Kinder, Geschwister, Partner sowie Geschwister der Partner.
- Anderer Haushaltstyp: Haushalt, der keinem der anderen definierten Haushaltstypen zugeordnet werden konnte.
- Generierung nicht möglich (fehlende Werte): Grundsätzlich alle Haushalte mit mindestens einem fehlenden (-1, -2, -4) oder unplausiblen Wert (-8) in der Oberkategorie in einer Beziehungsvariable oder bei der Altersangabe (Ausnahme: In Haushalten mit drei und weniger Mitgliedern wurde in eindeutigen Konstellationen auch bei fehlenden Altersangaben der Haushaltstyp gebildet.).

Literatur

Beckmann & Trometer (1991); Frick et al. (o.J.); Lengerer et al. (2005); Porst (1984)

Bedarfsgemeinschaftsnummer, Welle 3

<u>Variablenname</u>	<i>bgnr3</i>
<u>Variablenlabel</u>	Nummer der Bedarfsgemeinschaft in Welle 3
<u>AusgangsvARIABLEN</u>	Haushaltsinformationen zu Alter und Verwandtschaftsbeziehungen
<u>Typ / Datensatz</u>	Bedarfsgemeinschaft / Personenregister
<u>Bearbeitung</u>	Gerrit Müller
<u>Erläuterung</u>	<p>Die Variable <i>bgnr3</i> wird auf Personenebene gebildet und ordnet jedem Haushaltsmitglied eine Nummer zu, welche die Zugehörigkeit zu einer bestimmten Bedarfsgemeinschaft anzeigt. Das heißt, Haushaltsmitglieder mit gleicher Nummer bilden gemeinsam eine BG. Die Variable <i>bgnr3</i> setzt sich aus der bekannten Haushaltsnummer sowie einem zweistelligen BG-Indikator zusammen.</p> <p>Zur Identifikation der BG-Zugehörigkeit werden ausschließlich Angaben zu den Verwandtschaftsbeziehungen zwischen den verschiedenen Haushaltsmitgliedern aus dem HH-Grid sowie zum Alter der Personen laut Haushaltsinterview verwendet. Die so identifizierten Bedarfsgemeinschaften sind daher als „synthetische“ BGs zu verstehen. Angaben zu tatsächlichem Leistungsbezug, Erwerbsfähigkeit bzw. Ausbildungsstatus der einzelnen Personen werden bei der Bildung der BGs nicht berücksichtigt. Vielmehr geht es um die Identifikation von Personengruppen innerhalb eines Haushalts, die im Falle einer Bedürftigkeit gemeinsam eine Bedarfsgemeinschaft im Sinne des SGB II bilden bzw. bilden würden. Diese künstliche Zuordnung ist notwendig, da das Konzept der Bedarfsgemeinschaft (BG) und der entsprechenden Zugehörigkeit von Personen zu dieser im Rahmen einer Befragung nicht direkt zu erheben ist.</p> <p>Inhaltlich erfolgt die Zuordnung von Personen zu Bedarfsgemeinschaften auf Basis der aktuell gültigen Fassung des § 7 Abs. 3 SGB II (zuletzt geändert am 26. März 2007). Danach bilden grundsätzlich alle Personen, die das 25. Lebensjahr vollendet und das 65. Lebensjahr noch nicht vollendet haben, jeweils eine eigene BG; es sei denn, diese Personen leben in einer Partnerschaft und/oder haben ein Kind/Kinder, das/die wiederum selbst das 25. Lebensjahr noch nicht vollendet hat/haben und ohne Partner bzw. eigene Kinder sind. In letzterem Falle bilden Personen jeweils gemeinsam mit ihrem Partner und Kind(ern) eine BG. Sollten zwei Personen mit einem gemeinsamen Kind im gleichen Haushalt leben, aber im HH-Grid nicht angegeben, Partner voneinander zu sein, wird dies im Sinne des §7 Abs. (3a) trotzdem unterstellt und die entsprechenden Personen gemeinsam mit dem Kind/Kindern zu einer BG zugeordnet. Personen die das 15. Lebensjahr vollendet und das 25. Lebensjahr noch nicht vollendet haben werden grundsätzlich ihren Eltern zugeordnet, es sei denn sie leben bereits mit einem Partner (bzw. einem eigenen Kind) gemeinsam im Haushalt. Sollten Personen zwischen 15 und 25 Jahre ohne Eltern (oder Partner bzw. Kind) leben, so bilden sie jeweils eine eigene BG.</p> <p>Personen, die das 65. Lebensjahr vollendet haben, fallen nicht in den Geltungsbereich des SGB II und gehören daher keiner BG an (Code 0); es sei denn, sie leben gemeinsam mit einem Partner, der das 65. Lebensjahr (bzw. einem Kind, welches das 25 Lebensjahr) noch nicht vollendet hat in einem Haushalt. Kinder, die das 15. Lebensjahr noch nicht vollendet haben und ohne Eltern in einem Haushalt leben, gehören ebenfalls keiner BG an (Code 0). Sie fallen in den Geltungsbereich des SGB XII. Für Haushalte mit fehlenden Beziehungs- bzw. Altersangaben zu einzelnen Personen wurde keine BG-Zuordnung vorgenommen, sondern für alle Haushaltsmitglieder Code 99 vergeben. Diese Haushalte können approximativ als Ein-BG-Haushalte interpretiert werden.</p>
<u>Literatur</u>	Sozialgesetzbuch, Zweites Buch - Grundsicherung für Arbeitssuchende (SGB II)

Bedarfsgemeinschaftstypologie, Welle 3

<u>Variablenname</u>	<i>bgtyp3</i>
<u>Variablenlabel</u>	Typ der Bedarfsgemeinschaft in Welle 3
<u>Ausgangsvariablen</u>	Haushaltsinformationen zu Alter und Verwandtschaftsbeziehungen
<u>Typ / Datensatz</u>	Bedarfsgemeinschaft / Personenregister
<u>Bearbeitung</u>	Gerrit Müller
<u>Erläuterung</u>	<p>Der Bedarfsgemeinschaftstypologie liegt das gleiche Konzept der synthetischen Bedarfsgemeinschaft wie bei der Variable <i>bgnr3</i> zugrunde. Kinder gehören bis zur Vollendung des 25. Lebensjahres der BG ihrer Eltern an, sofern sie nicht selbst Partner oder Kind haben. Dies wird hier anders gehandhabt als in der BA-Statistik, wo oftmals noch Typisierungen auf Basis der Volljährigkeit (18. Lebensjahr) ausgewiesen werden. Alleinerziehenden-BGs, im Sinne unserer Typisierung, in denen das jüngste Kind zwischen 18 und 24 Jahre alt ist, werden in der BA-Statistik beispielsweise als alleinstehend ausgewiesen. Dieser Unterschied ist bei Vergleichen mit Zahlen der offiziellen Statistik unbedingt zu berücksichtigen.</p> <p>Code 0, keine BG, wurde vergeben, wenn eine oder mehrere Personen nicht in den Geltungsbereich des SGB II fallen (vgl. Code 0 bei <i>bgnr3</i>). Code -5, Generierung nicht möglich (fehlende Werte), wurde für Haushalte mit fehlenden Beziehungs- bzw. Altersangaben zu einzelnen Personen (vgl. Code 99 bei <i>bgnr2</i>) vergeben.</p>
<u>Literatur</u>	–

Arbeitslosengeld-II-Bezug der Bedarfsgemeinschaft zum Zeitpunkt der Stichprobenziehung, Welle 3

<u>Variablenname</u>	<i>bgbez3</i>
<u>Variablenlabel</u>	ALG2-Bezug der Bedarfsgemeinschaft bei Stichprobenziehung in Welle 3 (2007/2008)
<u>Ausgangsvariablen</u>	Neubefr. HH: <i>HH49, HH50, HH52, HH53, HH62, sample, hnr, bgnr2, hhgr</i> Wiederbefr. HH: <i>HH91, HH92, HH93, HH95, sample, hnr, bgnr2, hhgr</i>
<u>Typ / Datensatz</u>	Bedarfsgemeinschaft / Personenregister
<u>Bearbeitung</u>	Mark Trappmann
<u>Erläuterung</u>	Diese Variable gibt für jede Bedarfsgemeinschaft, die entsprechend dem bei Variable <i>bgnr3</i> beschriebenen Vorgehen gebildet wurde, an, ob diese Bedarfsgemeinschaft zum Zeitpunkt der Stichprobenziehung für Welle 3 Arbeitslosengeld II bezogen hat oder nicht.
<u>Literatur</u>	–

Arbeitslosengeld-II-Bezug der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 3

<u>Variablenname</u>	<i>bgbezb3</i>
<u>Variablenlabel</u>	ALG2-Bezug der Bedarfsgemeinschaft zum Befragungszeitpunkt in Welle 3 (2007/2008)
<u>Ausgangsvariablen</u>	<i>AL20601, AL20701, zensiert (alg2_spells), sample, hhgr, bgnr3</i>
<u>Typ / Datensatz</u>	Bedarfsgemeinschaft / Personenregister
<u>Bearbeitung</u>	Daniel Gebhardt
<u>Erläuterung</u>	Diese Variable gibt für jede Bedarfsgemeinschaft, die entsprechend dem bei Variable <i>bgnr3</i> beschriebenen Vorgehen gebildet wurde, an, ob diese Bedarfsgemeinschaft zum Befragungszeitpunkt in Welle 3 Arbeitslosengeld II bezogen hat oder nicht.
<u>Literatur</u>	–

Anzahl der Bedarfsgemeinschaften im Haushalt

<u>Variablenname</u>	<i>anzbg</i>
<u>Variablenlabel</u>	Anzahl der synthetischen Bedarfsgemeinschaften im HH, generiert
<u>Ausgangsvariablen</u>	<i>bgnr3, hnr</i>
<u>Typ / Datensatz</u>	Bedarfsgemeinschaft / Haushaltsdatensatz
<u>Bearbeitung</u>	Daniel Gebhardt
<u>Erläuterung</u>	Diese Variable gibt die Anzahl Bedarfsgemeinschaften an, die innerhalb des Haushalts existieren. Die Identifizierung der Bedarfsgemeinschaften erfolgte dabei entsprechend der bei den Erläuterungen zur Variable <i>bgnr3</i> beschriebenen Vorgehensweise.
<u>Literatur</u>	–

Anzahl der zum Stichprobenzeitpunkt Leistungen beziehenden Bedarfsgemeinschaften im Haushalt

<u>Variablenname</u>	<i>nbgbezug</i>
<u>Variablenlabel</u>	Anzahl zum Stichprobenzeitpunkt beziehender BGs im HH
<u>Ausgangsvariablen</u>	<i>bgbezs3, bgnr3, hnr</i>
<u>Typ / Datensatz</u>	Bedarfsgemeinschaft / Haushaltsdatensatz
<u>Bearbeitung</u>	Daniel Gebhardt
<u>Erläuterung</u>	Diese Variable gibt die Anzahl der Bedarfsgemeinschaften innerhalb des Haushalts an, die zum Zeitpunkt der Stichprobenziehungen SGB-II-Leistungen bezogen haben. Der Wert wurde durch Aggregation der laut der Variablen <i>bgbezs3</i> aus dem Personenregister beziehenden Bedarfsgemeinschaften über die Haushaltsnummer gebildet.
<u>Literatur</u>	–

5 Datenaufbereitung

Die Datenaufbereitung wurde in Welle 3 erstmals nicht vom IAB selbst, sondern von infas durchgeführt³⁹. Um dennoch die Konsistenz der Datenaufbereitung im Längsschnitt zu gewährleisten, wurden infas die relevanten Syntaxfiles der Datenaufbereitung aus Welle 2 zusammen mit den notwendigen Ausgangs- und Zwischendatensätzen sowie einer Dokumentation der Arbeitsschritte zur Verfügung gestellt. Wichtige Entscheidungen, bspw. zur Bereinigung von Strukturproblemen in den befragten Haushalten oder zur Integration der Spelldatensätze, wurden in Abstimmung mit dem IAB getroffen. Auch darüber hinaus stand das IAB für die Beantwortung von Fragen während der Zeit der Datenaufbereitung zur Verfügung.

Die in den Interviews der 3. Welle erhobenen Informationen liegen bei TNS Infratest zunächst als ASCII-Daten vor. In einer ersten Stufe, erstellte TNS Infratest aus diesen Rohdaten zunächst die folgenden einheitlichen Datensätze (vgl. Büngeler et al 2010: 71ff.):

- Haushaltsdatensatz für wiederbefragte Haushalte
- Haushaltsdatensatz für Neubefragte und abgespaltene Haushalte
- Personendatensatz (Befragte von 15 bis einschl. 64 Jahre)
- Lückendatensatz (Informationen zu Lücken in den Erwerbsbiografien von mehr als drei Monaten Dauer seit Januar 2005)
- Seniorendatensatz (Befragte ab 65 Jahre)

Diese Datensätze wurden von TNS Infratest einer Basisprüfung der Filterführung unterzogen, wobei fälschlicherweise nicht gestellte Fragen mit einem Code markiert wurden. Die so vorbereiteten Datensätze wurden über das IAB an infas geliefert. Dort wurden sie in einer zweiten Stufe einer weiteren, ausführlicheren formalen und inhaltlichen Prüfung unterzogen und zum Scientific Use File aufbereitet. Darüber hinaus lieferte TNS Infratest Datensätze mit offenen Angaben (z.B. zur Art der beruflichen Tätigkeit), einen Bruttodatensatz sowie weitere Spezialdatensätze, die sich jedoch nicht direkt aus den eigentlichen Erhebungsinstrumenten ergeben.

Die sich anschließende Datenprüfung bei infas unterteilt sich in drei Schritte, die in den folgenden Abschnitten genauer dargestellt werden. Zunächst wurde in den wiederbefragten Haushalten die Haushaltsstruktur überprüft und ggf. Bereinigungen vorgenommen. Bei schweren Strukturproblemen wurden die entsprechenden Interviews entfernt (siehe dazu Kapitel 5.1). Im Anschluss erfolgte eine detaillierte Überprüfung und ggf. Bereinigung der Filterführung. Dabei wurden einerseits Filterfehler markiert und andererseits spezifische Codes für Missing Values gesetzt (siehe dazu Kapitel 5.2). Im Anschluss wurden bei ausgewählten Items inhaltliche Plausibilitätsprüfungen vorgenommen. Eindeutig unplausible bzw. widersprüchliche Angaben wurden durch einen spezifischen Missing Code als solche ge-

³⁹ infas übernimmt ab Welle 4 auch die Feldarbeit für PASS. Die Datenaufbereitung ab Welle 3 war Teil der Neuausschreibung, die durch den auf zunächst drei Wellen begrenzten Vertrag mit TNS Infratest notwendig geworden war.

kennzeichnet, wobei eine derartige Bereinigung der Daten allerdings sehr restriktiv gehandhabt wurde.

Die folgende Abbildung gibt eine Übersicht über alle im Rahmen der Datenaufbereitung durchgeführten Arbeitsschritte und ihre Abfolge:

Tabelle 26: Übersicht über die Schritte der Datenaufbereitung in der 3. Welle des PASS

Nr.	Arbeitsschritt
1	Transfer der von TNS Infratest gelieferten Datensätze ins STATA-Format
2	Prüfung der Haushaltsstruktur bei wiederbefragten Haushalten (siehe Kapitel 5.1)
3	Entfernen problematischer Interviews (Haushalts- und/oder Personenebene) (siehe Kapitel 5.1)
4	Integration Personen- und Seniorendatensatz
5	Korrektur der Haushaltsstruktur bei wiederbefragten Haushalten (siehe Kapitel 5.1)
6	Filterbereinigung Haushaltsebene (siehe Kapitel 5.2)
7	Erstellung und Plausibilitätsprüfung eines Haushaltsgriddatensatzes (siehe Kapitel 5.3)
8	Generierung der synthetischen Bedarfsgemeinschaften (siehe Variablenbeschreibung 4.5)
9	Neugenerierung von Steuerungsvariablen auf Basis der filtergeprüften Haushaltsdaten und des plausibilitätsgeprüften Haushaltsgriddatensatzes
10	Filterbereinigung Personenebene (siehe Kapitel 5.2)
11	Zucodung offener Angaben (siehe Kapitel 4.1)
12	Plausibilitätsprüfung der Haushalts- und Personendaten (ohne Spelldaten) (siehe Kapitel 5.3)
13	Aufbereitung, Plausibilitätsprüfung und Erstellung der Spelldatensätze (siehe Kapitel 5.6 bis 5.8 sowie auch Kapitel 5.3)
14	Einfache Generierungen (siehe Kapitel 4.4)
15	Komplexe Generierungen (siehe Kapitel 4.5)
16	Generierung der Datenstruktur für Scientific Use File (Haushalts-, Personen-, Registerdatensätze)
17	Anonymisierung (siehe Kapitel 5.5)

5.1 Strukturprüfungen und entfernte Interviews

Der Filterbereinigung vorgelagert, fand in der 3. Welle eine Strukturprüfung statt. Dabei sollten zunächst Interviews identifiziert werden, die im Sinne des PASS als nicht realisiert gelten und aus diesem Grund ggf. aus den Datensätzen entfernt wurden. Darüber hinaus erfolgte ein Abgleich der Struktur der wiederbefragten Haushalte mit der Vorwelle, um unplausible oder problematische Veränderungen der Haushaltszusammensetzung bzw. Fehler bei der Zuweisung der Personeninterviews erkennen und ggf. korrigieren zu können. Für die Betrachtung der Haushalte im Längsschnitt ist die konsistente Zuordnung und eindeutige Identifizierbarkeit der Personen über die Wellen hinweg zentral. Eine feste Personennummer darf nicht in verschiedenen Wellen unterschiedlichen Personen zugeordnet sein. War die korrekte Haushaltszusammensetzung unklar, wurden alle Interviews der 3. Welle dieses Haushalts entfernt. Wurde auf der Personenebene ein Interview mit der falschen Person geführt, ohne dass weiterführende Probleme in der Haushaltszusammensetzung erkennbar waren, so wurde lediglich das Personeninterview entfernt.

Um problematische Fälle zu identifizieren, wurden verschiedene Prüfungen durchgeführt:

- Durch einen Vergleich der Vornamen zwischen der aktuellen und der Vorwelle wurden Fälle erkannt, bei denen Veränderungen der Haushaltszusammensetzung nicht korrekt aufgenommen wurden. Statt Ein- und Auszüge an den entsprechenden Stellen des Haushaltsinterviews aufzunehmen, kam es vor, dass Interviewer Haushaltsmitglieder umbenannten sowie Veränderungen an Alter oder Geschlecht der Personen vornahmen. Einzelfallprüfungen wurden vorgenommen, wenn sich Veränderungen an den Vornamen nicht auf Korrekturen der Schreibweise zurückführen ließen, wenn sich das in der Vorwelle berichtete Geburtsjahr um mehr als ein Jahr von dem in der aktuellen Welle berichteten unterschied oder wenn das Geschlecht nicht mit der aus der Vorwelle bekannten Angabe übereinstimmte. Dabei wurde entschieden, ob es sich bei der Veränderung lediglich um eine Korrektur des Vornamens, des Alters oder des Geschlechts handelte oder aber ob der Interviewer eine unzulässige Veränderung der Haushaltsstruktur vorgenommen hatte. Die entsprechenden Fälle wurden in einem formalisierten Verfahren zwischen infas und dem IAB diskutiert und abschließend wurde eine endgültige Entscheidung im IAB getroffen, wie mit diesen Fällen zu verfahren ist.
- Weiterhin wurden geprüft, ob mehrere Personen mit demselben Geburtsdatum im Haushalt leben. Im Haushaltskontext der beiden Wellen wurde hier jeweils entschieden, ob es sich um plausible oder unplausible Fälle handelte. Anschließend wurden die übrigen Fälle einer weiteren Prüfung unterzogen. Dabei wurden Haushalte identifiziert, bei denen ein Geburtsdatum in der aktuellen Welle und in der Vorwelle von Personen auf unterschiedlichen Positionen der Haushaltsstruktur berichtet wurde. Hier lag die Vermutung nahe, dass das entsprechende Personeninterview in der aktuellen Welle von einer anderen Person als in der Vorwelle geführt wurde. Im Kontext der Haushalts- und Personeninformationen der aktuellen und der Vorwelle wurden jeweils Einzelfallentscheidungen in Bezug auf die jeweiligen Haushalts- und Personeninterviews getroffen.
- Um Haushalte zu identifizieren, die im Sinne des PASS als nicht realisiert gelten, wurden die Datensätze auf Haushalts- und Personenebene zusammengespielt. Personeninterviews ohne vollständiges Haushaltsinterview wurden ebenso markiert wie Haushaltsinterviews, für die kein Interview auf Personenebene vorlag⁴⁰.
- Breiten Raum nehmen auch die Prüfungen von Ein- und Auszügen ein. Panelhaushalte, in denen Auszüge berichtet wurden, wurden grundsätzlich auf ihren Haushaltskontext sightgeprüft und mit den realisierten abgesplitteten Haushalten in Bezug gesetzt. Es wurde geprüft, ob der verbleibende Haushaltskontext des Panelhaushalts lebensweltlich plausibel ist. Interviews aus Panelhaushalten, in denen alle Haushaltspersonen den Haushalt verlassen bis auf einzelne Kinder im Alter von unter 15 Jahre, wurden sowohl mit Blick auf den Panelhaushalt als auch mit Blick auf die abgesplitteten Haushalte verworfen. Auch wurde beim Auszug von mehr als einer Person geprüft, ob die Ausziehen-

⁴⁰ Bei neubefragten Haushalten, für die zwar ein Haushalts-, aber kein gültiges Personeninterview vorlag, wurden die Haushaltsinterviews analog zum Vorgehen in der 1. Welle entfernt. Bei wiederbefragten und abgespaltenen Haushalten, wurden die Haushaltsinterviews dagegen beibehalten.

den einen gemeinsamen Splithaushalt bilden oder mehrere verschiedene und ob das jeweils plausibel ist. Als unplausibel betrachtet wurden Fälle, in denen ein Partner zusammen mit jungen Kindern den Panelhaushalt verlässt, die Ausziehenden aber laut Feldinformationen mehrere verschiedene Splithaushalte bilden, sprich die jungen Kinder angeblich eigene Haushalte bilden. Hier wurde im Falle der Nicht-Realisierung des Splithaushalts entschieden, die Auszüge als plausibel anzuerkennen, alle Ausgezogenen jedoch nachträglich in einem gemeinsamen Splithaushalt zusammenzuführen.

- In Einzelfällen wurde beobachtet, dass laut Befragung im Panelhaushalt einzelne Personen einen abgesplitteten Haushalt bilden, sich im Splithaushalt dann aber alle Mitglieder des Panelhaushalts wiederfinden. In einer abgewandelten Variante finden sich im abgesplitteten Haushalt nicht alle Mitglieder des Panelhaushalts wieder, jedoch zumindest eine Person des Panelhaushalts, die im dortigen Interview als nicht ausgezogen bzw. als in einen anderen Splithaushalt als den betrachteten verzogen berichtet wurde. Auch hier wurde differenziert entschieden, welche berichteten Auszüge als gültig anerkannt werden und welche als unplausibel verworfen wurden. Wenn ein berichteter Auszug nachträglich als unplausibel verworfen wurde, wurde die angeblich ausgezogene Person nachträglich wieder in den Haushaltskontext des Panelhaushalts integriert.
- In Panelhaushalten, die ab Welle 2 von einem Auszug berichtet haben, kann es ab Welle 3 auch zu Wiedereinzügen von früher bereits einmal dem Haushalt angehörigen Personen kommen. Die Anforderung, diese Einziehenden als wiedereinziehend zu erkennen und wieder auf ihrer früheren Haushaltsposition einzuordnen statt ihnen eine neue, bisher noch nicht vergebene Haushaltsposition zuzuweisen, ist Bestandteil des Haushaltsgrids. Es wurde nachträglich geprüft, ob diese Anforderung in allen Fällen im Feld erfüllt worden ist. Bei Personen, die in der aktuellen Welle nachträglich anhand eines Abgleichs von Vorname, Alter und Geschlecht mit den zuvor aus dem Haushalt Ausgezogenen als wiedereinziehende Person identifiziert wurden, musste in die Haushaltsstruktur eingegriffen werden. Notwendig waren dann ein nachträgliches Ändern der Personennummer der zu verschiebenden Person wie auch ein Verschieben der personenbezogenen Angaben im Haushaltsinterview z.B. zur Kinderbetreuung oder zu den Gründen der ALG2-Kürzung auf die im Rahmen der Strukturprüfung als richtig definierten Position.
- Haushaltsstrukturprüfungen prüfen im Wesentlichen nicht die Haushaltsstruktur als solche auf Plausibilität, sondern sie betrachten die Veränderungen zwischen den Wellen. Daher kann die Haushaltsstruktur erstbefragter Haushalte nur im geringen Umfang geprüft werden. Bei erstbefragten Haushalten wird anhand der Angaben zu Vornamen, Alter und Geschlecht betrachtet, ob einzelne Haushaltsmitglieder mehrfach aufgelistet werden. In diesem Fall werden von den doppelt berichteten Personen nur die zuerst berichtete Haushaltsposition behalten, die weiteren Haushaltspositionen hingegen verworfen. Dies zieht ggf. weitere Eingriffe in die Haushaltsstruktur nach sich. Wenn z.B. in einem erstbefragten Haushalt mit laut Befragung vier Personen die Personen auf den Positionen 2 und 3 identisch sind, wird nicht nur Person 3 entfernt, sondern auch Person 4 nachträglich auf Position 3 verschoben. Denn in einem erstbefragten Haushalt mit X Haushaltsmitglieder sollen immer die Positionen 1 bis X durchgehend und ohne Lücken

besetzt sein. Wie auch bei den nachträglich als wiedereinziehend erkannten Personen erfordert ein nachträgliches Ändern der Personennummer der zu verschiebenden Person wie auch ein Verschieben der personenbezogenen Angaben im Haushaltsinterview.

Auch beim Umgang mit den Fällen, die sich in den Strukturprüfungen als problematisch herausgestellt haben, wurden Einzelfallentscheidungen getroffen. Maßgeblich dabei war jeweils, als wie schwerwiegend das betroffene Problem angesehen wurde. In Fällen, in denen die korrekte Haushaltszusammensetzung in der 3. Welle unklar war, wurden alle Interviews der 3. Welle entfernt. Sie werden in der 4. Welle wie Haushalte behandelt, die in der 3. Welle nicht teilgenommen haben. Wurden in nachträglich entfernten Haushaltsinterviews Auszüge gemeldet, wurden auch die abgesplitteten Haushalte verworfen. Dies betraf sowohl die in der aktuellen Welle realisierten Interviews in diesen Splithaushalten als auch die Stichprobe der Folgewelle. Splithaushalte, die aus einem verworfenen Interview eines Panelhaushalts heraus entstanden sind, werden nachträglich als nicht entstanden gewertet und zählen nicht zur Panelstichprobe der Folgewelle. Bestand lediglich ein Zuordnungsproblem auf Personenebene, d.h. wurde vermutet, dass ein Personeninterview in der 3. Welle mit einer falschen Person geführt wurde, dann wurde nur das betroffene Personen- bzw. Senioreninterview entfernt. Handelte es sich um ein Strukturproblem, das keine schwerwiegenden Auswirkungen hatte und bspw. durch das Entfernen eines Personeninterviews behoben werden konnte, wurden zusätzliche Korrekturen der Vornamen, Alters- und Geschlechtsangaben auf Haushaltsebene vorgenommen. Die betroffene fehlerhafte Information wurde dann auf den letzten gültigen Wert aus der Vorwelle zurückgesetzt bzw. im Falle des Alters auf den Vorwollenwert + die Zahl der Jahre seit der letzten gültigen Realisierung in diesem Haushalt.

Darüber hinaus wurden alle Interviews von Personen entfernt, für deren Haushalt kein vollständiges Haushaltsinterview vorlag. Im umgekehrten Fall, d.h. bei Haushalten, für die kein Interview auf Personenebene vorhanden war, wurde zwischen wiederbefragten Haushalten und Haushalten der Auffrischungsstichprobe unterschieden. Die als nicht realisiert geltenden Haushalte aus der Auffrischungsstichprobe wurden analog zum Vorgehen in den Vorwollen entfernt. Bei wiederbefragten Haushalten ohne Interviews auf Personenebene wurde das Haushaltsinterview dagegen nicht gelöscht.

Darüber hinaus meldete TNS Infratest mit dem Bruttodatensatz Verweise auf Haushalte, bei denen die Interviews nicht korrekt durchgeführt wurden. Dies betraf zum einen Haushaltsstrukturprobleme, wie die zuvor Beschriebenen, die bereits vor Übergabe der Rohdaten an das IAB entdeckt wurden, zum anderen Interviews mit technischen Problemen. In diesen Fällen wurden ebenfalls alle Interviews entfernt.

Einen Hinweis auf entfernte Interviews geben die *Nettovariablen* (*hnettok3*, *hnettod3*, *pnettok3*, *pnettod3*) in den Haushalts- und Personenregisterdatensätzen. Über die entsprechenden Variablen im Haushaltsregister ist nachvollziehbar für welche wiederbefragten Haushalte die Haushaltsinterviews nachträglich entfernt wurden. Mittels der Nettovariablen des Personenregisters kann überprüft werden, in welchen Fällen lediglich einzelne Interviews auf Personenebene oder alle Interviews des Haushalts gelöscht wurden. Bei Haushalten aus der Auffrischungsstichprobe der 3. Welle, für die nicht mindestens ein gültiges Haushalts- und Personeninterview vorliegt, sind gelöschte Interviews nicht in den Register-

datensätzen nachvollziehbar, da diese Haushalte nicht in die Datensätze aufgenommen wurden.

5.2 Filterbereinigung

In der Filterbereinigung wurde die in den jeweiligen Instrumenten vorgesehene Filterführung mithilfe eines Statistikprogramms nachvollzogen. Wurden bestimmte Fragen gestellt, obwohl der Wert der relevanten Filtervariable etwas anderes verlangt hätte (wurden also z.B. detaillierte Angaben zur Berufsausbildung erhoben, obwohl die/der Befragte angegeben hatte, über keinen Berufsabschluss zu verfügen), wurden die Variablen auf den Missing Code „-3“ (trifft nicht zu) gesetzt, den sie auch bei einer korrekten Ausführung der Filterführung erhalten hätten.⁴¹ Des Weiteren kam es vor, dass in Einzelfällen Items nicht erhoben wurden, obwohl dies laut der entsprechenden Filtervariable notwendig gewesen wäre (z.B. wenn keine weiteren Angaben zur Berufsausbildung erhoben wurden, obwohl die/der Befragte angegeben hatte, eine solche Ausbildung gemacht zu haben). In diesen Fällen wurde der spezifische Missing Code „-4“ (Frage irrtümlich nicht gestellt) vergeben. Die Vergabe des Codes „-4“ kann auch Folge der in Kapitel 5.1 beschriebenen Haushaltsstrukturprüfungen sein. Wenn der Auszug einer Person nachträglich als unplausibel verworfen wird und die Person nachträglich als noch zu ihrem bisherigen Haushalt gehörend gewertet wird, dann bedeutet dies auch, dass personenbezogene Angaben zu diesen Personen im Haushaltsinterview nachträglich als irrtümlich nicht erhoben codiert werden müssen. Der Code „-4“ weist also nicht in allen Fällen auf ein Problem des Erhebungsinstruments hin. Wenn Code „-4“ bei einer Frage vergeben wird, die filterrelevant für Folgefragen ist, dann werden die Folgefragen ebenfalls mit „-4“ vercodet, sofern diese Folgefragen tatsächlich nicht erhoben wurden. Wurden Folgefragen hingegen dennoch erhoben, weil z.B. mehrere Filterfragen auf diese Folgefrage verweisen und eine andere Filterfrage die Folgefrage korrekt ausgelöst hat, bleibt der dort erhobene Wert hingegen erhalten.

Daneben wurden im Rahmen der Filterprüfung auch die vom Feldinstitut vergebenen Missing Codes sowie die System Missings durch für alle Variablen einheitliche Werte ersetzt. Die vergebenen Werte sind in Tabelle 27 im Überblick dargestellt. „-1“ und „-2“ sind die einheitliche Umcodierung der in der Befragung erhobenen „WN“- und „KA“-Werte. „-3“ ist der generelle „TNZ“-Code für aufgrund der Filterführung nicht gestellte Fragen. Der Code „-4“ wurde, wie bereits erläutert, vergeben, wenn eine Frage durch einen Filterfehler nicht gestellt wurde. Bei „-5“ bis „-7“ handelt es sich um fragespezifische Codes. Hierbei kann es sich sowohl um spezifische Missing Codes (z.B. „TNZ, steht dem Arbeitsmarkt nicht zur Verfügung“) als auch um Sonderkategorien für gültige Werte (z.B. eine Kategorie für Einkommen oberhalb von € 99.999,- bei der offenen Einkommensangabe) handeln. Diese Codes wurden nur im Bedarfsfall vergeben.

⁴¹ Die Bereinigung erfolgte dabei, wie in solchen Fällen üblich, immer von den zuerst erhobenen zu den später erfragten Items.

Tabelle 27: Übersicht über die vergebenen Missing Codes

Code	Erläuterung
-1	„weiß nicht“
-2	„keine Angabe“
-3	„trifft nicht zu (Filter)“ (Frage wurde aufgrund der Filterführung nicht gestellt)
-4	„Frage irrtümlich nicht gestellt“ (Frage hätte jedoch gestellt werden sollen)
-5	Fragespezifischer Code Nr. 1, wurde nur bei Bedarf vergeben
-6	Fragespezifischer Code Nr. 2, wurde nur bei Bedarf vergeben
-7	Fragespezifischer Code Nr. 3, wurde nur bei Bedarf vergeben
-8	„unplausibler Wert“
-9	„Item in Welle nicht erhoben“
-10	„Item in Fragebogenversion nicht erhoben“

Der Wert „-8“ ist ein spezifischer Missing Code, der im Rahmen der Plausibilitätsprüfung vergeben wurde (zur Plausibilitätsprüfung vgl. Kapitel 5.3). Der Missing Code „-9“ wird seit der 2. Welle erstmals benötigt. Er wird immer dann vergeben, wenn ein Item in einer bestimmten Welle nicht erhoben wurde. Aufgrund der bereits dargestellten Aufbereitung der Daten im Long-Format erhalten also Variablen, die ab der 2. Welle nicht mehr erfragt wurden, für die Beobachtungen dieser Welle den Wert „-9“. Ebenso verhält es sich mit Beobachtungen der 1. Welle. Erstmals nach der 1. Welle erhobene Variablen werden für Beobachtungen der Wellen ohne Erhebung nachträglich mit „-9“ vercodet. Mit dem Code „-10“ können Unterschiede zwischen den Fragebogenversionen berücksichtigt werden, also zwischen Personen- und Seniorenbogen bzw. zwischen den beiden Versionen des Haushaltsfragebogens.

5.3 Plausibilitätsprüfungen

Im Rahmen der Plausibilitätsprüfung wurde eine umfangreiche Liste theoretisch möglicher Widersprüche in den Angaben der Befragten kontrolliert. Dazu wurde die Liste der in den Vorwelle durchgeführten Prüfungen für die aktuelle Welle angepasst und sinnvoll erweitert. Daneben wurde auch die Haushaltsstruktur auf Plausibilität geprüft. Darüber hinaus wurden auch die Spelldaten plausibilitätsgeprüft – insbesondere auf unzulässige Überschneidungen innerhalb der einzelnen Spelldaten. Dabei wurden grundsätzlich nur die im Querschnitt der 3. Welle erhobenen Daten für die Prüfung herangezogen. Eine Prüfung im Längsschnitt, also zwischen den in der aktuellen und der Vorwelle gemachten Angaben, wurde nicht durchgeführt.

Im Einzelnen wurden die folgenden Schritte ausgeführt:

1. Widerspruchsprüfung: Generell wurden Widersprüche nur dann korrigiert, wenn entweder die Unplausibilität als besonders gravierend oder bzw. und wenn der Eingriff als vergleichsweise mild einzustufen war. Letzteres galt z.B., wenn nur eine geringe Zahl von Fällen betroffen war oder weil nur ein Missing Code (z. B. „-3“) durch einen neuen („-8“) ersetzt wurde. Für die Bereinigung unplausibler Angaben wurden zwei Strategien ver-

wendet: eine direkte Korrektur der unplausiblen Angaben oder die Vergabe eines spezifischen Missing Codes.

- Eine Korrektur unplausibler Angaben wurde ausschließlich dann vorgenommen, wenn mit an Sicherheit grenzender Wahrscheinlichkeit eine Fehleingabe des Interviewers vorlag. Ein Beispiel ist die Angabe einer monatlichen Gesamtmiete von € 9.998,-. Hier wurde im Rahmen der Datenbereinigung davon ausgegangen, dass es sich um eine Fehleingabe des fünfstellig erhobenen Missing Codes „99998“ (WN) handelte. Diese und ähnliche Angaben wurden in die entsprechenden Missingkategorien umcodiert. Hätten die umcodierten Missingkategorien einen Filter bei Folgefragen ausgelöst, wie dies z.B. bei der kategorialen Abfrage von Einkommensangaben der Fall ist, so wurden die kategorialen Nachfragen nachträglich auf Code „-4“ (Frage irrtümlich nicht gestellt) umgesetzt.
 - Allerdings war es eher selten der Fall, dass sich ein Wert mit hinreichender Sicherheit als Fehleingabe erkennen ließ. In den meisten Fällen konnte lediglich ein Widerspruch zwischen zwei Angaben festgestellt werden, ohne jedoch spezifische Fehleingaben oder ähnliches identifizieren zu können, die zu der unplausiblen Angabe geführt hatten. Daher wurde in diesen Fällen keine Korrektur vorgenommen, sondern der spezifische Missing Value Code „-8“ vergeben. Ob der Code dabei an eine oder beide der am Widerspruch beteiligten Variablen vergeben wurde, ist im Einzelfall entschieden worden.
2. Plausibilitätsprüfung der Haushaltsstruktur: Auf Grundlage der im Haushaltsinterview erhobenen Information zu den Verwandtschaftsbeziehungen der Haushaltsmitglieder untereinander sowie der Angaben zu Alter, Geschlecht und Vornamen wurde eine Plausibilitätsprüfung der Haushaltsstruktur durchgeführt. Im Vorfeld dieser Prüfung wurden die Beziehungsinformationen im Haushalt durch die im Personeninterview berichteten Partnerschaften ergänzt.
- Um unplausible Haushaltskonstellationen zu identifizieren, wurden zunächst die Beziehungsinformationen mit den demografischen Angaben zu den einzelnen Haushaltsmitgliedern kombiniert. Für die in diesen Prüfungen als unplausibel identifizierten Haushalte wurden Einzelfallentscheidungen unter Einbeziehung der gesamten Haushaltsstruktur und anderer Interviewinformationen (z.B. zum Familienstand im Personeninterview) getroffen. Unplausible Beziehungskonstellationen wurden dabei als solche markiert („-8“) bzw. auf Basis hinzugezogener Informationen über den Haushaltskontext korrigiert, sofern es sich mit an Sicherheit grenzender Wahrscheinlichkeit um eine fehlerhafte Angabe handelte. Ein Beispiel: Bei gleichgeschlechtlichen Personen, die beide leibliche Eltern einer dritten Person im Haushalt waren, wurde auf Basis des Vornamens das Geschlecht korrigiert. Deuteten auch die Vornamen darauf hin, dass beide Personen das gleiche Geschlecht haben und lagen keine weiteren relevanten Hinweise vor, wurde vor dem Hintergrund der Haushaltskonstellation die Verwandtschaftsbeziehung als unplausibel gekennzeichnet.

- In einem zweiten Teil wurden Prüfungen durchgeführt, bei denen jeweils drei Verwandtschaftsverhältnisse in Beziehung zueinander gesetzt und auf Plausibilität geprüft wurden. Ein Beispiel für eine in dieser Prüfung als unplausibel klassifizierte Beziehungskonstellation wäre: Person A ist Ehepartner von Person B. Person A ist leibliches Elternteil von Person C. Person C ist Geschwister von Person B. Wurde im Rahmen dieser Prüfung eine solche oder ähnliche unplausible Kombination von Beziehungen identifiziert, so wurde auch hier versucht, die Beziehung im Rahmen des Haushaltskontextes zu plausibilisieren. Im genannten Fall wurde die Beziehungsinformation korrigiert, indem Person C als nicht weiter spezifiziertes Kind von B codiert wurde. Ziel ist es, möglichst viele der so identifizierten Unplausibilitäten inhaltlich zu korrigieren, da eine plausible und vollständig vorliegende Beziehungskonstellation notwendige Voraussetzung für die Generierung der BG ist.
3. Auch die Spelldatensätze wurden einer Reihe von Plausibilitätsprüfungen unterzogen, welche in Kapitel 5.6 bis 5.8 vertieft dargestellt werden.

5.4 Rückwirkende Veränderungen der 1. Welle und 2. Welle

Im Rahmen der Datenaufbereitung für den Scientific Use File der 3. Welle wurden auch Veränderungen an den bereits ausgelieferten Wellen des PASS vorgenommen. Dabei handelt es sich um Korrekturen von Fehlern, die seit der Fertigstellung des Scientific Use File der 2. Welle festgestellt wurden. Tabelle 28 bis Tabelle 33 geben einen Überblick über die rückwirkenden Veränderungen an bereits ausgelieferten Wellen des PASS.⁴²

⁴² Anpassungen an Werte- oder Variablenlabels sind hier nur berücksichtigt, wenn sie dadurch die Interpretation von Variablen oder Werten verändert.

Tabelle 28: Übersicht über rückwirkende Veränderungen im Haushaltsdatensatz (HHENDDAT)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>HW1900</i>	<i>HHENDDAT</i>	2	Korrektur	Der gleichzeitige Bezug von Wohngeld und Arbeitslosengeld 2 ist nicht möglich. Im Rahmen der Plausibilitätsprüfungen wurde eine entsprechende Prüfung für Haushalte mit nur einer synthetischen Bedarfsgemeinschaft vorgenommen. Bezog ein Haushalt gleichzeitig beide Leistungen, sollte sowohl der Indikator, ob Wohngeld bezogen wird (<i>HW1800</i>), als auch die Angabe zur monatlichen Höhe (<i>HW1900</i>) auf -8 (unplausibler Wert) gesetzt werden. In der Datenaufbereitung der 2. Welle wurde in diesen Fällen jedoch fälschlich nur die Variable <i>HW1800</i> auf -8 gesetzt. Die Angaben zum Betrag in <i>HW1900</i> blieben erhalten. Dieser Fehler wurde korrigiert – für Fälle, für die der Bezug von Wohngeld an sich (<i>HW1800</i>) unplausibel ist, wurde nun auch der monatliche Betrag (<i>HW1900</i>) auf „unplausibel“ gesetzt.
<i>HW0880a-i</i> <i>HW0890</i> <i>HW0900</i> <i>HW0910</i> <i>HW0920</i> <i>HW1000</i> <i>HW1100</i> <i>HW1200</i> <i>HW1300</i> <i>HW1400</i> <i>HW1500</i> <i>HW1600</i> <i>HW1700</i> <i>HW1800</i> <i>HW1900</i> <i>HW2000</i> <i>HW2100</i> <i>einzugj</i> <i>umzug</i>	<i>HHENDDAT</i>	2	Korrektur	In H64 (<i>HW0880a-i</i>) wurde in Item I der offene Grund für einen Umzug erhoben. Diese Angabe wurde vercodet (<i>HW0881a-j</i>). Wurde bei der Vercodung festgestellt, dass kein Umzug vorlag und die Frage nach den Umzugsgründen fälschlich gestellt wurde, fand eine Korrektur der Steuerungsvariable <i>umzug</i> statt. Die von dieser Variable direkt oder indirekt abhängigen Variablen wurden in der Folge erneut auf Basis der korrigierten Variable filterbereinigt. Durch einen Fehler wurden diese Korrekturen an der Steuerungsvariable <i>umzug</i> und an H64 bis H80 (HHalt) nicht in den Datensatz übernommen. Dieser Fehler wurde korrigiert. Die auf Basis der offenen Angaben bereinigte Steuervariable <i>umzug</i> sowie die entsprechend korrigierten Variablen, die im Eingangsfiler von ihr abhängig sind, sind nun im Haushaltsdatensatz enthalten.

Tabelle 28: Übersicht über rückwirkende Veränderungen im Haushaltsdatensatz (HHENDDAT) (Fortsetzung)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>HEK1600</i>	<i>HHENDDAT</i>	2	Korrektur	Bei der Plausibilitätsprüfung sollte die Anzahl der Kinder, für die Kindergeld bezogen wird, gegen die Anzahl der im HH lebenden + die Anzahl der außerhalb des Haushalts lebenden Kinder geprüft werden. Unplausible Angaben wurden im Zuge der Datenaufbereitung mit dem Wert „-8“ codiert. Da nicht mehr alle für die Prüfung notwendigen Informationen in der aktuellen Welle erhoben werden, wurde entschieden, die Prüfung ab Welle 2 nicht mehr analog zu Welle 1 durchzuführen. Ab Welle 2 werden nur noch Fälle in <i>HEK1600</i> (Anzahl der Kinder, für die der HH Kindergeld bezieht) auf -8 gesetzt, die bei der Filterfrage <i>HEK1500</i> angeben, Kindergeld zu beziehen, dann aber in <i>HEK1600</i> angeben für „0“ Kinder zu beziehen.
<i>HEK1810</i>	<i>HHENDDAT</i>	2	Korrektur	In <i>HEK1810</i> wurde die Höhe des Unterhaltsvorschlusses erhoben, den ein Haushalt für Kinder unter 15 Jahre erhalten kann. Durch einen Fehler in der Plausibilitätsprüfung dieser Variable wurden Fälle, die in der Filterfrage, ob der Haushalt überhaupt Unterhaltsvorschuss erhalten hat (<i>HEK1800</i>), angaben, keine solche Zahlung zu erhalten oder denen die Frage laut Filter nicht gestellt wurde, in <i>HEK1810</i> fälschlich auf -8 (unplausibler Wert) gesetzt. Dieser Fehler wurde korrigiert – Fälle, die in <i>HEK1800</i> angeben, keinen Unterhaltsvorschuss zu erhalten oder denen die Frage nicht gestellt wurde, wurden in der Nachfrage zur monatlichen Höhe (<i>HEK1810</i>) auf -3 (Trifft nicht zu) gesetzt.
<i>depindg</i>	<i>HHENDDAT</i>	1,2	Korrektur	Im Datensatz wird neben einer ungewichteten auch eine präferenzgewichtete Version des Deprivationsindex bereitgestellt. Die Präferenzen zur Gewichtung wurden im Rahmen der Personeninterviews in Welle 1 und dann sowohl für Welle 1 und Welle 2 genutzt. Um für die Gesamtbevölkerung repräsentative Gewichte zu erhalten, ist es jedoch nicht ausreichend, nur die in der Stichprobe erhobenen Präferenzen zu berücksichtigen. Dieser Fehler wurde korrigiert – die erhobenen Präferenzen wurden zunächst auf die Gesamtbevölkerung hochgerechnet. Mit diesen für die Gesamtbevölkerung repräsentativen Präferenzgewichten wurde dann der Deprivationsindex für Welle 1 und Welle 2 neu gewichtet.

Tabelle 29: Übersicht über rückwirkende Veränderungen im Personendatensatz (PENDDAT)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>PET0700</i>	<i>PENDDAT</i>	2	Korrektur	Im Seniorenfragebogen wurde erhoben, ob es sich bei einer aktuellen Erwerbstätigkeit um einen Minijob handelt oder nicht (<i>PET0500</i>). Die Arbeitszeit von Senioren mit Minijob und von Senioren mit einer anderen Art von Erwerbstätigkeit wurde dabei zwei verschiedenen Variablen zugeordnet (<i>PET0700</i> für die Minijobber, <i>PET1300</i> für andere Erwerbstätigkeiten). Bei der Codierung von <i>PET0700</i> ist es zu einem Fehler bei der Zuweisung der Missing Codes gekommen. Senioren ohne Minijob wurden fälschlich auf -10 codiert. Dieser Fehler wurde behoben. Die Senioren, die zum Befragungszeitpunkt der Welle 2 keinen Minijob haben, sind nun in <i>PET0700</i> mit -3 codiert.
<i>PEK0100</i> <i>PEK0100a</i> <i>PEK0100b</i> <i>PEK0700</i> <i>PEK0700a</i> <i>PEK0700b</i> <i>PEK1300</i> <i>PEK1360a</i> <i>PEK1360b</i> <i>PEK1415</i> <i>PEK1425</i> <i>PEK1435</i> <i>PEK1445</i> <i>PEK1455</i> <i>PEK1500</i> <i>PEK1700</i> <i>PEK1900</i> <i>PEK2100</i>	<i>PENDDAT</i>	2	Korrektur	Die offenen Einkommensangaben im Personendatensatz enthalten teilweise einen Sondercode für Einkommen, die einen bestimmten Wert übersteigen (Top-Coding). Je nachdem, um welche Einkommensvariable es sich handelt, ist ein Top-Coding enthalten oder nicht. Ist ein Top-Coding enthalten, unterscheidet sich dessen Höhe (bspw. „mehr als 99.999 Euro“ oder „mehr als 9.999 Euro“). Für einige Einkommensvariablen waren hinsichtlich des Top-Codings falsche Variablenlabel vergeben. Dieser Fehler wurde korrigiert. Es wurden neue Wertelabels erstellt und die Variablen mit dem jeweils zutreffenden Label versehen.

Tabelle 29: Übersicht über rückwirkende Veränderungen im Personendatensatz (PENDDAT) (Fortsetzung 1)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<p>PEK1600 PEK1700 PEK1800 PEK1900 PEK2000 (neu) PEK2100 (neu)</p>	PENDDAT	2	Korrektur	<p>In Welle 1 wurde im Personeninterview der Bezug von Erziehungs-/Elterngeld erhoben. PEK1600 enthält die Angaben auf die Filterfrage, ob die Person diese Leistungen bezieht. PEK1700 enthält die Angaben zur monatlichen Bezugshöhe.</p> <p>Ab Welle 2 wird der Bezug von Erziehungs-/Elterngeld nicht mehr auf Personenebene, sondern im Rahmen des Haushaltinterviews erhoben (HEK1610, HEK1620). Weiterhin wurde ab Welle 2 das Spektrum der erhobenen Einkommenskomponenten deutlich erweitert. Ab Welle 2 wird so bspw. im Personeninterview gefragt, ob BaföG/Berufsausbildungsbeihilfe/Stipendium bezogen werden und wenn ja, wie hoch der monatliche Betrag ausfällt. Diese Informationen wurden für Welle 2 fälschlich nicht in neuen Variablen abgelegt, sondern in PEK1600 bzw. PEK1700 übernommen.</p> <p>Dieser Fehler wurde korrigiert – PEK1600 und PEK1700 wurden für die Beobachtungen der Welle 2 mit -9 gefüllt (Item in Welle nicht erhoben). Um die richtige Reihenfolge der Fragen einzuhalten, mussten zunächst zwei Variablen umbenannt werden, die ebenfalls nur in Welle 2 erhoben wurden: PEK1800 wurde zu PEK2000 (Indikator für Erhalt staatlicher Zahlungen für Erwerbstätige), PEK1900 wurde zu PEK2100 (Höhe der monatlichen staatlichen Zahlungen für Erwerbstätige). Die Information, ob BaföG (oder eine ähnliche Leistung) bezogen wird, wurde in PEK1800 abgelegt. In PEK1900 wurde die monatliche Höhe der BaföG-Zahlungen gespeichert.</p>
<p>PSK0400a PSK0400b PSK0400c PSK0400d PSK0400e</p>	PENDDAT	2	Korrektur	<p>Bei PSK0400a-e handelt es sich in Welle 1 und 2 um eine Frage mit Mehrfachantwortmöglichkeit. Die global für alle Items gegebenen WN- und KA-Werte werden zusammen mit dem Sondercode „Nein, nicht aktiv“ in alle Einzelitems übernommen.</p> <p>Für Welle 2 wurde die Information aus dem Sondercode „Nein, nicht aktiv“ fälschlich nicht in die Einzelitems übertragen.</p> <p>Dieser Fehler wurde korrigiert – gab der Befragte an, nicht aktiv zu sein, wurden alle Einzelitems analog zu Welle 1 auf den entsprechenden Sondercode (-5) gesetzt.</p>

Tabelle 29: Übersicht über rückwirkende Veränderungen im Personendatensatz (PENDDAT) (Fortsetzung 2)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<p>PAS0900a-f PAS0901a-f PAS1000a-d PAS1110 PAS1200 PAS1300 PAS1400a-f PAS1600 PAS1800 PAS1900 PAS2000 PAS2100 PAS2200 PAS2300</p>	PENDDAT	2	Korrektur	<p>In PAS0100 werden Erwerbstätige zur Arbeitssuche befragt, Neben den drei vorgelesenen Kategorien (1-3) hat die Frage eine zusätzliche Kategorie (4 „sowohl eine zusätzliche als auch eine andere Erwerbstätigkeit gesucht“), die allerdings nicht vorlesen wurde (da sie „unter dem Strich“) steht. Da es sich bei dieser zusätzlichen Kategorie um einen Sondercode handelt, wurde in der Filterbereinigung vorübergehend der Code „-5“ vergeben. PAS0100 ist für die Arbeitssuche der Erwerbstätigen eine zentrale Filtervariable, weshalb PAS0100 in den folgenden Fragen steuerungsrelevant ist. Bei der Prüfung von Eingangsfiltren in den Folgefragen wurde diese Umcodierung allerdings fälschlicherweise nicht berücksichtigt. Entsprechend kam es zu Fehlern in der Filterbereinigung der Folgefragen. Fälle, die in PAS0100 angegeben hatten, dass sie sowohl eine zusätzliche als auch eine andere Erwerbstätigkeit gesucht haben, wurden in den angegebenen Variablen fälschlich auf -3 gesetzt. Dieser Fehler in der Filterbereinigung wurde korrigiert. Die von Befragten gemachten Angaben sind jetzt enthalten.</p>

Tabelle 29: Übersicht über rückwirkende Veränderungen im Personendatensatz (PENDDAT) (Fortsetzung 3)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
PTK0200 PTK0500 PTK0900a PTK1000a PTK1100a	PENDDAT	2	Korrektur	<p>In Welle 1 wurde im Trägerkontaktemodul bei einigen Items (erstmals bei PTK0200) der Sondercode „TNZ, steht dem Arbeitsmarkt derzeit nicht zur Verfügung“ erhoben. Wurde eine entsprechende Angabe vom Befragten gemacht, so wurde diese in der Datenaufbereitung in allen folgenden Items des Trägerkontaktemoduls übernommen. Ab Welle 2 wurde dieser Sondercode nicht mehr erhoben. Stattdessen wurden mit PTK0310 und PTK0320* neue Items eingefügt, in denen erhoben wurde, ob der Befragte Arbeit suchen und aus welchem Grund ggf. keine Arbeit gesucht werden muss.</p> <p>Bei der Vergabe von Sondercodes ab Welle 2 musste deshalb berücksichtigt werden, dass der Code -5 nicht belegt werden darf – dieser war in Welle 1 mit dem TNZ-Code belegt worden. Weiterhin musste sichergestellt werden, dass die Sondercodes analog zu Welle 1 codiert werden. Dabei ist es in Welle 2 zu Fehlern gekommen. Die in Welle 2 erhobenen Sondercodes in den Variablen PTK0200, PTK0500, PTK0900a, PTK1000a und PTK1100a wurden nicht mit -6 beginnend vercodet, sondern es wurde auch der Sondercode -5 genutzt. In PTK0200 wurde zudem der Sondercode 997 (noch nie) in Welle 2 nicht analog zu Welle 1 auf „0“ codiert.</p> <p>Diese Fehler wurden korrigiert. Die Codierung der Sondercodes für Welle 2 berücksichtigt jetzt, dass der Code -5 nicht belegt werden darf, da in Welle 1 bereits eine Bedeutung festgelegt wurde.</p> <p>Darüber hinaus wurde in den Wertelabels kenntlich gemacht, in welchen Wellen welche Sondercodes erhoben wurden.</p>
PTK0321f PTK0321g	PENDDAT	2	Korrektur	<p>In der 2. Welle wurden Gründe erhoben, warum ein Befragter keine Arbeit suchen muss. Die dabei erfassten offenen Angaben wurden in PTK0321a-f vercodet. Dabei wurden mit PTK0321f und PTK0321g auch zwei neue Kategorien angelegt. Bei diesen Variablen ist es in der Datenaufbereitung zu einem Fehler in den Variablenlabels gekommen – sie wurden zwischen den beiden Variablen vertauscht.</p> <p>Dieser Fehler wurde korrigiert. PTK0321f (keine Arbeitssuche, da in Ausbildung) und PTK0321g (keine Arbeitssuche, da erwerbstätig) sind jetzt richtig gelabelt.</p>

Tabelle 29: Übersicht über rückwirkende Veränderungen im Personendatensatz (PENDDAT) (Fortsetzung 4)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>PTK0321f</i> <i>PTK0321g</i>	<i>PENDDAT</i>	2	Korrektur	Bei <i>PTK0321f</i> und <i>PTK0321g</i> handelt es sich um zwei neue Kategorien, die bei der Vercodung der offenen Angaben auf ein Item mit Mehrfachantwort erstellt wurden. Dabei kam es zu einem Fehler in der Datenaufbereitung. Die WN-/KA-Angaben wurden nicht richtig in diese neu erstellten Variablen übertragen. Dieser Fehler wurde korrigiert. Wurden im Rahmen des Mehrfachantwortitems keine weiteren inhaltlichen Angaben gemacht, wurden die WN-/KA-Angaben auch in die im Rahmen der Vercodung der offenen Angaben erstellten Variablen für die neuen Kategorien übertragen.
<i>PEO0600*</i>	<i>PENDDAT</i>	1, 2	Korrektur	Die Variablen <i>PEO0600a-o</i> enthalten die Schulabschlüsse, die der Befragte für seine Kinder erwartet. Der erwartete Schulabschluss für ein Kind ist dabei immer in der Variable abgelegt, die der Position des Kindes in der Haushaltsstruktur der jeweiligen Welle entspricht. Bspw. ist die Information zum ersten Kind der ZP in <i>PEO0600c</i> abgelegt, wenn das erste Kind auf der dritten Position der Haushaltsstruktur einsortiert ist. Die Position entspricht dabei der <i>zplfd</i> (laufende Nummer der Person in der HH-Struktur in der jeweiligen Welle). Die Variablenlabel der Variablen waren hier missverständlich, da sie suggerierten, dass die Informationen zum ersten Kind in der ersten Variable zu finden wären (<i>PEO0600a</i>), während sie tatsächlich in der Variable abgelegt sind, die der Position des Kindes in der HH-Struktur entspricht (in Bsp. <i>PEO0600c</i> , da sich das Kind auf der dritten Position befindet). Die Variablenlabels wurden korrigiert, um einen eindeutigeren Hinweis auf die Belegung der Variablen zu geben.
<i>alg1s05</i>	<i>PENDDAT</i>	2	Korrektur	Die Variable <i>alg1s05</i> wurde für Welle 2 generiert. Für Welle 1 wurde eine analoge, harmonisierte Variable auf Basis von <i>PA0400</i> erstellt. Bei der Generierung der Variable in Welle 1 wurde die Kategorie „nein“ nicht wie in Welle 2 als „0“, sondern fälschlich als „2“ codiert. Dieser Fehler wurde behoben. Die Kategorie „nein“ ist jetzt für Welle 1 und 2 mit „0“ codiert.

Tabelle 29: Übersicht über rückwirkende Veränderungen im Personendatensatz (PENDDAT) (Fortsetzung 5)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>arbzeit</i>	<i>PENDDAT</i>	2	Korrektur	In die generierte Variable <i>arbzeit</i> werden die offenen und kategorialen Angaben zur Arbeitszeit integriert. Für den Fall unregelmäßiger Arbeitszeiten (<i>ET2100</i> = -5) und der kategorialen Angabe einer Arbeitszeit von 40h und mehr (<i>ET2200</i> = 5) wird der Median der offenen Angaben in <i>arbzeit</i> imputiert. Allerdings wurde dabei fälschlich der Median aller gültigen Werte ermittelt und imputiert. Dieser Fehler wurde korrigiert – es wird nun der Median der gültigen offenen Angaben von 40h und mehr imputiert.
<i>erwerb, erwerb2</i>	<i>PENDDAT</i>	1	Korrektur	Bei der Generierung der Erwerbsstatusvariable für Welle 1 (<i>erwerb</i>) werden abschließend unplausible Kombinationen auf -8 gesetzt. So wird bei Personen, die in <i>PB0100</i> angegeben haben, Schüler, Student oder Azubi zu sein bei gleichzeitiger Angabe, aktuell einer öffentlich geförderten Beschäftigung nachzugehen (<i>PET0400</i> =1), die generierte Variable <i>erwerb</i> auf -8 gesetzt, da es sich dabei um eine Unplausibilität handelt, die nicht auflösbar ist. Durch einen Fehler wurden allerdings nicht die Schüler (<i>PB0100</i> =1) mit gleichzeitiger öffentlicher Beschäftigung auf <i>erwerb</i> =-8 gesetzt, sondern die Fälle, in denen die Zielperson keine Angabe zu ihrem Status als Schüler, Student oder Auszubildende/r gemacht hatte (<i>PB0100</i> =-2). Dieser Fehler wurde behoben. Die Schüler mit gleichzeitiger öffentlicher Beschäftigung werden nun auf -8 gesetzt, die Fälle mit KA zur Angabe nach ihrem Status (<i>PB0100</i> =-2) bleiben auf <i>erwerb</i> =3 (öffentlich geförderte Beschäftigung), da für sie ja keine Unplausibilität vorliegt. Die auf <i>erwerb</i> basierende harmonisierte Variable <i>erwerb2</i> ändert sich in der Folge analog.
<i>siops, isei, mps</i>	<i>PENDDAT</i>	1	Korrektur	In Welle 1 wurden die MV-Codes aus der von GESIS erstellten ISCO88-Vercodung direkt in die auf dieser Variable basierenden Variablen <i>siops</i> , <i>isei</i> und <i>mps</i> übernommen. Ab Welle 2 wurden die MV-Codes -1, -2, -5, -6 und -8 aus der ISCO-Vercodung in <i>siops</i> , <i>isei</i> und <i>mps</i> nicht mehr ausdifferenziert, sondern einheitlich auf -5 (nicht codierbar) gesetzt. Einheitlich, werden die MV-Codes jetzt auch für Welle 1 nicht mehr ausdifferenziert, sondern in <i>siops</i> , <i>isei</i> und <i>mps</i> ebenfalls einheitlich auf -5 gesetzt.

Tabelle 29: Übersicht über rückwirkende Veränderungen im Personendatensatz (PENDDAT) (Fortsetzung 6)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>isei</i> <i>siops</i> <i>iseilewt</i> <i>siopslewt</i> <i>iseieewt</i> <i>siopseewt</i>	PENDDAT	2	Korrektur	<p>Die Variablen <i>isco88</i>, <i>isei</i>, <i>siops</i> und <i>mps</i> werden von ZUMA geliefert. Für <i>isco88=110</i> (Sondercode für Streitkräfte) wird von ZUMA keine <i>siops</i>- und <i>isei</i>-Zuordnung erzeugt. Dies wurde nun nachträglich ergänzt. Diese Korrektur für Streitkräfte in der aktuellen Erwerbstätigkeit hat auch Auswirkungen auf Prestigeskalen der ersten und letzten Erwerbstätigkeit.</p> <p>Die Prestigeskalen der ebenfalls auf der ISCO88-Verkodung basierenden Berufe der Väter und Mütter sind davon nicht betroffen, da deren Zuordnung ordnungsgemäß erfolgte.</p>
<i>migration</i>	PENDDAT	2	Korrektur	<p>Die generierte Variable <i>migration</i> enthält Angaben zum Migrationshintergrund des Befragten. In Welle 1 war die Variable für Senioreninterviews nicht generierbar, da nur Angaben zur eigenen Migration vorlagen, Informationen zum Migrationshintergrund von Eltern und Großeltern aber nicht erhoben wurden. Ab Welle 2 werden auch diese Angaben im Rahmen des Senioreninterviews erhoben. Die Generierung von <i>migration</i> ist also für Senioreninterviews ab Welle 2 möglich. Allerdings wurden diese Fälle für die in Welle 2 neubefragten Senioren fälschlich auf -10 (Item in Welle nicht erhoben) gesetzt. Weiterhin wäre die Variable für die wiederbefragten Senioren in W2 generierbar, wurde aber fälschlich ebenfalls mit -10 generiert.</p>
<i>vegp</i> <i>megp</i> <i>egplewt</i> <i>egpeewt</i>	PENDDAT	2	Korrektur	<p>Bei der Generierung der EGP-Werte für die erste und letzte Erwerbstätigkeit der Zielperson sowie für die Mutter und den Vater der Zielperson ist es zu einem Generierungsfehler gekommen. In diesen Variablen sollte (unter anderem) der Wert „-5“ (nicht generierbar) gesetzt werden, wenn die jeweils korrespondierende Angabe zur Stellung im Beruf „-5“ (d.h. nicht generierbar) war. Statt der jeweils korrespondierenden Stellung im Beruf wurde fälschlich in allen Fällen auf die Stellung im Beruf der aktuellen Erwerbstätigkeit zurückgegriffen. Dieser Fehler wurde behoben.</p>

Tabelle 29: Übersicht über rückwirkende Veränderungen im Personendatensatz (PENDDAT) (Fortsetzung 7)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>vkldb</i> <i>visco_it</i>	<i>PENDDAT</i>	2	Korrektur	<p>Ab Welle 2 wird auch die Bezeichnung der beruflichen Tätigkeit der Mutter und des Vaters erhoben, als der Befragte 15 Jahre alt war. Die in Welle 2 erhobenen Angaben wurden von TNS Infratest zunächst nach KIdB 92 vercodet (<i>vkldb</i>, <i>mkldb</i>). In einem zweiten Schritt erfolgte ein Umstieg von KIdB 92 auf ISCO 88 (<i>visco_it</i>, <i>misco_it</i>). Daneben wurde durch GESIS eine direkte Vercodung vorgenommen (<i>visco</i>, <i>misco</i>). In den von Infratest vercodeten Angaben taucht ein Fall mit einem ungültigen Code für <i>vkldb</i> auf (<i>vkldb</i>=7670). Von diesem Code erfolgte auch ein Umstieg auf einen ISCO-Code (<i>visco_it</i>=3449). Die offenen Angaben reichen hier eigentlich nicht zur Vercodung aus. Entsprechend wurde der Fall auch von GESIS nicht vercodet (<i>visco</i>=5).</p> <p>Der Code 7670 ist in der genutzten Klassifikation des Statistischen Bundesamts nicht vorgesehen. Weiterhin hätte kein KIdB-Code (und in der Folge auch kein ISCO-Code) vergeben werden dürfen, da die Angaben für die Vercodung nicht ausreichend waren. Die Werte in <i>vkldb</i> und <i>visco_it</i> wurden für diesen Fall mit „-5“ (nicht codierbar) überschrieben.</p>

Tabelle 30: Übersicht über rückwirkende Veränderungen in den Spelldaten auf Haushaltsebene (*alg2_spells*)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>AL21201a-e</i> <i>AL21202a-e</i> <i>AL21850a-e</i> <i>AL21851a-e</i> <i>AL21900a-e</i> <i>AL21901a-e</i> <i>AL22150a-e</i> <i>AL22170a-e*</i>	<i>alg2_spells</i>	1,2	Korrektur	<p>Innerhalb der ALG2-Episoden wurde erhoben, ob eine Kürzung des ALG2 stattgefunden hat. Dabei wurden unter anderem auch die Gründe für die Kürzung sowie Informationen, für welche Haushaltsmitglieder gekürzt wurde, erhoben. Nicht alle Informationen wurden allerdings in allen Welle erhoben, so dass bei gefüllten Kürzungsepisoden „-9“-Werte (Item in Welle nicht erhoben) auftreten können. Fälschlich wurden die „-9“ allerdings auch dann vergeben, wenn Kürzungsepisoden nicht gefüllt waren. Dieser Fehler wurde korrigiert. Die Variablen nicht belegter Kürzungsepisoden stehen jetzt einheitlich auf „-3“ (trifft nicht zu (Filter)), statt vereinzelt auch „-9“-Werte aufzuweisen.</p>

Tabelle 31: Übersicht über rückwirkende Veränderungen in den Spelldaten auf Personenebene (*et_spells*; *al_spells*; *lu_spells*; *mn_spells*)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>arbeitszeit</i>	<i>et_spells</i>	2	Korrektur	In die generierte Variable <i>arbeitszeit</i> werden die offenen und kategorialen Angaben zur Arbeitszeit integriert. Für den Fall unregelmäßiger Arbeitszeiten (<i>ET2100=5</i>) und der kategorialen Angabe einer Arbeitszeit von 40h und mehr (<i>ET2200=5</i>) wird der Median der offenen Angaben in <i>arbeitszeit</i> imputiert. Allerdings wird dabei fälschlich der Median aller gültigen Werte ermittelt und imputiert. Dieser Fehler wurde korrigiert – es wird nun der Median der gültigen offenen Angaben von 40h und mehr imputiert.
<i>isei</i> <i>siops</i>	<i>et_spells</i>	2	Korrektur	Die Variablen <i>isco88</i> , <i>isei</i> , <i>siops</i> und <i>mps</i> werden von ZUMA geliefert. Für <i>isco88=110</i> (Sondercode für Streitkräfte) wird von ZUMA keine <i>siops</i> - und <i>isei</i> -Zuordnung erzeugt. Dies wurde nun nachträglich ergänzt.
<i>MN0200f</i> <i>MN0201f</i>	<i>mn_spells</i>	2	Korrektur	Das Variablenlabel von <i>MN0200f</i> war fehlerhaft („Teil d. Progr.: Beschäftigung in Übergangsfirma“). Dieser Fehler wurde korrigiert, das Label lautet jetzt richtig „Teil d. Progr.: Beschäftigung in Übungsfirma“.

Tabelle 32: Übersicht über rückwirkende Veränderungen in Registerdatensätzen (*hh_register*; *p_register*)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>pnettok2</i> <i>pnettod2</i>	<i>p_register</i>	2	Korrektur	Im Rahmen des Widerspruchsverfahrens fiel auf, dass ein Haushalt doppelt in der Strichprobe war, und zwar als <i>hnr=12002409</i> sowie als <i>hnr=21006023</i> . Der Haushalt <i>hnr=21006023</i> ist ein Split von <i>hnr=12002409</i> in Welle 2, der offenbar nicht korrekt ist. Der SplitHH wurde in Welle 2 nicht befragt. Der SplitHH wurde entfernt, da die Splitbildung falsch war. Da keine rückwirkenden Änderungen den Haushaltsstrukturen vorgenommen werden, bleibt der UrHH in seiner in Welle 2 berichteten Form erhalten. Die fälschlich ausgezogene Person wird im Personenregister mit einem Sondercode markiert.

Tabelle 33: Übersicht über rückwirkende Veränderungen in den Gewichtungsdatensätzen (hweights; pweights)

Geänderte Variable	Betroffener Datensatz	Geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>prop_t0</i>	hweights	2	Korrektur	Infratest hatte in den Eingangsgewichten in der Variable <i>gew_ges</i> das Produkt aus 1/Designgewicht und Teilnahmepropensity geliefert. Irrtümlich ist dieses Produkt in die Variable <i>prop_t0</i> (Teilnahmepropensity) geschrieben worden. In der Korrektur wird nun zuvor durch das Designgewicht dividiert.

5.5 Anonymisierung

Vom IAB als besondere Dienststelle der Bundesagentur für Arbeit (BA) erhobene Daten stellen immer Sozialdaten dar, weshalb die Ansprüche an den Datenschutz entsprechend hoch sind. Daher wurde es notwendig, einige Variablen nur in vergrößerter Form in den Scientific Use File aufzunehmen. Diese Variablen sind im Variablenlabel generell mit dem Zusatz „anonymisiert“ gekennzeichnet. Aus dem gleichen Grund war es auch notwendig, auf verfügbare Regionalinformationen mit Ausnahme der Bundesländer und einer davon abgeleiteten Ost-/West-Information zu verzichten. Auch die Angaben zu den Verwandtschaftsbeziehungen im Haushalt sowie die Vornamen der Haushaltsmitglieder sind aus Datenschutzgründen nicht Teil des Scientific Use Files. Hinweise auf die Haushaltsstruktur geben allerdings generierte Variablen, bspw. zum Haushalts- und Bedarfsgemeinschaftstyp (*hhtyp*⁴³, *bgtyp*⁴⁴), Indikatorvariablen zu Partnern im Haushalt (*apartner*, *epartner*⁴⁵), Zeigervariablen auf Eltern und Partner im Haushalt (*zmhh*; *zvhh*; *zparthh*⁴⁶) sowie verschiedene Indikatorvariablen, die anzeigen ob Eltern (*mhh*; *vhh*⁴⁷) oder Kinder der Zielperson (z. B. *ekind*⁴⁸) im Haushalt leben.

Tabelle 34 gibt einen Überblick über die betroffenen Variablen und die vorgenommenen Anonymisierungsschritte⁴⁹ im Personendatensatz. In Tabelle 35 sind die anonymisierten Variablen des Erwerbstätigkeitsspelldatensatzes dargestellt.

⁴³ Enthalten im Haushaltsdatensatz (*HHENDDAT*), siehe Kapitel 4.5.2.

⁴⁴ Wellenspezifische Variablen enthalten im Personenregister (*p_register*), siehe Kapitel 4.4.

⁴⁵ Enthalten im Personendatensatz (*PENDDAT*), siehe Kapitel 4.4.

⁴⁶ Wellenspezifische Variablen enthalten im Personenregister (*p_register*), siehe Kapitel 4.4.

⁴⁷ Enthalten im Personendatensatz (*PENDDAT*), siehe Kapitel 4.4.

⁴⁸ Enthalten im Personendatensatz (*PENDDAT*), siehe Kapitel 4.4.

⁴⁹ Falls für ihr Forschungsvorhaben die Verwendung der nicht anonymisierten Versionen einer oder mehrere dieser Variablen unabdingbar sein sollte, möchten wir Sie bitten, sich an das Forschungsdatenzentrum zu wenden, um eine geeignete Datenzugangsmöglichkeit zu finden. Wie diese aussehen kann, wird vom Forschungsvorhaben und den dafür benötigten Variablen abhängen.

Tabelle 34: Übersicht über die in Welle 3 anonymisierten Variablen im Personendatensatz (PENDDAT)

Varname	Frage­nummer		Variablen­label	Vorgehen
	Personen­bogen	Senioren­bogen		
<i>PD0100</i>	P1	P1	Geburtsjahr (Geburtsdatum, anonymisiert)	Das genaue Geburtsdatum wurde zum Geburtsjahr verkürzt.
<i>gebhalbj</i>	generiert	generiert	Geburtshalb­jahr, generiert	Das genaue Geburtsdatum wurde zu einem Indikator für das 1. bzw. 2. Halb­jahr verkürzt.
<i>PET1210</i>	P84	n. in Fbvers	Letzte Stellung im Beruf, einfache Einteilung (vor Januar 2005) (anon.)	In der Befragung wurden aus erhebungstechnischen Gründen Berufs- und Zeitsoldaten separat erhoben. Aufgrund der geringen Fallzahlen und da diese ohnehin üblicherweise nicht im Rahmen der Stellung im Beruf erhoben werden, wurde diese Gruppe mit den Beamten und Richtern zu einer gemeinsamen Gruppe zusammengeführt.
<i>PET1250</i>	P87, P88	n. in Fbvers	Letzte StiB Beamter: det. Erfassung, inkl. Soldaten (vor Januar 2005) (anon.)	Diese Variable enthält zusätzliche Fälle. Die Berufs- und Zeitsoldaten aus P87 wurden der entsprechenden Beamtenkategorie zugeschlagen. Die Variable für Berufs- und Zeitsoldaten (P87) wird nicht ausgeliefert.
<i>PET1211</i>	generiert	n. in Fbvers	Letzte Stellung im Beruf, einfache Einteilung (inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu <i>PET1210</i> .
<i>PET1251</i>	generiert	n. in Fbvers	Letzte StiB Beamter: det. Erfassung, inkl. Soldaten (inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten (P87) wird nicht ausgeliefert.
<i>stiblewt</i>	generiert	n. in Fbvers	Stellung im Beruf, letzte Erwerbstätigkeit, Kennziffer, generiert	Bei der Generierung der Stellung im Beruf, werden die Berufs- und Zeitsoldaten der entsprechenden Beamtenkategorie zugeordnet.
<i>PET1510</i>	generiert	P12	Jetzige Stellung im Beruf, einf. Eint., Erh. ab W2 (anon.)	Vorgehen analog zu <i>PET1210</i> .

Tabelle 34: Übersicht über die in Welle 3 anonymisierten Variablen im Personendatensatz (PENDDAT) (Fortsetzung 1)

Varname	Fragennummer		Variablenlabel	Vorgehen
	Personenbogen	Seniorenbogen		
<i>PET1900</i>	generiert	P15, P16	Jetzige StiB Beamter: det. Erfassung, inkl. Soldaten (anon.)	Vorgehen analog zu <i>PET1250</i> . Die in den Senioreninterviews erhobene Variable für Berufs- und Zeitsoldaten (P15) wird nicht ausgeliefert. Für die Personeninterviews wird aus den Erwerbstätigkeitsspells (P47) keine generierte Variable für Berufs- und Zeitsoldaten in den Personendatensatz übernommen.
<i>stibkz</i>	generiert	generiert	Jetzige Stellung im Beruf, einf. Eint., harmonisiert (anonymisiert)	Bei der Generierung der Stellung im Beruf werden die Berufs- und Zeitsoldaten der entsprechenden Beamtenkategorie zugeordnet.
<i>stib</i>	generiert	generiert	Stellung im Beruf, Kennziffer, generiert	Vorgehen analog zu <i>stiblewt</i> .
<i>PET3300</i>	P93	n. in Fbvers	Erste Stellung im Beruf, einfache Einteilung (anon.)	Vorgehen analog zu <i>PET1210</i> .
<i>PET3700</i>	P96, P97	n. in Fbvers	Erste StiB Beamter: det. Erfassung, inkl. Soldaten	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten (P96) wird nicht ausgeliefert.
<i>PET3301</i>	generiert	n. in Fbvers	Erste Stellung im Beruf, einf. Eint. (zus.gef., inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu <i>PET1210</i> .
<i>PET3701</i>	generiert	n. in Fbvers	Erste StiB Beamter: det. Erf., inkl. Soldaten, (zus.gef., inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten (P96) wird nicht ausgeliefert.
<i>stibeewt</i>	generiert	n. in Fbvers	Stellung im Beruf, erste Erwerbstätigkeit, Kennziffer, generiert	Vorgehen analog zu <i>stiblewt</i> .

Tabelle 34: Übersicht über die in Welle 3 anonymisierten Variablen im Personendatensatz (PENDDAT) (Fortsetzung 2)

Varname	Fragennummer	Personen-		Variablenlabel	Vorgehen
		bogen	Senioren- bogen		
<i>PSH0320</i>	P281	n. in Fbvers		Damalige Stellung im Beruf der Mutter, einfache Einteilung (anon.)	Vorgehen analog zu <i>PET1210</i> .
<i>PSH0360</i>	P284, P285	n. in Fbvers		Damalige StiB d. Mutter Beamte, inkl. Soldaten: det. Erfassung (anon.)	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten (P284) wird nicht ausgeliefert.
<i>mstib</i>	generiert	n. in Fbvers		Stellung im Beruf der Mutter, Kennziffer, generiert	Vorgehen analog zu <i>stiblewt</i> .
<i>PSH0620</i>	P292	n. in Fbvers		Damalige Stellung im Beruf des Vaters, einfache Einteilung (anon.)	Vorgehen analog zu <i>PET1210</i> .
<i>PSH0660</i>	P295, P296	n. in Fbvers		Damalige StiB d. Vaters Beamter, incl. Soldaten: det. Erfassung (anon.)	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten (P284) wird nicht ausgeliefert.
<i>vstib</i>	generiert	n. in Fbvers		Stellung im Beruf des Vaters, Kennziffer, generiert	Vorgehen analog zu <i>stiblewt</i> .
<i>PMI0200</i>	P264		P73	Nicht in D geboren: Geburtsland	Länder mit zu geringen Fallzahlen wurden zu größeren Kategorien zusammengefasst.
<i>ogebland</i>	generiert		generiert	Geburtsland, inkl. offene Angaben, Kategorien (anonymisiert)	Vorgehen analog zu <i>PMI0200</i> .
<i>PMI0500</i>	P267		P76	Keine D Staatsangehörigkeit: Welche Staatsangehörigkeit? (anonymisiert)	Staatsangehörigkeiten von Ländern mit zu geringen Fallzahlen wurden zu größeren Kategorien zusammengefasst.
<i>ostaatan</i>	generiert		generiert	Staatsangehörigkeit, inkl. offene Angaben, Kategorien (anonymisiert)	Vorgehen analog zu <i>PMI0500</i> .

Tabelle 34: Übersicht über die in Welle 3 anonymisierten Variablen im Personendatensatz (PENDDAT) (Fortsetzung 3)

Varname	Fragennummer		Variablenlabel	Vorgehen
	Personenbogen	Seniorenbogen		
<i>PMI1000a</i>	P274a	P80a	Vater: Zuzugsland (anonymisiert)	Zuzugsländer mit zu geringen Fallzahlen wurden zu größeren Kategorien zusammengefasst.
<i>PMI1000b</i>	P274b	P80b	Mutter: Zuzugsland (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .
<i>PMI1000c</i>	P274c	P80c	Vater des Vaters: Zuzugsland (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .
<i>PMI1000d</i>	P274d	P80d	Mutter des Vaters: Zuzugsland (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .
<i>PMI1000e</i>	P274e	P80e	Vater der Mutter: Zuzugsland (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .
<i>PMI1000f</i>	P274f	P80f	Mutter der Mutter: Zuzugsland (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulanda</i>	generiert	generiert	Vater: Zuzugsland, inkl. offene Angaben, Kategorien (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulandb</i>	generiert	generiert	Mutter: Zuzugsland, inkl. offene Angaben, Kategorien (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulandc</i>	generiert	generiert	Vater des Vaters: Zuzugsland, inkl. offene Angaben, Kategorien (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulandd</i>	generiert	generiert	Mutter des Vaters: Zuzugsland, inkl. offene Angaben, Kategorien (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .

Tabelle 34: Übersicht über die in Welle 3 anonymisierten Variablen im Personendatensatz (PENDDAT) (Fortsetzung 4)

Varname	Fragennummer		Variablenlabel	Vorgehen
	Personenbogen	Seniorenbogen		
<i>ozulande</i>	generiert	generiert	Vater der Mutter: Zuzugsland, inkl. offene Angaben, Kategorien (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulandf</i>	generiert	generiert	Mutter der Mutter: Zuzugsland, inkl. offene Angaben, Kategorien (anonymisiert)	Vorgehen analog zu <i>PMI1000a</i> .

Tabelle 35: Übersicht über die in Welle 3 anonymisierten Variablen im Erwerbstätigkeitsspelldatensatz (*et_spells*)

Varname	Fragennummer		Variablenlabel	Vorgehen
	Personenbogen	Seniorenbogen		
<i>ET0601</i>	P44		Stellung im Beruf, einfache Einteilung (anon.)	Vorgehen analog zu <i>PET1210</i> .
<i>ET1001</i>	P47, P48		StiB Beamter: detaillierte Erfassung (anon.)	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten (P47) wird nicht ausgeliefert.
<i>stib</i>	generiert		Stellung im Beruf, Kennziffer, generiert	Vorgehen analog zu <i>stiblewt</i> .

5.6 Arbeitslosengeld-II-Bezüge

Bereits in der 1. und 2. Welle wurden die Arbeitslosengeld-II-Bezüge auf Haushaltsebene in Episodenform erfasst. Dieses Konzept wurde in Welle 3 mit einem leicht überarbeiteten Frageprogramm fortgeführt. Neben Formulierungsänderungen wurde die Frage nach den Gründen für das Ende des Arbeitslosengeld-II-Bezugs neu aufgenommen (Z 1 in beiden Versionen des Haushaltsbogens; Variablen *AL22200a* bis *AL22200f* bzw. mit zugedodeten offenen Angaben *AL22201a* bis *AL22201f*).

5.6.1 Konzept für die Fortschreibung der in der Vorwelle aktuell andauernden Arbeitslosengeld-II-Spells

Für die Fortschreibung, der in der Vorwelle andauernden und damit im Spelldatensatz rechtszensierten Arbeitslosengeld-II-Bezüge sind in beiden Versionen des Haushaltsfragebogens Dependent-Interviewing-Fragen enthalten (HH91 in Haushaltsfragebogen für wiederbefragte Haushalt; HH48 im Haushaltsfragebogen für abgespaltene und Neubefragte Haushalte). Im Falle einer Aufspaltung des Vorwellenhaushalts werden die zensierten Arbeitslosengeld-II-Spells jeweils durch den Teil des Haushalts fortgeschrieben, in dem sich die Person befindet, mit der in der Vorwelle das Haushaltsinterview geführt wurde (im Folgenden kurz „HBV“⁵⁰ genannt).

Ist der HBV Teil des Haushalts, der zuerst unter der alten Adresse/Telefonnummer erreicht wird, erfolgt die Fortschreibung über die Angabe im Haushaltsbogen für wiederbefragte Haushalte (kurz „HHalt“). Anders jedoch, wenn der unter der alten Adresse/Telefonnummer erreichte Teil des Haushalts die Auskunft gibt, der HBV sei ausgezogen/ein Jahr oder länger nicht anwesend oder am Wohnort der Vorwelle verblieben. In diesen Fällen wird der abgespaltene Teil des Haushalts als eigener Befragungshaushalt verstanden und mit dem Fragebogen für Neubefragte Haushalte (kurz „HHneu“) interviewt. Ist der HBV der Vorwelle Teil dieses abgespaltenen Haushalts, wird der in der Vorwelle andauernde Arbeitslosengeld-II-Bezug des Ursprungshaushalts durch die Angaben des HBV im abgespaltenen Haushaltsteil fortgeschrieben.

Unterschiede bestehen in wiederholt befragten bzw. abgespaltenen Haushalten auch hinsichtlich des Zeitraums, für den in der 3. Welle Arbeitslosengeld-II-Bezüge erhoben werden. Auch hier ist wieder von Bedeutung, ob der HBV der Vorwelle im Haushalt lebt. Lebt der HBV der Vorwelle im Haushalt, werden Arbeitslosengeld-II-Bezüge seit dem Interviewdatum der Vorwelle erhoben. Lebt der HBV der Vorwelle nicht im Haushalt, werden nur Arbeitslosengeld-II-Bezüge seit dem Auszugsdatum des HBV bzw. dem Datum des Auszugs aus dem gemeinsamen Haushalt mit dem HBV erhoben.

Die in Welle 3 erstmals befragten Haushalte der Auffrischungsstichprobe wurden nach dem Bezug von Arbeitslosengeld II in der Zeit seit der letzten Veränderung des Haushalts gefragt. Lag diese Veränderung vor Januar 2007 oder wurde dazu keine Angabe gemacht, wurden die Arbeitslosengeld-II-Bezüge des Haushalts seit Januar 2007 erhoben.

5.6.2 Struktur des Arbeitslosengeld-II-Spelldatensatzes

Durch die Integration der in Welle 3 berichteten Arbeitslosengeld-II-Bezüge verändern sich die Struktur und der Inhalt des Arbeitslosengeld-II-Spelldatensatzes. Dabei ist zu unterscheiden zwischen (1) neuen Variablen, die auf eine bestimmte Welle bezogen sind, (2) neuen Variablen, die nicht auf eine bestimmte Welle bezogen sind und (3) in Welle 3 nicht mehr erhobenen Variablen.

⁵⁰ HBV steht als Abkürzung für „Haushaltsbevollmächtigter“.

1. Auch in Welle 3 wurden mit *AL20602*, *AL20702a bis AL20702o*, *AL20802* und *AL20902* neue, wellenspezifische Querschnittvariablen in den Arbeitslosengeld-II-Spelldatensatz eingefügt. Diese Variablen sind auf den Interviewzeitpunkt der 3. Welle bezogen. Für die Befragungszeitpunkte der Vorwellen existieren ebenfalls Querschnittvariablen, in denen die auf die jeweilige Welle bezogenen, analogen Informationen abgelegt sind. Tabelle 36 gibt einen Überblick über die im Arbeitslosengeld-II-Spelldatensatz enthaltenen Querschnittinformationen.

Tabelle 36: Querschnittvariablen im ALG2-Spelldatensatz (*alg2_spells*)

	Querschnittvariable mit Information bezogen auf ...		
	Welle 1	Welle 2	Welle 3
Erhält der HH ALG2 für alle HH-Mitglieder?	<i>AL20600</i>	<i>AL20601</i>	<i>AL20602</i>
Erhält der HH ALG2 für die Personen 1 bis 15?	<i>AL20700a bis AL20700o</i>	<i>AL20701a bis AL20701o</i>	<i>AL20702a bis AL20702o</i>
Höhe des monatlichen ALG2-Bezugs?	<i>AL20800</i>	<i>AL20801</i>	<i>AL20802</i>
Hat eine Kürzung des ALG2 begonnen?	<i>AL20900</i>	<i>AL20901</i>	<i>AL20902</i>

2. Eingebettet in den Arbeitslosengeld-II-Bezugsepisoden finden sich Informationen zu Zeiten, in denen Kürzungen des Leistungsbezugs stattgefunden haben. Bis Welle 2 kamen innerhalb einer Bezugsepisode bis zu fünf Kürzungen vor. Im Rahmen der Fortschreibung der Arbeitslosengeld-II-Bezüge, die in der Vorwelle zensiert waren, werden auch Informationen zu seit dem letzten Interview neu begonnenen Kürzungen erhoben. Diese neuen Kürzungen werden in die bestehenden, fortzuschreibenden Arbeitslosengeld-II-Episoden übertragen. Da die bestehende maximale Anzahl von fünf Kürzungen pro Bezugsepisode durch die erneute Fortschreibung nicht mehr ausreichte, wurde eine zusätzliche, sechste Kürzung eingefügt, die das Kürzel „f“ trägt⁵¹. Darüber hinaus entspricht die Datenstruktur der Datenstruktur aus Welle 2.
3. Der Kürzungsgrund *AL21900a bis AL21900e* wurde auch in Welle 3 nicht erhoben. Den Variablen *AL21901a bis AL21901f* wurden entsprechend auch keine offenen Angaben mehr zugecodet.

⁵¹ Die zu einer Kürzung gehörenden Variablen sind über einen Buchstaben am Ende der Variable zu erkennen. Kürzungsvariablen die sich auf die erste Kürzung beziehen, enden mit einem „a“, die auf die zweite Kürzung bezogenen mit einem „b“ usw.

5.6.3 Plausibilitätsprüfung und –bereinigung des Arbeitslosengeld-II-Spelldatensatzes

Wie in Welle 1 und 2 wurden die Information zu Arbeitslosengeld-II-Bezügen auch in der 3. Welle einer Reihe von Plausibilitätsprüfungen unterzogen. Unzulässige Überschneidungen und Datierungen der Arbeitslosengeld-II-Spells bzw. der Kürzungsepisoden wurden dabei ggf. bereinigt. Veränderungen wurden dabei grundsätzlich nur in den generierten Datumvariablen (*bmonat*; *bjahr*; *emonat*; *ejahr*) des Arbeitslosengeld-II-Spells, der Kürzungsepisoden (*alg2kbm*; *alg2kbj*; *alg2kem*; *alg2kej*) sowie im Zensierungsindikator des Arbeitslosengeld-II-Spells (zensiert) vorgenommen. Konnten Unplausibilitäten nicht durch Korrektur der Datumvariablen bereinigt werden, wurden in geringem Umfang Arbeitslosengeld-II-Spells zusammengeführt bzw. Arbeitslosengeld-II-Spells bzw. Kürzungsepisoden gänzlich gelöscht.

5.6.4 Fortschreibung des Arbeitslosengeld-II-Spelldatensatzes

Im Anschluss an die Umstrukturierung der in Welle 3 berichteten Arbeitslosengeld-II-Episoden in das Spellformat und die Plausibilitätsprüfungen und -bereinigungen, im Zuge derer unzulässige Überschneidungen und Spells mit unplausiblen Datierungen bereinigt wurden, werden die zum Befragungszeitpunkt der Vorwelle andauernden Arbeitslosengeld-II-Bezüge mit den in Welle 3 erhobenen Informationen fortgeschrieben. Dabei sind drei Varianten zu unterscheiden, bei denen in den ersten beiden (1) und (2) lediglich der Zensierungsindikator zensiert verändert wird. Die dritte Variante (3) stellt eine Fortschreibung des in der Vorwelle zensierten Spells mit Welle 3 Informationen im engeren Sinne dar. Dabei werden neben dem Zensierungsindikator auch die generierten und erhobenen Enddaten, die wellenspezifischen Querschnittinformationen (siehe oben) sowie Informationen zu neuen Kürzungsepisoden in den in der Vorwelle andauernden Arbeitslosengeld-II-Spell integriert. Neben der Fortschreibung von in der Vorwelle zensierten Spells werden dem Spelldatensatz neue, in Welle 3 berichtete Spells zugespielt (4). Diese vier Varianten sollen im Folgenden kurz dargestellt werden:

1. *Fälle, in denen der HBV der Vorwelle nicht mehr im Haushalt lebt und auch kein Mitglied eines abgespaltenen, in der aktuellen Welle befragten Haushaltsteils ist.*

Um zu vermeiden, dass die aus der Vorwelle stammenden zensierten Arbeitslosengeld-II-Spells des Ursprungshaushalts auch weiterhin als aktueller Bezug dieses Haushalts gewertet werden, wurde der Zensierungsindikator in diesen Fällen auf zensiert „-5“ (HBV der Vorwelle nicht im HH und nicht interviewt) gesetzt. Auch in Fällen, in denen der HBV der Vorwelle verstorben war, wurde zensiert auf „-5“ gesetzt. Die berichteten und generierten Variablen zum Enddatum des Spells (*AL20300*, *AL20400* sowie *emonat*, *ejahr*) sowie die Nachfrage, ob ein Spell noch andauert (*AL20500*), bleiben dabei unverändert.⁵²

⁵² Das berichtete Enddatum ist also weiterhin mit dem Interviewdatum der Welle, in dem der Spell zensiert war, bzw. dem Sondercode „0“ für andauernde Spells gefüllt. Auch die Nachfrage, ob der Spell noch andauert (für den Fall, dass das Enddatum dem Interviewdatum entspricht), wird nicht verändert. Die generierten Datumvariablen enthalten weiterhin die letzte gültige Information, die hier dem Interviewdatum der Welle entspricht, in welcher der Spell zensiert war.

2. *Fälle, in denen der Haushalt in Welle 3 einem andauernden Arbeitslosengeld-II-Bezug zum Befragungszeit der Vorwelle widerspricht.*

Hat der Haushalt der Information, dass zum Zeitpunkt der Vorwelle ein aktuell andauernder Arbeitslosengeld-II-Bezug vorlag, in der Fortschreibungsfrage (HH91 in HHalt; HH48 in HHneu) explizit bzw. durch Angabe eines vor dem Interviewdatum der Vorwelle liegenden Enddatums implizit widersprochen, so wurde *zensiert* auf „2“ (nein) gesetzt. Es wird davon ausgegangen, dass die zum Interviewdatum der Vorwelle gegebene Information korrekt war. Da über das Interviewdatum der Vorwelle hinaus keine verlässlichen Aussagen über die weitere Dauer des Bezugs möglich sind, wird angenommen, dass dieser im Interviewmonat der Vorwelle geendet hat. Die berichteten und generierten Variablen zum Enddatum des Spells (*AL20300*, *AL20400* sowie *emonat*, *ejahr*) sowie die Nachfrage, ob ein Spell noch andauert (*AL20500*), bleiben dabei unverändert.⁵³ Das generierte Enddatum des Arbeitslosengeld-II-Spells (*emonat*; *ejahr*) wurde bereits in der Vorwelle auf das Interviewdatum der Vorwelle gesetzt.

3. *Fälle, in denen der Haushalt Angaben zum Enddatum eines in der Vorwelle aktuell andauernden Bezugs macht.*

Liegen in Welle 3 Angaben zum Enddatum eines in der Vorwelle zensierten Arbeitslosengeld-II-Bezugs vor, so wurde der in der Vorwelle zensierte Spell mit den aktuellen Angaben fortgeschrieben. Zunächst wurden das erhobene Enddatum (*AL20300*; *AL20400*), das generierte Enddatum (*emonat*; *ejahr*), die Nachfrage, ob der Arbeitslosengeld-II-Bezug aktuell noch andauert (*AL20500*) sowie der Zensierungsindikator (zensiert) mit den in der Vorwelle erhobenen Informationen überschrieben. Weiterhin wurden die in der 3. Welle berichteten Kürzungsepisoden sowie die auf die Welle 3 bezogenen Querschnittinformationen (*AL20602*; *AL20702a bis AL20702o*, *AL20802*, *AL20902*) übernommen.

4. *In Welle 3 neu berichtete Arbeitslosengeld-II-Spells, die keine in der Vorwelle zensierte Episode fortschreiben.*

In Welle 3 neu berichtete Spells wurden dem Arbeitslosengeld-II-Spelldatensatz angehängt. Danach wurde der Spellzähler neu generiert, um eine lückenlose Variable *spellnr* zu schaffen.

5.7 Erwerbsbiografien

Bereits in der 2. Welle wurden die Erwerbstätigkeits-, Arbeitslosigkeits- und Lückenphasen auf Personenebene in Episodenform erfasst. Dieses Konzept wurde in Welle 3 fortgeführt. Neben den für die Fortschreibung der zum Interviewdatum der 2. Welle aktuell andauernden Erwerbstätigkeitsspells (ET-Spells) und Arbeitslosigkeitsspells (AL-Spells) notwendigen Anpassungen wurden inhaltliche Erweiterungen sowie kleinere Korrekturen und Erweiterungen

⁵³ Hier gilt dasselbe. Verändert wird nur der Zensierungsindikator. Das berichtete Enddatum, die Nachfrage für andauernde Spells und das generierte Enddatum bleiben unverändert.

in Reaktion auf die Erfahrungen der 2. Welle vorgenommen. Für Personen, die in Welle 3 erstmals nach ihrer Erwerbsbiografie gefragt wurden, wurde das Referenzdatum für den Beginn des Retrospektivintervalls angepasst. In Welle 3 sollten hier alle Erwerbstätigkeits- und Arbeitslosigkeitsepisoden seit Januar 2006 berichtet werden (in Welle 2: Januar 2005). Personen, die bereits in der Vorwelle zu ihrer Erwerbsbiografie befragt wurden, sollten dagegen alle neuen Episoden seit dem Datum des letzten Interviews berichten.

5.7.1 Konzept für die Fortschreibung der in der Vorwelle aktuell andauernden Spells

In der 3. Welle wurden andauernde ET- und AL-Spells aus Welle 2 fortgeschrieben. Nicht fortgeschrieben wurden Lückenspells (LU-Spells), da diese die Lücken zum jeweiligen Interviewzeitpunkt abbilden sollen und somit in der Fragebogenkonzeption nicht wellenübergreifend angelegt sind. Für die Fortschreibung der in der Vorwelle andauernden und damit im Spelldatensatz rechtszensierten ET- und AL-Spells sind in den Personenfragebögen Dependent-Interviewing-Fragen enthalten (E38 für ET-Spells und A106 für AL-Spells). Es konnten bis zu zwei ET-Spells und ein AL-Spell aus der Vorwelle fortgeschrieben werden. Bei Befragten mit mehr als zwei zum Interviewzeitpunkt andauernden ET-Spells wurden jeweils die Erwerbstätigkeiten mit dem höchsten Stundenumfang fortgeschrieben.

5.7.2 Struktur der Spelldatensätze

Der LU-Spelldatensatz bleibt hinsichtlich der Struktur und der enthaltenen Variablen gegenüber Welle 2 unverändert. Durch die Integration der in Welle 3 berichteten Erwerbstätigkeits- und Arbeitslosigkeitsphasen in die Spelldatensätze der Vorwelle werden der ET- und der AL-Spelldatensatz um neue Variablen erweitert. Dabei ist zu unterscheiden zwischen (1) neuen Variablen, die auf eine bestimmte Welle bezogen sind, und (2) neuen Variablen, die nicht auf eine bestimmte Welle bezogen sind.

1. Im ET-Spelldatensatz werden die Variablen *ET0600 bis ET2200* als wellenspezifische, auf Welle 2 bezogene Querschnittinformationen verstanden. Für die in Welle 3 erhobenen Querschnittinformationen wurden analog die neuen Variablen *ET0601 bis ET2201* in den ET-Spelldatensatz eingefügt. Tabelle 37 gibt einen Überblick über die im ET-Spelldatensatz enthaltenen Querschnittinformationen.

Tabelle 37: Querschnittvariablen im ET-Spelldatensatz (*et_spells*)

Querschnittvariable mit Information bezogen auf ...			
	Welle 1	Welle 2	Welle 3
Stellung im Beruf (einfache und detaillierte Einteilung)	<i>(keine ET-Spells)</i>	ET0600 ET0700 ET0800 ET1000 ET1100 ET1200	ET0601 ET0701 ET0801 ET1001 ET1101 ET1201
Aufsichtungsfunktion; Zahl der beaufsichtigten Arbeitskräfte	<i>(keine ET-Spells)</i>	ET1300 ET1400	ET1301 ET1401
Entfristung einer anfangs unbe- fristeten Tätigkeit	<i>(keine ET-Spells)</i>	ET1700	ET1701
Arbeitszeit (vertraglich; tatsächlich; durch- schnittlich bei unregelmäßiger Arbeitszeit)	<i>(keine ET-Spells)</i>	ET2000 ET2100 ET2200	ET2001 ET2101 ET2201

Im AL-Spelldatensatz wird die Variable *AL 1300* als wellenspezifische, auf Welle 2 bezogene Querschnittinformationen verstanden. Für die in Welle 3 erhobene Querschnittinformation wurde analog die neue Variable *AL 1301* in den AL-Spelldatensatz eingefügt. Tabelle 38 gibt einen Überblick über die im Spelldatensatz enthaltenen Querschnittinformationen.

Tabelle 38: Querschnittvariablen im AL-Spelldatensatz (*al_spells*)

Querschnittvariable mit Information bezogen auf ...			
	Welle 1	Welle 2	Welle 3
Höhe des monatlichen ALG1-Bezugs?	<i>(keine Spells)</i>	AL1300	AL1301

- Die nicht wellenspezifische Variable *ET2400* (Wie hat die Person zuerst von der neuen Stelle erfahren?) und die dazugehörige Variable inklusive Vercodung *ET2401* sind in Welle 3 erstmals erhoben und in den ET-Spelldatensatz integriert worden.

5.7.3 Plausibilitätsprüfung und –bereinigung der Spelldatensätze

Im Lückenmodul konnten vom Befragten verschiedene Arten von Angaben gemacht werden, um Lücken von mehr als drei Monaten in der Erwerbsbiografie zu schließen. Die Datumangaben bereits aufgenommener Episoden konnten entweder korrigiert werden, oder der Befragte konnte Episoden ergänzen, die zuvor vergessen worden waren (Erwerbstätigkeiten oder Arbeitslosigkeiten) oder die erst im Rahmen des Lückenmoduls erhoben wurden (Nichterwerbstätigkeitsperioden).

Die im Lückenmodul berichteten Datumkorrekturen wurden zunächst in den jeweiligen Spelldaten umgesetzt. Anschließend wurden im Lückenmodul berichtete ET- oder AL-Spells dem ET- bzw. AL-Spelldatensatz zugeordnet und dort weiter verarbeitet. Neben den ET- und AL-Spells, die im Lückenmodul direkt als solche berichtet wurden (kategoriale Angabe), wurden bei der Vercodung der offenen Angaben weitere Lückenepisoden als ET bzw. AL-Spells identifiziert und dem jeweiligen Spelldatensatz zugeordnet.

Auf Personenebene wurden die Spelldatensätze zu Erwerbstätigkeiten, Arbeitslosigkeiten und der Lückendatensatz plausibilitätsgeprüft und ggf. bereinigt. Prüfungen fanden dabei nur innerhalb eines Spelldatensatzes statt. Datensatzübergreifende Prüfungen wurden nicht durchgeführt. Bereinigungen und Umcodierungen wurden, wie bei den Arbeitslosengeld-II-Spelldaten, nur in den generierten Datumvariablen vorgenommen. Auch hier wurden Jahreszeitangaben in konkrete Monatsangaben umcodiert, „-8“ Werte für un plausible Angaben gesetzt oder Datumsangaben ersetzt oder plausibilisiert. Da lediglich die generierten Datumvariablen editiert wurden, stehen dem Nutzer mit den Datumvariablen *ET0100-ET0400*, *AL0100-AL0400* bzw. *AL0800-AL1100* und *LU0200-LU0500* die ursprünglich erhobenen Informationen zur Verfügung und erlauben so die Durchführung eigener Prüfungen und Bereinigungen.

Darüber hinaus erschien es in einigen Fällen geboten, ganze Episoden zu löschen. Diese Löschungen sind zu einem Großteil auf Fehlfunktionen des Lückenmoduls zurückzuführen. So wurden bspw. im Lückenmodul weitere Episoden für eine Person aufgenommen, obwohl bereits der gesamte retrospektive Zeitraum abgedeckt war. Oder Interviewer haben durch Bedienungsfehler, anstatt die vorhandene Korrekturfunktion zu nutzen, inhaltlich nahezu identische Episoden mehrfach aufgenommen. Ebenfalls gelöscht wurden Episoden, die komplett außerhalb des abgefragten Zeitraums liegen und trotzdem erhoben wurden.

5.7.4 Fortschreibung der Spelldatensätze ET und AL

Im Anschluss an die Umstrukturierung der in Welle 3 berichteten Episoden ins Spellformat und die Plausibilitätsprüfungen und -bereinigungen, im Zuge derer unzulässige Überschneidungen und Spells mit un plausiblen Datierungen bereinigt wurden, werden die zum Befragungszeitpunkt der Vorwelle andauernden ET- bzw. AL-Spells mit den in Welle 3 erhobenen Informationen fortgeschrieben.

Dabei sind drei Varianten zu unterscheiden, bei denen in der ersten (1) lediglich der Zensierungsindikator *zensiert* verändert wird. Die zweite Variante (2) stellt eine Fortschreibung des in der Vorwelle zensierten Spells mit Welle 3-Informationen im engeren Sinne dar. Dabei werden neben dem Zensierungsindikator auch die generierten und erhobenen Enddaten und die wellenspezifischen Querschnittinformationen (siehe oben) in den in der Vorwelle andauernden Spell integriert.

Neben der Fortschreibung von in der Vorwelle zensierten Spells werden dem Spelldatensatz neue, in Welle 3 berichtete Spells zugespielt (3). Diese drei Varianten sollen im Folgenden kurz dargestellt werden:

1. *Fälle, in denen die Person in Welle 3 einem andauernden Spell zum Befragungszeitpunkt der Welle 2 widerspricht.*

Hat die Person der Information, dass zum Zeitpunkt der Vorwelle ein aktuell andauernder ET- oder AL-Spell vorlag, in der Fortschreibungsfrage (E38 für ET-Spells und A106 für AL-Spells) explizit bzw. durch Angabe eines vor dem Interviewdatum der Welle 2 liegenden Enddatums implizit widersprochen, so wurde der Zensierungsindikator *zensiert* auf „2“ (nein) gesetzt. Es wird davon ausgegangen, dass die zum Interviewdatum der Vorwelle gegebene Information korrekt war. Da über das Interviewdatum der Welle 2 hinaus keine verlässlichen Aussagen über die weitere Dauer des Spells möglich sind, wird angenommen, dass dieser im Interviewmonat der Welle 2 geendet hat. Die berichteten und generierten Variablen zum Enddatum des Spells (*ET0300*, *ET0400* bzw. *AL0300*, *AL0400* sowie *emonat*, *ejahr*) sowie die Nachfrage, ob ein Spell noch andauert (*ET0500* bzw. *AL0500*), bleiben dabei unverändert⁵⁴. Das generierte Enddatum des Spells (*emonat*; *ejahr*) wurde bereits in der Vorwelle auf das Interviewdatum der 2. Welle gesetzt.

2. *Fälle, in denen die Person Angaben zum Enddatum eines in der Vorwelle aktuell andauernden Spells macht.*

Liegen in Welle 3 Angaben zum Enddatum eines in Welle 2 zensierten ET- oder AL-Spells vor, so wurde der in Welle 2 zensierte Spell mit den aktuellen Angaben fortgeschrieben. Bei ET-Spells wurden das erhobene Enddatum (*ET0300*; *ET0400*), das generierte Enddatum (*emonat*; *ejahr*), die Nachfrage, ob der Spell aktuell noch andauert (*ET0500*), der Grund für die Auflösung des Arbeitsvertrags (*ET2300*), die generierten Variablen zur Stellung im Beruf und der wöchentlichen Arbeitszeit (*stib*, *arbzeit*) sowie der Zensierungsindikator (zensiert) mit den in Welle 3 erhobenen Informationen überschrieben. Weiterhin wurden die auf die Welle 3 bezogenen Querschnittinformationen (*ET0601* bis *ET2201*) übernommen.

⁵⁴ Das berichtete Enddatum ist also weiterhin mit dem Interviewdatum der Welle, in dem der Spell zensiert war bzw. dem Sondercode „0“ für andauernde Spells gefüllt. Auch die Nachfrage, ob der Spell noch andauert (für den Fall, dass das Enddatum dem Interviewdatum entspricht), wird nicht verändert. Die generierten Datumvariablen enthalten weiterhin die letzte gültige Information, die hier dem Interviewdatum der Welle entspricht, in welcher der Spell zensiert war.

Bei AL-Spells wurden das erhobene Enddatum (*AL0300; AL0400*), das generierte Enddatum (*emonat; ejahr*), die Nachfrage, ob der Spell aktuell noch andauert (*AL0500*), der Grund für die Beendigung der Arbeitslosigkeit (*AL0600, AL0601*) sowie der Zensierungsindikator (zensiert) mit den in Welle 3 erhobenen Informationen überschrieben. Weiterhin wurden die auf die Welle 3 bezogene Querschnittinformation (*AL1301*) übernommen. Bei den AL-Spelldaten gibt es zudem noch die Besonderheit, dass innerhalb des AL-Spells die Arbeitslosengeld-I-Bezugsepisode (*ALG1-Bezug*) erhoben wird. Welche Informationen fortgeschrieben werden, ist abhängig davon, ob in diesem AL-Spell zu einem früheren Zeitpunkt schon einmal *ALG1* bezogen wurde und ob dieser Bezug in der Vorwelle noch angedauert hat:

- Lag in der Vorwelle in dem fortzuschreibenden AL-Spell auch ein andauernder Bezug von *ALG1* vor, wurde das Ende des Bezugs erhoben⁵⁵. In diesem Fall wurden das erhobene Enddatum des Bezugs (*AL1000, AL1100*), der Indikator, ob der Bezug noch andauert (*AL1200*), das generierte Enddatum des Bezugs (*alg1em, alg1ej*) sowie der Zensierungsindikator des Bezugs (*alg1akt*) mit den in Welle 3 erhobenen Informationen überschrieben.
- War in dem fortzuschreibenden AL-Spell noch nie *ALG1* bezogen worden, wurde gefragt, ob seit dem Interviewdatum der Vorwelle ein *ALG1*-Bezug stattgefunden hat. Lag ein Bezug vor, wurden alle Informationen zum *ALG1*-Bezug mit den in Welle 3 erhobenen Informationen überschrieben. Neben dem Indikator, ob im AL-Spell schon einmal *ALG1* bezogen wurde (*AL0700*), wurden das berichtete Beginn- und Enddatum (*AL0800, AL0900, AL1000, AL1100*), der Indikator für andauernden Bezug (*AL1200*) sowie die entsprechenden generierten Variablen (*alg1bm, alg1bj, alg1em, alg1ej, alg1akt*) mit den neu erhobenen Informationen ersetzt.
- Hat in dem fortzuschreibenden AL-Spell in der Vergangenheit ein *ALG1*-Bezug stattgefunden, der aber bereits in der Vorwelle beendet war oder hat die Person in der Vorwelle keine Angaben zum *ALG1*-Bezug gemacht, wurde innerhalb dieses AL-Meldungsspells nicht noch einmal nach dem Bezug von *ALG1* gefragt. Entsprechend wurden in diesen Spells auch keine Veränderungen vorgenommen.

3. In Welle 3 neu berichtete ET- oder AL-Spells, die keine in Welle 2 zensierte Episode fortzuschreiben.

In Welle 3 neu berichtete Spells wurden dem jeweiligen Spelldatensatz angehängt. Danach wurde der Spellzähler neu generiert, um eine lückenlose Variable *spellnr* zu schaffen.

Das Fortschreiben der Spelldatensätze hat keine Auswirkung auf die Spellnummer der bereits im SUF zu Welle 2 enthaltenen Spells. Diese Spells behalten ihre Spellnummer. Die neuen Spells aus Welle 3 werden an den entsprechenden Datensatz angehängt und die Spellnummer fortgeschrieben.

⁵⁵ Sofern der Befragte den Bezug in der Vorwelle bestätigt hat (siehe A 112a im Personenbogen Welle 3).

5.7.5 Fortschreibung des LU-Spelldatensatzes

Bei der Fortschreibung der LU-Spells wurden im Gegensatz zu ET- und AL-Spells keine Aufsatzspells aus Welle 2 in Welle 3 fortgeführt. Die Integration der in Welle 3 berichteten Lückenspells zu den Lückenspells aus Welle 2 erfolgte deshalb modifiziert. Lag im LU-Spelldatensatz der Welle 2 eine zensierte Episode vor, so wurde geprüft, ob in Welle 3 eine Episode gleichen Typs berichtet wurde, in der das Interviewdatum der Vorwelle liegt. Falls dies der Fall war, wurden die Spells verschmolzen, d.h. der Spell aus Welle 2 übernahm das generierte Enddatum (*emonat und ejahr*) und das berichtete Enddatum (*LU0400 und LU0500*) sowie den Zensierungsindikator zensiert des Spells aus Welle 3. Bei zensierten Lückenspells aus Welle 2, für die in Welle 3 nur ein Senioreninterview vorlag, wurde die Zensierungsvariable zensiert auf den Sondercode -5 („Zensierung in Datenaufbereitung aufgehoben“) gesetzt. In Welle 3 neu berichtete Lückenspells wurden dem jeweiligen Spelldatensatz angehängt. Danach wurde der Spellzähler neu generiert, um eine lückenlose Variable *spellnr* zu schaffen.

5.8 Maßnahmeteilnahmen

In der 2. Welle wurde das Konzept zur Erhebung der Maßnahmeteilnahme grundlegend überarbeitet. Dieses Konzept wurde in Welle 3 fortgeführt. Als Referenzzeitpunkt, ab dem Maßnahmen erfragt wurden, diente Januar 2007 für Neubefragte Personen bzw. das Interviewdatum der letzten Befragung für Altbefragte.

Der MN-Spelldatensatz weist im Vergleich zu den anderen Spelldatensätzen des PASS einige Besonderheiten auf, die sich aus dem Erhebungskonzept ergeben. Auf diese soll im Folgenden kurz eingegangen werden.

Für die Teilnahme an arbeitsmarktpolitischen Maßnahmen, die in den MN-Spells abgelegt sind, ist keine Fortschreibung von in der Vorwelle zensierten Maßnahmeepisoden vorgesehen. Das Maßnahmemodul des Personenfragebogens sollte in Welle 2 und 3 nicht nur Programme von vergleichsweise langer Dauer (bspw. 1-Euro-Jobs) erfassen; die Befragten sollten auch deutlich kürzere Maßnahmen (wie bspw. Bewerbungstrainings etc.) berichten. Die Fortschreibung zensierter Episoden in der Folgewelle wäre deshalb problematisch gewesen, weil davon auszugehen ist, dass sich an kürzere Maßnahmen schlecht erinnert wird.

Aus diesem Grund wurden für jeden neu angelegten MN-Spell neben dem Beginndatum auch Informationen zur Dauer bzw. zum Enddatum erhoben. Für bereits abgeschlossene Maßnahmen wurde dabei die tatsächliche Dauer oder das tatsächliche Enddatum erfragt. Für zum Befragungszeitpunkt andauernde Maßnahmen sollten die Befragten die geplante Dauer oder das geplante Enddatum berichten.

Weil für zensierte MN-Spells aus der Vorwelle keine Fortschreibung vorgesehen ist und weil für das Ende der Maßnahme Informationen unterschiedlicher „Qualität“ vorliegen (tatsächliches vs. geplantes Ende), ergeben sich Schwierigkeiten für die Fortschreibung des MN-Spelldatensatzes. Statt die Zensierung von Spells aus Welle 2 nach einer Regel (bspw. auf

Basis des geplanten Endes der Maßnahme) bei der Integration mit Welle 3 zu schließen, wurde anders vorgegangen.

In Welle 3 wurde mit *spwelle* eine Wellenkennung in den MN-Spelldatensatz aufgenommen, mittels derer die Welle identifiziert werden kann, in welcher ein Spell angelegt wurde. Der Zensierungsindikator der MN-Spells (*zensiert*) wurde für Spells aus der Vorwelle nicht ediert. Auch dann nicht, wenn das Interviewdatum der Welle 3 nach dem geplanten Ende einer in der Vorwelle berichteten Maßnahme lag. Welche Fälle im MN-Spelldatensatz der Welle 3 vorkommen können und welche Informationen in diesen Fällen zum Ende des Spells vorhanden sind, ist in Tabelle 39 im Überblick dargestellt.

Tabelle 39: Überblick über die Informationen zum Enddatum im integrierten MN-Spelldatensatz der Welle 2 und 3 (*mn_spells*)

	Wert der Variable in MN-Spells Welle 3		Berichtete/s Dauer/ Enddatum	Generiertes Enddatum
	<i>spwelle</i>	<i>zensiert</i>	Typ Variablen	
Spell berichtet in W2 und damals abgeschlossen	2	2	<i>tatsächlich</i>	<i>MN0600</i> <i>MN0700</i> <i>MN0800</i> <i>emonat</i> <i>ejahr</i>
Spell berichtet in W2 und damals andauernd	2	1	<i>geplant</i>	<i>MN1100</i> <i>MN1200</i> <i>MN1300</i> <i>emonat</i> <i>ejahr</i>
Spell berichtet in W3 und damals abgeschlossen	3	2	<i>tatsächlich</i>	<i>MN0600</i> <i>MN0700</i> <i>MN0800</i> <i>emonat</i> <i>ejahr</i>
Spell berichtet in W3 und damals andauernd	3	1	<i>geplant</i>	<i>MN1100</i> <i>MN1200</i> <i>MN1300</i> <i>emonat</i> <i>ejahr</i>

Der Zensierungsindikator (*zensiert*) zeigt im MN-Spelldatensatz – im Gegensatz zu den übrigen Spelldaten des PASS – also nicht die Zensierung des Spells zum Zeitpunkt der letzten Befragung an, sondern die Zensierung des Spells in der Welle, in der er berichtet wurde. Auf welche Welle sich die Zensierung bezieht, ist der Wellenkennung (*spwelle*) zu entnehmen.

In Welle 3 erfolgte also keine Anpassung der in Welle 2 erhobenen Informationen, d.h. zensierte Spells aus Welle 2 bleiben auch nach Integration von Welle 3 immer noch zensiert.

5.8.1 Struktur des MN-Spelldatensatzes

Im MN-Spelldatensatz bleibt die Struktur gegenüber Welle 2 weitestgehend unverändert. Lediglich die Variable Name des Programms (MN1500), die in Welle 2 ohnehin nur mit Missingwerten belegt war, wurde aus dem Datensatz entfernt.

5.8.2 Plausibilitätsprüfung und –bereinigung des MN-Spelldatensatzes

Der MN-Spelldatensatz zur Teilnahme an arbeitsmarktpolitischen Maßnahmen wurde plausibilitätsgeprüft und bereinigt. Bereinigungen und Umcodierungen wurden nur in den generierten Datumvariablen vorgenommen. Jahreszeitangaben wurden in konkrete Monatsangaben umkodiert, „-8“ Werte für unplausible Angaben gesetzt oder Datumangaben ersetzt oder plausibilisiert.

In Welle 3 neu berichtete Maßnahmespells wurden dem MN-Spelldatensatz der Welle 2 angehängt. Danach wurde der Spellzähler neu generiert, um eine lückenlose Variable *spellnr* zu generieren.

6 Gewichtung Welle 3

Die Konstruktion der Gewichte für die dritte Welle erfolgte im Wesentlichen analog zur zweiten Welle (vgl. Gebhardt et al. 2009, Kapitel 9, dort findet sich auch ein schematischer Überblick über das Gewichtungskonzept von PASS). Unterschiede ergeben sich durch die erstmalig notwendige Integration von temporären Ausfällen, d.h. Haushalten, die zwar in der Welle zwei Jahre zuvor (Welle 1) teilgenommen, in der Vorwelle (Welle 2) jedoch nicht teilgenommen haben und durch ein anderes Verfahren bei der Kalibrierung. Anstelle von GREG wurde nun ein Raking-Verfahren (IPF) verwendet, was für die Gewichte allerdings keine wesentliche Bedeutung hat⁵⁶.

Ausgangspunkt für die Gewichtung der dritten Welle und des Längsschnitts von Welle 2 auf Welle 3 waren nunmehr die Querschnittsgewichte aus Welle 2 für Haushalte bzw. Personen. Die zwei Gewichte jedes Haushalts und die zwei Gewichte jeder Person wurden erneut fortgeschrieben.

6.1 Designgewichte für die Bestandshaushalte der 3. Welle

Unter Berücksichtigung von Zuzügen aus dem Inland wurden für die dritte Welle nunmehr aus den Haushaltsquerschnittsgewichten der zweiten Welle neue „Haushaltsdesignngewichte“ für die 3. Welle generiert. Dies erfolgte erneut über das für Welle 2 beschriebene „Weight-

⁵⁶ Die Gewichte bei beiden Verfahren sind in der Regel sehr ähnlich. Beim Raking lassen sich aber im Gegensatz zu GREG keine korrigierten Standardfehler schätzen. Allerdings ist die Berücksichtigung der Effekte der Kalibrierung auf den Standardfehler in Stata wie in den meisten gängigen Statistikprogrammen ohnehin nicht möglich.

Share“-Verfahren. Geburten, Todesfälle oder Auszüge haben keinen Einfluss auf das Gewicht, Zuzüge aus dem Inland erhöhen jedoch die Inklusionswahrscheinlichkeit eines Haushalts, da auch die zugezogenen Personen die Chance hatten, in Welle 1 oder Welle 2 (nur Auffrischungsstichprobe BA) ins Sample zu gelangen. Das neue Designgewicht für Teilstichprobe i dw_{i, hh_3} errechnet sich also aus dem alten Querschnittgewicht w_{i, hh_2} :

$$1 / dw_{i, hh_3} = 1 / w_{i, hh_2} + (n_{\text{sample } i} / n_{\text{population } i})$$

Das neue Designgewicht ist nur ein Zwischenschritt und daher in der Datenlieferung der dritten Welle nicht enthalten.

6.2 Designgewichte für die Auffrischungsstichprobe der 3. Welle

In der dritten Welle wurde das Panel erneut durch Nachziehung von Neuzugängen zum Leistungsbezug aufgefrischt. Eine Ziehungschance hatten alle Haushalte, die zwar im Juli 2008 Leistungen bezogen haben, im selben Monat 2007 und im selben Monat 2006 jedoch keine Auswahlwahrscheinlichkeit für die Prozessdatenstichprobe hatten. Diese Auffrischung ließ sich realisieren, indem nur Bedarfsgemeinschaften ausgewählt wurden, in denen kein Mitglied im Juli der vorangegangenen beiden Jahre Leistungen bezogen hatte. Nachgezogen wurde wiederum in den 300 Points der ersten Welle. Die Größe der Stichprobe war in Analogie zum speziellen pps-Verfahren bei der Ziehung der ersten Prozessdatenstichprobe, das in Rudolph und Trappmann (2007) beschrieben ist, proportional zum Anteil der Neuzugänge an der Bevölkerung im Point (zum Zeitpunkt der Auswahl der Points). Die Berechnung der Designgewichte ist dort ebenfalls beschrieben. Das Designgewicht der Auffrischungsstichprobe befindet sich für Fälle mit $sample=4$ in der Variable dw_ba .

6.3 Wiederteilnahme-Propensity Haushalte

In diesem Schritt wurde wiederum analog zum Vorgehen für Welle 2 für jeden Haushalt, der in der zweiten Welle teilgenommen hatte, die Wahrscheinlichkeit zur Wiederteilnahme über Logit-Modelle für Panelbereitschaft, Erreichbarkeit und Teilnahme geschätzt. Haushalte, die nur in Welle 1, nicht aber in Welle 2 teilgenommen hatten (temporäre Ausfälle), wurden hier in der Modellierung nicht berücksichtigt. Neben Variablen aus dem Haushaltsinterview und dem Personeninterview des Haushaltsvorstands der Vorwelle gingen hier auch Variablen ein, die mit der Feldarbeit zusammenhängen, z.B. Interviewmodus, Zahl der Kontaktversuche, Umzug, Haushaltsgröße, Anzahl der teilnahmebereiten Personen. Die vorhergesagten Wahrscheinlichkeiten aller drei Modelle wurden multipliziert. Der Kehrwert dieses Produkts ist für jede Welle in der Variable $hpbleib$ zu finden. Das Längsschnittgewicht eines Haushalts aus einer der Stichproben der ersten Welle für den gesamten möglichen Zeitraum $[t_1; t_2; t_3]$ zwischen allen drei Wellen lässt sich dann als Produkt aus Querschnittgewicht zu t_1 , $hpbleib$ (Welle 1 zu Welle 2) und $hpbleib$ (Welle 2 zu Welle 3) berechnen.

Tabelle 40: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholerhaushalte

Variablenkürzel und Referenzkategorie	Erläuterung
alter29	Haushaltsauskunftsperson (HBV) jünger als 30 Jahre
alter3039	HBV 30 – 39 Jahre alt
alter4049	HBV 40 – 49 Jahre alt
alter65	HBV älter als 65 Jahre
Referenzkategorie	HBV 50 – 64 Jahre
Mann	HBV männlich
Referenzkategorie	HBV weiblich
staatandere	HBV hat nicht-deutsche Staatsangehörigkeit
Referenzkategorie	HBV hat deutsche Staatsangehörigkeit oder fehlende Angabe
arbeitszeit2	HBV: Wochenarbeitszeit > 0 und < 40 Stunden
arbeitszeit3	HBV: Wochenarbeitszeit >= 40 Stunden
Referenzkategorie	HBV: Wochenarbeitszeit = 0 Stunden
Dschul1	Schulabschluss HBV: noch Schüler, anderer deutscher Schulabschluss, ausländischer Abschluss, fehlende Angabe
Dschul2	Schulabschluss HBV: Schule beendet ohne Abschluss, Sonder-/Förderschulabschluss
Dschul3	Schulabschluss HBV: Volks-/Hauptschulabschluss, POS 8. Klasse
Dschul5	Schulabschluss HBV: Fachhochschulreife, allgemeine oder fachgebundene Hochschulreife
Referenzkategorie	Schulabschluss HBV: Mittlere Reife, Realschulabschluss
DhealthKAWN	Subjektive Beurteilung des Gesundheitszustands HBV: fehlende Angabe
DhealthZu	Subjektive Beurteilung des Gesundheitszustands HBV: zufriedenstellend
DhealthSchlecht	Subjektive Beurteilung des Gesundheitszustands HBV: weniger gut, schlecht
Referenzkategorie	Subjektive Beurteilung des Gesundheitszustands HBV: sehr gut, gut
DZufrKAWN	generelle Lebenszufriedenheit HBV: fehlende Angabe
DZufr04	generelle Lebenszufriedenheit HBV: Skalenwert 0 – 4
DZufr910	generelle Lebenszufriedenheit HBV: Skalenwert 9 – 10
Referenzkategorie	generelle Lebenszufriedenheit HBV: Skalenwert 5 - 8
eigentum	Art des Wohneigentums: Eigentümer
Referenzkategorie	Art des Wohneigentums: Mieter, fehlende Angabe
anz_0_3	Anzahl Personen im Alter 0 – 3 Jahre
anz_4_6	Anzahl Personen im Alter 4 – 6 Jahre
anz_7_14	Anzahl Personen im Alter 7 – 14 Jahre
anz_15_64	Anzahl Personen im Alter 15 – 64 Jahre
anz_65	Anzahl Personen im Alter 65 Jahre und älter

Tabelle 40: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholerhaushalte (Fortsetzung 1)

Variablenkürzel und Referenzkategorie	Erläuterung
DinvalidAge	nicht auswertbare Altersangaben: Ja
Referenzkategorie	nicht auswertbare Altersangaben: Nein
in_hh2	Zahl Personeninterviews im HH: 2
in_hh3	Zahl Personeninterviews im HH: 3 und mehr
Referenzkategorie	Zahl Personeninterviews im HH: 1
Dhincom	Haushaltseinkommen: fehlende oder unplausible Angabe
hhincom1	Haushaltseinkommen: bis 871 Euro
hhincom3	Haushaltseinkommen: 1401 - 2200 Euro
hhincom4	Haushaltseinkommen: mehr als 2200 Euro
Referenzkategorie	Haushaltseinkommen: 872 - 1400 Euro
alg2abez	ALG2-Bezug des Haushalts: aktueller Bezug
Referenzkategorie	ALG2-Bezug des Haushalts: kein aktueller Bezug
halg2st3	ALG2-Bezug Haushalt in den drei Jahren vor Befragung: Bezug
Dhalg2st3	ALG2-Bezug Haushalt in den drei Jahren vor Befragung: fehlende Angabe
Referenzkategorie	ALG2-Bezug Haushalt in den drei Jahren vor Befragung: kein Bezug
Dhpwnka0	Anzahl w.n.- und k.A.-Angaben in Haushalts- und Personeninterview des HBV: keine
Dhpwnka2	Anzahl w.n.- und k.A.-Angaben in Haushalts- und Personeninterview des HBV: 11 und mehr
Referenzkategorie	Anzahl w.n.- und k.A.-Angaben in Haushalts- und Personeninterview des HBV: 1 – 10
eastwest	alte oder neue Bundesländer: neue Bundesländer
Referenzkategorie	alte oder neue Bundesländer: alte Bundesländer
Dbundesl1	Bundesland: Schleswig-Holstein
Dbundesl2	Bundesland: Hamburg
Dbundesl3	Bundesland: Niedersachsen
Dbundesl4	Bundesland: Bremen
Dbundesl6	Bundesland: Hessen
Dbundesl7	Bundesland: Rheinland-Pfalz
Dbundesl8	Bundesland: Baden-Württemberg
Dbundesl9	Bundesland: Bayern
Dbundesl10	Bundesland: Saarland
Dbundesl11	Bundesland: Berlin
Dbundesl12	Bundesland: Brandenburg
Dbundesl13	Bundesland: Mecklenburg-Vorpommern

Tabelle 40: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholerhaushalte (Fortsetzung 2)

Variablenkürzel und Referenzkategorie	Erläuterung
Dbundesl14	Bundesland: Sachsen
Dbundesl15	Bundesland: Sachsen-Anhalt
Dbundesl16	Bundesland: Thüringen
Referenzkategorie	Bundesland: Nordrhein-Westfalen
Dbik_2	BIK-Größenklasse der Gemeinde: 2000 – unter 5000 EW
Dbik_3	BIK-Größenklasse der Gemeinde: 5000 – unter 20000 EW
Dbik_4	BIK-Größenklasse der Gemeinde: 20000 – unter 50000 EW
Dbik_5	BIK-Größenklasse der Gemeinde: 50000 – unter 100000 EW Rand
Dbik_6	BIK-Größenklasse der Gemeinde: 10000 – unter 100000 EW Kern
Dbik_7	BIK-Größenklasse der Gemeinde: 100000 – unter 500000 EW Rand
Dbik_8	BIK-Größenklasse der Gemeinde: 100000 – unter 500000 EW Kern
Dbik_9	BIK-Größenklasse der Gemeinde: 500000 und mehr EW Rand
Dbik_10	BIK-Größenklasse der Gemeinde: 500000 und mehr EW Kern
Referenzkategorie	BIK-Größenklasse der Gemeinde: unter 2000 EW
kcati0	Kontaktversuche CATI: 0
kcati2	Kontaktversuche CATI: 5 – 10
kcati3	Kontaktversuche CATI: 11 – 27
kcati4	Kontaktversuche CATI: 28 und mehr
Referenzkategorie	Kontaktversuche CATI: 1 – 4
kcapi0	Kontaktversuche CAPI: 0 oder fehlende Angabe
kcapi2	Kontaktversuche CAPI: 3
kcapi3	Kontaktversuche CAPI: 4
kcapi4	Kontaktversuche CAPI: 5 und mehr
Referenzkategorie	Kontaktversuche CAPI: 1 – 2

Tabelle 41: Logit-Modelle zur Wiederteilnahme für Panelbereitschaft, Erreichbarkeit und Teilnahme

	Panelbereitschaft		Kontakt		Teilnahme	
	Coef.	p	Coef.	p	Coef.	p
alter29	-.4179966	0.191	-.8311556	0.000	-.5715283	0.000
alter3039	-.6758633	0.017	-.6455897	0.000	-.2683763	0.045
alter4049	-.2119312	0.427	-.2243778	0.046	-.1361424	0.217
alter65	.5006297	0.313	.0803519	0.764	.1780036	0.465
Mann	.0951685	0.619	-.0361669	0.656	-.0940206	0.271
staatandere	-.7545756	0.009	-.3118517	0.024	-.2040702	0.232
arbeitszeit	.227652	0.448	.2733963	0.044	-.0801597	0.520

Tabelle 41: Logit-Modelle zur Wiederteilnahme für Panelbereitschaft, Erreichbarkeit und Teilnahme (Fortsetzung 1)

	Panelbereitschaft		Kontakt		Teilnahme	
	Coef.	p	Coef.	p	Coef.	p
arbzeit3	.3782709	0.211	.0284119	0.824	.0635809	0.624
Dschul1	-.0029973	0.995	-.0959908	0.644	.0328029	0.894
Dschul2	-.1588747	0.687	-.3502206	0.024	-.1364375	0.492
Dschul3	-.23506	0.291	-.1238202	0.199	-.1683959	0.092
Dschul5	-.0376717	0.885	.1295561	0.248	.0028406	0.980
DhealthKAWN	2.937874	0.024	-.8767879	0.196	-.942785	0.159
DhealthZu	.0708809	0.732	.1378718	0.134	.1871987	0.050
DhealthSchlecht	.260788	0.301	-.1212947	0.237	-.1009442	0.353
DZufrKAWN	-1.844352	0.006	.6504756	0.348	.1695471	0.795
DZufr04	-.4867122	0.032	-.0084285	0.934	-.072271	0.526
DZufr910	.0658411	0.834	.2003547	0.152	.0495304	0.702
eigentum	-.2927264	0.227	.2691123	0.028	-.0191214	0.859
anz_0_3	.0684786	0.837	-.0999625	0.423	.1360783	0.388
anz_4_6	.3885585	0.266	-.1451466	0.192	-.1446253	0.252
anz_7_14	-.0008379	0.996	.0708032	0.317	.0613971	0.407
anz_15_64	-.0186249	0.884	-.2013297	0.000	-.477008	0.000
anz_65	-.5701717	0.040	-.0753248	0.645	-.3791891	0.007
DinvalidAge	-1.200804	0.160	-.7285887	0.266	-1.648609	0.006
in_hh2	1.369934	0.000	.4284563	0.000	.5352592	0.000
in_hh3	2.13238	0.004	.7714587	0.000	1.161619	0.000
Dhheincom	-1.175575	0.004	.4580934	0.234	-.4479153	0.101
hheincom1	.021228	0.936	-.193863	0.073	.2231407	0.072
hheincom3	-.4689878	0.070	.0460383	0.687	.163656	0.167
hheincom4	.3694011	0.339	.3583555	0.020	.3563863	0.016
alg2abez	.3348003	0.226	.1252726	0.276	-.0393806	0.763
halg2st3	.1390143	0.624	-.0921112	0.460	.2290695	0.084
Dhalg2st3			-1.477168	0.003	.466441	0.581
Dhpwnka0	.5728716	0.010	.2676548	0.001	.1530721	0.071
Dhpwnka2	-1.084382	0.000	.0414075	0.785	.001363	0.994
eastwest	-.0897827	0.664				
Dbundesl1			-.2056056	0.272	.1275223	0.556
Dbundesl2			-.7086271	0.009	-.2949669	0.402
Dbundesl3			-.0139394	0.922	-.01159	0.938
Dbundesl4			.7596985	0.156	.0246193	0.958
Dbundesl6			.5715236	0.005	.1964544	0.279
Dbundesl7			-.2165509	0.287	.011608	0.958

Tabelle 41: Logit-Modelle zur Wiederteilnahme für Panelbereitschaft, Erreichbarkeit und Teilnahme (Fortsetzung 2)

	Panelbereitschaft		Kontakt		Teilnahme	
	Coef.	p	Coef.	p	Coef.	p
Dbundesl8			.173082	0.290	-.0377617	0.806
Dbundesl9			.1098784	0.437	.0428446	0.766
Dbundesl10			-.3480398	0.276	.5340396	0.278
Dbundesl11			-.2323356	0.147	-.1222737	0.532
Dbundesl12			.0672585	0.747	.1917281	0.371
Dbundesl13			.1826805	0.499	.3863675	0.183
Dbundesl14			.173065	0.345	.4142374	0.031
Dbundesl15			.2912595	0.159	.21407	0.302
Dbundesl16			.3309171	0.190	.6795506	0.015
Dbik_2			.4852294	0.014	-.0985176	0.587
Dbik_3			.1629304	0.185	-.0916656	0.513
Dbik_4			.2484895	0.088	-.0364019	0.812
Dbik_5			.5679636	0.032	-.3075491	0.225
Dbik_6			.0964047	0.564	-.0833082	0.651
Dbik_7			.2507438	0.127	-.1531106	0.350
Dbik_8			.3224649	0.070	-.2950502	0.090
Dbik_9			.2409303	0.326	-.3722877	0.113
Dbik_10			.8106134	0.067	-.8256749	0.004
kcati0					2.595051	0.000
kcati2					.4144359	0.001
kcati3					.4314559	0.000
kcati4					.4958591	0.000
kcapi0					.9932981	0.000
kcapi2					.1224835	0.552
kcapi3					-.1604692	0.473
kcapi4					-.4160623	0.017
cons	4.099317	0.000	2.325814	0.000	1.277427	0.000
n	8440		8304		7487	
Log likelihood	-624.93564		-2489.6614		-2263.2766	
Pseudo R²	0.1025		0.0675		0.1000	

6.4 Teilnahme-Propensity erstbefragte Split-Haushalte

In diesem Schritt wurden die Teilnahme-Propensities für neue Split-Haushalte, d.h. Haushalte, die aufgrund des Auszugs einer Person der Panelstichprobe in einen neuen Haushalt in das Panel aufgenommen werden, berechnet. Hierbei wurden nur solche Split-Haushalte berücksichtigt, die in den beiden vorausgegangenen Wellen noch nicht befragt wurden. Modelliert wurden also die Teilnahmepropensities für die erstbefragten Split-Haushalte. Die

Wahrscheinlichkeit zur Wiederteilnahme wurde über Logit-Modelle für Erreichbarkeit und Teilnahme geschätzt. Fehlende zeitstabile Informationen über die Haushaltsauskunftsperson (HBV) wurden, falls notwendig, aus der Vorwelle ergänzt. Die vorhergesagten Wahrscheinlichkeiten der beiden Modelle wurden multipliziert. Der Kehrwert des Produkts ist für die Split-Haushalte ebenfalls in der Variable *hpleib* zu finden.

Tabelle 42: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der erstbefragten Split-Haushalte

Variablenkürzel und Referenzkategorie	Erläuterung
alter3039	HBV 30 – 39 Jahre alt
alter4049	HBV 40 – 49 Jahre alt
alter5064	HBV 50 – 64 Jahre
alter65	HBV älter als 65 Jahre
Referenzkategorie	HBV jünger als 30 Jahre
Mann	HBV männlich
Referenzkategorie	HBV weiblich
staatandere	HBV hat nicht-deutsche Staatsangehörigkeit
staatsymis	HBV hat fehlende Angabe bei Staatsangehörigkeit
Referenzkategorie	HBV hat deutsche Staatsangehörigkeit
Dschul2	Schulabschluss HBV: Schule beendet ohne Abschluss, Sonder-/Förderschulabschluss
Dschul3	Schulabschluss HBV: Volks-/Hauptschulabschluss, POS 8. Klasse
Dschul4	Schulabschluss HBV: Mittlere Reife, Realschulabschluss
Dschul5	Schulabschluss HBV: Fachhochschulreife, allgemeine oder fachgebundene Hochschulreife
Referenzkategorie	Schulabschluss HBV: noch Schüler, anderer deutscher Schulabschluss, ausländischer Abschluss, fehlende Angabe
sample_BA	Aus BA-Stichprobe Welle 1
sample_Auffrischer	Aus BA-Auffrischerstichprobe Welle 2
Referenzkategorie	Aus Microm-Stichprobe
kcati0	Kontaktversuche CATI: 0
kcati1	Kontaktversuche CATI: 1 – 7
kcati2	Kontaktversuche CATI: 8 – 19
kcati4	Kontaktversuche CATI: 49 und mehr
Referenzkategorie	Kontaktversuche CATI: 20 – 48
kcapi0	Kontaktversuche CAPI: 0
kcapi2	Kontaktversuche CAPI: 3
kcapi3	Kontaktversuche CAPI: 4
kcapi4	Kontaktversuche CAPI: 5 und mehr
kon_capi_sysmis	Kontaktversuche CAPI: fehlende Angabe
Referenzkategorie	Kontaktversuche CAPI: 1 – 2

Tabelle 43: Logit-Modelle zur Erstteilnahme von Split-Haushalten für Erreichbarkeit und Teilnahme

	Kontakt		Teilnahme	
	Coef.	p	Coef.	p
alter3039	-.0457938	0.882	-.4974757	0.232
alter4049	.1541013	0.685	.435613	0.414
alter5064	-.2011846	0.584	-.0742023	0.890
alter65	1.03412	0.099	-.6649824	0.332
Mann	.0829381	0.680	.2489844	0.360
staatandere	-.6301316	0.258	-1.874751	0.030
staatsystemis	-1.014355	0.010	-.7183244	0.128
Dschul2	-.5599909	0.348	.1770058	0.830
Dschul3	-.8941441	0.048	-.3839034	0.496
Dschul4	-.5311538	0.205	-.2538346	0.608
Dschul5	-.750205	0.089	-.2666337	0.624
sample_BA	.0285283	0.889	-.1333173	0.625
sample_Auffrischer	-1.319607	0.017	-.8828261	0.365
kcati0	.678541	0.101		
kcati1	1.313406	0.002		
kcati2	2.137653	0.000		
kcati4	-.0279659	0.955		
kon_capi_sysmis	-.4791278	0.134		
kcapi0	.0868798	0.827		
kcapi2	-.0524601	0.882		
kcapi3	.1001297	0.821		
kcapi4	-.233398	0.487		
cons	.2971993	0.591	.897202	0.045
n	483		262	
Log likelihood	-305.3388		-168.78392	
Pseudo R²	0.0832		0.0410	

6.5 Ausfallgewichtung für Haushalte aus Auffrischungsstichprobe der 3. Welle

Für die Haushalte aus der Auffrischungsstichprobe wurde erneut eine zweistufige Ausfallmodellierung (Erreichbarkeit und Teilnahme) analog zur Auffrischungsstichprobe der zweiten Welle vorgenommen. Die daraus abgeleitete Teilnahmewahrscheinlichkeit befindet sich in der Variable *prop_t0*.

Tabelle 44: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Auffrischungsstichprobe Welle 3

Variablenkürzel und Referenzkategorie	Erläuterung
alter3039	HBV 30 – 39 Jahre alt
alter4049	HBV 40 – 49 Jahre alt
alter5059	HBV 50 – 59 Jahre alt
alter60	HBV 60 Jahre und älter
Referenzkategorie	HBV jünger als 30 Jahre
sex_w	HBV weiblich
Referenzkategorie	HBV männlich
staat_rge	HBV Staatsangehörigkeit: europäisch (nicht deutsch), russisch, ehemalige GUS-Staaten
staat_son	HBV Staatsangehörigkeit: türkisch, asiatisch, australisch, afrikanisch, amerikanisch
Referenzkategorie	HBV hat deutsche Staatsangehörigkeit oder fehlende Angabe
Schul_Kein	Schulabschluss HBV: kein Abschluss
Schul_HS	Schulabschluss HBV: Hauptschulabschluss
Schul_MR	Schulabschluss HBV: Mittlere Reife
Schul_FA	Schulabschluss HBV: Fachhochschulreife
Referenzkategorie	Schulabschluss HBV: Hochschulreife
typ_alleinerz	BG-Typ: alleinerziehend
typ_paarokind	BG-Typ: Paar ohne Kind
typ_paarmkind	BG-Typ: Paar mit Kind
typ_sonst	BG-Typ: sonstige
Referenzkategorie	BG-Typ: alleinstehend
erw_NEF	HBV nicht erwerbsfähig
erw_KA	HBV ohne Bestimmung der Erwerbsfähigkeit
Referenzkategorie	HBV erwerbsfähig
catifeld	Haushalt war (ursprünglich) im CATI-Feld
Referenzkategorie	Haushalt war (ursprünglich) im CAPI-Feld
tranche2	Tranche: 2
tranche3	Tranche: 3
tranche4	Tranche: 4

Tabelle 44: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Auffrischungstichprobe Welle 3 (Fortsetzung 1)

Variablenkürzel und Referenzkategorie	Erläuterung
tranche5	Tranche: 5
tranche6	Tranche: 6
Referenzkategorie	Tranche: 1
anz_verwf	Anzahl der Erwerbsfähigen in der BG
kontcati	Anzahl CATI-Kontakte
kontcati_ka	fehlende Angabe zur Zahl der CATI-Kontakte: ja
Referenzkategorie	fehlende Angabe zur Zahl der CATI-Kontakte: nein
kontcapi	Anzahl CAPI-Kontakte
kontcapi_ka	fehlende Angabe zur Zahl der CAPI-Kontakte: ja
Referenzkategorie	fehlende Angabe zur Zahl der CAPI-Kontakte: nein
bula1	Bundesland: Schleswig-Holstein
bula2	Bundesland: Hamburg
bula3	Bundesland: Niedersachsen
bula4	Bundesland: Bremen
bula6	Bundesland: Hessen
bula7	Bundesland: Rheinland-Pfalz
bula8	Bundesland: Baden-Württemberg
bula9	Bundesland: Bayern
bula10	Bundesland: Saarland
bula11	Bundesland: Berlin
bula12	Bundesland: Brandenburg
bula13	Bundesland: Mecklenburg-Vorpommern
bula14	Bundesland: Sachsen
bula15	Bundesland: Sachsen-Anhalt
bula16	Bundesland: Thüringen
Referenzkategorie	Bundesland: Nordrhein-Westfalen
bik_1	BIK-Größenklasse der Gemeinde: unter 2000 EW
bik_2	BIK-Größenklasse der Gemeinde: 2000 – unter 5000 EW
bik_3	BIK-Größenklasse der Gemeinde: 5000 – unter 20000 EW
bik_4	BIK-Größenklasse der Gemeinde: 20000 – unter 50000 EW
bik_5	BIK-Größenklasse der Gemeinde: 50000 – unter 100000 EW Rand
bik_6	BIK-Größenklasse der Gemeinde: 10000 – unter 100000 EW Kern
bik_7	BIK-Größenklasse der Gemeinde: 100000 – unter 500000 EW Rand
bik_8	BIK-Größenklasse der Gemeinde: 100000 – unter 500000 EW Kern
bik_9	BIK-Größenklasse der Gemeinde: 500000 und mehr EW Rand
Referenzkategorie	BIK-Größenklasse der Gemeinde: 500000 und mehr EW Kern

Tabelle 45: Logit-Modelle zur Erstteilnahme für Erreichbarkeit und Teilnahme

	Kontakt		Teilnahme	
	Coef.	p	Coef.	p
alter3039	.3090823	0.003	-.1314074	0.249
alter4049	.3885464	0.000	-.1144351	0.314
alter5059	.3887638	0.001	-.2671515	0.035
alter60	.4672857	0.035	-.3456237	0.138
sex_w	.023271	0.777	-.012965	0.884
Schul_Kein	-.5021477	0.000	-.2747929	0.114
Schul_HS	-.0455601	0.651	-.1582463	0.144
Schul_MR	.0568066	0.635	-.0463256	0.703
Schul_FA	.2362025	0.094	.1239275	0.389
staat_rge	-.5830485	0.000	.13027	0.445
staat_son	-.8213009	0.000	-.4579772	0.013
typ_alleinerz	.473528	0.005	.3056752	0.094
typ_paarokind	.4321966	0.002	-.0737591	0.658
typ_paarmkind	.6248467	0.001	.1028116	0.665
typ_sonst	.3926251	0.022	.179789	0.488
erw_NEF	-.0766154	0.614	-.0265557	0.867
erw_KA	-.7834184	0.005	.2289636	0.527
catifeld	.8830739	0.000		
anz_verwf			-.1363106	0.183
kontcati			.0017048	0.326
kontcati_ka			.2885493	0.213
kontcapi			.0815515	0.003
kontcapi_ka			-3.742772	0.000
tranche2	.1046926	0.428	.0937907	0.498
tranche3	.17847	0.179	.0056628	0.967
tranche4	-.0669418	0.599	-.1135456	0.421
tranche5	-.0269351	0.833	-.3703891	0.010
tranche6	.1719382	0.196	-.2572793	0.066

Tabelle 45: Logit-Modelle zur Erstteilnahme für Erreichbarkeit und Teilnahme (Fortsetzung)

	Kontakt		Teilnahme	
	Coef.	p	Coef.	p
bik_1	.2860343	0.114	-.0870541	0.653
bik_2	.1278021	0.262	.0802627	0.533
bik_3	.4756256	0.002	-.1189641	0.446
bik_4	.4653795	0.065	.1358901	0.580
bik_5	.2334757	0.180	.1972223	0.290
bik_6	.3774596	0.016	-.0071909	0.964
bik_7	.3402342	0.068	-.0200336	0.918
bik_8	.2427334	0.398	-.5795322	0.052
bik_9	-.2359266	0.602	.2676086	0.601
bula1	.0358522	0.864	-.0866758	0.686
bula2	-.0826558	0.751	-.0283367	0.925
bula3	-.0031247	0.983	-.0724963	0.640
bula4	-.4717033	0.164	-.1457921	0.763
bula6	-.1771432	0.280	.373867	0.043
bula7	-.2555386	0.205	.0989945	0.658
bula8	-.2784098	0.047	.0020443	0.990
bula9	.1582747	0.288	.1193608	0.439
bula10	-.1490653	0.641	.2069501	0.540
bula11	.0950723	0.540	-.2387027	0.177
bula12	.0297586	0.895	-.5428247	0.023
bula13	-.157555	0.559	.409411	0.139
bula14	-.0000678	1.000	.000055	1.000
bula15	.0489525	0.823	.0454987	0.831
bula16	.3197248	0.243	.0673526	0.773
cons	-.3548499	0.054	.0960608	0.605
n	3785		2704	
Log likelihood	-2138.9356		-1754.3689	
Pseudo R²	0.0553		0.0543	

6.6 Wiederteilnahme-Propensity Personen

Das entscheidende Längsschnittgewicht ist nicht dasjenige auf Haushalts-, sondern das auf Personenebene, da es sich hierbei um die über die Zeit stabilen Einheiten handelt. Erneut wurden wie in Welle 2 Teilnahmepropensities für Personen unter Einschluss von zusätzlichen Personenmerkmalen über Logit-Modelle für Erreichbarkeit und Teilnahme geschätzt. Die über den Haushaltszusammenhang vermittelte Abhängigkeit der Personenstichprobe und die dadurch notwendige Korrektur der Schätzung der Standardfehler wurde bei diesen Modellen berücksichtigt, indem die Fehlerterme auf Haushaltsebene geclustert wurden. Die vorhergesagten Wahrscheinlichkeiten der Modelle wurden wiederum multipliziert. Der Kehrwert dieses Produkts ist in der Variable *ppbleib* zu finden. Das Längsschnittgewicht einer Person für den Zeitraum $[t_1; t_2; t_3]$ über alle drei Wellen lässt sich dann als Produkt aus Querschnittgewicht zu t_1 , *ppbleib* (Welle 1 zu Welle 2) und *ppbleib* (Welle 2 zu Welle 3) berechnen.

Tabelle 46: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholer

Variablenkürzel und Referenzkategorie	Erläuterung
alter29	Person jünger als 30 Jahre
alter3039	Person 30 – 39 Jahre alt
alter4049	Person 40 – 49 Jahre alt
alter65	Person älter als 65 Jahre
Referenzkategorie	Person 50 – 64 Jahre
Mann	Person männlich
Referenzkategorie	Person weiblich
staatandere	Person hat nicht-deutsche Staatsangehörigkeit (oder fehlende Angabe)
staatsymis	Person hat fehlende Angabe bei Staatsangehörigkeit
Referenzkategorie	HBV hat deutsche Staatsangehörigkeit (oder fehlende Angabe, falls staatsymis nicht im Modell)
arbeits1	Wochenarbeitszeit <= 32 Stunden
arbeits2	Wochenarbeitszeit > 32 und <= 40 Stunden
arbeits3	Wochenarbeitszeit > 40 Stunden
Referenzkategorie	HBV: Wochenarbeitszeit = 0 Stunden
Dschul1	Schulabschluss: noch Schüler, anderer deutscher Schulabschluss, ausländischer Abschluss, fehlende Angabe
Dschul2	Schulabschluss: Schule beendet ohne Abschluss, Sonder-/Förderschulabschluss
Dschul4	Schulabschluss: Mittlere Reife, Realschulabschluss
Dschul5	Schulabschluss: Fachhochschulreife, allgemeine oder fachgebundene Hochschulreife
Referenzkategorie	Schulabschluss: Volks-/Hauptschulabschluss, POS 8. Klasse
DhealthKAWN	Subjektive Beurteilung des Gesundheitszustands HBV: fehlende Angabe

Tabelle 46: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholer (Fortsetzung 1)

Variablenkürzel und Referenzkategorie	Erläuterung
DhealthZu	Subjektive Beurteilung des Gesundheitszustands HBV: zufriedenstellend
DhealthSchlecht	Subjektive Beurteilung des Gesundheitszustands HBV: weniger gut, schlecht
Referenzkategorie	Subjektive Beurteilung des Gesundheitszustands HBV: sehr gut, gut
DZufrKAWN	generelle Lebenszufriedenheit HBV: fehlende Angabe
DZufr04	generelle Lebenszufriedenheit HBV: Skalenwert 0 – 4
DZufr910	generelle Lebenszufriedenheit HBV: Skalenwert 9 – 10
Referenzkategorie	generelle Lebenszufriedenheit HBV: Skalenwert 5 - 8
Dpwnka0	Anzahl der w.n. oder k.A. Angaben in Haushalts- und Personeninterview: 0
Dpwnka2	Anzahl der w.n. oder k.A. Angaben in Haushalts- und Personeninterview: mehr als 10
Referenzkategorie	Anzahl der w.n. oder k.A. Angaben in Haushalts- und Personeninterview: 1 - 10
Dhsprache	Hauptsprache: Deutsch
Referenzkategorie	Hauptsprache: nicht Deutsch
eigentum	Art des Wohneigentums: Eigentümer
Referenzkategorie	Art des Wohneigentums: Mieter, fehlende Angabe
anz_0_3	Anzahl Personen im Alter 0 – 3 Jahre
anz_4_6	Anzahl Personen im Alter 4 – 6 Jahre
anz_7_14	Anzahl Personen im Alter 7 – 14 Jahre
anz_65	Anzahl Personen im Alter 65 Jahre und älter
Danz_15_64_1	Anzahl Personen im Alter 15 – 64: 1
Danz_15_64_2	Anzahl Personen im Alter 15 – 64: 2
Danz_15_64_3	Anzahl Personen im Alter 15 – 64: 3
Danz_15_64_4	Anzahl Personen im Alter 15 – 64: 4 und mehr
Referenzkategorie	Anzahl Personen im Alter 15 – 64: keine
DinvalidAge	nicht auswertbare Altersangaben: ja
Referenzkategorie	nicht auswertbare Altersangaben: nein
in_hh2	Zahl Personeninterviews im HH: 2
in_hh3	Zahl Personeninterviews im HH: 3 und mehr
Referenzkategorie	Zahl Personeninterviews im HH: 1
Dhheincom	Haushaltseinkommen: fehlende oder unplausible Angabe
hhincom1	Haushaltseinkommen: bis 986 Euro
hhincom3	Haushaltseinkommen: 1501 - 2500 Euro
hhincom4	Haushaltseinkommen: mehr als 2500 Euro
Referenzkategorie	Haushaltseinkommen: 987 - 1500 Euro
alg2abez	ALG2-Bezug des Haushalts: aktueller Bezug

Tabelle 46: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholer (Fortsetzung 2)

Variablenkürzel und Referenzkategorie	Erläuterung
Referenzkategorie	ALG2-Bezug des Haushalts: kein aktueller Bezug
halg2st3	ALG2-Bezug Haushalt in den drei Jahren vor Befragung: Bezug
Referenzkategorie	ALG2-Bezug Haushalt in den drei Jahren vor Befragung: kein Bezug
sample_BA	BA-Stichprobe
Referenzkategorie	Microm-Stichprobe
Dbundesl1	Bundesland: Schleswig-Holstein
Dbundesl2	Bundesland: Hamburg
Dbundesl3	Bundesland: Niedersachsen
Dbundesl4	Bundesland: Bremen
Dbundesl6	Bundesland: Hessen
Dbundesl7	Bundesland: Rheinland-Pfalz
Dbundesl8	Bundesland: Baden-Württemberg
Dbundesl9	Bundesland: Bayern
Dbundesl10	Bundesland: Saarland
Dbundesl11	Bundesland: Berlin
Dbundesl12	Bundesland: Brandenburg
Dbundesl13	Bundesland: Mecklenburg-Vorpommern
Dbundesl14	Bundesland: Sachsen
Dbundesl15	Bundesland: Sachsen-Anhalt
Dbundesl16	Bundesland: Thüringen
Referenzkategorie	Bundesland: Nordrhein-Westfalen
Dbik_2	BIK-Größenklasse der Gemeinde: 2000 – unter 5000 EW
Dbik_3	BIK-Größenklasse der Gemeinde: 5000 – unter 20000 EW
Dbik_4	BIK-Größenklasse der Gemeinde: 20000 – unter 50000 EW
Dbik_5	BIK-Größenklasse der Gemeinde: 50000 – unter 100000 EW Rand
Dbik_6	BIK-Größenklasse der Gemeinde: 10000 – unter 100000 EW Kern
Dbik_7	BIK-Größenklasse der Gemeinde: 100000 – unter 500000 EW Rand
Dbik_8	BIK-Größenklasse der Gemeinde: 100000 – unter 500000 EW Kern
Dbik_9	BIK-Größenklasse der Gemeinde: 500000 und mehr EW Rand
Dbik_10	BIK-Größenklasse der Gemeinde: 500000 und mehr EW Kern
Referenzkategorie	BIK-Größenklasse der Gemeinde: unter 2000 EW
kcati0	Kontaktversuche CATI: 0
kcati2	Kontaktversuche CATI: 6 – 14
kcati3	Kontaktversuche CATI: 14 – 42 bzw. 14 und mehr
kcati4	Kontaktversuche CATI: 43 und mehr
Referenzkategorie	Kontaktversuche CATI: 1 – 5

Tabelle 46: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholer (Fortsetzung 3)

Variablenkürzel und Referenzkategorie	Erläuterung
kcapi0	Kontaktversuche CAPI: 0 oder fehlende Angabe
kcapi2	Kontaktversuche CAPI: 3
kcapi3	Kontaktversuche CAPI: 4
kcapi4	Kontaktversuche CAPI: 5 und mehr
Referenzkategorie	Kontaktversuche CAPI: 1 – 2

Tabelle 47: Logit-Modelle zur Wiederteilnahme für Erreichbarkeit und Teilnahme

	Kontakt		Teilnahme	
	Coef.	p	Coef.	p
alter29	-.3194031	0.007	-.9574916	0.000
alter3039	-.1641501	0.208	-.4442976	0.000
alter4049	.0162689	0.896	-.0822184	0.329
alter65	.2934335	0.379	-.3358986	0.047
Mann	-.04211	0.533	-.2058298	0.000
staatandere	-.3855001	0.022	-.3378695	0.005
staatsystemis	-.0434743	0.972		
arbeitszeit1	.213965	0.169	.0032377	0.975
arbeitszeit2	.1397969	0.272	-.0777185	0.369
arbeitszeit3	.0705923	0.606	.040118	0.651
Dschul1	.0264277	0.870	.3279346	0.003
Dschul2	-.5820174	0.000	-.0636041	0.633
Dschul4	.0247343	0.807	.1842116	0.008
Dschul5	.0261321	0.828	.2079732	0.009
DhealthKAWN	-.2165862	0.772	-1.963675	0.001
DhealthZu	.1195256	0.179	.1115784	0.075
DhealthSchlecht	-.0363983	0.717	-.0250116	0.746
DZufrKAWN	.9939802	0.146	-.5768915	0.352
DZufr04	.041639	0.698	.0073365	0.932
DZufr910	.176946	0.173	.0829721	0.279

Tabelle 47: Logit-Modelle zur Wiederteilnahme für Erreichbarkeit und Teilnahme (Forts. 1)

	Kontakt		Teilnahme	
	Coef.	p	Coef.	p
Dpwnka0	.2501217	0.002	.1773289	0.002
Dpwnka2	.0653905	0.696	-.1595383	0.185
Dhsprache	.0024693	0.993	-.1833932	0.289
eigentum	.3327693	0.026	.0403194	0.616
anz_0_3	-.2668426	0.080	.303347	0.014
anz_4_6	-.1438577	0.299	-.0555842	0.539
anz_7_14	.0557589	0.546	.0848012	0.117
Danz_15_64_1	.1226313	0.782	-.5429105	0.000
Danz_15_64_2	-.2948062	0.527	-2.556655	0.000
Danz_15_64_3	-.337792	0.502	-.1068527	0.000
Danz_15_64_4	-.3430471	0.516	.2323787	0.000
anz_65	-.3113932	0.089	-.383435	0.000
DinvalidAge	-.2810615	0.793	-.2419069	0.002
in_hh2	.3757126	0.002	-.0890333	0.192
in_hh3	.6023466	0.011	.1173713	0.062
Dhincom	.2755263	0.493	.0533518	0.115
hhincom1	-.4698099	0.007	.0801471	0.040
hhincom2	-.2008494	0.168	.05335	0.331
hhincom4	.2859489	0.151	2.084369	0.213
alg2abez	.0990292	0.476	.5554232	0.583
halg2st3	-.2049675	0.237	.271659	0.461
sample_BA	.1647116	0.180	-.4007131	0.538
kcati0	2.156775	0.000	.6827928	0.000
kcati2	.3595989	0.075	-.1429823	0.000
kcati3	-1.001132	0.000	-.3206767	0.007
kcati4			-.5170584	0.000
kcapi0	1.811629	0.000	-.1676071	0.000
kcapi2	.5007148	0.013	-.6044384	0.297

Tabelle 47: Logit-Modelle zur Wiederteilnahme für Erreichbarkeit und Teilnahme (Forts. 2)

	Kontakt		Teilnahme	
	Coef.	p	Coef.	p
kcapi3	.0638927	0.764	.0435586	0.042
kcapi4	-.3761208	0.015	.4082219	0.000
Dbundesl1	-.2458053	0.255	.0656108	0.283
Dbundesl2	-.7702539	0.025	-.0115265	0.045
Dbundesl3	-.0309503	0.866	-.1124558	0.710
Dbundesl4	.3374801	0.590	.1536802	0.365
Dbundesl6	.3380317	0.147	.2855574	0.645
Dbundesl7	-.5305397	0.037	-.3208774	0.949
Dbundesl8	.1868234	0.343	-.0324926	0.343
Dbundesl9	.0400736	0.818	-.1450802	0.178
Dbundesl10	-.4245764	0.308	.3532369	0.358
Dbundesl11	.0253143	0.903	-.0259127	0.046
Dbundesl12	-.0930031	0.740	.5016664	0.826
Dbundesl13	.2009278	0.491	-.0585923	0.448
Dbundesl14	.1927373	0.388	-.2191701	0.021
Dbundesl15	.6268441	0.008	.1047016	0.849
Dbundesl16	.286547	0.311	-.0202333	0.004
Dbik_2	.7123127	0.002	-.2138616	0.683
Dbik_3	.0246498	0.872	-.1596903	0.040
Dbik_4	.2873936	0.118	-.3368483	0.369
Dbik_5	.759438	0.019	-.2437077	0.921
Dbik_6	.1190219	0.559	-.375665	0.106
Dbik_7	.406349	0.044	2.400581	0.195
Dbik_8	.4490884	0.058	-.5429105	0.014
Dbik_9	.3872928	0.197	-2.556655	0.166
Dbik_10	.7727135	0.171	-.1068527	0.130
cons	.9859379	0.096	.2323787	0.000
n	12259		11203	
Log pseudolikelihood	-2754.1834		-4782.647	
Pseudo R²	0.2346		0.1114	

Anmerkung: Die Korrektur der Standardfehler erfolgt mittels einer über Haushalte geclusterten Schätzung.

6.7 Integration der Gewichte zum Gesamtgewicht vor Kalibrierung

In diesem Schritt ging es erneut darum, die durch die Ausfallmodellierung modifizierten Haushaltsgewichte der neuen Auffrischungs- und der Bestandsstichprobe (einschließlich der Auffrischer aus Welle 2) zu kombinieren. Die doppelte Auswahlwahrscheinlichkeit, die ein nun gezogener Leistungsempfänger, der in einem der beiden vorausgegangenen Jahre in einem gemeinsamen Haushalt mit Leistungsempfängern gelebt hat, ohne Mitglied der Bedarfsgemeinschaft gewesen zu sein, besitzt, wurde wiederum ignoriert. Die neuen Designgewichte der Leistungsempfängerstichprobe rechnen im Querschnitt hoch auf alle Personen, die in 7/2006, in 7/2007 oder in 7/2008 in einem Haushalt, der mindestens eine BG enthält, gewohnt haben. Erst bei der Berechnung neuer Gewichte für die Gesamtstichprobe besteht Bedarf an einer Anpassung der Gewichte für alle Leistungsempfängerhaushalte mit Bezug in 7/2008. Dazu wurde die Inklusionswahrscheinlichkeit in der jeweils anderen Stichprobe für Fälle aus der Microm-Stichprobe (Welle 1) und der neuen Auffrischungsstichprobe (Welle 3) geschätzt. Für Fälle aus der Auffrischungsstichprobe wurde die mittlere Welle 1-Auswahlwahrscheinlichkeit in der Microm-Stichprobe in der jeweiligen Postleitzahl und jeweils die durchschnittliche Teilnahmewahrscheinlichkeit (für W1, W2 und W3) in dieser Stichprobe unterstellt. Für Fälle aus der Microm-Stichprobe wurde, falls es sich (laut Befragungsdaten) um Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug zwischen den beiden letzten Stichprobenziehungszeitpunkten (W2, W3) handelt, die durchschnittliche Auswahlwahrscheinlichkeit eines Auffrischungshaushalts in der jeweiligen PLZ und die durchschnittliche Teilnahmewahrscheinlichkeit in dieser Stichprobe unterstellt. Danach wurden die beiden Gewichte zu einem neuen Gesamtgewicht integriert.

6.8 Integration temporärer Ausfälle (Haushalte)

In Welle 3 konnten zum ersten Mal Haushalte wieder teilnehmen, die eine Welle ausgesetzt, d.h. nicht teilgenommen hatten (temporäre Ausfälle). Längsschnittgewichte für diese Haushalte werden nicht berechnet, d.h. (gewichtete) Längsschnittauswertungen können nur mit Teilnehmern über alle in Frage kommenden Wellen durchgeführt werden. Die Nichtteilnahme eines Haushalts kann nur über eine Welle stattfinden, fällt ein Haushalt über zwei aufeinander folgende Wellen aus, wird er nicht mehr kontaktiert. Für die Berechnung von gemeinsamen Querschnittsgewichten einschließlich der temporären Ausfälle erfolgte eine Konvexkombination der modifizierten Haushaltsgewichte der temporären Ausfälle und der modifizierten Haushaltsgewichte der Bestandsstichprobe (nicht der Auffrischerstichprobe) vor Kalibrierung. Die Konvexkombination der Haushaltsgewichte erfolgte also vor der Kalibrierung, diese wurde dann mit den neuen kombinierten Haushaltsgewichten durchgeführt.

Während mit den durch Ausfallmodellierung modifizierten Haushaltsgewichten bereits Hochrechnungsfaktoren für die Bestands- und Auffrischerstichprobe existierten, war es für die temporären Ausfälle notwendig, solche modifizierten Haushaltsgewichte als Schätzer für die jeweilige Population neu zu berechnen. Ausgangspunkt waren dabei die kalibrierten Haushaltsgewichte der ersten Welle.

Für die temporären Ausfälle wurde dann die Wahrscheinlichkeit der Nichtteilnahme in Welle 2 gegeben die Teilnahme in Welle 1 (Nichtteilnahme-Propensities W_2) und die Wahrscheinlichkeit der Teilnahme in Welle 3 gegeben die Nichtteilnahme in Welle 2 (Teilnahme-Propensities W_3) bestimmt. Die Wahrscheinlichkeit der Nichtteilnahme in Welle 2 ergibt sich über 1-Teilnahmewahrscheinlichkeit in Welle 2. Da die Wahrscheinlichkeit zur Wiederteilnahme in Welle 2 über Logit-Modelle für Panelbereitschaft, Erreichbarkeit und Teilnahme bereits geschätzt wurde, konnte auf die Ergebnisse dieser Modelle zurückgegriffen werden. Die Teilnahme-Propensities für die Teilnahme in Welle 3 gegeben Nichtteilnahme in Welle 2 mussten jedoch neu geschätzt werden. Verwendet wurden hierbei Logit-Modelle für Erreichbarkeit und Teilnahme mit nur wenigen Variablen (Ausfallgrund, Methode des Kontaktversuchs, Stichprobe und Kennung, ob es sich um einen Split-Haushalt der ersten Welle handelt). Die Modelle reproduzieren im Ergebnis die mittlere Teilnahmewahrscheinlichkeit der temporären Ausfälle bei einer vergleichsweise geringen Varianz der Propensities.

Das Produkt der vorhergesagten Wahrscheinlichkeiten der beiden Modelle wurde mit der Wahrscheinlichkeit der Nichtteilnahme in Welle 2 multipliziert. Das modifizierte Haushaltsgewicht der temporären Ausfälle ergibt sich dann über die Multiplikation der kalibrierten Haushaltsgewichte der ersten Welle mit dem Kehrwert dieses Produkts.

Tabelle 48: Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der temporären Ausfälle

Variablenkürzel und Referenzkategorie	Erläuterung
end_ne	Ausfallgrund in Welle 2: HH nicht erreicht
end_nidl	Ausfallgrund in Welle 2: HH nicht in der Lage, ein Interview zu geben
end_aw	Ausfallgrund in Welle 2: HH umgezogen, ungültige Telefonnummer
Referenzkategorie	Ausfallgrund in Welle 2: Verweigerung
cati1	Kontaktversuch über CATI oder CAPI: zuerst über CATI
cati2	Kontaktversuch über CATI oder CAPI: erneut über CATI
Referenzkategorie	Kontaktversuch über CATI oder CAPI: zuerst über CAPI oder Recherche
sample_microm	Teilstichprobe Microm
Referenzkategorie	Teilstichprobe BA
split	Split-Haushalt
Referenzkategorie	kein Split-Haushalt
uv_imp	Werte der unabhängigen Variablen imputiert
Referenzkategorie	Werte der unabhängigen Variablen nicht imputiert

Tabelle 49: Logit-Modelle zur Wiederteilnahme in Welle 3 gegeben Nichtteilnahme in Welle 2 für Erreichbarkeit und Teilnahme

	Kontakt		Teilnahme	
	Coef.	p	Coef.	p
end_ne	-.2280403	0.034	.2571192	0.056
end_nidl	-.4508862	0.024	-.0650028	0.805
end_aw	-.0852084	0.370	.3998799	0.001
cati1	.564227	0.000	-.2624226	0.015
cati2	-.3230414	0.128	.3277073	0.271
sample_microm	.4433349	0.000	-.1261313	0.178
split	-.3645094	0.009	.1304105	0.514
uv_imp	-.6368384	0.309	-.535854	0.567
cons	.2587525	0.008	.6734514	0.000
n	3773		2453	
Log likelihood	-2394.156		-1541.6122	
Pseudo R²	0.0198		0.0056	

Die Konvexkombination der Gewichte der Teilnehmer über alle Wellen (Bestandsstichprobe) und der temporären Ausfälle erfolgte für die Gewichte aller drei Teilstichproben i (Microm, BA und Gesamt) über die Multiplikation der jeweiligen modifizierten Haushaltsgewichte mit dem Anteil der Bestandsstichprobe bzw. der temporären Ausfälle am Gesamtstichprobenumfang, d.h. der Summe aus Bestandsstichprobe und temporären Ausfällen:

$dw_{i, hh_{temp.Ausfall}} * (n_{temp.Ausfall\ i} / (n_{temp.Ausfall\ i} + n_{Bestand\ i}))$ für die temporären Ausfälle und

$dw_{i, hh_{Bestand}} * (n_{Bestand\ i} / (n_{temp.Ausfall\ i} + n_{Bestand\ i}))$ für die Bestandsstichprobe.

6.9 Kalibrierung zum Haushaltsgewicht 3. Welle Querschnitt

Anschließend fand wieder eine Kalibrierung der modifizierten Designgewichte inklusive der Ausfallgewichtung auf Haushaltsebene mittels Raking an die Eckwerte des Statistischen Bundesamts für 2008 und für Haushalte mit Leistungsbezug an die Eckwerte der Statistik der Bundesagentur für Arbeit für Juli 2008 statt. Ebenso wie im Vorjahr wurde dabei als Eckwert auch die Zahl der Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug seit dem Vorjahr auf Ebene von Bedarfsgemeinschaften (477.034) verwendet. Dabei werden diejenigen Fälle der bisherigen Stichproben aus Welle 1 und Welle 2 auf die Eckwerte zum Arbeitslosengeld-II-Bezug der Bundesagentur für Arbeit hochgerechnet, die laut Welle 3 des Surveys im Juli 2008 Arbeitslosengeld II bezogen haben.

Das Hauptziel der Gewichtung ist der Ausgleich von Verzerrungen, die durch das Stichprobendesign (mit ungleichen Auswahlwahrscheinlichkeiten) und durch selektive Teilnahme bzw. Nichtteilnahme entstehen. Unter Verwendung der Gewichte lassen sich Populationswerte aus der Stichprobe erwartungstreu schätzen. Weisen die Gewichte allerdings eine starke Streuung auf, so kann dies zu einer großen Varianz der Schätzfunktionen führen. Hier tritt der in der Statistik immer wieder anzutreffende Trade-off zwischen Bias und Varianz auf. Die Gewichtung reduziert den Bias; eine allzu große Steigerung der Varianz durch die Gewichtung sollte allerdings auch vermieden werden. Aus diesem Grunde wird versucht, besonders große Gewichtungsfaktoren (und nachrangig auch besonders kleine Faktoren) möglichst zu umgehen und ggf. geeignete Korrekturen an den Gewichten vorzunehmen. Im Rahmen der vorliegenden Kalibrierung erfolgte dies an zwei Punkten:

- Die Eingangsgewichte für die Kalibrierung (die modifizierten Designgewichte nach Berücksichtigung der Ausfallanalysen) wurden vor der Kalibrierung getrimmt, d.h. sie wurden durch neue Eingangsgewichte ersetzt. Unter- bzw. Obergrenze der getrimmten Designgewichte wurden durch bestimmte Perzentile der Verteilung, in Abhängigkeit von der Verteilung der Designgewichte, ermittelt.
- Bei der Kalibrierung wurde zudem das Intervall der Gewichte begrenzt, d.h. es wurde eine obere und eine untere Grenze für die Gewichte festgelegt. Festgelegt wurde dabei die Gesamtbreite der Gewichte; die Spannbreite der reinen Kalibrierungsgewichte ergibt sich aus dem Verhältnis von ursprünglichen Gewichten zum getrimmten Eingangsgewicht. Zu beachten war dabei, dass engere Grenzen für die Gewichte in geringerer Streuung und damit geringerer Varianz der Schätzungen resultieren, bei zu eng gesetzten Grenzen kann allerdings die Kalibrierung aller Eckwerte unmöglich werden.

Zur Beurteilung der Gewichte wird im Folgenden neben dem Mittelwert und der Standardabweichung auch das Effektivitätsmaß angegeben (E). Das Effektivitätsmaß E basiert auf der Varianz des Gewichtungsfaktors. Das Effektivitätsmaß gibt in Prozent der realisierten Fallzahl an, wie groß bei Verwendung des Gewichts die effektive Fallzahl bei einem passiven Merkmal ist, das mit den aktiven Merkmalen nicht korreliert. Die effektive Fallzahl entspricht der Anzahl Befragter, die bei einer uneingeschränkten Zufallsauswahl gegeben die Varianz des Merkmals in der Stichprobe den gleichen Stichprobenfehler produziert hätte. Das Effektivitätsmaß drückt das Verhältnis von n zu n' als Prozentwert aus.

6.9.1 Kalibrierung der BA-Stichprobe

Die Grundgesamtheit der kumulierten BA-Stichprobe der ersten drei Wellen bilden alle Haushalte in Deutschland mit mindestens einer Bedarfsgemeinschaft mit Leistungen nach dem SGB II zu einem der bisher drei Ziehungszeitpunkte (im Juli 2006, Juli 2007 oder Juli 2008). Eine Kalibrierung findet in Welle 3 dagegen nur an die Eckwerte der BA-Statistik vom Juli 2008 statt. Damit beeinflusst die Kalibrierung nur die Gewichte der Haushalte aus der BA-Stichprobe, in denen auch im Juli 2008 noch mindestens eine Bedarfsgemeinschaft mit Leistungen nach dem SGB II lebt. Ausgangspunkt für die Kalibrierung waren die modifizierten Designgewichte inklusive der Ausfallgewichtung. Die modifizierten Designgewichte wurden am 5%-Perzentil und am 95%-Perzentil ihrer Verteilung getrimmt und anschließend so

reskaliert, dass ihre Summe wieder die Summe der ungetrimmten Designgewichte ergab. Die Hochrechnungsfaktoren der getrimmten Designgewichte reichen von 260,62 bis 1887,11 (Gewichtungsfaktoren von 0,32 bis 2,34). Das Intervall der Gesamthochrechnungsfaktoren wurde nach unten auf 82,78 und nach oben auf 3311,18 begrenzt, was einer Beschränkung der Gesamtgewichtungsfaktoren auf den Bereich von 0,01 bis 4,0 entspricht.

Für folgende Merkmale wurde kalibriert:

Bedarfsgemeinschaften Grundlage BA-Statistik:

- Zahl der Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug
- Anzahl BG im Leistungsbezug nach SGB II nach Bundesländern
- Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Personen unter 65 Jahre in der Bedarfsgemeinschaft, nach West/Ost
- Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Kinder unter 15 Jahre in der Bedarfsgemeinschaft, nach West/Ost
- Anzahl BG im Leistungsbezug nach SGB II, die aus einer/einem Alleinerziehenden mit Kindern besteht, nach West/Ost

Daneben wurde wie im Vorjahr ein zusätzlicher Eckwert mit aufgenommen, und zwar die Zahl der Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug seit dem Vorjahr auf Ebene von Bedarfsgemeinschaften (477.034).

Um die Kalibrierung durchzuführen, muss jede Eckwertvariable für jeden Haushalt einen gültigen Wert besitzen. Deshalb wurde der sehr geringe Item-Nonresponse vor der Kalibrierung imputiert. Die Imputation erfolgte über den Mittelwert bzw. den häufigsten Wert der jeweiligen Variable.

Da die Imputation einzig der Durchführbarkeit der Kalibrierung dient, wurden die imputierten Werte nach der Kalibrierung wieder auf fehlende Werte zurückgesetzt. Eine Hochrechnung mit den kalibrierten Gewichten ohne Berücksichtigung des Item-Nonresponse führt daher zu geringfügigen Abweichungen von den im Folgenden dargestellten Werten.

Tabelle 50: Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Haushalte)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik	Verteilung mit kalibrierten Gewichten
Anzahl BG im Leistungsbezug nach SGB II nach Bundesländern (16 Kategorien)	Zahl BGs Schleswig-Holstein	181	123.366	123.366
	Zahl BGs Hamburg	70	107.745	107.746
	Zahl BGs Niedersachsen	436	332.147	332.147
	Zahl BGs Bremen	49	50.367	50.367
	Zahl BGs Nordrhein-Westfalen	931	812.279	812.279
	Zahl BGs Hessen	271	216.808	216.808
	Zahl BGs Rheinland-Pfalz	153	119.932	119.932

Tabelle 50: Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Haushalte) (Fortsetzung 1)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik	Verteilung mit kalibrierten Gewichten
	Zahl BGs Baden-Württemberg	289	234.615	234.615
	Zahl BGs Bayern	372	259.709	259.709
	Zahl BGs Saarland	65	43.632	43.632
	Zahl BGs Berlin	372	331.568	331.555
	Zahl BGs Brandenburg	234	179.084	179.081
	Zahl BGs Mecklenburg-Vorpommern	148	138.195	138.195
	Zahl BGs Sachsen	326	294.155	294.149
	Zahl BGs Sachsen-Anhalt	307	196.771	196.781
	Zahl BGs Thüringen	172	136.945	136.949
Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Personen unter 65 Jahre in der Bedarfsgemeinschaft (1, 2, 3, 4 und „5 oder mehr“) und nach West/Ost (10 Kategorien)	Zahl BGs mit 1 Person unter 65 (West)	1.147	1.176.271	1.176.263
	Zahl BGs mit 2 Personen unter 65 (West)	698	486.072	486.070
	Zahl BGs mit 3 Personen unter 65 (West)	491	309.659	309.662
	Zahl BGs mit 4 Personen unter 65 (West)	271	193.824	193.828
	Zahl BGs mit 5 oder mehr Personen unter 65 (West)	210	134.775	134.779
	Zahl BGs mit 1 Person unter 65 (Ost)	678	702.562	702.548
	Zahl BGs mit 2 Personen unter 65 (Ost)	388	291.769	291.769
	Zahl BGs mit 3 Personen unter 65 (Ost)	273	156.596	156.589
	Zahl BGs mit 4 Personen unter 65 (Ost)	137	83.122	83.118
	Zahl BGs mit 5 oder mehr Personen unter 65 (Ost)	83	42.669	42.686
Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Kinder unter 15 Jahre in der Bedarfsgemeinschaft (0, 1, 2, 3, „4 oder mehr“) und nach West/Ost (10 Kategorien)	Zahl BGs ohne Kinder unter 15 Jahre (West)	1.821	1.520.533	1.520.536
	Zahl BGs mit 1 Kind unter 15 Jahre (West)	542	420.221	420.220
	Zahl BGs mit 2 Kindern unter 15 Jahre (West)	312	242.250	242.250
	Zahl BGs mit 3 Kindern unter 15 Jahre (West)	99	83.664	83.663

Tabelle 50: Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Haushalte) (Fortsetzung 2)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik	Verteilung mit kalibrierten Gewichten
	Zahl BGs mit 4 oder mehr Kindern unter 15 Jahre (West)	43	33.933	33.933
	Zahl BGs ohne Kinder unter 15 Jahre (Ost)	1.125	933.101	933.111
	Zahl BGs mit 1 Kind unter 15 Jahre (Ost)	273	209.901	209.891
	Zahl BGs mit 2 Kindern unter 15 Jahre (Ost)	124	96.928	96.921
	Zahl BGs mit 3 Kindern unter 15 Jahre (Ost)	25	26.552	26.552
	Zahl BGs mit 4 oder mehr Kindern unter 15 Jahre (Ost)	12	10.236	10.236
Anzahl BG im Leistungsbezug nach SGB II, die aus einer/einem Alleinerziehenden mit Kindern besteht, nach West/Ost (2 Kategorien)	Zahl BGs mit einem/einer Alleinerziehenden (West)	720	459.676	459.676
	Rest BGs ohne Alleinerziehenden (West)	2.097	1.840.924	1.840.926
	Zahl BGs mit einem/einer Alleinerziehenden (Ost)	302	202.694	202.688
	Rest BGs ohne Alleinerziehenden (Ost)	1.257	1.074.024	1.074.022

Tabelle 51: Kenngrößen der Verteilung der Gewichte

1%-Perzentil	129.9741
5%-Perzentil	222.0341
10%-Perzentil	301.1921
25%-Perzentil	411.9451
50%-Perzentil	630.5933
75%-Perzentil	1163.411
90%-Perzentil	1657.794
95%-Perzentil	1963.006
99%-Perzentil	2336.944
Mittelwert	827.7941
Standardabweichung	557.7468
Minimum	82.78
Maximum	3311.18
Fallzahl	4231
Effektivitätsmaß	68,8%

6.9.2 Microm-Stichprobe

Die Grundgesamtheit bilden alle Privathaushalte in Deutschland. Ausgangspunkt für die Kalibrierung waren die modifizierten Designgewichte inklusive der Ausfallgewichtung. Die modifizierten Designgewichte wurden am 5%-Perzentil und am 95%-Perzentil ihrer Verteilung getrimmt und anschließend so reskaliert, dass ihre Summe wieder die Summe der ungetrimmten Designgewichte ergab. Die Hochrechnungsfaktoren der getrimmten Designgewichte reichen von 1794,10 bis 26840,67 (Gewichtungsfaktoren von 0,18 bis 2,66). Das Intervall der Gesamthochrechnungsfaktoren wurde nach unten auf 101,09 und nach oben auf 73799,05 begrenzt, was einer Beschränkung der Gesamtgewichtungsfaktoren auf den Bereich von 0,01 bis 7,3 entspricht.

Für folgende Merkmale wurde kalibriert:

Bedarfsgemeinschaften: Grundlage BA-Statistik:

- Anzahl BG im Leistungsbezug nach SGB II nach Bundesländern
- Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Personen unter 65 Jahre in der Bedarfsgemeinschaft, nach West/Ost
- Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Kinder unter 15 Jahre in der Bedarfsgemeinschaft, nach West/Ost
- Anzahl BG im Leistungsbezug nach SGB II, die aus einer/einem Alleinerziehenden mit Kindern besteht, nach West/Ost

Haushalte: Grundlage Mikrozensus 2008:

- Anzahl der Haushalte nach Bundesland und BIK-Typ
- Anzahl der Haushalte nach Haushaltsgröße und West/Ost
- Anzahl der Haushalte nach „Kinder unter 15 Jahre im Haushalt ja/nein“ und West/Ost

Um die Kalibrierung durchzuführen, muss jede Eckwertvariable für jeden Haushalt einen gültigen Wert besitzen. Deshalb wurde der sehr geringe Item-Nonresponse vor der Kalibrierung imputiert. Die Imputation erfolgte über den Mittelwert bzw. den häufigsten Wert der jeweiligen Variable.

Da die Imputation einzig der Durchführbarkeit der Kalibrierung dient, wurden die imputierten Werte nach der Kalibrierung wieder auf fehlende Werte zurückgesetzt. Eine Hochrechnung mit den kalibrierten Gewichten ohne Berücksichtigung des Item-Nonresponse führt daher zu geringfügigen Abweichungen von den im Folgenden dargestellten Werten.

Tabelle 52: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
Anzahl BG im Leistungsbezug nach SGB II nach Bundesländern (16 Kategorien)	Zahl BGs Schleswig-Holstein	16	123.366	123.695
	Zahl BGs Hamburg	2	107.745	107.745
	Zahl BGs Niedersachsen	47	332.147	332.417
	Zahl BGs Bremen	7	50.367	50.367
	Zahl BGs Nordrhein-Westfalen	86	812.279	811.044
	Zahl BGs Hessen	17	216.808	216.808
	Zahl BGs Rheinland-Pfalz	9	119.932	119.931
	Zahl BGs Baden-Württemberg	15	234.615	234.615
	Zahl BGs Bayern	46	259.709	259.857
	Zahl BGs Saarland	13	43.632	43.632
	Zahl BGs Berlin	13	331.568	331.576
	Zahl BGs Brandenburg	25	179.084	178.620
	Zahl BGs Mecklenburg-Vorpommern	6	138.195	138.180
	Zahl BGs Sachsen	19	294.155	294.282
	Zahl BGs Sachsen-Anhalt	27	196.771	196.759
	Zahl BGs Thüringen	12	136.945	136.941
Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Personen unter 65 Jahre in der Bedarfsgemeinschaft (1, 2, 3, 4 und „5 oder mehr“) und nach West/Ost (10 Kategorien)	Zahl BGs mit 1 Person unter 65 (West)	73	1.176.271	1.176.265
	Zahl BGs mit 2 Personen unter 65 (West)	60	486.072	484.356
	Zahl BGs mit 3 Personen unter 65 (West)	55	309.659	310.563
	Zahl BGs mit 4 Personen unter 65 (West)	41	193.824	194.156
	Zahl BGs mit 5 oder mehr Personen unter 65 (West)	29	134.775	134.772
	Zahl BGs mit 1 Person unter 65 (Ost)	25	702.562	702.376
	Zahl BGs mit 2 Personen unter 65 (Ost)	33	291.769	291.724
	Zahl BGs mit 3 Personen unter 65 (Ost)	29	156.596	156.308
Zahl BGs mit 4 Personen unter 65 (Ost)	7	83.122	83.215	

Tabelle 52: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 1)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl BGs mit 5 oder mehr Personen unter 65 (Ost)	8	42.669	42.735
Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Kinder unter 15 Jahren in der Bedarfsgemeinschaft (0, 1, 2, 3, „4 oder mehr“) und nach West/Ost (10 Kategorien)	Zahl BGs ohne Kinder unter 15 Jahre (West)	139	1.520.533	1.520.053
	Zahl BGs mit 1 Kind unter 15 Jahre (West)	47	420.221	420.220
	Zahl BGs mit 2 Kindern unter 15 Jahre (West)	48	242.250	242.243
	Zahl BGs mit 3 Kindern unter 15 Jahre (West)	13	83.664	83.664
	Zahl BGs mit 4 oder mehr Kindern unter 15 Jahre (West)	11	33.933	33.933
	Zahl BGs ohne Kinder unter 15 Jahre (Ost)	70	933.101	932.784
	Zahl BGs mit 1 Kind unter 15 Jahre (Ost)	16	209.901	209.862
	Zahl BGs mit 2 Kindern unter 15 Jahre (Ost)	8	96.928	96.924
	Zahl BGs mit 3 Kindern unter 15 Jahre (Ost)	7	26.552	26.550
	Zahl BGs mit 4 oder mehr Kindern unter 15 Jahre (Ost)	1	10.236	10.237
Anzahl BG im Leistungsbezug nach SGB II, die aus einer/einem Alleinerziehenden mit Kindern besteht, nach West/Ost (2 Kategorien)	Zahl BGs mit einem/einer Alleinerziehenden (West)	68	459.676	459.676
	Rest BGs ohne Alleinerziehenden (West)	190	1.840.924	1.840.436
	Zahl BGs mit einem/einer Alleinerziehenden (Ost)	22	202.694	202.690
	Rest BGs ohne Alleinerziehenden (Ost)	80	1.074.024	1.073.667
Anzahl der Haushalte nach Bundesland und BIK-Typ (Schreibweise: „Bundesland.BIK-Größenkategorie“)	1.1 bis 1.4	21	316.000	315.870
	1.5	22	81.000	80.841
	1.6	15	66.000	65.856
	1.7	26	139.000	139.204
	1.8	19	374.000	374.233
	1.9	38	170.000	169.963
	1.10	27	208.000	208.214

Tabelle 52: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 2)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	2.10	42	957.000	957.075
	3.1 bis 3.2	13	151.000	151.034
	3.3	52	318.000	318.236
	3.4	30	458.000	458.229
	3.5	31	406.000	406.175
	3.7	126	906.000	906.204
	3.8	71	541.000	541.041
	3.9	71	640.000	640.022
	3.10	41	380.000	380.148
	4.8 bis 4.10	30	352.000	352.150
	5.2 bis 5.3	43	355.000	354.910
	5.4	69	1.037.000	1.037.127
	5.5	80	643.000	643.176
	5.6	31	304.000	304.040
	5.7	101	867.000	867.164
	5.8	205	2.551.000	2.551.411
	5.9	49	318.000	318.236
	5.10	237	2.446.000	2.446.515
	6.1 bis 6.2	30	67.000	67.039
	6.3	45	330.000	330.067
	6.4	32	237.000	237.001
	6.5 bis 6.7	64	627.000	627.008
	6.8	32	462.000	462.172
	6.9	59	363.000	363.192
	6.10	55	776.000	776.071
	7.1	17	208.000	208.214
	7.2	10	97.000	97.009
	7.3	28	189.000	188.891
	7.4	19	150.000	149.851
	7.5	18	166.000	166.019
	7.6	14	79.000	78.869
	7.7	36	399.000	399.077

Tabelle 52: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 3)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	7.8	18	326.000	326.123
	7.9 bis 7.10	12	257.000	257.113
	8.1 bis 8.3	39	645.000	645.148
	8.4	71	516.000	516.197
	8.5 bis 8.6	43	466.000	466.116
	8.7	77	1.058.000	1.058.027
	8.8	74	786.000	785.929
	8.9	22	355.000	354.910
	8.10	102	1.064.000	1.064.337
	9.1	10	93.000	93.065
	9.2	15	263.000	263.028
	9.3	70	500.000	500.029
	9.4	89	668.000	668.020
	9.5 bis 9.6	43	455.000	455.074
	9.7	117	1.007.000	1.007.157
	9.8	64	631.000	630.952
	9.9	117	721.000	721.257
	9.10	157	1.478.000	1.478.399
	10.3 bis 10.5	20	154.000	154.189
	10.7 bis 10.8	26	333.000	333.221
	11.10	112	1.949.000	1.949.458
	12.1 bis 12.3	18	262.000	261.874
	12.4	20	250.000	250.042
	12.5 bis 12.6	18	138.000	138.036
	12.7	21	122.000	121.866
	12.8	20	142.000	141.979
	12.9 bis 12.10	30	323.000	323.005
	13.1 bis 13.3	9	192.000	192.067
	13.4	12	179.000	179.054
	13.5 bis 13.6	14	156.000	156.178
	13.7	19	104.000	104.119
	13.8	14	217.000	216.913

Tabelle 52: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 4)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	14.1	24	19.000	18.931
	14.2	46	137.000	136.853
	14.3	16	183.000	182.998
	14.4	10	259.000	259.112
	14.5	13	172.000	171.953
	14.6	17	132.000	132.120
	14.7 bis 14.8	10	373.000	373.090
	14.9	7	230.000	229.928
	14.10	48	663.000	663.360
	15.1 bis 15.3	28	205.000	205.082
	15.4	28	161.000	160.910
	15.5 bis 15.6	17	313.000	313.145
	15.7	34	236.000	236.238
	15.8	35	281.000	281.198
	16.1 bis 16.3	16	233.000	233.084
	16.4	35	264.000	264.241
	16.5 bis 16.6	40	307.000	307.229
	16.7 bis 16.8	35	317.000	317.091
Anzahl der Haushalte nach Haushaltsgröße (1,2,3,4,„5 und mehr Personen“) und West/Ost (10 Kategorien)	Anzahl Haushalte mit 1 Person (West)	764	11.753.010	11.753.015
	Anzahl Haushalte mit 2 Personen (West)	1.115	10.484.510	10.484.721
	Anzahl Haushalte mit 3 Personen (West)	525	4.043.850	4.043.870
	Anzahl Haushalte mit 4 Personen (West)	503	3.354.560	3.354.511
	Anzahl Haushalte mit 5 oder mehr Pers. (West)	228	1.279.700	1.279.730
	Anzahl Haushalte mit 1 Person (Ost)	214	3.566.830	3.566.699
	Anzahl Haushalte mit 2 Personen (Ost)	304	3.023.290	3.023.250
	Anzahl Haushalte mit 3 Personen (Ost)	135	1.178.940	1.178.760

Tabelle 52: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Haushalte) (Fortsetzung 5)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Anzahl Haushalte mit 4 Personen (Ost)	81	598.880	599.038
	Anzahl Haushalte mit 5 oder mehr Pers. (Ost)	32	153.350	153.406
Anzahl der Haushalte nach „Kinder unter 15 Jahre im Haushalt ja/nein“ und West/Ost	Anzahl Haushalte mit Kindern unter 15 (West)	878	5.799.000	5.799.211
	Anzahl Haushalte ohne Kinder unter 15 (West)	2.257	25.116.000	25.116.637
	Anzahl Haushalte mit Kindern unter 15 (Ost)	142	1.215.000	1.215.054
	Anzahl Haushalte ohne Kinder unter 15 (Ost)	624	7.306.000	7.306.099

Tabelle 53: Kenngrößen der Verteilung der Gewichte

1%-Perzentil	867.5521
5%-Perzentil	1756.647
10%-Perzentil	2241.188
25%-Perzentil	4117.769
50%-Perzentil	7659.663
75%-Perzentil	13893.96
90%-Perzentil	22045.54
95%-Perzentil	25891.3
99%-Perzentil	34531.22
Mittelwert	10109.46
Standardabweichung	8008.667
Minimum	106.1988
Maximum	73080.57
Fallzahl	3901
Effektivitätsmaß	61,4%

6.9.3 Gesamtstichprobe

Die Grundgesamtheit bilden alle Privathaushalte in Deutschland. Ausgangspunkt für die Kalibrierung waren die modifizierten Designgewichte inklusive der Ausfallgewichtung. Die modifizierten Designgewichte wurden am 5%-Perzentil und am 95%-Perzentil ihrer Verteilung getrimmt und anschließend so reskaliert, dass ihre Summe wieder die Summe der

ungetrimmten Designgewichte ergab. Die Hochrechnungsfaktoren der getrimmten Designgewichte reichen von 211,64 bis 19028,16 (Gewichtungsfaktoren von 0,05 bis 4,60). Das Intervall der Gesamthochrechnungsfaktoren wurde nach unten auf 41,36 und nach oben auf 35156,21 begrenzt, was einer Beschränkung der Gesamtgewichtungsfaktoren auf den Bereich von 0,01 bis 8,5 entspricht.

Für folgende Merkmale wurde kalibriert:

Bedarfsgemeinschaften Grundlage BA-Statistik:

- Anzahl BG im Leistungsbezug nach SGB II nach Bundesländern
- Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Personen unter 65 Jahre in der Bedarfsgemeinschaft, nach West/Ost
- Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Kinder unter 15 Jahre in der Bedarfsgemeinschaft, nach West/Ost
- Anzahl BG im Leistungsbezug nach SGB II, die aus einer/einem Alleinerziehenden mit Kindern besteht, nach West/Ost

Haushalte Grundlage Mikrozensus 2008:

- Anzahl der Haushalte nach Bundesland und BIK-Typ
- Anzahl der Haushalte nach Haushaltsgröße und West/Ost
- Anzahl der Haushalte nach „Kinder unter 15 Jahre im Haushalt ja/nein“ und West/Ost

Daneben wurde auch bei der Gesamtstichprobe als zusätzlicher Eckwert die Zahl der Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug seit dem Vorjahr auf Ebene von Bedarfsgemeinschaften (477.034) mit aufgenommen.

Um die Kalibrierung durchzuführen, muss jede Eckwertvariable für jeden Haushalt einen gültigen Wert besitzen. Deshalb wurde der sehr geringe Item-Nonresponse vor der Kalibrierung imputiert. Die Imputation erfolgte über den Mittelwert bzw. den häufigsten Wert der jeweiligen Variable.

Da die Imputation einzig der Durchführbarkeit der Kalibrierung dient, wurden die imputierten Werte nach der Kalibrierung wieder auf fehlende Werte zurückgesetzt. Eine Hochrechnung mit den kalibrierten Gewichten ohne Berücksichtigung des Item-Nonresponse führt daher zu geringfügigen Abweichungen von den im Folgenden dargestellten Werten.

Tabelle 54: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
Anzahl BG im Leistungsbezug nach SGB II nach Bundesländern (16 Kategorien)	Zahl BGs Schleswig-Holstein	197	123.366	123.970
	Zahl BGs Hamburg	72	107.745	105.131
	Zahl BGs Niedersachsen	483	332.147	331.190
	Zahl BGs Bremen	56	50.367	47.804
	Zahl BGs Nordrhein-Westfalen	1.017	812.279	813.795
	Zahl BGs Hessen	288	216.808	216.588
	Zahl BGs Rheinland-Pfalz	162	119.932	118.016
	Zahl BGs Baden-Württemberg	304	234.615	234.657
	Zahl BGs Bayern	418	259.709	258.818
	Zahl BGs Saarland	78	43.632	40.953
	Zahl BGs Berlin	385	331.568	338.139
	Zahl BGs Brandenburg	259	179.084	179.032
	Zahl BGs Mecklenburg-Vorpommern	154	138.195	133.426
	Zahl BGs Sachsen	345	294.155	299.359
	Zahl BGs Sachsen-Anhalt	334	196.771	198.354
Zahl BGs Thüringen	184	136.945	129.345	
Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Personen unter 65 Jahre in der Bedarfsgemeinschaft (1, 2, 3, 4 und „5 oder mehr“) und nach West/Ost (10 Kategorien)	Zahl BGs mit 1 Person unter 65 (West)	1.220	1.176.271	1.173.311
	Zahl BGs mit 2 Personen unter 65 (West)	758	486.072	487.197
	Zahl BGs mit 3 Personen unter 65 (West)	546	309.659	303.253
	Zahl BGs mit 4 Personen unter 65 (West)	312	193.824	190.407
	Zahl BGs mit 5 oder mehr Personen unter 65 (West)	239	134.775	136.754
	Zahl BGs mit 1 Person unter 65 (Ost)	703	702.562	702.518
	Zahl BGs mit 2 Personen unter 65 (Ost)	421	291.769	293.264
	Zahl BGs mit 3 Personen unter 65 (Ost)	302	156.596	153.588
	Zahl BGs mit 4 Personen unter 65 (Ost)	144	83.122	87.070

Tabelle 54: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 1)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl BGs mit 5 oder mehr Personen unter 65 (Ost)	91	42.669	41.215
Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Kinder unter 15 Jahre in der Bedarfsgemeinschaft (0, 1, 2, 3, „4 oder mehr“) und nach West/Ost (10 Kategorien)	Zahl BGs ohne Kinder unter 15 Jahre (West)	1.960	1.520.533	1.514.073
	Zahl BGs mit 1 Kind unter 15 Jahre (West)	589	420.221	414.400
	Zahl BGs mit 2 Kindern unter 15 Jahre (West)	360	242.250	244.125
	Zahl BGs mit 3 Kindern unter 15 Jahre (West)	112	83.664	83.664
	Zahl BGs mit 4 oder mehr Kindern unter 15 Jahre (West)	54	33.933	34.660
	Zahl BGs ohne Kinder unter 15 Jahre (Ost)	1.195	933.101	931.962
	Zahl BGs mit 1 Kind unter 15 Jahre (Ost)	289	209.901	211.236
	Zahl BGs mit 2 Kindern unter 15 Jahre (Ost)	132	96.928	97.669
	Zahl BGs mit 3 Kindern unter 15 Jahre (Ost)	32	26.552	26.552
	Zahl BGs mit 4 oder mehr Kindern unter 15 Jahre (Ost)	13	10.236	10.236
Anzahl BG im Leistungsbezug nach SGB II, die aus einer/einem Alleinerziehenden mit Kindern besteht, nach West/Ost (2 Kategorien)	Zahl BGs mit einem/einer Alleinerziehenden (West)	788	459.676	455.311
	Rest BGs ohne Alleinerziehenden (West)	2.287	1.840.924	1.835.611
	Zahl BGs mit einem/einer Alleinerziehenden (Ost)	324	202.694	202.410
	Rest BGs ohne Alleinerziehenden (Ost)	1.337	1.074.024	1.075.245
Anzahl der Haushalte nach Bundesland und BIK-Typ (Schreibweise: „Bundesland.BIK-Größenkategorie“)	1.1 bis 1.4	77	316.000	316.059
	1.5	33	81.000	81.015
	1.6	37	66.000	66.012
	1.7	45	139.000	139.026
	1.8	100	374.000	374.070
	1.9	66	170.000	170.032
	1.10	68	208.000	208.039

Tabelle 54: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 2)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	2.10	126	957.000	957.179
	3.1 bis 3.2	20	151.000	151.028
	3.3	111	318.000	318.060
	3.4	74	458.000	458.086
	3.5	48	406.000	406.076
	3.7	299	906.000	906.170
	3.8	211	541.000	541.101
	3.9	127	640.000	640.120
	3.10	105	380.000	380.071
	4.8 bis 4.10	91	352.000	352.066
	5.2 bis 5.3	103	355.000	355.066
	5.4	135	1.037.000	1.037.194
	5.5	180	643.000	643.120
	5.6	63	304.000	304.057
	5.7	220	867.000	867.162
	5.8	557	2.551.000	2.551.477
	5.9	94	318.000	318.060
	5.10	653	2.446.000	2.446.458
	6.1 bis 6.2	47	67.000	67.013
	6.3	96	330.000	330.062
	6.4	68	237.000	237.044
	6.5 bis 6.7	118	627.000	627.117
	6.8	99	462.000	462.086
	6.9	95	363.000	363.068
	6.10	142	776.000	776.145
	7.1	23	208.000	208.039
	7.2	17	97.000	97.018
	7.3	40	189.000	189.035
	7.4	31	150.000	150.028
	7.5	45	166.000	166.031
	7.6	57	79.000	79.015
	7.7	61	399.000	399.075

Tabelle 54: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 3)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	7.8	48	326.000	326.061
	7.9 bis 7.10	45	257.000	257.048
	8.1 bis 8.3	70	645.000	645.121
	8.4	108	516.000	516.097
	8.5 bis 8.6	54	466.000	466.087
	8.7	137	1.058.000	1.058.198
	8.8	160	786.000	786.147
	8.9	37	355.000	355.066
	8.10	233	1.064.000	1.064.199
	9.1	12	93.000	93.017
	9.2	28	263.000	263.049
	9.3	106	500.000	500.094
	9.4	133	668.000	668.125
	9.5 bis 9.6	66	455.000	455.085
	9.7	164	1.007.000	1.007.188
	9.8	188	631.000	631.118
	9.9	150	721.000	721.135
	9.10	325	1.478.000	1.478.277
	10.3 bis 10.5	49	154.000	154.029
	10.7 bis 10.8	72	333.000	333.062
	11.10	572	1.949.000	1.949.506
	12.1 bis 12.3	117	262.000	262.068
	12.4	47	250.000	250.065
	12.5 bis 12.6	87	138.000	138.036
	12.7	60	122.000	122.032
	12.8	35	142.000	142.037
	12.9 bis 12.10	78	323.000	323.084
	13.1 bis 13.3	49	192.000	192.050
	13.4	55	179.000	179.047
	13.5 bis 13.6	60	156.000	156.041
	13.7	54	104.000	104.027
	13.8	45	217.000	217.056

Tabelle 54: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 4)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	14.1	36	19.000	19.005
	14.2	119	137.000	137.036
	14.3	29	183.000	183.048
	14.4	53	259.000	259.067
	14.5	42	172.000	172.045
	14.6	109	132.000	132.034
	14.7 bis 14.8	54	373.000	373.097
	14.9	23	230.000	230.060
	14.10	155	663.000	663.173
	15.1 bis 15.3	76	205.000	205.053
	15.4	68	161.000	161.042
	15.5 bis 15.6	67	313.000	313.082
	15.7	161	236.000	236.061
	15.8	158	281.000	281.073
	16.1 bis 16.3	41	233.000	233.061
	16.4	113	264.000	264.069
	16.5 bis 16.6	101	307.000	307.080
	16.7 bis 16.8	104	317.000	317.083
Anzahl der Haushalte nach Haushaltsgröße (1,2,3,4,„5 und mehr Personen“) und West/Ost (10 Kategorien)	Anzahl Haushalte mit 1 Person (West)	2.158	11.753.010	11.753.070
	Anzahl Haushalte mit 2 Personen (West)	2.053	10.484.510	10.484.563
	Anzahl Haushalte mit 3 Personen (West)	1.183	4.043.850	4.043.869
	Anzahl Haushalte mit 4 Personen (West)	888	3.354.560	3.354.575
	Anzahl Haushalte mit 5 oder mehr Pers. (West)	485	1.279.700	1.279.706
	Anzahl Haushalte mit 1 Person (Ost)	1.053	3.566.830	3.566.802
	Anzahl Haushalte mit 2 Personen (Ost)	836	3.023.290	3.023.265
	Anzahl Haushalte mit 3 Personen (Ost)	494	1.178.940	1.178.928

Tabelle 54: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) (Fortsetzung 5)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Anzahl Haushalte mit 4 Personen (Ost)	264	598.880	598.873
	Anzahl Haushalte mit 5 oder mehr Pers. (Ost)	121	153.350	153.348
Anzahl der Haushalte nach „Kinder unter 15 Jahre im Haushalt ja/nein“ und West/Ost	Anzahl Haushalte mit Kindern unter 15 (West)	2.161	5.799.000	5.799.147
	Anzahl Haushalte ohne Kinder unter 15 (West)	4.606	25.116.000	25.116.637
	Anzahl Haushalte mit Kindern unter 15 (Ost)	683	1.215.000	1.215.031
	Anzahl Haushalte ohne Kinder unter 15 (Ost)	2.085	7.306.000	7.306.185

Tabelle 55: Kenngrößen der Verteilung der Gewichte

1%-Perzentil	68.68383
5%-Perzentil	120.0191
10%-Perzentil	169.9333
25%-Perzentil	353.8263
50%-Perzentil	801.0488
75%-Perzentil	5528.988
90%-Perzentil	14697.51
95%-Perzentil	19157.05
99%-Perzentil	23630.28
Mittelwert	4136.025
Standardabweichung	6211.77
Minimum	41.3601
Maximum	35155.65
Fallzahl	9535
Effektivitätsmaß	30,7%

6.10 Kalibrierung zum Personengewicht 3. Welle Querschnitt

Die Kalibrierung der Personengewichte wurde – wie in den beiden Vorgängerwellen – unter der Restriktion möglichst geringer Abweichung von den kalibrierten Haushaltsgewichten vorgenommen und basiert damit nicht direkt auf den Personengewichten der ersten Welle. Vielmehr wurden die kalibrierten Haushaltsgewichte an die einzelnen Haushaltsmitglieder

gewissermaßen vererbt. Anschließend wurden diese Eingangsgewichte auf Personenebene kalibriert.

Ebenso wie im Vorjahr wurde dabei als Eckwert auch die Zahl der Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug seit dem Vorjahr auf Ebene von Personen zwischen 15 und 64 Jahre (648.988) verwendet. Es werden erneut diejenigen Fälle der Stichproben aus Welle 1 und Welle 2 auf die Eckwerte zum Arbeitslosengeld-II-Bezug der Bundesagentur für Arbeit hochgerechnet, die laut Welle 3 des Surveys im Juli 2008 Arbeitslosengeld II bezogen haben.

Vor der Kalibrierung wurden die kalibrierten Haushaltsgewichte, die das Eingangsgewicht bilden, ebenfalls getrimmt. Auch bei der Kalibrierung der Personengewichte wurde zudem der Bereich der Gewichte auf ein bestimmtes Intervall festgelegt.

6.10.1 BA-Stichprobe

Die Grundgesamtheit der kumulierten BA-Stichprobe der ersten drei Wellen bilden alle Personen im Alter ab 15 Jahre, die in einem Haushalt leben, in dem es zu einem der bisher drei Ziehungszeitpunkte (im Juli 2006, Juli 2007 oder Juli 2008) mindestens eine Bedarfsgemeinschaft mit Bezug von Leistungen nach dem SGB II gab. Für die Kalibrierung wurden lediglich all diejenigen Personen im Alter ab 15 Jahre, die in einer Bedarfsgemeinschaft mit Leistungen nach dem SGB II im Juli 2009 leben, berücksichtigt. Personen, die in Haushalten leben, die den Bezug inzwischen komplett verlassen haben und Personen, die zwar noch in einem Haushalt mit mindestens einer Bedarfsgemeinschaft nach SGB II leben, selbst aber nicht Teil einer Bedarfsgemeinschaft sind, wurden für die Kalibrierung aus dem Datensatz entfernt. Die Gewichte für diese Personen wurden über einen anderen Weg berechnet (s.u.).

Ausgangspunkt für die Kalibrierung sind die kalibrierten Haushaltsgewichte der BA-Stichprobe. Diese wurden am 5%-Perzentil und am 95%-Perzentil ihrer Verteilung getrimmt und anschließend so reskaliert, dass ihre Summe wieder die Summe der ungetrimmten kalibrierten Haushaltsgewichte ergab. Die getrimmten Hochrechnungsfaktoren reichen von 187,22 bis 1927,14 (Gewichtungsfaktoren von 0,24 bis 2,51). Das Intervall der Gesamthochrechnungsfaktoren wurde nach unten auf 144,25 und nach oben auf 5289,12 begrenzt, was einer Beschränkung der Gesamtgewichtungsfaktoren auf den Bereich von 0,15 bis 5,5 entspricht.

Für folgende Merkmale wurde kalibriert:

Leistungsempfänger Grundlage BA-Statistik:

- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Bundesländern
- Anzahl der Personen in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Alter (15-24 und 25-64)

- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Geschlecht, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach „Alleinerziehend ja/nein“, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Staatsangehörigkeit (deutsch/nicht deutsch)

Daneben wurde wie im Vorjahr als zusätzlicher Eckwert die Zahl der Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug seit dem Vorjahr auf Ebene von Personen zwischen 15 und 64 Jahre (648.988) mit aufgenommen.

Um die Kalibrierung durchzuführen, muss jede Eckwertvariable für jede Person einen gültigen Wert besitzen. Deshalb wurde der sehr geringe Item-Nonresponse vor der Kalibrierung imputiert. Die Imputation erfolgte über den Mittelwert bzw. den häufigsten Wert der jeweiligen Variable.

Da die Imputation einzig der Durchführbarkeit der Kalibrierung dient, wurden die imputierten Werte nach der Kalibrierung wieder auf fehlende Werte zurückgesetzt. Eine Hochrechnung mit den kalibrierten Gewichten ohne Berücksichtigung des Item-Nonresponse führt daher zu geringfügigen Abweichungen von den im Folgenden dargestellten Werten.

Tabelle 56: Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Personen)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik	Verteilung mit kalibrierten Gewichten
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Bundesländern (16 Kategorien)	Zahl Personen in BGs Schleswig-Holstein	223	175.064	175.064
	Zahl Personen in BGs Hamburg	72	147.471	147.471
	Zahl Personen in BGs Niedersachsen	528	481.832	481.831
	Zahl Personen in BGs Bremen	61	70.213	70.213
	Zahl Personen in BGs Nordrhein-Westfalen	1.141	1.184.500	1.184.502
	Zahl Personen in BGs Hessen	323	316.240	316.239
	Zahl Personen in BGs Rheinland-Pfalz	179	174.363	174.363
	Zahl Personen in BGs Baden-Württemberg	336	330.928	330.928
	Zahl Personen in BGs Bayern	452	354.691	354.691
	Zahl Personen in BGs Saarland	78	61.668	61.668

**Tabelle 56: Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Personen)
(Fortsetzung)**

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik	Verteilung mit kalibrierten Gewichten
	Zahl Personen in BGs Berlin	454	454.535	454.535
	Zahl Personen in BGs Brandenburg	288	254.167	254.167
	Zahl Personen in BGs Mecklenburg-Vorpommern	170	195.559	195.559
	Zahl Personen in BGs Sachsen	405	414.878	414.878
	Zahl Personen in BGs Sachsen-Anhalt	384	282.916	282.916
	Zahl Personen in BGs Thüringen	201	192.959	192.959
Anzahl der Personen in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Alter (15-24 und 25-64; 2 Kategorien)	Zahl Personen in BGs im Alter 15-24	874	1.004.739	1.004.740
	Zahl Personen in BGs im Alter 25-64	4.421	4.087.245	4.087.244
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Geschlecht und nach West/Ost (4 Kategorien)	Zahl Männer in BGs (West)	1.491	1.565.168	1.565.174
	Zahl Frauen in BGs (West)	1.902	1.731.802	1.731.797
	Zahl Männer in BGs (Ost)	896	903.032	903.033
	Zahl Frauen in BGs (Ost)	1.006	891.982	891.981
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach „Alleinerziehend ja/nein“, nach Geschlecht und nach West/Ost (8 Kategorien)	Zahl alleinerziehende in BGs (West)	690	459.676	459.676
	Zahl alleinerziehende in BGs (Ost)	290	202.694	202.694
	Zahl nicht alleinerziehende in BGs (West)	2.703	2.837.294	2.837.295
	Zahl nicht alleinerziehende in BGs (Ost)	1.612	1.592.320	1.592.319
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Staatsangehörigkeit (deutsch/ nicht deutsch) und nach West/Ost	Zahl deutscher Personen in BGs	4.736	4.116.358	4.127.427
	Zahl nicht deutscher Personen in BGs	559	961.971	964.557

Tabelle 57: Kenngrößen der Verteilung der Gewichte

1%-Perzentil	164.9536
5%-Perzentil	230.2621
10%-Perzentil	308.6981
25%-Perzentil	404.0884
50%-Perzentil	735.0017
75%-Perzentil	1280.256
90%-Perzentil	1943.159
95%-Perzentil	2435.064
99%-Perzentil	3829.328
Mittelwert	961.6589
Standardabweichung	762.4775
Minimum	144.25
Maximum	5289.12
Fallzahl	5295
Effektivitätsmaß	61,4%

6.10.2 Microm-Stichprobe

Die Grundgesamtheit bilden alle über 14-jährigen Personen in Privathaushalten in Deutschland. Ausgangspunkt für die Kalibrierung waren die kalibrierten Haushaltsgewichte der Microm-Stichprobe. Diese wurden am 10%-Perzentil und am 90%-Perzentil ihrer Verteilung getrimmt und anschließend so reskaliert, dass ihre Summe wieder die Summe der ungetrimmten kalibrierten Haushaltsgewichte ergab. Die getrimmten Hochrechnungsfaktoren reichen von 2768,07 bis 26258,62 (Gewichtungsfaktoren von 0,24 bis 2,31). Das Intervall der Gesamthochrechnungsfaktoren wurde nach unten auf 567,55 und nach oben auf 147563,72 begrenzt, was einer Beschränkung der Gesamtgewichtungsfaktoren auf den Bereich von 0,05 bis 13,0 entspricht.

Für folgende Merkmale wurde kalibriert:

Leistungsempfänger Grundlage BA-Statistik:

- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Bundesländern
- Anzahl der Personen in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Alter (15-24 und 25-64)
- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Geschlecht, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach „Alleinerziehend ja/nein“, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Staatsangehörigkeit (deutsch/nicht deutsch)

Bevölkerung Grundlage Mikrozensus 2008:

- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Bundesland
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Alter, Geschlecht und West/Ost
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Haushaltsgröße, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach höchstem Schulabschluss, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Familienstand und West/Ost
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Staatsangehörigkeiten

Bevölkerung Grundlage BA Statistik:

- Anzahl der Arbeitslosen inklusive Maßnahmeteilnehmer nach West/Ost
- Anzahl der sozialversicherungspflichtig Beschäftigten nach West/Ost

Als Quelle für den Eckwert des Erwerbsstatus wurde die BA-Statistik verwendet, da die Definition der Arbeitslosigkeit und der sozialversicherungspflichtigen Beschäftigung in PASS nicht dem ILO-Konzept des statistischen Bundesamtes entspricht, den Statistiken der BA aber entnommen werden kann.

Um die Kalibrierung durchzuführen, muss jede Eckwertvariable für jede Person einen gültigen Wert besitzen. Deshalb wurde der sehr geringe Item-Nonresponse vor der Kalibrierung imputiert. Die Imputation erfolgte über den Mittelwert bzw. den häufigsten Wert der jeweiligen Variable.

Da die Imputation einzig der Durchführbarkeit der Kalibrierung dient, wurden die imputierten Werte nach der Kalibrierung wieder auf fehlende Werte zurückgesetzt. Eine Hochrechnung mit den kalibrierten Gewichten ohne Berücksichtigung des Item-Nonresponse führt daher zu geringfügigen Abweichungen von den im Folgenden dargestellten Werten.

Tabelle 58: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Bundesländern (16 Kategorien)	Zahl Personen in BGs Schleswig-Holstein	21	175.064	174.610
	Zahl Personen in BGs Hamburg	1	147.471	147.302
	Zahl Personen in BGs Niedersachsen	58	481.832	480.589
	Zahl Personen in BGs Bremen	9	70.213	69.984

Tabelle 58: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen) (Fortsetzung 1)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Personen in BGs Nordrhein-Westfalen	119	1.184.500	1.181.446
	Zahl Personen in BGs Hessen	23	316.240	315.281
	Zahl Personen in BGs Rheinland-Pfalz	16	174.363	173.807
	Zahl Personen in BGs Baden-Württemberg	24	330.928	329.967
	Zahl Personen in BGs Bayern	62	354.691	353.784
	Zahl Personen in BGs Saarland	17	61.668	61.467
	Zahl Personen in BGs Berlin	18	454.535	453.123
	Zahl Personen in BGs Brandenburg	34	254.167	253.690
	Zahl Personen in BGs Mecklenburg-Vorpommern	8	195.559	195.016
	Zahl Personen in BGs Sachsen	25	414.878	413.743
	Zahl Personen in BGs Sachsen-Anhalt	41	282.916	282.270
	Zahl Personen in BGs Thüringen	19	192.959	192.264
Anzahl der Personen in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Alter (15-24 und 25-64; 2 Kategorien)	Zahl Personen in BGs im Alter 15-24	77	1.004.739	1.003.208
	Zahl Personen in BGs im Alter 25-64	418	4.087.245	4.075.136
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Geschlecht und nach West/Ost (4 Kategorien)	Zahl Männer in BGs (West)	151	1.565.168	1.561.331
	Zahl Frauen in BGs (West)	199	1.731.802	1.726.907
	Zahl Männer in BGs (Ost)	59	903.032	900.280
	Zahl Frauen in BGs (Ost)	86	891.982	889.826

Tabelle 58: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen) (Fortsetzung 2)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach „Alleinerziehend ja/nein“, nach Geschlecht und nach West/Ost (8 Kategorien)	Zahl alleinerziehende in BGs (West)	68	459.676	458.183
	Zahl alleinerziehende in BGs (Ost)	20	202.694	201.955
	Zahl nicht alleinerziehende in BGs (West)	282	2.837.294	2.830.054
	Zahl nicht alleinerziehende in BGs (Ost)	125	1.592.320	1.588.151
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Staatsangehörigkeit (deutsch/nicht deutsch) und nach West/Ost	Zahl deutscher Personen in BGs	419	4.116.358	4.116.250
	Zahl nicht deutscher Personen in BGs	76	961.971	962.094
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Bundesland (16 Kategorien)	Zahl Personen in Privathaushalten Schleswig-Holstein	272	2.400.000	2.399.916
	Zahl Personen in Privathaushalten Hamburg	51	1.543.000	1.542.600
	Zahl Personen in Privathaushalten Niedersachsen	695	6.743.000	6.742.834
	Zahl Personen in Privathaushalten Bremen	41	573.000	572.983
	Zahl Personen in Privathaushalten Nordrhein-Westfalen	1.270	15.327.000	15.327.031
	Zahl Personen in Privathaushalten Hessen	513	5.189.000	5.189.462
	Zahl Personen in Privathaushalten Rheinland-Pfalz	286	3.441.000	3.441.713
	Zahl Personen in Privathaushalten Baden-Württemberg	699	9.123.000	9.123.633
	Zahl Personen in Privathaushalten Bayern	1.136	10.611.000	10.610.649
	Zahl Personen in Privathaushalten Saarland	71	889.000	889.149
	Zahl Personen in Privathaushalten Berlin	155	3.007.000	3.006.483

Tabelle 58: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen) (Fortsetzung 3)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Personen in Privathaushalten Brandenburg	195	2.241.000	2.240.645
	Zahl Personen in Privathaushalten Mecklenburg-Vorpommern	101	1.483.000	1.483.095
	Zahl Personen in Privathaushalten Sachsen	304	3.731.000	3.730.605
	Zahl Personen in Privathaushalten Sachsen-Anhalt	208	2.132.000	2.132.164
	Zahl Personen in Privathaushalten Thüringen	210	2.023.000	2.023.008
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Alter (in 5-Jahres-Klassen), Geschlecht und West/Ost (56 Kategorien)	Zahl Männer in PH (West), 15-19 Jahre	183	1.978.000	1.968.446
	Zahl Männer in PH (West), 20-24 Jahre	129	1.933.000	1.930.497
	Zahl Männer in PH (West), 25-29 Jahre	109	1.928.000	1.927.047
	Zahl Männer in PH (West), 30-34 Jahre	136	1.872.000	1.872.007
	Zahl Männer in PH (West), 35-39 Jahre	180	2.363.000	2.363.064
	Zahl Männer in PH (West), 40-44 Jahre	270	2.942.000	2.942.498
	Zahl Männer in PH (West), 45-49 Jahre	271	2.605.000	2.605.497
	Zahl Männer in PH (West), 50-54 Jahre	215	2.255.000	2.255.853
	Zahl Männer in PH (West), 55-59 Jahre	205	2.067.000	2.068.040
	Zahl Männer in PH (West), 60-64 Jahre	170	1.697.000	1.698.207
	Zahl Männer in PH (West), 65-69 Jahre	194	1.970.000	1.971.883
	Zahl Männer in PH (West), 70-74 Jahre	142	1.611.000	1.612.766
	Zahl Männer in PH (West), 75-79 Jahre	65	1.029.000	1.030.183

Tabelle 58: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen) (Fortsetzung 4)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Männer in PH (West), 80+ Jahre	48	947.000	948.170
	Zahl Frauen in PH (West), 15-19 Jahre	173	1.836.000	1.825.491
	Zahl Frauen in PH (West), 20-24 Jahre	136	1.833.000	1.830.958
	Zahl Frauen in PH (West), 25-29 Jahre	138	1.959.000	1.958.785
	Zahl Frauen in PH (West), 30-34 Jahre	174	1.902.000	1.902.171
	Zahl Frauen in PH (West), 35-39 Jahre	262	2.334.000	2.334.396
	Zahl Frauen in PH (West), 40-44 Jahre	335	2.850.000	2.850.380
	Zahl Frauen in PH (West), 45-49 Jahre	335	2.578.000	2.578.703
	Zahl Frauen in PH (West), 50-54 Jahre	269	2.287.000	2.288.177
	Zahl Frauen in PH (West), 55-59 Jahre	239	2.115.000	2.116.401
	Zahl Frauen in PH (West), 60-64 Jahre	174	1.746.000	1.747.779
	Zahl Frauen in PH (West), 65-69 Jahre	205	2.172.000	2.174.532
	Zahl Frauen in PH (West), 70-74 Jahre	136	1.832.000	1.834.340
	Zahl Frauen in PH (West), 75-79 Jahre	81	1.341.000	1.342.971
	Zahl Frauen in PH (West), 80+ Jahre	60	1.858.000	1.860.730
	Zahl Männer in PH (Ost), 15-19 Jahre	31	434.000	432.517
	Zahl Männer in PH (Ost), 20-24 Jahre	25	568.000	567.605
	Zahl Männer in PH (Ost), 25-29 Jahre	28	541.000	540.857
	Zahl Männer in PH (Ost), 30-34 Jahre	26	475.000	474.789

Tabelle 58: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen) (Fortsetzung 5)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Männer in PH (Ost), 35-39 Jahre	30	579.000	578.898
	Zahl Männer in PH (Ost), 40-44 Jahre	55	728.000	727.928
	Zahl Männer in PH (Ost), 45-49 Jahre	70	691.000	691.109
	Zahl Männer in PH (Ost), 50-54 Jahre	61	639.000	639.043
	Zahl Männer in PH (Ost), 55-59 Jahre	46	580.000	580.077
	Zahl Männer in PH (Ost), 60-64 Jahre	36	436.000	436.236
	Zahl Männer in PH (Ost), 65-69 Jahre	46	574.000	574.399
	Zahl Männer in PH (Ost), 70-74 Jahre	31	453.000	453.322
	Zahl Männer in PH (Ost), 75-79 Jahre	20	266.000	266.247
	Zahl Männer in PH (Ost), 80+ Jahre	6	198.000	198.158
	Zahl Frauen in PH (Ost), 15-19 Jahre	38	387.000	386.086
	Zahl Frauen in PH (Ost), 20-24 Jahre	35	519.000	518.663
	Zahl Frauen in PH (Ost), 25-29 Jahre	32	497.000	496.944
	Zahl Frauen in PH (Ost), 30-34 Jahre	28	419.000	418.910
	Zahl Frauen in PH (Ost), 35-39 Jahre	52	526.000	525.883
	Zahl Frauen in PH (Ost), 40-44 Jahre	62	688.000	687.960
	Zahl Frauen in PH (Ost), 45-49 Jahre	66	672.000	672.024
	Zahl Frauen in PH (Ost), 50-54 Jahre	81	632.000	632.081
	Zahl Frauen in PH (Ost), 55-59 Jahre	82	613.000	613.137

Tabelle 58: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen) (Fortsetzung 6)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Frauen in PH (Ost), 60-64 Jahre	49	461.000	461.254
	Zahl Frauen in PH (Ost), 65-69 Jahre	59	649.000	649.536
	Zahl Frauen in PH (Ost), 70-74 Jahre	38	552.000	552.496
	Zahl Frauen in PH (Ost), 75-79 Jahre	23	373.000	373.364
	Zahl Frauen in PH (Ost), 80+ Jahre	17	466.000	466.476
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Haushaltsgröße (1, 2, 3, 4, „5 oder mehr Personen“) und West/Ost (10 Kategorien)	Zahl Personen in PH mit 1 Person (West)	756	11.753.000	11.752.588
	Zahl Personen in PH mit 2 Personen (West)	1.751	20.499.000	20.509.147
	Zahl Personen in PH mit 3 Personen (West)	974	9.987.000	9.985.175
	Zahl Personen in PH mit 4 Personen (West)	999	9.335.000	9.331.354
	Zahl Personen in PH mit 5 oder mehr Personen (West)	554	4.265.000	4.261.707
	Zahl Personen in PH mit 1 Person (Ost)	211	3.567.000	3.566.706
	Zahl Personen in PH mit 2 Personen (Ost)	476	5.867.000	5.869.110
	Zahl Personen in PH mit 3 Personen (Ost)	255	2.978.000	2.976.941
	Zahl Personen in PH mit 4 Personen (Ost)	155	1.707.000	1.706.347
	Zahl Personen in PH mit 5 oder mehr Personen (Ost)	76	497.000	496.895
Anzahl der Personen ab 15 Jahre in Privathaushalten nach höchstem Schulabschluss und West/Ost (12 Kategorien)	Zahl Personen in PH mit höchstem Schulabschluss: noch Schüler (West)	222	2.350.000	2.379.775
	Zahl Personen in PH mit höchstem Schulabschluss: ohne Abschluss (West)	142	1.880.000	1.904.762
	Zahl Personen in PH mit höchstem Schulabschluss: Hauptschule (West)	1.705	23.781.000	24.094.794

Tabelle 58: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen) (Fortsetzung 7)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Personen in PH mit höchstem Schulabschluss: Realschule/POS (West)	1.426	13.221.000	13.392.817
	Zahl Personen in PH mit höchstem Schulabschluss:(Fach-) Hochschulreife (West)	1.539	13.889.000	14.067.822
	Zahl Personen in PH mit höchstem Schulabschluss: noch Schüler (Ost)	37	442.000	445.272
	Zahl Personen in PH mit höchstem Schulabschluss: ohne Abschluss (Ost)	19	269.000	270.954
	Zahl Personen in PH mit höchstem Schulabschluss: Hauptschule (Ost)	288	3.842.000	3.871.263
	Zahl Personen in PH mit höchstem Schulabschluss: Realschule/POS (Ost)	523	6.473.000	6.521.235
	Zahl Personen in PH mit höchstem Schulabschluss: (Fach-) Hochschulreife (Ost)	306	3.481.000	3.507.276
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Familienstand und West/Ost (10 Kategorien)	Zahl Personen in PH mit Familienstand: ledig (West)	1.170	9.714.000	9.713.419
	Zahl Personen in PH mit Familienstand: verheiratet, eingetragene Lebensgemeinschaft. (West)	3.163	36.549.000	36.548.855
	Zahl Personen in PH mit Familienstand: geschieden (West)	405	4.729.000	4.728.653
	Zahl Personen in PH mit Familienstand: verwitwet (West)	296	4.848.000	4.849.044
	Zahl Personen in PH mit Familienstand: ledig (Ost)	275	3.246.000	3.245.413
	Zahl Personen in PH mit Familienstand: verheiratet, eingetragene Lebensgemeinschaft. (Ost)	687	8.578.000	8.577.406
	Zahl Personen in PH mit Familienstand: geschieden (Ost)	116	1.499.000	1.498.954

Tabelle 58: Soll-Verteilungen und Verteilungen nach Kalibrierung (Microm-Stichprobe, Personen) (Fortsetzung 8)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Personen in PH mit Familienstand: verwitwet (Ost)	95	1.294.000	1.294.227
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Staatsangehörigkeiten und West/Ost	Zahl Personen in PH Deutsch	5.947	64.137.000	64.138.315
	Zahl Personen in PH Nicht-Deutsch	260	6.318.000	6.317.656
Arbeitslose inkl. Maßnahmeteilnehmer West/Ost	Arbeitslose inkl. Maßnahmeteilnehmer (West)	322	3.213.295	3.213.970
	Arbeitslose inkl. Maßnahmeteilnehmer (Ost)	142	1.618.732	1.619.007
Sozialversicherungspflichtig Beschäftigte West/Ost	Sozialversicherungspflichtig Beschäftigte (West)	2.042	22.205.091	22.205.091
	Sozialversicherungspflichtig Beschäftigte (Ost)	480	5.178.108	5.178.108

Tabelle 59: Kenngrößen der Verteilung der Gewichte

1%-Perzentil	1269.838
5%-Perzentil	2001.988
10%-Perzentil	2597.447
25%-Perzentil	4170.006
50%-Perzentil	7965.598
75%-Perzentil	14576.53
90%-Perzentil	24197.6
95%-Perzentil	30944.07
99%-Perzentil	53203.17
Mittelwert	11351.05
Standardabweichung	11179.41
Minimum	567.55
Maximum	147563.7
Fallzahl	6207
Effektivitätsmaß	50,8%

6.10.3 Gesamt-Stichprobe

Die Grundgesamtheit bilden wie bei der Microm-Stichprobe alle 15-jährigen und älteren Personen in Privathaushalten in Deutschland. Ausgangspunkt für die Kalibrierung waren die kalibrierten Haushaltsgewichte der Gesamtstichprobe. Diese wurden am 10%-Perzentil und am 90%-Perzentil ihrer Verteilung getrimmt und anschließend so reskaliert, dass ihre Summe wieder die Summe der ungetrimmten kalibrierten Haushaltsgewichte ergab. Die getrimmten Hochrechnungsfaktoren reichen von 222,42 bis 19648,46 (Gewichtungsfaktoren von 0,04 bis 3,75). Das Intervall der Gesamthochrechnungsfaktoren wurde nach unten auf 52,43 und nach oben auf 52426,52 begrenzt, was einer Beschränkung der Gesamtgewichtungsfaktoren auf den Bereich von 0,01 bis 10,0 entspricht.

Für folgende Merkmale wurde kalibriert:

Leistungsempfänger Grundlage BA-Statistik:

- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Bundesländern
- Anzahl der Personen in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Alter (15-24 und 25-64)
- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Geschlecht, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach „Alleinerziehend ja/nein“, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Staatsangehörigkeit (deutsch/nicht deutsch)

Bevölkerung Grundlage Mikrozensus 2008:

- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Bundesland
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Alter, Geschlecht und West/Ost
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Haushaltgröße, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach höchstem Schulabschluss, nach West/Ost
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Familienstand und West/Ost
- Anzahl der Personen ab 15 Jahre in Privathaushalten nach Staatsangehörigkeiten

Bevölkerung Grundlage BA Statistik:

- Anzahl der Arbeitslosen inklusive Maßnahmeteilnehmer nach West/Ost
- Anzahl der sozialversicherungspflichtig Beschäftigten nach West/Ost

Als Quelle für den Eckwert des Erwerbsstatus wurde die BA-Statistik verwendet, da die Definition der Arbeitslosigkeit und der sozialversicherungspflichtigen Beschäftigung in PASS nicht dem ILO-Konzept des statistischen Bundesamtes entspricht, den Statistiken der BA aber entnommen werden kann.

Daneben wurde auch bei der Gesamtstichprobe als zusätzlicher Eckwert die Zahl der Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug seit dem Vorjahr auf Ebene von Personen zwischen 15 und 64 Jahren (648.988) mit aufgenommen.

Um die Kalibrierung durchzuführen, muss jede Eckwertvariable für jede Person einen gültigen Wert besitzen. Deshalb wurde der sehr geringe Item-Nonresponse vor der Kalibrierung imputiert. Die Imputation erfolgte über den Mittelwert bzw. den häufigsten Wert der jeweiligen Variable.

Da die Imputation einzig der Durchführbarkeit der Kalibrierung dient, wurden die imputierten Werte nach der Kalibrierung wieder auf fehlende Werte zurückgesetzt. Eine Hochrechnung mit den kalibrierten Gewichten ohne Berücksichtigung des Item-Nonresponse führt daher zu geringfügigen Abweichungen von den im Folgenden dargestellten Werten.

Tabelle 60: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Bundesländern (16 Kategorien)	Zahl Personen in BGs Schleswig-Holstein	244	175.064	174.668
	Zahl Personen in BGs Hamburg	73	147.471	147.243
	Zahl Personen in BGs Niedersachsen	586	481.832	480.636
	Zahl Personen in BGs Bremen	70	70.213	70.004
	Zahl Personen in BGs Nordrhein-Westfalen	1,260	1.184.500	1.181.743
	Zahl Personen in BGs Hessen	346	316.240	315.493
	Zahl Personen in BGs Rheinland-Pfalz	195	174.363	173.824
	Zahl Personen in BGs Baden-Württemberg	360	330.928	330.083
	Zahl Personen in BGs Bayern	514	354.691	353.830
	Zahl Personen in BGs Saarland	95	61.668	61.507
Zahl Personen in BGs Berlin	472	454.535	453.559	

Tabelle 60: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 1)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Personen in BGs Brandenburg	322	254.167	253.593
	Zahl Personen in BGs Mecklenburg-Vorpommern	178	195.559	195.000
	Zahl Personen in BGs Sachsen	430	414.878	413.963
	Zahl Personen in BGs Sachsen-Anhalt	425	282.916	282.237
	Zahl Personen in BGs Thüringen	220	192.959	192.510
Anzahl der Personen in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Alter (15-24 und 25-64; 2 Kategorien)	Zahl Personen in BGs im Alter 15-24	951	1.004.739	1.004.157
	Zahl Personen in BGs im Alter 25-64	4,839	4.087.245	4.075.739
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Geschlecht und nach West/Ost (4 Kategorien)	Zahl Männer in BGs (West)	1,642	1.565.168	1.561.349
	Zahl Frauen in BGs (West)	2,101	1.731.802	1.727.684
	Zahl Männer in BGs (Ost)	955	903.032	900.819
	Zahl Frauen in BGs (Ost)	1,092	891.982	890.044
Anzahl der Personen ab 15 Jahren in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach „Alleinerziehend ja/nein“, nach Geschlecht und nach West/Ost (8 Kategorien)	Zahl alleinerziehende in BGs (West)	758	459.676	458.214
	Zahl alleinerziehende in BGs (Ost)	310	202.694	202.082
	Zahl nicht alleinerziehende in BGs (West)	2,985	2.837.294	2.830.818
	Zahl nicht alleinerziehende in BGs (Ost)	1,737	1.592.320	1.588.781
Anzahl der Personen ab 15 Jahre in Bedarfsgemeinschaften im Leistungsbezug nach SGB II nach Staatsangehörigkeit (deutsch/nicht deutsch) und nach West/Ost	Zahl deutscher Personen in BGs	5,155	4.116.358	4.117.533
	Zahl nicht deutscher Personen in BGs	635	961.971	962.362

Tabelle 60: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 2)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Bundesland (16 Kategorien)	Zahl Personen in Privathaushalten Schleswig-Holstein	604	2.400.000	2.400.106
	Zahl Personen in Privathaushalten Hamburg	143	1.543.000	1.542.235
	Zahl Personen in Privathaushalten Niedersachsen	1,416	6.743.000	6.743.028
	Zahl Personen in Privathaushalten Bremen	122	573.000	572.991
	Zahl Personen in Privathaushalten Nordrhein-Westfalen	2,802	15.327.000	15.326.658
	Zahl Personen in Privathaushalten Hessen	951	5.189.000	5.189.359
	Zahl Personen in Privathaushalten Rheinland-Pfalz	530	3.441.000	3.441.734
	Zahl Personen in Privathaushalten Baden-Württemberg	1,179	9.123.000	9.123.526
	Zahl Personen in Privathaushalten Bayern	1,795	10.611.000	10.611.113
	Zahl Personen in Privathaushalten Saarland	166	889.000	889.149
	Zahl Personen in Privathaushalten Berlin	722	3.007.000	3.006.510
	Zahl Personen in Privathaushalten Brandenburg	580	2.241.000	2.240.660
	Zahl Personen in Privathaushalten Mecklenburg-Vorpommern	341	1.483.000	1.482.956
	Zahl Personen in Privathaushalten Sachsen	870	3.731.000	3.730.596
	Zahl Personen in Privathaushalten Sachsen-Anhalt	712	2.132.000	2.132.291
	Zahl Personen in Privathaushalten Thüringen	506	2.023.000	2.022.986

Tabelle 60: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 3)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Alter (in 5-Jahres-Klassen), Geschlecht und West/Ost (56 Kategorien)	Zahl Männer in PH (West), 15-19 Jahre	366	1.978.000	1.966.053
	Zahl Männer in PH (West), 20-24 Jahre	308	1.933.000	1.928.487
	Zahl Männer in PH (West), 25-29 Jahre	313	1.928.000	1.925.612
	Zahl Männer in PH (West), 30-34 Jahre	320	1.872.000	1.872.046
	Zahl Männer in PH (West), 35-39 Jahre	388	2.363.000	2.363.314
	Zahl Männer in PH (West), 40-44 Jahre	526	2.942.000	2.942.956
	Zahl Männer in PH (West), 45-49 Jahre	529	2.605.000	2.606.021
	Zahl Männer in PH (West), 50-54 Jahre	413	2.255.000	2.256.273
	Zahl Männer in PH (West), 55-59 Jahre	407	2.067.000	2.068.481
	Zahl Männer in PH (West), 60-64 Jahre	302	1.697.000	1.698.692
	Zahl Männer in PH (West), 65-69 Jahre	250	1.970.000	1.972.389
	Zahl Männer in PH (West), 70-74 Jahre	153	1.611.000	1.613.207
	Zahl Männer in PH (West), 75-79 Jahre	68	1.029.000	1.030.409
	Zahl Männer in PH (West), 80+ Jahre	50	947.000	948.361
	Zahl Frauen in PH (West), 15-19 Jahre	381	1.836.000	1.823.687
	Zahl Frauen in PH (West), 20-24 Jahre	340	1.833.000	1.829.468
	Zahl Frauen in PH (West), 25-29 Jahre	427	1.959.000	1.958.518
	Zahl Frauen in PH (West), 30-34 Jahre	503	1.902.000	1.902.386
	Zahl Frauen in PH (West), 35-39 Jahre	566	2.334.000	2.334.731

Tabelle 60: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 4)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Frauen in PH (West), 40-44 Jahre	672	2.850.000	2.850.792
	Zahl Frauen in PH (West), 45-49 Jahre	657	2.578.000	2.579.148
	Zahl Frauen in PH (West), 50-54 Jahre	509	2.287.000	2.288.701
	Zahl Frauen in PH (West), 55-59 Jahre	426	2.115.000	2.116.905
	Zahl Frauen in PH (West), 60-64 Jahre	304	1.746.000	1.748.322
	Zahl Frauen in PH (West), 65-69 Jahre	235	2.172.000	2.175.227
	Zahl Frauen in PH (West), 70-74 Jahre	148	1.832.000	1.834.938
	Zahl Frauen in PH (West), 75-79 Jahre	83	1.341.000	1.343.455
	Zahl Frauen in PH (West), 80+ Jahre	64	1.858.000	1.861.318
	Zahl Männer in PH (Ost), 15-19 Jahre	113	434.000	432.594
	Zahl Männer in PH (Ost), 20-24 Jahre	160	568.000	567.021
	Zahl Männer in PH (Ost), 25-29 Jahre	169	541.000	540.223
	Zahl Männer in PH (Ost), 30-34 Jahre	144	475.000	474.602
	Zahl Männer in PH (Ost), 35-39 Jahre	129	579.000	578.804
	Zahl Männer in PH (Ost), 40-44 Jahre	174	728.000	727.931
	Zahl Männer in PH (Ost), 45-49 Jahre	238	691.000	691.170
	Zahl Männer in PH (Ost), 50-54 Jahre	200	639.000	639.158
	Zahl Männer in PH (Ost), 55-59 Jahre	191	580.000	580.274
	Zahl Männer in PH (Ost), 60-64 Jahre	94	436.000	436.368

Tabelle 60: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 5)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Männer in PH (Ost), 65-69 Jahre	59	574.000	574.637
	Zahl Männer in PH (Ost), 70-74 Jahre	38	453.000	453.501
	Zahl Männer in PH (Ost), 75-79 Jahre	20	266.000	266.350
	Zahl Männer in PH (Ost), 80+ Jahre	8	198.000	198.266
	Zahl Frauen in PH (Ost), 15-19 Jahre	120	387.000	385.876
	Zahl Frauen in PH (Ost), 20-24 Jahre	151	519.000	517.979
	Zahl Frauen in PH (Ost), 25-29 Jahre	193	497.000	496.651
	Zahl Frauen in PH (Ost), 30-34 Jahre	156	419.000	418.771
	Zahl Frauen in PH (Ost), 35-39 Jahre	171	526.000	525.879
	Zahl Frauen in PH (Ost), 40-44 Jahre	219	688.000	687.950
	Zahl Frauen in PH (Ost), 45-49 Jahre	258	672.000	672.202
	Zahl Frauen in PH (Ost), 50-54 Jahre	247	632.000	632.232
	Zahl Frauen in PH (Ost), 55-59 Jahre	216	613.000	613.344
	Zahl Frauen in PH (Ost), 60-64 Jahre	104	461.000	461.462
	Zahl Frauen in PH (Ost), 65-69 Jahre	72	649.000	649.834
	Zahl Frauen in PH (Ost), 70-74 Jahre	41	552.000	552.732
	Zahl Frauen in PH (Ost), 75-79 Jahre	24	373.000	373.517
	Zahl Frauen in PH (Ost), 80+ Jahre	22	466.000	466.672

Tabelle 60: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 6)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Haushaltsgröße (1, 2, 3, 4, „5 oder mehr Personen“) und West/Ost (10 Kategorien)	Zahl Personen in PH mit 1 Person (West)	2,113	11.753.000	11.748.501
	Zahl Personen in PH mit 2 Personen (West)	2,960	20.499.000	20.513.222
	Zahl Personen in PH mit 3 Personen (West)	1,930	9.987.000	9.985.555
	Zahl Personen in PH mit 4 Personen (West)	1,652	9.335.000	9.331.486
	Zahl Personen in PH mit 5 oder mehr Personen (West)	1,053	4.265.000	4.261.133
	Zahl Personen in PH mit 1 Person (Ost)	1,030	3.567.000	3.566.369
Anzahl der Personen ab 15 Jahre in Privathaushalten nach höchstem Schulabschluss und West/Ost (12 Kategorien)	Zahl Personen in PH mit höchstem Schulabschluss: noch Schüler (West)	439	2.350.000	2.379.250
	Zahl Personen in PH mit höchstem Schulabschluss: ohne Abschluss (West)	488	1.880.000	1.904.834
	Zahl Personen in PH mit höchstem Schulabschluss: Hauptschule (West)	3,588	23.781.000	24.095.400
	Zahl Personen in PH mit höchstem Schulabschluss: Realschule/POS (West)	2,779	13.221.000	13.392.814
	Zahl Personen in PH mit höchstem Schulabschluss:(Fach-) Hochschulreife (West)	2,414	13.889.000	14.067.601
	Zahl Personen in PH mit höchstem Schulabschluss: noch Schüler (Ost)	117	442.000	445.187

Tabelle 60: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 7)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
	Zahl Personen in PH mit höchstem Schulabschluss: ohne Abschluss (Ost)	119	269.000	270.922
	Zahl Personen in PH mit höchstem Schulabschluss: Hauptschule (Ost)	939	3.842.000	3.872.147
	Zahl Personen in PH mit höchstem Schulabschluss: Realschule/POS (Ost)	1,850	6.473.000	6.520.965
	Zahl Personen in PH mit höchstem Schulabschluss: (Fach-) Hochschulreife (Ost)	706	3.481.000	3.506.779
	Zahl Personen in PH mit Familienstand: ledig (West)	3,100	9.714.000	9.713.491
	Zahl Personen in PH mit Familienstand: verheiratet, eingetragene Lebensgemeinschaft. (West)	4,773	36.549.000	36.548.722
	Zahl Personen in PH mit Familienstand: geschieden (West)	1,420	4.729.000	4.728.660
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Familienstand und West/Ost (10 Kategorien)	Zahl Personen in PH mit Familienstand: verwitwet (West)	415	4.848.000	4.849.026
	Zahl Personen in PH mit Familienstand: ledig (Ost)	1,470	3.246.000	3.245.419
	Zahl Personen in PH mit Familienstand: verheiratet, eingetragene Lebensgemeinschaft. (Ost)	1,506	8.578.000	8.577.426
	Zahl Personen in PH mit Familienstand: geschieden (Ost)	585	1.499.000	1.498.907
	Zahl Personen in PH mit Familienstand: verwitwet (Ost)	170	1.294.000	1.294.248
Anzahl der Personen ab 15 Jahre in Privathaushalten nach Staatsangehörigkeiten und West/Ost	Zahl Personen in PH Deutsch	12,464	64.137.000	64.138.414
	Zahl Personen in PH Nicht-Deutsch	975	6.318.000	6.317.484

Tabelle 60: Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) (Fortsetzung 8)

Eckwert	Ausprägung Eckwert	ungewichtete Verteilung	Soll-Werte aus der BA-Statistik bzw. MZ 2008	Verteilung mit kalibrierten Gewichten
Arbeitslose inkl. Maßnahmeteilnehmer West/Ost	Arbeitslose inkl. Maßnahmeteilnehmer (West)	2.435	3.213.295	3.213.962
	Arbeitslose inkl. Maßnahmeteilnehmer (Ost)	1.418	1.618.732	1.619.109
Sozialversicherungspflichtig Beschäftigte West/Ost	Sozialversicherungspflichtig Beschäftigte (West)	3.389	22.205.091	22.205.091
	Sozialversicherungspflichtig Beschäftigte (Ost)	1.285	5.178.108	5.178.108

Tabelle 61: Kenngrößen der Verteilung der Gewichte

1%-Perzentil	61.5697
5%-Perzentil	114.8264
10%-Perzentil	165.983
25%-Perzentil	324.2362
50%-Perzentil	1287.922
75%-Perzentil	7187.478
90%-Perzentil	15795.81
95%-Perzentil	21377.3
99%-Perzentil	40675.11
Mittelwert	5242.644
Standardabweichung	8205.737
Minimum	52.43
Maximum	52426.52
Fallzahl	13439
Effektivitätsmaß	29,0%

6.11 Schätzung der BA-Querschnittsgewichte für Haushalte und Personen ohne Arbeitslosengeld-II-Bezug

Schließlich verblieben auch in der Welle 3 noch Haushalte und Personen, denen durch Kalibrierung kein BA-Haushaltsquerschnittsgewicht oder BA-Personenquerschnittsgewicht zugewiesen werden konnte. Die Anzahl dieser Haushalte ist nun in Welle 3 größer als in Welle 2, da ein größerer Teil der BA-Stichprobe der Welle 1 mittlerweile aus dem Leistungsbezug ausgeschieden ist. Es handelt sich um die folgenden drei Gruppen ohne Leistungsbezug im Juli 2008, die aber zur Grundgesamtheit der BA-Stichprobe (Haushalte mit Arbeitslosengeld-II-Bezug in 7/2006 oder 7/2007 oder 7/2008 bzw. Personen in Haushalten mit Arbeitslosengeld-II-Bezug in 7/2006 oder 7/2007 oder 7/2008) gehören.

- Aus der Auffrischungstichprobe: Nicht-BG-Mitglieder im Haushalt: Das Personengewicht wurde hier aus dem BA-Haushaltsgewicht in Welle 3 nach Kalibrierung (*wqbahh*) durch Division durch den Anteil dieser Personen, die – vorausgesetzt ihr Haushalt hat teilgenommen – ein Personen- oder Senioreninterview gegeben haben, gewonnen.
- Bestandshaushalte, in denen im Juli 2008 niemand mehr Arbeitslosengeld II bezogen hat: Der Haushalt behält das BA-Gewicht vor Kalibrierung. Personen in diesen Haushalten mit Interviews in beiden Wellen wurde als neues BA-Personengewicht ihr altes BA-Personengewicht der Vorwelle multipliziert mit der reziproken Bleibewahrscheinlichkeit *ppbleib* zugewiesen. Personen in diesen Haushalten ohne Personeninterview in Welle 2 erhalten als neues BA-Personengewicht das BA-Haushaltsgewicht ihres Haushalts für Welle 3 dividiert durch den Anteil solcher Personen, die teilnehmen, vorausgesetzt ihr Haushalt nimmt teil.
- Nicht-BG-Mitglieder in Bestandshaushalten, die im Juli 2008 noch Arbeitslosengeld II beziehen: Personen in diesen Haushalten mit Interviews in beiden Wellen wurde als neues BA-Personengewicht ihr BA-Personengewicht der Vorwelle multipliziert mit der reziproken Bleibewahrscheinlichkeit *ppbleib* zugewiesen.

7 Anhang: Kurzbeschreibung des Datensatzes

Inhaltliche Charakteristika

Kategorien	Erläuterungen
Themen/ Merkmalsgruppen	<p>Soziodemographische Merkmale: systemfreie Personennummer; Geschlecht; Geburtsjahr; Alter; Familienstand; Kinderzahl im und außerhalb des Haushalts; Staatsangehörigkeit; Herkunftsland und Migrationshintergrund; Schul- und Berufsbildung (incl. generierter Skalen: CASMIN, ISCED-97, Bildung in Jahren); Schul- und Berufsbildung der Eltern; Gesundheitsindikatoren; Konfession; soziale Kontakte; Betreuung und Schulbesuch von Kindern; Haushaltseinkommen (incl. Einzelkomponenten und bedarfsgewichtetes Haushaltseinkommen); Basisinformationen zu Vermögen und Schulden; Ausstattung des Haushalts (Deprivationsindex); Wohnung und Wohnumfeld; detaillierte Informationen zum Thema Altersvorsorge (nur Welle 3);</p> <p>Beschäftigungsbezogene Merkmale: Erwerbsstatus/Status Nichterwerbstätigkeit; Mini-Job; Arbeitszeit; Stellung im Beruf (detailliert); berufliche Tätigkeit (ISCO-88 und KldB-92); ISCO-basierte Status- und Prestigemaße (ISEI, SIOPS, MPS, EGP, ESeC); Erwerbseinkommen (brutto & netto); Erwerbsbiografien mit Erwerbstätigkeiten/Arbeitslosigkeiten und Zeiten der Nichterwerbstätigkeit seit 01/2005 (ab Welle 2); Befristung der Tätigkeit; Aufsichtsfunktion; Arbeitgeber: öfftl. Dienst/Privatwirtschaft; Arbeitgeber: Zahl der Beschäftigten; weitere Erwerbstätigkeit; Zusammenfassende Informationen zur Erwerbs- und Arbeitslosigkeitshistorie; detaillierte Informationen zum Thema Arbeitssuche; Reservationslohn;</p> <p>Merkmale zum Leistungsbezug: <u>ALG I:</u> Beginn und Ende der Leistungsbezugsepisode(n) seit 01/2005 (nur Welle 1); Informationen zu ALG-I-Bezugszeiten im Rahmen gemeldeter Arbeitslosigkeiten seit 01/2005 (ab Welle 2); Höhe der Leistung; Beendigungsgrund; <u>ALG II:</u> Beginn und Ende der Leistungsbezugsepisode(n) seit 01/2005; Beendigungsgrund; Identifizierung beziehende HH-Mitglieder; Höhe der Leistungen; Leistungskürzungen (Anfang, Dauer, Gründe, HH-Mitglieder mit Kürzung); <u>Maßnahmeteilnahme:</u> Maßnahmetyp; Beginn und Ende der Maßnahme; Indikator vorzeitige Beendigung; Gründe für vorzeitige Beendigung; Art Maßnahmezugang; Maßnahmebewertung; Arbeitszeit in Maßnahme; Vergleich der Tätigkeit zu regulärer Beschäftigung; Branche; <u>Kontakte zum ALGII-Träger:</u> Anzahl und Art der Kontakte; Gesprächsinhalte; Angebote; Eingliederungsvereinbarung; Bewertung des Trägers;</p> <p>Subjektive Indikatoren: Zufriedenheit; Ängste und Sorgen; Erwerbsorientierungen; Bildungsaspiration; Geschlechterrollenorientierungen; subjektive soziale Position (Oben-Unten-Skala); subjektive Bewertung des Gesundheitszustands</p>

Kategorien	Erläuterungen
Untersuchungseinheit	<p>Personen und Haushalte mit ALG II – Leistungsbezug in 7/2006 (Sample I)</p> <p>Personen und Haushalte der Wohnbevölkerung der BRD (Sample II)</p> <p>Personen und Haushalte mit ALG II – Leistungsbezug in 7/2007 aber ohne Bezug in 7/2006 (Sample III; Zugangsstichprobe 1)</p> <p>Personen und Haushalte mit ALG II – Leistungsbezug in 7/2008 aber ohne Bezug in 7/2006 oder 7/2007 (Sample IV; Zugangsstichprobe 2)</p> <p>Anmerkung: Befragung von Personen ab 65 Jahren nur mittels eines Kurzfragebogens</p>
Fallzahlen	<p>Welle 1:</p> <p>Sample I: 9.386 Personen (in 6.804 Haushalten)</p> <p>Sample II: 9.568 Personen (in 5.990 Haushalten)</p> <p>Welle 2:</p> <p>Sample I: 4.753 Personen (in 3.491 Haushalten)</p> <p>Sample II: 6.392 Personen (in 3.897 Haushalten)</p> <p>Sample III: 1.342 Personen (in 1.041 Haushalten)</p> <p>Welle 3:</p> <p>Sample I: 4.913 Personen (in 3.754 Haushalten)</p> <p>Sample II: 6.207 Personen (in 3.901 Haushalten)</p> <p>Sample III: 898 Personen (in 694 Haushalten)</p> <p>Sample IV: 1.421 Personen (in 1.186 Haushalten)</p>
Erhebungsmodus	<p>CATI und CAPI</p> <p>CAPI Interviews wurden dann durchgeführt, wenn ein Stichprobenhaushalt nicht telefonisch erreichbar war oder persönliche Befragung gewünscht wurde</p> <p>Welle 1:</p> <p>N (CATI): 12.414 Personen (8.445 Haushalte)</p> <p>N (CAPI): 6.540 Personen (4.339 Haushalte)</p> <p>Welle 2:</p> <p>N (CATI): 7.888 Personen (5.378 Haushalte)</p> <p>N (CAPI): 4.599 Personen (3.051 Haushalte)</p> <p>Welle 3:</p> <p>N (CATI): 7776 Personen (5664 Haushalte)</p> <p>N (CAPI): 5663 Personen (3871 Haushalte)</p>

Kategorien	Erläuterungen
Erhebungssprachen	<p>Welle 1: Deutsch: 18.205 Personen (12.347 Haushalte) Russisch: 432 Personen (275 Haushalte) Türkisch: 305 Personen (163 Haushalte) Englisch: 12 Personen (9 Haushalte)</p> <p>Welle 2: Deutsch: 12.237 Personen (8.234 Haushalte) Russisch: 219 Personen (156 Haushalte) Türkisch: 31 Personen (39 Haushalte) Englisch: ab Welle 2 nicht mehr angeboten, da zu geringe Fallzahlen in Welle 1</p> <p>Welle 3: Deutsch: 13.000 Personen (9.256 Haushalte) Russisch: 330 Personen (210 Haushalte) Türkisch: 109 Personen (69 Haushalte)</p>
Response Rate	<p>Welle 1: Sample I: 35,1 % Sample II: 26,6 % Gesamt: 30,5 %</p> <p>Welle 2: Sample I (nur panelbereite HH): 51,1 % Sample II (nur panelbereite HH): 64,7 % Sample III: 26,3 % Abgespaltene Haushalte (aus Sample I und II): 13,4 % Gesamt: 45,0 %</p> <p>Welle 3: Sample I (nur panelbereite HH): 64,5 % Sample II (nur panelbereite HH): 76,4 % Sample III (nur panelbereite HH): 69,0 % Sample IV: 31,3% Gesamt: 60,7 %</p>

Kategorien	Erläuterungen
Ausschöpfung innerhalb der Haushalte	<p>Welle 1: Sample I: 85,6 % Sample II: 84,3 % Gesamt: 85,0 %</p> <p>Welle 2: Sample I (nur wiederbefragte HH): 85,5 % Sample II (nur wiederbefragte HH): 85,1 % Sample III: 86,2 % Abgespaltene Haushalte (aus Sample I und II): 88,3 % Gesamt: 85,4 %</p> <p>Welle 3: Sample I (nur wiederbefragte HH): 83,1 % Sample I (nur wiederbefragte HH): 83,6 % Sample III (nur wiederbefragte HH): 84,3 % Sample IV: 84,2 % Gesamt: 83,5 %</p>
Feldzeit:	<p>Welle 1: Dezember 2006-Juni 2007 Welle 2: Dezember 2007-Juli 2008 Welle 3: Dezember 2008-August 2009</p>
Zeitraum	<p>Welle 1: Feldzeit sowie Spelldaten retrospektiv ab 01/2005 Welle 2: Feldzeit sowie Spelldaten retrospektiv ab 01/2005 bzw. jeweiligem Referenzzeitraum des Spelldatentyps Welle 3: Feldzeit sowie Spelldaten retrospektiv ab 01/2006 bzw. jeweiligem Referenzzeitraum des Spelldatentyps</p>
Zeitbezug	Wiederholungsbefragung (Haushaltspanel)
Regionale Gliederung	Bundesland, Ost/West (Weitere regionale Informationen sind vorhanden, aber aus Datenschutzgründen nicht im Scientific Use File enthalten. Detaillierte Informationen hierzu auf Anfrage)
Gebietsstand	Zum Befragungszeitpunkt

Methodische Charakteristika

Kategorien	Erläuterungen
Erhebungsdesign	<p>Ursprungsstichprobe Welle 1: Zweistufige Zufallsstichprobe mit zwei Teilpopulationen</p> <p>1. Stufe: Auswahl von 300 Postleitzahlbezirken als Primary Sampling Unit (PSU) für beide Teilstichproben. Die Ziehungswahrscheinlichkeit der einzelnen Postleitzahlbezirke hing dabei von der jeweiligen Größe (im Sinne der Einwohnerzahl) ab (probability proportional to size/pps).</p> <p>2. Stufe, Sample I: Ziehung von Bedarfsgemeinschaften aus den Registerdaten der Bundesagentur für Arbeit. Die Zahl der pro PSU gezogenen Bruttostichprobe war abhängig von der ‚Größe‘ der PSU im Sinne des relativen Anteils der Leistungsempfänger im jeweiligen Postleitzahlbezirk (probability proportional to size/pps). Die durchschnittliche Größe der Bruttostichprobe betrug N=100 pro Postleitzahlbezirk.</p> <p>2. Stufe, Sample II: Bei Sample II wurde zunächst eine Gebäudestichprobe aus einer kommerziellen Datenbank (Microm mosaic) gezogen. Diese wurde nach einem in der Datenbank enthaltenen Schichtindex im Verhältnis 4:2:1 für Haushalte mit niedrigem, mittleren bzw. hohem Status geschichtet. Die ausgewählten Gebäude wurden durch einen Interviewer des Erhebungsinstituts begangen. Handelte es sich um Gebäude mit mehreren Parteien, wurden diese notiert und anschließend wurde im Erhebungsinstitut ein zu befragender Haushalt ausgewählt. Die Bruttostichprobe umfasste dabei N=100 Haushalte pro Postleitzahlbezirk.</p> <p>Zugangsstichprobe 1 für Sample I in Welle 2 (Sample III): Zusätzlich zur Fortführung der für Welle 1 gezogenen Samples I wurde in der 2. Welle eine Auffrischungsstichprobe aus den Registerdaten der Bundesagentur für Arbeit gezogen. Dabei wurden Bedarfsgemeinschaften ausgewählt, die im Juli 2007, aber nicht im Juli 2006 Arbeitslosengeld II bezogen haben, die also Neuzugänge in den Bezug darstellen. Die Ziehung erfolgte in für Welle 1 ausgewählten Postleitzahlbezirken analog zum Vorgehen in der 1. Welle.</p> <p>Zugangsstichprobe 2 für Sample I in Welle 3 (Sample IV): Auch in Welle 3 wurde wieder eine Zugangsstichprobe zum Sample I aus den Registerdaten der Bundesagentur für Arbeit gezogen. Dabei wurden Bedarfsgemeinschaften ausgewählt, die im Juli 2008, aber nicht im Juli 2006 und Juli 2007 Arbeitslosengeld II bezogen haben, die also Neuzugänge in den Bezug darstellen. Die Ziehung erfolgte in für Welle 1 ausgewählten Postleitzahlbezirken analog zum Vorgehen in der 1. Welle.</p>
An der Erhebung beteiligte Institutionen	Institut für Arbeitsmarkt- und Berufsforschung; TNS Infratest Sozialforschung, infas Institut für angewandte Sozialwissenschaft GmbH (Datenaufbereitung und –dokumentation Welle 3)
Frequenz der Datensammlung	Jährlich (Panel)
Dateiformat/-größe	STATA, SPSS (mehrere Files)

Kategorien	Erläuterungen
Dateiorganisation	Haushaltsdatensatz: HHENDDAT.dta/.sav Personendatensatz: <i>PENDDAT.dta/.sav</i> Spelldaten ALG I: alg1_spells.dta/.sav (nur Welle 1) Spelldaten ALG II: alg2_spells.dta/.sav Spelldaten Arbeitslosigkeiten: al_spells.dta/.sav (ab Welle 2) Spelldaten Erwerbstätigkeiten: et_spells.dta/.sav (ab Welle 2) Spelldaten Lücken: lu_spells.dta/.sav (ab Welle 2) Spelldaten Maßnahmen: mn_spells.dta/.sav (ab Welle 2) Spelldaten Maßnahmeteiln.: massnahmespells.dta/.sav (nur Welle 1) Registerdaten Haushaltsebene: hh_register.dta/.sav Registerdaten Personenebene: p_register.dta/.sav Gewichtungsdaten Haushaltsebene: hweights.dta/.sav Gewichtungsdaten Personenebene: pweights.dta/.sav Altersvorsorge Haushaltsebene: HAVDAT.dta/.sav (nur Welle 3) Altersvorsorge Personenebene: PAVDAT.dta/.sav (nur Welle 3)

Datenzugang

Kategorien	Erläuterungen
Zugangswege	Scientific Use File (SUF)
Anonymisierungsgrad	faktisch anonymisiert
Sensible Merkmale	keine

Literatur

AAPOR (2006), Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 4th Edition. AAPOR, Lanexa.

Achatz, Juliane, Hirsland, Andreas und Promberger, Markus (2007), „IAB-Panelbefragung von Haushalten im Niedrigeinkommensbereich. Entwurf für ein Rahmenkonzept“, S. 11-32 in Promberger, Markus (Hrsg.), Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung 'Arbeitsmarkt und Soziale Sicherung', IAB Forschungsbericht 12/2007, Nürnberg.

Andersen, H. H., Mühlbacher, A., Nübling, M., (2007a): Die SOEP-Version des SF-12 als Instrument gesundheitsökonomischer Analysen. SOEPPapers 6/2007, 46 S., Berlin: DIW.

Andersen, H. H., Mühlbacher, A., Nübling, M., Schupp, J., Wagner, G. G., (2007b): Computation of Standard Values for Physical and Mental Health Scale Scores Using the SOEP-Version of SF-12v2. Schmollers Jahrbuch 127: 171-182.

Andreß Hans-Jürgen, Burkatzki, Eckhard, Lipsmeier, Gero, Salentin, Kurt, Schulte, Katja und Strengmann-Kuhn, Wolfgang (1996): Leben in Armut. Analysen der Verhaltensweisen armer Haushalte mit Umfragedaten. Endbericht des DFG-Projekts „Versorgungsstrategien privater Haushalte im unteren Einkommensbereich (VuE)“. Bielefeld.

Andreß, Hans-Jürgen und Lipsmeier, Gero (1995): Was gehört zum notwendigen Lebensstandard und wer kann ihn sich leisten? Ein neues Konzept zur Armutsmessung. Aus Politik und Zeitgeschichte, B 31-32/1995: 35-49.

Andreß, Hans-Jürgen und Lipsmeier, Gero (2001): Armut und Lebensstandard. Gutachten im Rahmen des Armuts- und Reichtumsberichts der Bundesregierung. BMAS. Bonn.

Beckmann, Petra und Trometer, Reiner (1991): Neue Dienstleistungen des ALLBUS: Haushalts- und Familientypologien, Klassenschema nach Goldthorpe. ZUMA-Nachrichten 28: 7-17.

Brauns, Hildegard und Steinmann, Susanne (1999): Educational Reform in France, West-Germany and the United Kingdom: Updating the Casmin Classification. ZUMA-Nachrichten 44: 7-45

Bundesministerium für Bildung und Forschung [BMBF] (2003): Berufsausbildung sichtbar gemacht. Schaubildsammlung. 4. Aufl. BMBF, Bonn.

Büngeler, Kathrin, Gensicke, Miriam, Hartmann, Josef, Jäckle, Robert, Tschersich, Nikolai (2009): IAB-Haushaltspanel m Niedrigeinkommensbereich Welle 2 (2007/2008). Methoden- und Feldbericht. FDZ-Methodenreport, 08/2009, Nürnberg.

Büngeler, Kathrin, Gensicke, Miriam, Hartmann, Josef, Jäckle, Robert, Tschersich, Nikolai (im Erscheinen): IAB-Haushaltspanel im Niedrigeinkommensbereich Welle 3 (2008/2009). Methoden- und Feldbericht. FDZ-Methodenreport, Nürnberg.

Christoph, Bernhard (2005): Zur Messung des Berufsprestiges: Aktualisierung der Magnitude-Prestigeskala auf die Berufsklassifikation ISCO88. ZUMA-Nachrichten 57: 79-127.

Europäische Gemeinschaften (EG) (2002): Verordnung (EG) Nr. 29/2002 der Kommission vom 19. Dezember 2001 zur Änderung der Verordnung (EWG) Nr. 3037/90 des Rates betreffend die statistische Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft. Amtsblatt der Europäischen Gemeinschaften, L6/3-L6-33, Brüssel.

Erikson, Robert und Goldthorpe, John H. (1992): *The Constant Flux. A Study of Class Mobility in Industrial Society*. Clarendon Press, Oxford.

Erikson, Robert, Goldthorpe, John H. und Portocarero, Lucienne (1979): *Intergenerational Class Mobility in Three Western Societies: England, France and Sweden*. *British Journal of Sociology* 30: 415-441.

Erikson, Robert, Goldthorpe, John H. und Portocarero, Lucienne (1982): *Social Fluidity in Industrial Nations: England, France and Sweden*. *British Journal of Sociology* 33: 1-34.

Fischer, Andreas und Wirth, Heike (2007): *Constructing Version 4 of ESEC Classes from 3-digit ISCO (Stata-do file)*. Gesis-ZUMA, Mannheim

Frick, Joachim R.; Göbel, Jan und Krause, Peter (o. J.): *\$HGEN: Generated Household-Level Variables*. Download: [<http://www.diw.de/documents/dokumentenarchiv/17/60053/hgen.pdf> (8.11.2007)].

Ganzeboom, Harry B. G. und Treiman, Donald (1996): *Internationally Comparable Measures for Occupational Status for the 1988 International Standard Classification of Occupations*. *Social Science Research*, 25: 201-239.

Ganzeboom, Harry B. G. und Treiman, Donald (2003): *Three Internationally Standardised Measures for Comparative Research on Occupational Status*. In Jürgen H. P. Hoffmeyer-Zlotnik und Wolf, Christof (Hg.): *Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables* (S. 159-193). Kluwer Academic/Plenum Publishers, New York et al.

Ganzeboom, Harry B. G.; De Graaf, Paul M. und Treiman, Donald J. (1992). *A Standard International Socio-Economic Index of Occupational Status*. *Social Science Research*, 21: 1-56.

Gebhardt, Daniel; Müller, Gerrit; Bethmann, Arne; Trappmann, Mark; Christoph, Bernhard; Gayer, Christine; Müller, Bettina; Tisch, Anita; Siflinger, Bettina; Kiesl, Hans; Huyer-May, Bernadette; Achatz, Juliane; Wenzig, Claudia; Rudolph, Helmut; Graf, Tobias; Biedermann, Anika (2009): *Codebuch und Dokumentation des 'Panel Arbeitsmarkt und soziale Sicherung' (PASS) Welle 2 (2007/2008)*, (FDZ Datenreport, 06/2009 (de), Nürnberg

Granato, Nadia (2000): *CASMIN-Bildungsklassifikation. Eine Umsetzung mit dem Mikrozensus 1996*. ZUMA-Technischer Bericht 2000/12.

Hagenaars, Aldi, J. M. de Vos, Klaas und Zaidi, M. Ashgar (1994): *Poverty Statistics in the Late 1980s: Research Based on Micro-data*. Office for Official Publications of the European Communities, Luxembourg.

Halleröd, Björn (1995): *The Truly Poor: Direct and Indirect Consensual Measurement of Poverty in Sweden*. *Journal of European Social Policy*. Vol. 5, Nr. 2: 111-129.

Harrison, Eric und Rose, Richard (2006): *ESeC User Guide, Appendix 6 (SPSS-Syntax: Esec Full)* [<http://www.iser.essex.ac.uk/esec/guide/docs/Appendix6.sps>]

Hartmann, Josef, Brink, Kathrin, Jäckle, Robert und Tschersich, Niko (2008). *IAB-Haushaltspanel im Niedrigeinkommensbereich: Methoden- und Feldbericht*. München.

Hauser, Richard (1996): *Zur Messung individueller Wohlfahrt und Ihrer Verteilung*. S. 13-38 in Statistisches Bundesamt (Hg.): *Wohlfahrtsmessung. Aufgabe der Statistik im gesellschaftlichen Wandel*. Metzler-Poeschel, Stuttgart.

Helberger, Christof (1988): Eine Überprüfung der Linearitätsannahme der Humankapitaltheorie. S. 151-170 in: H.-J. Bodenhöfer (Hrsg.): Bildung, Beruf, Arbeitsmarkt. Duncker & Humblot, Berlin.

International Labour Office [ILO] (1990): International Standard Classification of Occupations. ISCO-88. International Labour Office, Genf.

Jäckle, Annette (2008): The Causes of Seam Effects in Panel Surveys. ISEP Working Paper Series, No. 2008-14, University of Essex.

König, Wolfgang, Lüttinger, Paul und Müller, Walter (1987): Eine vergleichende Analyse der Entwicklung und Struktur von Bildungssystemen. Methodologische Grundlagen und Konstruktion einer vergleichbaren Bildungsskala. CASMIN-Projekt. Arbeitspapier Nr. 12.

Lechert, Yvonne; Schroedter, Julia; Lüttinger, Paul (2006): Die Umsetzung der Bildungs-klassifikation CASMIN für die Volkszählung 1970, die Mikrozensus- Zusatzerhebung 1971 und die Mikrozensusen 1976-2004. ZUMA-Methodenbericht 2006/12.

Lengerer, Andrea; Bohr, Jeanette und Janßen, Andrea (2005): Haushalte, Familien und Lebensformen im Mikrozensus – Konzepte und Typisierungen. ZUMA-Arbeitsbericht Nr. 2005/05; ZUMA; Mannheim.

Lipsmeier, Gero (1999): Die Bestimmung des notwendigen Lebensstandards – Einschätzungsunterschiede und Entscheidungsprobleme. Zeitschrift für Soziologie, Jg. 28, Heft 4: 281-300.

Müller, Walter, Wirth, Heike, Bauer, Gerrit, Pollak, Reinhard und Weiss, Felix (2006): ESeC – Kurzbericht zur Validierung und Operationalisierung einer europäischen sozioökonomischen Klassifikation. ZUMA-Nachrichten 59: 111–119.

Müller, Walter, Wirth, Heike, Bauer, Gerrit, Pollak, Reinhard und Weiss, Felix (2007): Entwicklung einer europäischen sozioökonomischen Klassifikation. Wirtschaft und Statistik, 5/2007: 527-530.

Nolan, Brian und Whelan, Christopher T. (1996). Measuring Poverty Using Income and Deprivation Indicators: Alternative Approaches. Journal of European Social Policy, Vol. 6, No. 3: 225-240

Nübling, M., Andersen, H. H., Mühlbacher, A., (2006): Entwicklung eines Verfahrens zur Berechnung der körperlichen und psychischen Summenskalen auf Basis der SOEP – Version des SF 12 (Algorithmus). Data Documentation 16, 14 S., Berlin: DIW.

Organisation for Economic Co-Operation and Development [OECD] (1999): Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries. 1999 Edition. OECD, Paris.

Organisation for Economic Co-Operation and Development [OECD] (Hrsg.) (1982): The OECD-List of Social Indicators. OECD, Paris.

Porst, Rolf (1984): Haushalt und Familien 1982. Zur Erfassung und Beschreibung von Haushalts- und Familienstrukturen mit Hilfe repräsentativer Bevölkerungsumfragen. Zeitschrift für Soziologie, 13(2): 164-175.

Rendtel, Ulrich, Harms, Torsten (2009): Weighting and calibration for household panels, S. 265-286 in Lynn, Peter (Hrsg.), Methodology of Longitudinal Surveys. Chichester, Wiley.

Ringen, Stein (1988): Direct and Indirect Measurement of Poverty. Journal of Social Policy 17, 3: 351-365.

Rose, Richard und Harrison, Eric (2007): The European Socio-Economic Classification: A New Social Class Schema for Comparative European Research. *European Societies* 9, 3: 459-490.

Sozialgesetzbuch Zweites Buch [SGB II]: Grundsicherung für Arbeitsuchende.

Stata Corp (2007): Survey Data Reference Manual Release 10. Stata Press, College Station.

Statistisches Bundesamt [StBA] (1992): Klassifizierung der Berufe. Systematisches und alphabetisches Verzeichnis der Berufsbenennungen. Statistisches Bundesamt, Wiesbaden.

Statistisches Bundesamt [StBA] (2002): Klassifikation der Wirtschaftszweige, Ausgabe 2003 (WZ2003). Statistisches Bundesamt, Wiesbaden.

Trappmann, Mark, Christoph, Bernhard, Achatz, Juliane, Wenzig, Claudia, Müller, Gerrit, Gebhardt, Daniel (2009): Design and stratification of PASS. A New Panel Study for Research on Long Term Unemployment. IAB-Discussion Paper, 5/2009, Nürnberg.

Treiman, Donald (1977): Occupational Prestige in Comparative Perspective. Academic Press, New York.

Ware JE, Kosinski M, Turner-Bowker DM, Gandek B (2002): How to score Version 2 of the SF-12® Health Survey. Lincoln, RI: Qualitymetric Incorporated,

Wegener, Bernd (1985): Gibt es Sozialprestige? *Zeitschrift für Soziologie*, 14: 209-235.

Wegener, Bernd (1988): Kritik des Prestiges. Westdeutscher Verlag, Opladen.

Impressum

FDZ-Datenreport 6/2010

Herausgeber

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

Redaktion

Stefan Bender, Iris Dieterich

Technische Herstellung

Iris Dieterich

Rechte

Nachdruck - auch auszugsweise - nur mit
Genehmigung des FDZ gestattet

Bezugsmöglichkeit

http://doku.iab.de/fdz/reporte/2010/DR_06-10_I.pdf

Internet

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

Rückfragen zum Inhalt an:

Arne Bethmann, Institut für Arbeitsmarkt- und
Berufsforschung, Regensburger Str. 104,
90478 Nürnberg; Telefon: 0911/179-2307,
E-Mail: arne.bethmann@iab.de

Mark Trappmann, Institut für Arbeitsmarkt- und
Berufsforschung, Regensburger Str. 104,
90478 Nürnberg; Telefon: 0911/179-3096,,
E-Mail: mark.trappmann@iab.de