

FDZ-Datenreport

Dokumentation zu Arbeitsmarktdaten

06/2018
DE

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

Datenreport Welle 11

Marco Berg,
Ralph Cramer,
Christian Dickmann,
Reiner Gilberg,
Birgit Jesske,
Martin Kleudgen,
Jonas Beste,
Sandra Dummert,
Corinna Frodermann,
Stefan Schwarz,
Mark Trappmann



Codebuch und Dokumentation des Panel „Arbeitsmarkt und soziale Sicherung“ (PASS)

Datenreport Welle 11

Dokumentation: PASS-SUF0617 DE v1 dok1
DOI: 10.5164/IAB.FDZD.1806.de.v1
Datensatz: PASS-SUF0617, Version 1
DOI: 10.5164/IAB.PASS-SUF0617.de.en.v1

Marco Berg, Ralph Cramer, Christian Dickmann, Reiner Gilberg, Birgit Jesske,
Martin Kleudgen (infas Institut für angewandte Sozialwissenschaft GmbH)
Jonas Beste, Sandra Dummert, Corinna Frodermann, Stefan Schwarz, Mark
Trappmann (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 Welle 11 von PASS. Das Dokument basiert auf dem Datenreport der Welle 10: Marco Berg, Ralph Cramer, Christian Dickmann, Reiner Gilberg, Birgit Jesske, Martin Kleudgen (alle infas Institut für angewandte Sozialwissenschaft GmbH), Jonas Beste, Sandra Dummert, Corinna Frodermann, Benjamin Fuchs, Stefan Schwarz, Mark Trappmann, Simon Trenkle (alle Institut für Arbeitsmarkt- und Berufsforschung (IAB)): Codebuch und Dokumentation des Panel „Arbeitsmarkt und soziale Sicherung“ (PASS) Band I: Datenreport Welle 10, FDZ Datenreport, 07/2017 (de), Nürnberg.

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 PASS wave 11 and is based upon the tenth wave's data report: Marco Berg, Ralph Cramer, Christian Dickmann, Reiner Gilberg, Birgit Jesske, Martin Kleudgen (all infas Institut für angewandte Sozialwissenschaft GmbH), Jonas Beste, Sandra Dummert, Corinna Frodermann, Benjamin Fuchs, Stefan Schwarz, Mark Trappmann, Simon Trenkle (alle Institut für Arbeitsmarkt- und Berufsforschung (IAB)): Codebuch und Dokumentation des Panel „Arbeitsmarkt und soziale Sicherung“ (PASS) Band I: Datenreport Welle 10, FDZ Datenreport, 07/2017 (de), Nürnberg.

Datenverfügbarkeit

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

Tabellenanhang

Der diesem Beitrag zugrunde liegende Tabellenanhang ist unter http://doku.iab.de/fdz/pass/Tabellenanhang_PASS_Welle11.xlsx verfügbar.

Inhaltsverzeichnis

1	Einleitung	8
1.1	Ziele und Fragestellungen des Panel „Arbeitsmarkt und soziale Sicherung“ .	8
1.2	Instrumente und Befragungsprogramm	9
1.3	Besonderheiten und Neuerungen der Welle 11	10
1.3.1	Personenfragebogen	10
1.3.2	Seniorenfragebogen	12
1.3.3	Haushaltsfragebogen	13
1.3.4	Stichprobe und Datenaufbereitung	13
2	Kennwerte	15
2.1	Stichprobenumfang	16
2.2	Response-Raten	18
2.3	Panelbereitschaft, Zuspielbereitschaft, Verknüpfung mit Prozessdaten . . .	19
2.4	Abgespaltene Haushalte	20
3	Datensatzstruktur	21
4	Generierte Variablen	24
4.1	Vercodung offener Angaben	24
4.1.1	Offene Restkategorien und offene Items	24
4.1.2	Berufs- und Branchencodierung	28
4.2	Harmonisierung	33
4.3	Dependent Interviewing	35
4.4	Einfache Generierungen	39
4.5	Konstruktvariablen	75
4.5.1	Personenebene	75
4.5.2	Haushalts- oder Bedarfsgemeinschaftsebene	98
5	Datenaufbereitung	112
5.1	Strukturprüfungen und entfernte Interviews	113
5.2	Filterbereinigung	119
5.3	Plausibilitätsprüfungen	121
5.4	Rückwirkende Veränderungen der Wellen 1 bis 10	123
5.5	Anonymisierung	127
5.6	Arbeitslosengeld-II-Bezüge	131
5.6.1	Konzept für die Fortschreibung der in der Vorwelle aktuell andauernden Arbeitslosengeld-II-Spells	131
5.6.2	Struktur des Arbeitslosengeld-II-Spelldatensatzes	132
5.6.3	Plausibilitätsprüfung und –bereinigung des Arbeitslosengeld-II Spell-datensatzes	132
5.6.4	Fortschreibung des Arbeitslosengeld-II-Spelldatensatzes	133
5.7	Erwerbsbiografien	134
5.7.1	Variablen zum (Nicht-)Erwerbsstatus im PENDDAT	134
5.7.2	Einkommensvariablen und Arbeitszeit im PENDDAT und in BIO-Spells	144
5.7.3	Konzept für die Fortschreibung der in der Vorwelle aktuell andauernden Spells	149
5.7.4	Struktur des BIO-Spelldatensatzes	150

5.7.5	Plausibilitätsprüfung und –bereinigung der Spelldatensätze	151
5.7.6	Fortschreibung der Spelldatensätze	152
5.8	Ein-Euro-Spelldatensatz (<i>ee_spells</i>)	154
5.8.1	Konzept für die Fortschreibung der in der Vorwelle aktuell andauernden Spells	154
5.8.2	Struktur des Ein-Euro-Spelldatensatzes	154
5.8.3	Plausibilitätsprüfung und –bereinigung des EEJ-Spelldatensatzes	155
6	Gewichtung Welle 11	156
6.1	Designgewichte für die Bestandsauffrischung (EWO) der Welle 11	156
6.2	Integration der Designgewichte Bestandsauffrischung (EWO) mit den bestehenden Gewichten der Bevölkerungsstichprobe	157
6.3	Designgewichte für die Bestands Haushalte der Welle 11	157
6.4	Designgewichte für die Zugangsstichprobe der Welle 11	157
6.5	Wiederteilnahme-Propensity Haushalte	158
6.6	Teilnahme-Propensity erstbefragte Split-Haushalte	158
6.7	Ausfallgewichtung für Haushalte aus der Auffrischungsstichprobe der BA-Neuzugänge Welle 11	159
6.8	Ausfallgewichtung für Haushalte aus der Bestandsauffrischung (EWO) Welle 11	159
6.9	Wiederteilnahme-Propensity Personen	160
6.10	Integration der Gewichte zum Gesamtgewicht vor Kalibrierung	160
6.11	Integration temporärer Ausfälle (Haushalte)	161
6.12	Kalibrierung zum Haushaltsgewicht Welle 11 Querschnitt	162
6.13	Kalibrierung der BA-Stichprobe	163
6.14	Kalibrierung der Bevölkerungsstichprobe	164
6.15	Kalibrierung der Gesamtstichprobe	165
6.16	Kalibrierung zum Personengewicht Welle 11 Querschnitt	166
6.17	BA-Stichprobe	167
6.18	Bevölkerungsstichprobe	168
6.19	Gesamtstichprobe	169
6.20	Schätzung der BA-Querschnittsgewichte für Haushalte und Personen ohne Arbeitslosengeld-II-Bezug	171
7	Anhang: Kurzbeschreibung des Datensatzes	173

Abbildungsverzeichnis

Abbildung 1	Realisierte Panelstichprobe Haushalte und Personen nach Erhebungswellen	18
Abbildung 2	Datensatzstruktur des PASS in Welle 11	23
Abbildung 3	Übersicht über generierte Variablen auf Personenebene in Welle 11	73

Tabellenverzeichnis

Tabelle 1	Vercodung offener Angaben auf Haushaltsebene in Welle 11	25
Tabelle 2	Vercodung offener Angaben auf Personenebene in Welle 11	26
Tabelle 3	Codeschema der im PASS genutzten Zusatzvariablen	30

Tabelle 4	Harmonisierte Variablen im Personendatensatz (<i>PENDDAT</i>)	34
Tabelle 5	Wellenübergreifend generierte, aber nicht vollständig harmonisierte Variablen im Personendatensatz (<i>PENDDAT</i>)	35
Tabelle 6	Fortgeschriebene Vorwelleninformationen in Welle 11, Haushaltsfragebogen	37
Tabelle 7	Fortgeschriebene Vorwelleninformationen in Welle 11, Personenfragebogen	38
Tabelle 8	Typen der einfachen generierten Variablen in den Querschnittsdatensätzen (<i>HHENDDAT</i> ; <i>PENDDAT</i>) bei Haushalten bzw. Personen, die schon in einer früheren Welle Angaben zum jeweiligen Thema gemacht haben	40
Tabelle 9	Einfache Generierungen für Welle 11 im Haushaltsdatensatz (<i>HHENDDAT</i>) und im <i>KINDER</i> -Datensatz (alphabetische Reihenfolge)	43
Tabelle 10	Einfache Generierungen für Welle 11 im Personendatensatz (<i>PENDDAT</i>) (alphabetische Reihenfolge)	46
Tabelle 11	Einfache Generierungen für Welle 11 im Arbeitslosengeld-II-Datensatz (<i>alg2_spells</i>) (Reihenfolge analog zur Reihenfolge im Datensatz)	58
Tabelle 12	Einfache Generierungen für Welle 11 im BIO-Spellldatensatz (<i>bio_spells</i>) (Reihenfolge analog zur Reihenfolge im Datensatz)	62
Tabelle 13	Einfache Generierungen für Welle 11 im Ein-Euro-Spellldatensatz (<i>ee_spells</i>) (Reihenfolge analog zur Reihenfolge im Datensatz)	67
Tabelle 14	Einfache Generierungen für Welle 11 im Personenregisterdatensatz (<i>p_register</i>) (alphabetische Reihenfolge)	69
Tabelle 15	Bildung in Jahren	75
Tabelle 16	Bildung in Jahren, Mutter	76
Tabelle 17	Bildung in Jahren, Vater	77
Tabelle 18	CASMIN	77
Tabelle 19	MCASMIN	78
Tabelle 20	VCASMIN	79
Tabelle 21	ISCED 97	80
Tabelle 22	MISCED 97	81
Tabelle 23	VISCED 97	82
Tabelle 24	International Standard Classification of Occupations 1988 (ISCO88)	83
Tabelle 25	International Standard Classification of Occupations 2008 (ISCO08)	84
Tabelle 26	Klassifizierung der Berufe 1992 (KIdB1992)	86
Tabelle 27	Klassifizierung der Berufe 2010 (KIdB2010)	87
Tabelle 28	Klassenschema nach Erikson, Goldthorpe und Portocarrero (EGP)	88
Tabelle 29	European Socio-economic Classification (ESeC)	90
Tabelle 30	Magnitude-Prestigeskala (MPS)	91
Tabelle 31	Standard International Occupational Prestige Scale (SIOPS/Treiman-Skala)- Basis ISCO-88	91
Tabelle 32	Standard International Occupational Prestige Scale (SIOPS/Treiman-Skala) – Basis ISCO-08	92
Tabelle 33	International Socio-Economic Index (ISEI) – Basis ISCO-88	93
Tabelle 34	International Socio-Economic Index (ISEI) – Basis ISCO-08	94
Tabelle 35	Klassifikation der Wirtschaftszweige 2003 (WZ2003)	95
Tabelle 36	Klassifikation der Wirtschaftszweige 2008 (WZ2008)	95
Tabelle 37	Körperliche Skala des SF12v2 (SOEP-Version, NBS)	96

Tabelle 38	Psychische Skala des SF12v2 (SOEP-Version, NBS)	97
Tabelle 39	Bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung .	98
Tabelle 40	Bedarfsgewichtetes Haushaltseinkommen, neue OECD-Gewichtung	98
Tabelle 41	Deprivationsindex, ungewichtet	99
Tabelle 42	Deprivationsindex, gewichtet	100
Tabelle 43	Haushaltstypologie	102
Tabelle 44	Bedarfsgemeinschaftsnummer, Welle 11	104
Tabelle 45	Bedarfsgemeinschaftstypologie, Welle 11	106
Tabelle 46	Arbeitslosengeld II-Bezug der Bedarfsgemeinschaft zum Zeitpunkt der Stichprobenziehung, Welle 11	107
Tabelle 47	Arbeitslosengeld II-Bezug der Bedarfsgemeinschaft zum Befragungs- zeitpunkt, Welle 11	107
Tabelle 48	Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 10	108
Tabelle 49	Flag zur Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsge- meinschaft zum Befragungszeitpunkt, Welle 10	108
Tabelle 50	Anzahl der Bedarfsgemeinschaften im Haushalt	110
Tabelle 51	Anzahl der zum Stichprobenzeitpunkt Leistungen beziehenden Be- darfsgemeinschaften im Haushalt	111
Tabelle 52	Übersicht über die Schritte der Datenaufbereitung in der Welle 11 des PASS	113
Tabelle 53	Übersicht über die vergebenen Missing Codes	121
Tabelle 54	Übersicht über rückwirkenden Veränderungen im Haushaltsdaten- satz (<i>HHENDDAT</i> , <i>KINDER</i>)	123
Tabelle 55	Übersicht über rückwirkenden Veränderungen im Personendaten- satz (<i>PENDDAT</i>)	124
Tabelle 56	Übersicht über rückwirkenden Veränderungen in Spelldatensätzen (<i>bio_spells</i> , <i>alg2_spells</i> , <i>ee_spells</i>)	126
Tabelle 57	Übersicht über rückwirkenden Veränderungen in Registerdatensät- zen (<i>hh_register</i> ; <i>p_register</i>)	126
Tabelle 58	Übersicht über rückwirkenden Veränderungen in den Gewichtungs- datensätzen (<i>hweights</i> ; <i>pweights</i>)	127
Tabelle 59	Übersicht über die in Welle 11 anonymisierten Variablen im Perso- nendatensatz (<i>PENDDAT</i>)	128
Tabelle 60	Übersicht über die in Welle 11 anonymisierten Variablen im BIO- Spelldatensatz (<i>bio_spells</i>)	131
Tabelle 61	Übersicht über die in Welle 11 anonymisierten Variablen im Kinder- Datensatz (<i>KINDER</i>)	131
Tabelle 62	Querschnittvariablen im ALG II-Spelldatensatz (<i>alg2_spells</i>)	132
Tabelle 63	Generierungslogik <i>erwerb</i> , <i>erwerb2</i> , <i>nichterw</i> , <i>nichterw2</i>	136
Tabelle 64	Überarbeitung <i>erwerb</i> , <i>erwerb2</i> , <i>nichterw</i> , <i>nichterw2</i>	140
Tabelle 65	Grobzuweisung - Spell höherer Priorität schlägt Spell niedrigerer Priorität	143
Tabelle 66	Feinzuweisung für Sonderfälle	144
Tabelle 67	Überarbeitung Einkommensvariable	145
Tabelle 68	Überarbeitung Arbeitszeitvariablen	147
Tabelle 69	ET-spezifische Querschnittvariablen im BIO-Spelldatensatz (<i>bio_- spells</i>)	150

Tabelle 70	AL-spezifische Querschnittvariablen im BIO-Spelldatensatz (<i>bio_spells</i>)	151
Tabelle 71	Querschnittvariablen im Ein-Euro-Spelldatensatz (<i>ee_spells</i>)	154

1 Einleitung

1.1 Ziele und Fragestellungen des Panel „Arbeitsmarkt und soziale Sicherung“

Mit dem Panel „Arbeitsmarkt und soziale Sicherung“ (PASS) hat das Institut für Arbeitsmarkt- und Berufsforschung (IAB) 2006 einen Datensatz für die Arbeitsmarkt-, Sozialstaats- und Armutsforschung in Deutschland aufgebaut, 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 Befragungswelle 11, für die zwischen Februar 2017 und Oktober 2017 13.703 Personen in 9.420 Haushalten² befragt wurden. Darunter befanden sich 10.305 Personen und 7.165 Haushalte, mit denen bereits zum wiederholten Mal ein Interview im Rahmen von PASS geführt wurde.

Der vorliegende wellenspezifische Datenreport³ der Welle 11 dokumentiert die wellenbezogenen Aspekte der Studie. Im Kapitel 1 findet sich ein Überblick über die Ziele und

¹ Sozialgesetzbuch (SGB) Zweites Buch (II) - Grundsicherung für Arbeitsuchende

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

³ Mit der Dokumentation der Welle 3 wurde der Report erstmalig in zwei Teile aufgeteilt: ein wellenspezifischer Datenreport (inkl. Codebuch) und ein wellenübergreifender User Guide. Der wellenübergreifende User Guide wird in Verantwortung des PASS-Projektteams beim IAB erstellt. Die Dokumentation des wellenspezifischen Datenreports ab der Welle 3 wird von infas erstellt. Grundlage dafür bildete der Datenreport der Welle 2. Der wellenübergreifende User Guide soll der Dokumentation der Studie insgesamt dienen. Hier werden ausführlich die Ziele und das Design von 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.

Fragestellungen der Studie mit einer Kurzbeschreibung der Instrumente und des Befragungsprogramms im Kapitel 1.2, sowie den Besonderheiten und Neuerungen der Welle 11 im Kapitel 1.3. Die zentralen Kennzahlen zu Stichprobe und Ausschöpfungsquoten der Welle 11 werden in Kapitel 2 dargelegt. In den darauf folgenden Kapiteln geht es um die eigentlichen Daten und deren Aufbereitung. Ein Überblick über die Datenstruktur findet sich in Kapitel 3 und in Kapitel 4 werden die generierten Variablen beschrieben. Weiterhin werden die Schritte der Datenaufbereitung und die dabei getroffenen Entscheidungen in Kapitel 5 erläutert. In Kapitel 6 wird dann das Gewichtungungsverfahren vorgestellt. Am Ende findet sich in Kapitel 7 ein Gesamtüberblick über die Datensätze im PASS über alle Wellen. In den gesonderten Tabellenbänden werden die Häufigkeiten aller im Scientific Use File enthaltenen Variablen, unterteilt nach den jeweiligen Datensätzen, aufgelistet, die in Welle 11 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 Haushaltsinterview⁴ 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 Welle 11 basieren auf der Welle 10 des PASS, wobei jedoch Einzelfragen und -module überarbeitet oder neu entwickelt wurden (siehe für einen Überblick Kapitel 1.3).

Die Erhebungsinstrumente von PASS sind so angelegt, dass sie sowohl die wiederholte Befragung von Haushalten und Personen, die bereits an einer der Vorwellen teilgenommen haben, als auch die erstmalige Befragung ermöglichen⁵. Um bei der erneuten Befragung Seam-Effekte⁶ zu vermeiden und die Datenqualität zu erhöhen, wurden bereits seit

⁴ Die Zielperson für das Haushaltsinterview sollte sich möglichst gut mit allgemeinen Fragen zum Haushalt auskennen. Die Auswahl erfolgte nach bestimmten Regeln und ist ausführlich in den Methodenberichten dokumentiert (Jesske & Quandt (2011); Jesske & Schulz (2012); Jesske & Schulz (2013), Jesske & Schulz (2014); Jesske & Schulz(2015); Jesske et al. (2016); Jesske et al. (2017); Jesske & Schulz (2018 im Erscheinen)).

⁵ Zu den erstmals befragten Haushalten gehören: (1) Haushalte der Aufstockungsstichprobe und Zugangsstichprobe der aktuellen Welle sowie (2) Haushalte, die sich von bereits in den Vorwellen befragten Haushalten abgespalten haben (Split-Haushalte) (weitere Erläuterungen vgl. Methodenbericht der Welle 4, Jesske & Quandt (2011)).

⁶ 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)).

der Welle 3 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 ab Welle 4 - abweichend zum Vorgehen der Wellen 1 bis 3 - ein integrierter Fragebogen für wiederbefragte Haushalte (HHalt) und für erstmals befragte Haushalte (HHneu)⁷.

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 Welle 11.

1.3 Besonderheiten und Neuerungen der Welle 11

Nutzerinnen und Nutzern, die bereits mit den Daten der Panelwellen gearbeitet haben, möchten wir an dieser Stelle einen schnellen Überblick über Besonderheiten der PASS-Welle 11 geben.

Die Besonderheiten und Neuerungen in Welle 11 betreffen das Frageprogramm für den Haushalts- und Personenfragebogen (Änderung der Referenzzeiträume, Modifikation von Einzelfragen und neue Fragenmodule)⁸, die Stichprobe und die Aufbereitung der Daten.

1.3.1 Personenfragebogen

Im Personenfragebogen werden die seit Welle 2 erhobenen erwerbsbiografischen Informationen fortgeschrieben⁹. Ab Welle 4 ist eine geänderte Abfragelogik der chronologischen Retrospektiverhebung enthalten (s.a. Kap. 1.3.1 in Berg et al., FDZ Datenreport 08/2011).

Im Personenbogen wurden in Welle 11 Module bzw. Fragenblöcke neu entwickelt oder aus den bereits verwendeten Modulen der Vorwellen erneut zum Einsatz gebracht. Darüber hinaus wurden einzelne Module im Vergleich zur Vorwelle modifiziert bzw. gestrichen.

Gestrichen wurden folgende Module bzw. Einzelfragen:

Modul **Impulsivität** (I-8-Skala) (*PEO1400**)

Modul **Arbeitszeitsveränderungen** (*PET1460 - PET1480*)

⁷ Dabei werden Split-Haushalte wie neue Haushalte für die Abfrage behandelt.

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

⁹ Dies geschieht u.a. mit Hilfe der so genannten „dependent interviewing“-Methode. Beim Dependent Interviewing werden Angaben, die wiederholt befragte Personen im Vorwelleninterview gemacht haben, in den Interviewtext des aktuellen Interviews eingespielt, um zu prüfen, ob diese Informationen aktualisiert werden müssen.

Modul **Einstellungen (Beruf und Familie)** (*PEO0800a-b - PEO1100a-b*)

Modul **Einstellung institutioneller Kinderbetreuung** (*PEO1700**)

Modul **Einstellungen zum Mindestlohn** (*PML0100*)

Modul **Freizeit von Jugendlichen** (*PA1100-PA1300*)

Modul **Einstellungen (Freizeit von Kindern)** (*PEO1500**)

Modul **Einstellungen (Reziprozität)** (*PEO1600**)

Frage zur Scheinselbstständigkeit (*ET4300*)

Die Aufnahme neuer Module bzw. Fragen betreffen im Wesentlichen die folgenden Punkte:

Das Modul **Einstellungen (Selbstwirksamkeit)** (*PEO0100**) wird in Welle 11 wieder aufgenommen. Die Wiederaufnahme basierte auf den Fragen aus Welle 8.

Das Modul **Ausbildungsaspirationen** (*PAA0100- PAA1200*) wurde für Welle 11 neu entwickelt. Befragt werden hier zum einen im Teil I Schüler an allgemeinbildenden Schulen und Schüler an beruflichen Schulen/Fachschulen, die keine vollqualifizierenden Abschluss erwerben. Um die betreffenden Schüler identifizieren zu können, wurde im Bildungsmodul die Frage PB0205 aufgenommen. Zum anderen werden im Teil II unter 35-Jährige ohne beruflichen Abschluss befragt. Als Kernstück beinhaltet das Modul für Schüler die offene Abfrage des Wunschberufs (*PAA0100*). Es werden aber auch Zusatzinformationen erhoben, wie sicher der Berufswunsch ist (*PAA0200*), ob man Personen kennt, die diesen Beruf ausüben (*PAA0300, PAA0400*), inwieweit die Eltern sich bei der Wahl beteiligen (*PAA0500*) sowie Fragen zur Berufsfindung (*PA0600*) und wichtige Aspekte bei der Berufswahl (*PAA0700*). Unter 35-Jährige ohne beruflichen Abschluss werden ebenfalls kurz zu ihren Berufsvorstellungen gefragt. Dazu gehören die Fragen nach dem Erwerb eines Abschlusses (*PAA0900*), die offene Abfrage des Berufswunsches (*PAA0100*), die Informiertheit über Ausbildungsmöglichkeiten (*PAA01200*) sowie Gründe für den Nicht-Erwerb (*PAA1100*).

Das Modul **Einstellungen (Rollenbilder)** (*PEO0400a-d*) wird wieder aus der Welle 8 (und vorherigen Wellen) aufgenommen.

Das Modul **Soziale Integration** wurde in Welle 11 neu entwickelt. Dafür wurden neben der Frage zum Sozialen Vertrauen (*PA2000*) auch im Rahmen des Moduls **Netzwerke** Detailfragen zur Aktivität in Gewerkschaften, Vereinen und Parteien integriert (*PSK410-PSK0500*).

Das Modul **Religion** wird in Welle 11 wieder an alle Befragte gestellt.

Im Rahmen des Moduls **Pflege** wurden zur Kenntnis (*PP1500*) und Inanspruchnahme der Pflegezeit (*PP1600*) eine zusätzliche Frageformulierung und Filterung für Wiederholtbefragte aufgenommen.

Neu aufgenommen wurde ein Modul mit zwei Fragen zum **Smartphone-Besitz** (*PSM0100-PSM0200*).

Neu aufgenommen wurde eine Frage zum **Arbeitslosengeld 2-Bezug seit 2005** (*PA0980*), mit der abgefragt wurde, ob die Person jemals seit Einführung des Arbeitslosengeldes 2 eine solche Leistung erhalten hat.

Das Modul **Einstellung (Finanzen)** wurde neu entwickelt. Integriert wurde dabei die modifizierte Itematterie zum Umgang mit Finanzen (4 von 8 Items) aus Welle 8 (*PEF0100*). Außerdem wurden die finanzielle Bildung und die mathematischen Kompetenzen der Befragten mit einer Reihe von Wissens- und Rechenfragen erhoben (*PEF0200-PEF0500*). Die richtigen Antworten sind den Variablenlabels zu entnehmen.

Im Modul **Migration** wurden einerseits Anpassungen aufgrund der Neuerungen in Welle 10 vorgenommen. So wird *PMI1700* in Welle 11 nur noch an Neubefragte gestellt, die nicht in Deutschland geboren sind. Außerdem wurde die Kategorie 5 in *PMI1700* geringfügig geändert. Durch die neue Filterung beginnt das Modul **Migration** für zugezogene Altbefragte in Welle 11 bei *PMI1800*. Dadurch musste eine neue Textvariante mit Moduleinleitung in *PMI1800* integriert, sowie die Sonderkategorie „Befragter widerspricht“ in *PMI1700* gelöscht und bei *PMI1800* aufgenommen werden. Außerdem wurde der Besuch eines Deutschkurses für Wiederholbefragte in eine neue Frage integriert, um die Fortschreibung zu realisieren (*PMI2000*).

Andererseits wurde ein neues Teilmodul **Anerkennung ausländischer Abschlüsse** (*PMI2100-PMI2900*) entwickelt. Nach der Filterfrage an alle Zuwanderer, ob berufliche Abschlüsse im Ausland erworben wurden (*PMI2100*) wird erhoben, ob man die Anerkennung bereits beantragt hat (*PMI2200*), was das Ergebnis der Anerkennung war (*PMI2300-PMI2900*) und den Grund für die Nichtbeantragung der Anerkennung (*PMI3000*).

1.3.2 Seniorenfragebogen

Aufgrund der schrittweisen Erhöhung des Rentenalters erfolgt - um den Seniorenbogen als Kurzversion weiterhin bereits für 65-Jährige und älter vorzusehen - die Filterung für Befragte mit validen Angaben zum Geburtstag ab Welle 10 monatsgenau. Gemäß des Renteneintrittsalters wird die Altersgrenze für den Übergang vom Senioren bzw. Personenbogen in Welle 11 wie folgt angepasst: bei 65 Jahre und 5 Monate (Jahrgang 1951) bzw. bei 65 Jahre und 6 Monate (Jahrgang 1952).

Aus der Liste der Änderungen für den Personenfragebogen wurden folgende Änderungen auch im Seniorenfragebogen vorgenommen:

Die Module **Einstellungen Selbstwirksamkeit** (*PEO0100**) und **Einstellungen Rollenbilder** (*PEO0400a-d*) wurden aus Welle 8 wieder aufgenommen.

Das Modul **Religion** wird in Welle 11 wieder an alle Befragte gestellt.

Im Modul **Pflege** (*PP1500, PP1600*) wurden zusätzliche Frageformulierungen und Filterungen für Wiederholtbefragte aufgenommen.

Die Module **Einstellung (Finanzen)** (*PEF0100*-PEF0500*) und **Soziale Integration** (*PA2000*) wurden in Welle 11 neu entwickelt.

Im Rahmen des Moduls **Netzwerke** wurden Nachfragen zur Aktivität in Gewerkschaften, Vereinen und Parteien integriert (*PSK410-PSK0500*).

Neu aufgenommen wurde eine Frage zum **Arbeitslosengeld 2-Bezug seit 2005** (*PA0980*).

1.3.3 Haushaltsfragebogen

Im Haushaltsfragebogen wurden in Welle 11 nur wenige Änderungen vorgenommen.

Das Modul **Soziale Teilhabe von Kindern und Jugendlichen** (*HT0100-HT0510*) wurde komplett gestrichen.

Ebenfalls komplett gestrichen wurde das Modul **Bildungs- und Teilhabepaket** (*HBT0100-HBT0815*). Die Fragen zur Erstellung der Variablen HHBTP_Bez (Anspruch auf BuT-Leistungen) sollen weiter erhoben werden. Die Variable wird zwar zur Steuerung des Haushaltsfragebogens nicht mehr benötigt, soll aber im SUF weiterhin ausgeliefert werden.

Im Modul **Einkommen** wurden die Fragen zum Betreuungsgeld (*HEK1650, HEK1660*) gestrichen.

1.3.4 Stichprobe und Datenaufbereitung

Wie in den Vorwellen, wurde auch in Welle 11 eine sog. Zugangsstichprobe für die BA-Teilstichprobe¹⁰ gezogen. Ziel ist es, die Repräsentativität der BA-Stichprobe im Querschnitt sicherzustellen. Für die Auffrischung wurden Bedarfsgemeinschaften gezogen, die im Juli 2016 Arbeitslosengeld II bezogen haben, nicht jedoch zum Zeitpunkt der Stichprobenziehungen der Wellen 1 bis 10 (vgl. Kapitel 2.1 und zum Auffrischungskonzept Trappmann et al. (2009), S. 11 ff.). Alle in Welle 11 neubefragten Haushalte können wiederum über den Stichprobenindikator (*sample*) identifiziert werden.

Der aktuelle verstärkte Zuzug von Flüchtlingen nach Deutschland ist nicht ohne Folgen für den Zugang in die Gruppe der Leistungsempfänger des SGB II geblieben, so dass im PASS ab der Welle 10 arabisch als weitere Erhebungssprache eingesetzt wurde. Mit der

¹⁰ 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. Bevölkerungsstichprobe.

Erweiterung der Erhebungssprache ist sichergestellt, dass anerkannte Flüchtlinge aus den häufigsten Herkunftsländern (Syrien und Irak) im Rahmen der jährlichen Auffrischungstichproben der Grundsicherungsempfänger erreicht und dann auch im Panel weiterverfolgt werden können. Während in Welle 10 ein Oversampling von neu zugegangenen Bedarfsgemeinschaften mit Personen syrischer und irakischer Nationalität ins SGB II erfolgte, damit Flüchtlinge in ausreichender Zahl befragt werden konnten, wurde in Welle 11 eine Auffrischungstichprobe von Bedarfsgemeinschaften innerhalb der Sampling Points von PASS nach dem üblichen Verfahren gezogen (weitere Details s. Methodenbericht der Welle 11¹¹). In dieser Auffrischungstichprobe waren Haushalte mit Personen syrischer und irakischer Nationalität proportional vertreten. Vor dem Hintergrund, dass sich die Leistungsempfänger des SGB II mit syrischer und irakischer Nationalität deutlich von den üblichen Bedarfsgemeinschaften unterscheiden, wird weiterhin eine getrennte Ausweisung in den weiteren Darstellungen sowie im Datensatz vorgenommen. Als syrische / irakische Haushalte werden Haushalte gewertet in denen mindestens eine Person mit syrischer oder irakischer Nationalität lebt. In wenigen Fällen führt dies dazu, dass weitere Personen, die in diesen Haushalten leben aber nicht aus diesen Ländern stammen, dieser Gruppe zugeordnet werden. Um in der Gruppe der Personen aus der Teilstichprobe der syrischen und irakischen Haushalte Personen mit syrischer Staatsangehörigkeiten identifizieren zu können, wird ab dem Scientific-Use-File der Welle 11 die zusätzliche Variable *ostaatansyr* ausgeliefert. Diese Variable ist rückwirkend bereits ab Welle 10 gefüllt. Aufgrund der geringen Fallzahlen werden nur die beiden Kategorien „syrische Staatsangehörigkeit“ und „eine andere oder keine Staatsangehörigkeit“ ausgewiesen.

Zusätzlich fand in Welle 11 eine Bestandsauffrischung der Bevölkerungstichprobe statt. Dazu wurden die Ziehungsgemeinden der Welle 5 erneut ausgewählt und innerhalb der Gemeinden erfolgte die Ziehung aus dem Einwohnermelderegister. Die Ziehung der Personenadressen aus der Auswahlgesamtheit in den Gemeinden erfolgte über eine systematische Zufallsauswahl (Intervallziehung). Eine genaue Beschreibung zum Vorgehen findet sich im Kapitel 6.3.

Die Datenaufbereitung erfolgte in enger Abstimmung mit dem IAB. Grundsätzliche Vorgehensweisen z.B. zur Fortschreibung der Datensätze und zur Bereinigung von Problemen in den Haushaltsstrukturen wurden im Aufbereitungsprozess diskutiert und durch das IAB entschieden.

Das Konzept für die Integration der Spelldatensätze im **Modul Erwerbsbiografie** sowie die notwendigen Aufbereitungsschritte wurden gemeinsam mit dem IAB diskutiert und beschlossen. Das Vorgehen wird im Kapitel 5.7 dokumentiert.

¹¹ Jesske & Schulz. (2018) im Erscheinen

2 Kennwerte

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

- Sample I: Teilstichprobe 1 (BA-Stichprobe) bezeichnet dabei im Folgenden die Leistungsempfängerstichprobe aus den Prozessdaten der Bundesagentur für Arbeit.
- Sample II: Teilstichprobe 2 (MICROM-Stichprobe) verweist auf die geschichtete Bevölkerungsstichprobe.
- Sample III: Zugangsstichprobe 1 (BA-Stichprobe) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 1 und Welle 2.
- Sample IV: Zugangsstichprobe 2 (BA-Stichprobe) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 2 und Welle 3.
- Sample V: Zugangsstichprobe 3 (BA-Stichprobe) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 3 und Welle 4.
- Sample VI: Bestandsauffrischung/Aufstockung 1 (EWO-Stichprobe) benennt die Stichprobe aus den Zugängen der Einwohnermeldeamtsregister in 100 neuen PLZ-Bereichen (Welle 5).
- Sample VII: Bestandsauffrischung/Aufstockung 2 (BA-Stichprobe) benennt die Stichprobe aus den SGB II-Zugängen in 100 neuen PLZ-Bereichen (Welle 5).
- Sample VIII: Zugangsstichprobe 4 (BA-Stichprobe) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 4 und Welle 5.
- Sample IX: Zugangsstichprobe 5 (BA-Stichprobe) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 5 und Welle 6.
- Sample X: Zugangsstichprobe 6 (BA-Stichprobe) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 6 und Welle 7.
- Sample XI: Zugangsstichprobe 7 (BA-Stichprobe) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 7 und Welle 8.
- Sample XII: Zugangsstichprobe 8 (BA-Stichprobe) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 8 und Welle 9.
- Sample XIII: Zugangsstichprobe 9 (BA-Stichprobe) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 9 und Welle 10.
- Sample XIV: Zugangsstichprobe 10 (BA-Stichprobe syrische/irakische Haushalte) benennt die Stichprobe aus dem Oversampling syrischer/irakischer Haushalte.

- Sample XV: Bestandsauffrischung/Aufstockung 1 (EWO-Stichprobe) benennt die Stichprobe aus den Zugängen der Einwohnermeldeamtsregister in den PLZ-Bereichen der Welle 5 (Welle 11).
- Sample XVI: Zugangsstichprobe 11 (BA-Stichprobe) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 10 und Welle 11.
- Sample XVII: Zugangsstichprobe 12 (BA-Stichprobe syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/ irakischer Haushalte zwischen Welle 10 und Welle 11.

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 Welle 1, aber auch die erstbefragten Haushalte aus den Zugangsstichproben der Wellen 2 bis 10¹². Dabei muss berücksichtigt werden, dass nur solche erstbefragten Haushalte auch bei den Wiederholungsbefragungen zur Verfügung stehen, für welche die Panelbereitschaft¹³ 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 von PASS einen möglichst selektionsfreien Zugang zum Panel sicherstellen sollten, werden ausführlich im Methoden- und Feldbericht der Wellen 1 bis 11 beschrieben¹⁴.

PASS ist in der Welle 1 mit 12.794 realisierten Haushaltsinterviews gestartet, von denen 12.000 Haushalte ihre Panelbereitschaft erklärt haben. Die panelbereiten Haushalte der Welle 1 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 „Abgespaltene Haushalte“ näher beschrieben. 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

¹² Aus der Panelstichprobe im PASS wurde vor Welle 10 ein Teil der sog. „reinen Seniorenhaushalte“ verabschiedet. Aus den PASS-Haushalten, in denen in Welle 9 ausschließlich Personen über 67 Jahren lebten (reine Seniorenhaushalte) wurde zufällig die Hälfte ausgewählt und verabschiedet. Dies betraf insgesamt 420 Haushalte (s. dazu auch Datenreport der Welle 10 in Berg et al. (2017))

¹³ Die Panelbereitschaft wird durch den Haushaltsbevollmächtigten erteilt und zählt damit für alle Haushaltsmitglieder. Panelbereite Haushalte haben der Speicherung ihrer Adresse zum Zweck von Wiederholungsbefragungen im Rahmen der Studie zugestimmt.

¹⁴ Vgl. Hartmann et al. (2008); Büngeler et al. (2009); Büngeler et al. (2010); Jesske & Quandt (2011); Jesske & Schulz (2012); Jesske & Schulz (2013); Jesske & Schulz (2014); Jesske & Schulz (2015); Jesske et al. (2016); Jesske et al. (2017); Jesske & Schulz (2018 im Erscheinen)

Panelbrutto gezählt. Panelverluste entstehen ebenfalls dadurch, dass für einen Haushalt über zwei Wellen infolge kein Haushaltsinterview realisiert werden konnte.

Dieser Fall konnte erstmalig am Ende der Welle 3 eintreten und wirkt sich auf das Panelbrutto in den Wellen 4¹⁵ bis 11 aus. Die Bruttoeinsatzstichprobe, die im Feld für die Welle 11 von PASS bearbeitet wurde, umfasste insgesamt 9.418 Panelhaushalte. Dazu kommen erstbefragte Haushalte aus der üblichen Zugangsstichprobe (n=3.772, davon 1.325 syrische / irakische Haushalte) und den Splithaushalten, die sich in den Wellen 10¹⁶ (n=189) oder 11 (n=346) neu ergeben haben, sowie der zusätzlichen Bestandsauffrischung / Aufstockungsstichprobe der allgemeinen Bevölkerung (n=6.051).¹⁷

Die Fallzahlen für den Bruttostichprobenumfang der Panelhaushalte in den jeweiligen Erhebungswellen und Teilstichproben¹⁸ werden in → **Tabelle A1** ausgewiesen. In der Welle 11 konnte in 7.287 Haushalten der Panelstichprobe mindestens ein Interview geführt werden. Dazu kommen 484 erstbefragte Haushalte aus der üblichen Zugangsstichprobe, von denen 451 ihre Panelbereitschaft erklärt haben, sowie 466 Haushalte aus der Zugangsstichprobe syrischer/ irakischer Haushalte, von denen 456 panelbereit waren, sowie 1.183 erstbefragte der Bestandsauffrischung/ Aufstockungsstichprobe aus dem Einwohnermelderegister mit 1.112 panelbereiten Haushalten. Ebenfalls erstbefragte Haushalte der Welle 11 sind weitere 159 Splithaushalte, die sich aus den Teilstichproben der bisherigen Wellen 1 bis 10 ergeben und in den Haushalten der Panelstichprobe enthalten sind.

Den 9.420 realisierten Haushaltsinterviews in der Welle 11 entsprechen 13.703 realisierte Personeninterviews. Die Verteilung der befragten Personen über die Teilstichproben und die jeweiligen Erhebungswellen sind → **Tabelle A2** zu entnehmen.

Für Befragte ohne ausreichende Deutschkenntnisse wurden in den Wellen 1 bis 9 zwei weitere Interviewsprachen (Türkisch und Russisch) angeboten. Ab der Welle 10 wurde Arabisch als weitere Fremdsprache aufgenommen, damit auch syrische und irakische Haushalte befragt werden konnten. Türkisch als Erhebungssprache wurde ab der Welle 10 nicht mehr angeboten. → **Tabelle A3** gibt an, wie viele Haushalte bzw. Personen in den beiden Erhebungssprachen befragt wurden.

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

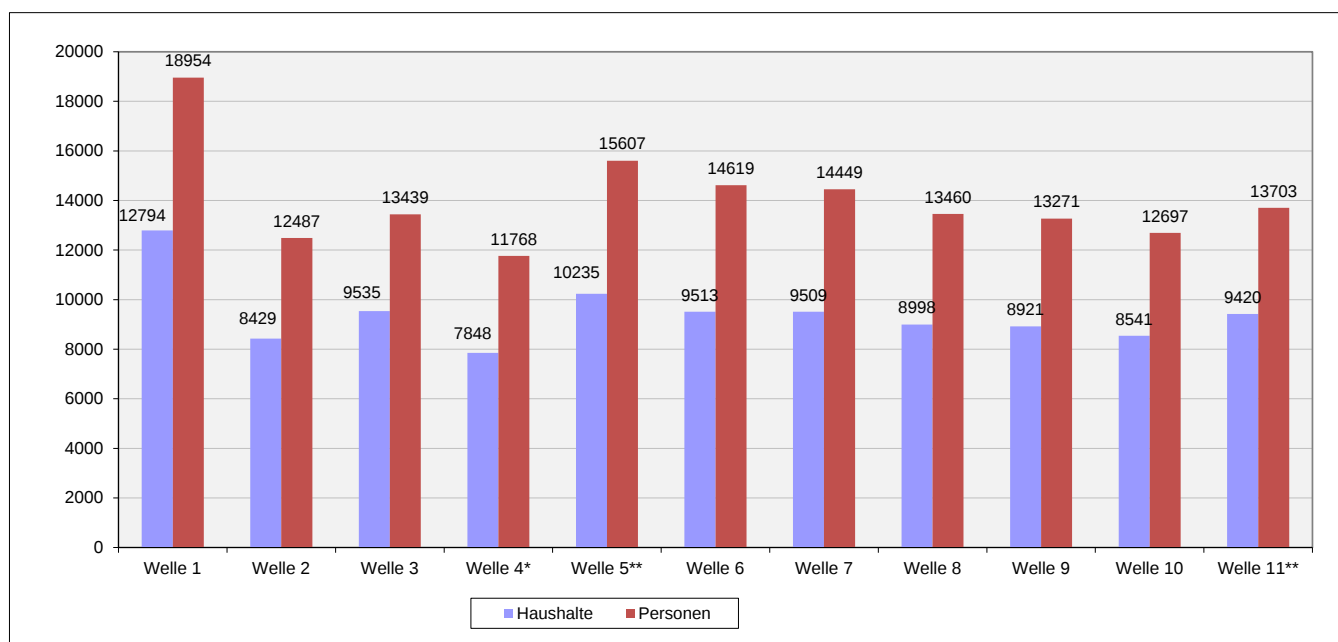
¹⁵ Zusätzlich auf das Panelbrutto in der Welle 4 wirkte sich der Wechsel des Erhebungsinstituts aus. Damit die Adressen der Panelteilnehmer vom IAB an infas übergeben werden konnten, war die Erlaubnis der Zielperson zur Weitergabe notwendig. Ausführlichere Erläuterungen zu diesem Verfahren und den Ergebnissen finden sich im Methodenbericht der Welle 4 (Jesske & Quandt, 2011).

¹⁶ Split-Haushalte, die in der Vorwelle nicht befragt werden konnten, werden als temporäre Ausfälle behandelt und in der Folgewelle versucht zu befragen. Fälle, die auch in der Folgewelle nicht realisiert werden konnten, werden als endgültige Ausfälle betrachtet.

¹⁷ Zahlen der Einsatzstichprobe s. Methodenbericht Welle 11 (Jesske & Schulz 2018, im Erscheinen).

¹⁸ Die Zahlen umfassen alle im Registerfile enthaltenen Fälle. Abweichungen zum Methodendatensatz sind aufgrund der nachträglichen Datenprüfungen und Bereinigungen möglich.

Abbildung 1: Realisierte Panelstichprobe Haushalte und Personen nach Erhebungswellen



2.2 Response-Raten

Die Berechnung der Response-Rate erfolgt nach AAPOR-Standards (AAPOR 2011). Berichtet wird die Ausschöpfungsquote RR1, in der auch alle Fälle ungeklärter Eligibilität in den Nenner eingehen und die somit den geringsten Wert aller Ausschöpfungsquoten annimmt¹⁹. 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 entweder alle Personen verstorben oder 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.

In → **Tabelle A4** sind die Ausschöpfungsquoten für die Welle 11 auf Haushaltsebene dargestellt.

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

¹⁹ 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.

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.

Die durchschnittlichen Realisierungsraten innerhalb der befragten Haushalte sind → **Tabelle A5** zu entnehmen.

Neben den Response-Raten auf der Haushaltsebene und innerhalb der Haushalte sind in → **Tabelle A6** 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.

2.3 Panelbereitschaft, Zuspielbereitschaft, Verknüpfung mit Prozessdaten

Zur Speicherung der Adressen zum Zweck 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 Haushalte²⁰ im PASS sind in → **Tabelle A7** abgebildet.

Die Panelbereitschaft wird bei erstmals befragten Haushalten in jeder Welle im Anschluss an das erste Personeninterview erfasst. Die Angabe dieser Person wird dann auf den Haushalt übertraübertragen. Ist die Person panelbereit, wird der Haushalt als panelbereit gewertet. Ist die Person nicht panelbereit, wird der Haushalt als nicht panelbereit gewertet (s. dazu auch Kap. 2.1)²¹. 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 Nachfrage, da sie im Seniorenfragebogen nicht enthalten ist. Die Zuspielbereitschaft wird dabei nicht in jeder Welle neu erhoben²².

²⁰ Bei allen Haushalten der Welle 1 handelt es sich um erstmals befragte Haushalte. Ab Welle 2 handelt es sich nur bei Haushalten der jeweiligen Auffrischungstichproben (in Welle 5 Haushalte der Aufstockungstichprobe) um erstmals befragte Haushalte sowie bei den erstmals teilnehmenden Split-Haushalten. Ab Welle 2 sind erstmals 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.

²¹ 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 ab Welle 2 für alle Folgewellen verfahren. 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.

²² Aufgrund von Umstellungen bei der Filterung konnte es vorkommen, dass in Welle 2 und 3 die Frage nach

→ **Tabelle A8** 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.

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. Auch Personen, die aus den Befragungshaushalten ausziehen bzw. ein Jahr oder länger nicht im Haushalt leben, sollen weiterhin befragt werden. Ihre neuen Haushalte werden als Abspaltungen des ursprünglichen Befragungshaushalts verstanden. Diese abgespaltenen Haushaltsteile (oder Splithaushalte) werden zu Befragungshaushalten des PASS. Alle in ihnen lebenden Personen ab 15 Jahren 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.

Aus den Befragungen der Welle 1 bis 11 liegen insgesamt 1.367 abgespaltene Haushalte vor, von denen in Welle 11 genau 670 befragt werden konnten. Darunter befanden sich 98 neue Splithaushalte aus Welle 11 sowie 61 erstbefragte Splithaushalte, die bereits in Welle 10 identifiziert werden konnten. Weitere Informationen zu den Splithaushalten finden sich im Methodenbericht der Erhebungswelle 11 (Jesske & Schulz 2018, im Erscheinen).

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.

der Zuspielbereitschaft erneut gestellt wurde, sofern die befragte Person ihre Zuspielbereitschaft in einer der Vorwellen noch nicht erklärt hatte. Ab Welle 6 erfolgt eine erneute Abfrage der Zuspielbereitschaft bei allen Personen, die in der Vorwelle ihre Bereitschaft nicht erteilt haben. Bei erneuter Verweigerung erfolgt dann im weiteren Verlauf keine weitere Abfrage.

3 Datensatzstruktur

Die übliche Struktur für die Aufbereitung eines Paneldatensatzes, wie sie z.B. in Erhebungen wie dem Sozio-oekonomischen Panel (SOEP) oder der 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 Konstrukte handelt. Für typische Längsschnittanalysen muss der erstellte, wellenübergreifende Datensatz schließlich in das 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.

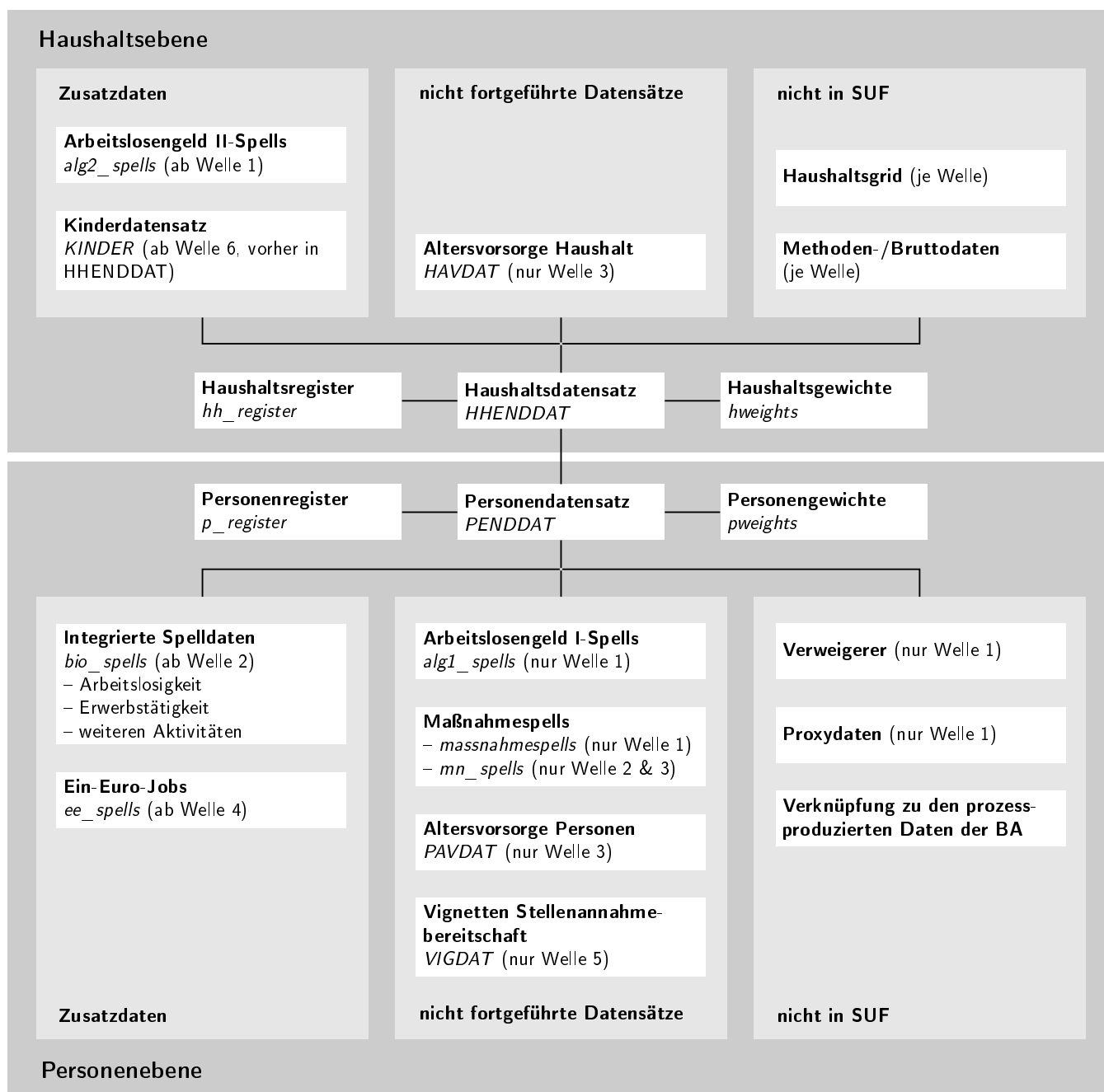
Auch die Aufbereitung im Long-Format ist mit Vor- und Nachteilen verbunden. Der entscheidende Vorteil dieser Variante ist, dass die Daten bereits in der für viele Längsschnittanalysen (wie z.B. Verweildaueranalysen) benötigten Struktur vorliegen. Der zusätzliche Aufwand für die Erstellung eines wellenübergreifenden 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 im Long-Format 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 Nutzerinnen und Nutzer überwiegen u. E. deutlich die Vorteile. Entsprechend wurden der Haushalts- und der Personendatensatz des PASS (*HHENDDAT*; *PENDDAT*) sowie die entsprechenden Gewichtungsdaten (*hweights*; *pweights*) und der ab Welle 6 neu hinzugekommene Kinderdatensatz (*KINDER*) im Long-Format aufbereitet.

Weiterhin enthält der Scientific Use File auf Haushaltsebene die in Spellform aufbereiteten Daten zum Arbeitslosengeld II-Bezug des Haushalts (*alg2_spells*). Auf Personenebene existiert ab Welle 4 ein integrierter biografischer Spelldatensatz (*bio_spells*), dieser integriert und ersetzt die bis Welle 3 vorliegenden Spelldatensätze *et_spells*, *al_spells* und *lu_spells*. Darüber hinaus liegt ab Welle 4 ein Ein-Euro-Spelldatensatz vor (*ee_spells*). Das Haushalts- und das Personenregister (*hh_register*; *p_register*) liegen im Wide-Format vor. In der Welle 5 wurde der Scientific Use File auf der Personenebene um einen Datensatz für das „Vignetten“ Modul (*VIGDAT*) erweitert, der nur in jener Welle erhoben wurde. Zudem wurde er ab der Welle 6 auf der Haushaltsebene um einen Datensatz für im Haushalt lebende Kinder (*KINDER*) ergänzt, in dem Informationen aus dem Haushaltsdatensatz ausgelagert wurden. Detaillierte Informationen zur Struktur der einzelnen Datensätze finden sich im User-Guide von PASS (Fuchs 2013).

Abbildung 2: Datensatzstruktur des PASS in Welle 11



4 Generierte Variablen

4.1 Vercodung offener Angaben

4.1.1 Offene Restkategorien und offene Items

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 generiert, 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 wurde. Die Items zu Geburtsland, Staatsangehörigkeit und den Zuzugsländern von Eltern/Großeltern wurden zudem anonymisiert und tragen entsprechende Variablenamen²³. Die nachfolgenden beiden Tabellen geben einen Überblick über die offenen Angaben, die in der Welle 11 vercodet wurden²⁴.

²³ *ogebland* (Geburtsland); *ostaatan* (Staatsangehörigkeit); *ozulanda* bis *ozulandf* (Zuzugsländer Eltern/Großeltern).

²⁴ Variablen, für die zwar in Vorwellen offene Angaben erhoben und vercodet wurden, jedoch nicht mehr in der aktuellen Welle, sind (mit Ausnahme des ALG II-Spelldatensatzes) nicht aufgeführt. Diese Variablen sind für die Beobachtungen in den Wellen ohne Erhebung dieser Variablen mit -9 gefüllt (Item in Welle nicht erhoben) und in den Datenreports der Erhebungswellen dokumentiert.

Tabelle 1: Vercodung offener Angaben auf Haushaltsebene in Welle 11

fester Variablen- name	zugecodet zu Variable	Datensatz	Bezeichnung
<i>HD1100a-o</i>	<i>HD1101a-o</i>	<i>HHENDDAT</i>	Anderer Erwerbsstatus der HH-Mitglieder, ggf. Proxy-Angabe
<i>HW0880a-i</i>	<i>HW0881a-j</i>	<i>HHENDDAT</i>	Anderer, nicht in Liste enthaltener Auszugsgrund
<i>AL20550a-h</i>	<i>AL20551a-h</i>	<i>alg2_spells</i>	Anderer Grund für den Beginn des ALG2-Bezugs
<i>AL21300a-h</i> bis <i>AL22100a-h</i>	<i>AL21301a-h</i> <i>AL21401a-h</i> <i>AL21501a-h</i> <i>AL21601a-h</i> <i>AL21701a-h</i> <i>AL21801a-h</i> <i>AL21851a-h</i> <i>AL21901a-h</i> <i>AL22001a-h</i> <i>AL22101a-h</i> <i>AL22102a-h</i> <i>AL22103a-h</i>	<i>alg2_spells</i>	Anderer, nicht in Liste enthaltener Grund für Kürzung
<i>AL22200a–</i> <i>AL22200h</i>	<i>AL22201a-h</i>	<i>alg2_spells</i>	Anderer, nicht in Liste enthaltener Grund, warum kein ALG II mehr bezogen wurde

Tabelle 2: Vercodung offener Angaben auf Personenebene in Welle 11

fester Variablen- name	zugecodet zu Variable	Datensatz	Bezeichnung
<i>PB0230 (Code 6)</i>	<i>PB0231</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tener deutscher Schulabschluss (Fortschreibung)
<i>PB0230 (Code 7)</i>	<i>PB0231</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tener ausländischer Schulab- schluss (Fortschreibung)
<i>PB0400 (Code 9)</i>	<i>PB0401</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tener deutscher Schulabschluss (Ersterhebung o. keine Angabe in VW)
<i>PB0400 (Code 10)</i>	<i>PB0401</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tener ausländischer Schulab- schluss (Ersterhebung o. keine Angabe in VW)
<i>PB1000</i>	<i>PB1001</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tener ausländischer Schulab- schluss (Ersterhebung o. keine Angabe in VW)
<i>PB1300a-j (Item I)</i>	<i>PB1301a-j</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tener deutscher Ausbildungsab- schluss (Ersterhebung o. keine Angabe in VW)
<i>PB1300a-j (Item J)</i>	<i>PB1301a-j</i>	<i>PENDDAT</i>	Anderer, nicht in Liste ent- haltener ausländischer Aus- bildungsabschluss (Erster- hebung o. keine Angabe in VW)
<i>PB1600</i>	<i>PB1601</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tener Ausbildungsabschluss, dem der ausländische Ab- schluss entspricht
<i>PAA1100</i>	<i>PAA1100</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tener Grund, keinen beruflichen Abschluss anzustreben
<i>AL0600</i>	<i>AL0601</i>	<i>bio_spells</i>	Anderer, nicht in Liste enthal- tener Grund, weshalb nicht mehr arbeitslos gemeldet

Tabelle 2: Vercodung offener Angaben auf Personenebene in Welle 11 (Fortsetzung)

fester Variablen- name	zugecodet zu Variable	Datensatz	Bezeichnung
<i>BIO0100</i>	<i>BIO0101</i>	<i>bio_spells</i>	Anderer, nicht in Liste enthalte- ner Aktivitätstyp
<i>ET2400</i>	<i>ET2401</i>	<i>bio_spells</i>	Andere Art, von der Stelle zu erfahren
<i>ET2420</i>	<i>ET2421</i>	<i>bio_spells</i>	Über anderes soziales Netz- werk von Stelle erfahren
<i>ET4020</i>	<i>ET4021</i>	<i>bio_spells</i>	Anderes Beziehungsverhältnis zur bei der Jobfindung wichti- gen Person
<i>EE0300a-h</i>	<i>EE0301a-h</i>	<i>ee_spells</i>	Anderer Grund, warum keine Teilnahme am Ein-Euro-Job
<i>EE1000a-e</i>	<i>EE1001a-e</i>	<i>ee_spells</i>	Anderer Grund, warum Ein- Euro- Job vorzeitig beendet
<i>PTK0320a-g</i>	<i>PTK0321a-g</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthalte- ner Grund, weshalb keine Ar- beit gesucht werden muss
<i>PTK1700a-i</i>	<i>PTK1701a-i</i>	<i>PENDDAT</i>	Anderes Förderangebot des Jobcenters
<i>PTK1800a-e</i>	<i>PTK1801a-e</i>	<i>PENDDAT</i>	Andere Forderung des Jobcen- ters
<i>PAS0900a-g</i>	<i>PAS0901a-g</i> <i>PAS0901i</i>	<i>PENDDAT</i>	Andere, nicht in Liste enthalte- ne Stellen, bei denen sich ZP über offene Stellen informiert hat
<i>PSP0200</i>	<i>PSP0201</i>	<i>PENDDAT</i>	Anderes Betriebssystem auf dem Smartphone
<i>PG0950a-i</i>	<i>PAS0951a-i</i>	<i>PENDDAT</i>	Andere Art der Behinderung/ Einschränkung
<i>PG1300</i>	<i>PG1301</i>	<i>PENDDAT</i>	Andere, nicht in Liste enthalte- ne Krankenversicherung
<i>PG1300a-e</i>	<i>PG1301a-e</i>	<i>PENDDAT</i>	Andere private Pflegetätigkei- ten
<i>PP1400a-f</i>	<i>PP1401a-f</i>	<i>PENDDAT</i>	Bei der Pflege unterstützende Stellen
<i>PMI0200</i>	<i>ogebland</i>	<i>PENDDAT</i>	Anderes, nicht in Liste enthal- tenes Geburtsland
<i>PMI0500</i>	<i>ostaatan</i>	<i>PENDDAT</i>	Andere, nicht in Liste enthalte- ne Staatsangehörigkeit
<i>PMI1000a-f</i>	<i>ozulanda-f</i>	<i>PENDDAT</i>	Anderes, nicht in der Liste ent- haltenes Zuzugsland des El- ternteils/Großelternteils

Tabelle 2: Vercodung offener Angaben auf Personenebene in Welle 11 (Fortsetzung)

fester Variablen- name	zugecodet zu Variable	Datensatz	Bezeichnung
<i>PMI1700</i>	<i>PMI1701</i>	<i>PENDDAT</i>	Rechtliche Grundlage des Zu- zugs
<i>PMI3000</i>	<i>PMI3001</i>	<i>PENDDAT</i>	Anderer Grund, die Anerken- nung eines im Ausland erwor- benen beruflichen Abschlus- ses in Deutschland nicht zu be- antragen
<i>PSH0200</i> (Code 9)	<i>PSH0201</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthalte- ner deutscher Schulabschluss der Mutter
<i>PSH0200</i> (Code 10)	<i>PSH0201</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tender ausländischer Schulab- schluss der Mutter
<i>PSH0300a-i</i> (Co- de 7)	<i>PSH0301a-i</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthalte- ner deutscher Berufsabschluss der Mutter
<i>PSH0300a-i</i> (Co- de 8)	<i>PSH0301a-i</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tender ausländischer Berufsab- schluss der Mutter
<i>PSH0500</i> (Code 9)	<i>PSH0501</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthalte- ner deutscher Schulabschluss des Vaters
<i>PSH0500</i> (Code 10)	<i>PSH0501</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tender ausländischer Schulab- schluss des Vaters
<i>PSH0600a-i</i> (Co- de 7)	<i>PSH0601a-i</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthalte- ner deutscher Berufsabschluss des Vaters
<i>PSH0600a-i</i> (Co- de 8)	<i>PSH0601a-i</i>	<i>PENDDAT</i>	Anderer, nicht in Liste enthal- tender ausländischer Berufsab- schluss des Vaters

4.1.2 Berufs- und Branchencodierung

Berufscodierungen finden nach ISCO (ISCO-88/ISCO-08) und KldB (1992/2010) statt, Branchencodierungen nach WZ (2003/2008). Die Berufscodierungen setzen ein spezielles Wissen voraus, das den Codierern in Schulungen vermittelt wird. Die Schulungen finden unter Einsatz standardisierter Schulungsunterlagen statt. Die erste Schulung neuer Codie-

rer beinhaltet eine Präsentation, in der die Grundregeln des Codierens und der ISCO/KIdB-Codierung vermittelt werden, sowie die Codierung und Besprechung von ausgewählten Testfällen unterschiedlichen Schwierigkeitsgrades. Die Schulung umfasst eineinhalb Tage.

Sofern Codierer länger als ein halbes Jahr keine Berufscodierungen durchgeführt haben, erfolgt zu Beginn eines neuen Projekts eine Auffrischung der Codierregeln sowie ein einmaliger Abgleich aller Codierer untereinander. Hierfür werden mindestens 500 randomisierte Fälle von allen beteiligten Personen codiert und die Abweichungen analysiert. Anhand dieses Vorgehens können systematische Fehler einzelner Codierer vor dem Codierungsprozess ermittelt und besprochen werden.

Im laufenden Prozess werden zur Sicherung der Qualität neben der Schulung regelmäßig Qualitätskontrollen durchgeführt. Während des Codierprozesses erhalten die Codierer ein individuelles Feedback über eventuell auftretende Abweichungen. Hierfür werden für alle Codierer Fälle aufgelistet, bei welchen ein vorgeschlagener Code abgelehnt wurde. Zeichnen sich systematische Fehler ab, werden diese mit dem jeweiligen Codierer besprochen.

Die Berufs- und Branchencodierung wird in den folgenden Prozessschritten durchgeführt:

1. Vorbereitung der Codierunterlage

Für Berufscodierungen sollen nicht nur die offenen Angaben zum ausgeübten Beruf aus dem Interview, sondern auch Zusatzvariablen genutzt werden. Vor Beginn der Codierung stimmen die hauptverantwortlichen Mitarbeiter der Codierung mit den Mitarbeitern der Datenaufbereitung ab, welche zusätzlichen Informationen im Frageprogramm zur Verfügung stehen und zusammen mit den offenen Angaben zum ausgeübten Beruf an die Codierung übergeben werden.

Bei PASS werden folgende Zusatzvariablen aus den erhobenen Angaben generiert und zusammen mit den offenen Angaben zum ausgeübten Beruf als Codierliste im Excelformat an die Codierung übergeben:

Tabelle 3: Codeschema der im PASS genutzten Zusatzvariablen

Abkürzung	Titel
StiB_g	Grobe Einteilung der Stellung im Beruf
ang	Angestellter
arb	Arbeiter
bea	Beamter oder Richter
selbst_f	Selbstständig im freien Beruf
selbst_H/DL	Selbstständig Handel/Gew./Ind./Dienstl.
landw	Selbstständiger Landwirt
mith_f	Mithelfende/r Familienangehörige/r
sol	Berufssoldat
k.A.	keine Angabe
wn	weiß nicht
StiB_f	Feine Einteilung der Stellung im Beruf
xxHektar	Landwirt mit xx Hektar
xxMitarbeiter	Selbstständiger oder Akademischer freier Beruf mit xx Mitarbeitern
40	Beamter im einfachen Dienst
41	Beamter im mittleren Dienst
42	Beamter im gehobenen Dienst
43	Beamter im höheren Dienst
45	Mannschaftsgrad
46	Unteroffiziersgrad
47	Offiziersgrad bis einschl. Hauptmann
48	Offiziersgrad ab Major
51	Angestellte mit einfacher Tätigkeit
52	Angestellte mit qualifizierter Tätigkeit
53	Angestellte mit hochqualifizierter Tätigkeit/ Leitungsfunktion
54	Angestellte mit höheren leitungsaufgaben/ umfassenden Entscheidungsbefugnissen
60	Ungelernter Arbeiter
61	Angelernter Arbeiter
62	Gelernter Arbeiter/ Facharb./ Geselle
63	Vorarbeiter/ Kolonnenführer
64	Meister/ Polier
k.A.	keine Angabe
wn	weiß nicht
Aufs,x	Aufsichtsfunktion, Zahl der beaufsichtigten Mitarbeiter
Aufs,x	Aufsichtsfunktion, Zahl der beaufsichtigten Mitarbeiter
k.Aufs	keine Aufsichtsfunktion

Tabelle 3: Codeschema der im PASS genutzten Zusatzvariablen (Fortsetzung)

Abkürzung	Titel
BerSchul	Höchster Schulabschluss
(fa)Abi, Eos12	Allgemeine oder fachgebundene Hochschulreife, EOS 12. Klasse
Fabi	Fachhochschulreife
Real, Pos.10	Mittlere Reife, realschulabschluss bzw. POS 10. Klasse
Haupt, Pos.8/9	Haupt-/Volkshochschule bzw. POS 8./9. Klasse
Sonder	Sonder-/ Förderschule
and	anderer Abschluss
Ausl	ausländischer Schulabschluss
kAB	kein Schulabschluss
Schüler	noch Schüler an allgemeinbildender Schule
k.A.	keine Angabe
wn	weiß nicht
Aus	Ausbildungsabschluss (Mehrfachangaben möglich)
Anlern/Tfach.	Beruflich-betriebliche Anlernzeit mit Abschlusszeugnis, aber keine Lehre; Teilfacharbeiterabschluss; Berufliches Praktikum, Volontariat; Keinen beruflichen Ausbildungsabschluss
Le	Lehre oder betriebliche Ausbildung
Ges	Schule des Gesundheitswesens
BerAk	Abschluss an einer Berufsakademie
BerUFab	Abschluss einer Berufsschule
Meist/Tech	Meister, Techniker, o.ä.
Dipl (FH), BA (Uni,FH)	Diplom an einer Fachhochschule oder Bachelor an Universität/ Fachhochschule
Dipl (Uni), BA + MA (Uni)	Diplom an Universität, Bachelor und Master an Universität oder FH
Prom/Hab	Promotion, Habilitation
Schüler	noch Schüler an allgemeinbildender Schule
and	anderer Abschluss
Ausl	ausländischer Schulabschluss
kAB	kein Schulabschluss
k.A.	keine Angabe
wn	weiß nicht
ÖD	Öffentlicher Dienst
ÖD	im öffentlichen Dienst beschäftigt
nÖD	nicht im Öffentlichen Dienst beschäftigt

Die Codiervorlage umfasst neben der Codierliste weitere Informationen wie u.a. Regelungen zur Codevergabe bei nicht eindeutigen Merkmalsausprägungen in Form einer beständig wachsenden Fallsammlung. Diese wird aus allen im Institut umge-

setzten Berufscodierungen fortlaufend gefüllt. Das Internet kann ebenfalls für die Recherche von Tätigkeiten genutzt werden (z.B. berufenet der Agentur für Arbeit; Klassifikationsserver des Statistischen Bundesamts, ILO, Statistik Austria für ISCO08).

Zu Beginn eines Projektes erfolgt, wenn nötig, eine Anpassung der allgemeinen Codierregeln bzw. das Aufstellen von Sonderregeln für das jeweils spezifische Projekt in Abhängigkeit der zur Verfügung gestellten Daten oder Regelungen aus Vorwahlen des Projektes. Diese angepassten Codierregeln werden dokumentiert und den Codierern mitgeteilt.

Die Spaltenbelegung der Codierlisten ist projektübergreifend standardisiert und darauf ausgelegt, nicht nur das finale Ergebnis, sondern auch alle im Folgenden beschriebenen Arbeitsschritte dauerhaft zu dokumentieren. Dokumentiert werden neben den Codes der einzelnen Codierschritte und den Codierernummern der Codierer auch die ggf. abgegebenen Kommentare zu Schwierigkeiten im Codierprozess.

2. Erstcodierung

Die Erstcodierung ist ein Prozessschritt, der aus zwei Teilschritten besteht: Einer maschinellen Vorcodierung sowie einer manuellen Vercodung. Die Daten werden in ein elektronisches Codier-System importiert und dort anhand eines umfassenden Wörterbuches maschinell vorcodiert. Rund 50 Prozent der Fälle können auf diese Weise maschinell vercodet werden. Im Anschluss erfolgt eine kritische Prüfung der Fälle der maschinellen Vorcodierung auf inhaltliche Plausibilität. Alle übrigen Fälle (ca. 50 Prozent) werden im Rahmen der Erstcodierung rein manuell vercodet.

3. Zweitcodierung

Für alle Nennungen erfolgt eine blinde Zweitcodierung. Dabei sieht der Zweitcodierer das Ergebnis der Erstcodierung nicht, erhält jedoch einen formelgestützt erzeugten Hinweis in einer gesonderten Problem-Spalte, ob die vergebenen Codes übereinstimmen oder nicht. Bei einer Abweichung kann der Zweitcodierer seinen vergebenen Code noch einmal überdenken, überprüfen und ggf. korrigieren. Stimmen die beiden vergebenen Codes überein, so wird der Code mithilfe einer Funktion in die Entscheidungsspalte übertragen.

4. Drittcodierung

Codeabweichungen zwischen Erst- und Zweitcodierung werden von einem dritten Codierer entschieden. Problemfälle werden in Diskussionsrunden besprochen und entschieden. Stimmt der Drittcodierer der Erst- oder der Zweitcodierung eindeutig zu, da die andere Codierung eindeutig falsch ist, überträgt er den richtigen Code in die Entscheidungsspalte. Kann sich der Drittcodierer nicht für einen der beiden Codes entscheiden oder schlägt einen weiteren Code vor, so wird dies über eine

Excelfunktion in der Problemspalte gekennzeichnet. Damit ist der Fall in der Problemfallbesprechung zu thematisieren. Zudem kann eine Kommentarspalte verwendet werden, um die Entscheidung zu begründen.

5. Problemfallbesprechung

Die Codierer treffen sich regelmäßig für die Besprechung und Entscheidung der Problemfälle.

6. Last Check

Abschließend werden die Richtigkeit der Codes, die Einhaltung der wichtigsten Codierregeln und die Eingabe der Codes (z.B. Zahlendreher) durch die hauptverantwortlichen Mitarbeiter der Codierung geprüft.

4.2 Harmonisierung

Bei einigen Variablen gab es über die Wellen hinweg Veränderungen an den Erhebungsinstrumenten. Vor allem die Integration des Moduls „Erwerbsbiografie“ 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 Welle 1. 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 oder Veränderungen an Kategorien sowie an den befragten Gruppen ergeben. Es wurde damit versucht, diese soweit möglich über die Wellen hinweg zu vereinheitlichen, bevor auf ihrer Basis die Generierung durchgeführt wurde.

Eine Harmonisierung wurde bisher für 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.

Während die ausdrücklich harmonisierten Variablen – neben Veränderungen am Erhebungskonzept – auch Veränderungen an Kategorien und den befragten Gruppen über die

Tabelle 4: Harmonisierte Variablen im Personendatensatz (*PENDDAT*)

Variable	thematischer Bereich	Bezeichnung
<i>stibkiz</i>	Erwerbstätigkeit	Jetzige Stellung im Beruf, einf. Eint., harmonisiert (anonymisiert)

Wellen berücksichtigen, werden von einem zweiten Variablentyp Veränderungen an den befragten Gruppen nicht explizit berücksichtigt. Diese Variablen werden zwar für alle Wellen 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 den eigentlichen Harmonisierungen auf 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 ist²⁵.

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

²⁵ Bspw. wurden in Welle 1 andere Befragtengruppen nach einer Erwerbstätigkeit gefragt als in den Folgewellen. Entsprechend haben auch jeweils andere Gruppen Angaben zur StB, zur beruflichen Tätigkeit, zur Arbeitszeit, zur Befristung etc. gemacht.

Tabelle 5: Wellenübergreifend generierte, aber nicht vollständig harmonisierte Variablen im Personendatensatz (PENDDAT)

Variable	thematischer Bereich	Bezeichnung
<i>isco88</i>	Erwerbstätigkeit	Intern. Standard Classification of Occupations 88, aktuelle Erwerbstätig., gen.
<i>kldb1992</i>	Erwerbstätigkeit	Klassifizierung der Berufe 1992, aktuelle Erwerbstätigkeit
<i>azhpt2</i>	Erwerbstätigkeit	Akt. tatsächl. Arb.zeit Haupterwerbst. (ohne Minijobs, inkl. kat. Ang.), gen.
<i>azges2</i>	Erwerbstätigkeit	Akt. tatsächl. Arb.zeit insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.
<i>befrist</i>	Erwerbstätigkeit	jetzige Tätigkeit: Befristung? generiert (alle Wellen)
<i>mps</i>	Erwerbstätigkeit	Magnitude-Prestigeskala, aktuelle Ewt., gen.
<i>siops1</i>	Erwerbstätigkeit	Standard Intern. Occupational Prestige Scale (Basis ISCO88), aktuelle Ewt., gen.
<i>isei1</i>	Erwerbstätigkeit	International Socio-Economic Index (Basis ISCO88), aktuelle Ewt., gen.
<i>egp</i>	Erwerbstätigkeit	Klassenschema nach Erikson, Goldthorpe und Portocarrero (EGP), akt. Beruf, gen.
<i>esec</i>	Erwerbstätigkeit	European Socio-economic Classification (ESeC), akt. Beruf, gen.
<i>stib</i>	Erwerbstätigkeit	Stellung im Beruf, Kennziffer, aktuelle Erwerbstätigkeit, generiert
<i>netges</i>	Erwerbstätigkeit	Akt. Nettoeinkommen insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.
<i>alg1abez</i>	Leistungsbezug	aktueller ALG 1-Bezug, generiert
<i>aktmassn</i>	Maßnahmeteilnahme	Akt. Teiln. an einem v. Arbeitsamt finanz./ geförd. Programm, 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 Fragetexts im aktuellen Interview eingeblendet.

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 be-

stimmten Fragekomplex vorlagen²⁶. 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 abzugrenzen²⁷. An anderen Stellen, insbesondere bei der Fortschreibung von Episodeninformationen²⁸, 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 infolge 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 sind²⁹.

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 Querschnittsdatsätzen (*HHENDDAT*, *PENDDAT*) werden dagegen generierte Variablen erstellt, in denen die Informationen der Vorwelle mit den erhobenen Veränderungen kombiniert werden.

Die beiden nachfolgenden Tabellen geben 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.

²⁶ 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.

²⁷ 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?“

²⁸ 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.

²⁹ 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 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 6: Fortgeschriebene Vorwelleninformationen in Welle 11, Haushaltsfragebogen

Konstrukt	Fragenr.	Anmerkung	Fortschreibung in Var.
Wohnsituation		Wohnform, Art des Mietverhältnisses und die Art des Wohnheims während des Interviews fortgeschrieben	<i>HHENDDAT: HW0200 bis HW0400</i>
Haushaltsstruktur		Haushaltsgröße während des Interviews fortgeschrieben	<i>HHENDDAT: HA0100</i>
		Geschlecht der Personen im HH während des Interviews ggf. korrigiert	<i>HHENDDAT: HD0100a bis HD0100o</i>
		Alter der Personen im HH während des Interviews fortgeschrieben	<i>HHENDDAT: HD0200a bis HD0200o</i>
		Verwandtschaftsbeziehungen während des Interviews fortgeschrieben	nicht im SUF ausgeliefert
Wohnfläche der Wohnung in qm	HW1000	in generierter Variable fortgeschrieben	<i>HHENDDAT: wohnfl</i>
Arbeitslosengeld II-Bezug	Modul „Arbeitslosengeld 2“	in Arbeitslosengeld II - Spell-Datensatz fortgeschrieben	<i>alg2_spells: Variablen des Arbeitslosengeld-II-Spelldatensatzes</i>
		Information zum aktuellen Arbeitslosengeld- II - Bezug des HH	<i>HHENDDAT: alg2abez</i>
		Information zum Arbeitslosengeld II - Bezug der Bedarfsgemeinschaft	<i>p_register: bgbez11; bgbezb11</i>

Tabelle 7: Fortgeschriebene Vorwelleninformationen in Welle 11, Personenfragebogen

Konstrukt	Fragenr.	Anmerkung	Fortschreibung in Var.
höchster allgemeiner Schulabschluss	PB0220-PB1100	in generierter Variable fortgeschrieben	<i>PENDDAT</i> : <i>schul1</i> (ohne offene Ang.) <i>schul2</i> (mit offenen Ang.)
Jahr des höchsten Schulabschlusses	PB0410	in generierter Variable fortgeschrieben	<i>PENDDAT</i> : <i>schulabj</i>
Berufsabschluss	PB1200-PB1600	höchster Berufsabschluss in generierter Variable fortgeschrieben	<i>PENDDAT</i> : <i>beruf1</i> (ohne offene Ang.) <i>beruf2</i> (mit offenen Ang.)
Jahr des Berufsabschlusses	PB1310a-k	in generierter Variable fortgeschrieben	<i>berabj</i>
Zeiten von fortgeschriebenen Aktivitäten im BIO-Spell Datensatz	BIO0600z1, BIO0600z2, BIO0400z, BIO0500z	in BIO-Spell Datensatz für Aufsatzspells fortgeschrieben	<i>bio_spells</i> : <i>BIO0400</i> , <i>BIO0500</i> , <i>BIO0600</i>
		in BIO-Spell Datensatz für Aufsatzspells fortgeschrieben	<i>bio_spells</i> : <i>ET2300</i> , <i>ET2700</i>
		Informationen zur aktuellen Erwerbstätigkeit in generierten Variablen fortgeschrieben	<i>PENDDAT</i> : <i>isco88</i> ; <i>isco08</i> ; <i>kldb1992</i> ; <i>kldb2010</i> ; <i>stib</i> ; <i>stibkz</i> ; <i>azhpt1</i> ; <i>azhpt2</i> ; <i>azges1</i> ; <i>azges2</i> ; <i>befrist</i> ; <i>mps</i> ; <i>siops1</i> ; <i>siops2</i> ; <i>isei1</i> ; <i>isei2</i> ; <i>egp</i> ; <i>esec</i> ; <i>branche1</i> ; <i>branche2</i>
		Informationen zum aktuellen (Nicht-) Erwerbstätigkeitsstatus in generierten Variablen fortgeschrieben	<i>PENDDAT</i> : <i>etakt</i> ; <i>alakt</i> ; <i>statakt</i>
Zeiten von Arbeitslosengeld-I-Bezug bei fortgeschriebenen Arbeitslosigkeits-Spells		Informationen zum aktuellen Arbeitslosengeld I - Bezug	<i>bio_spells</i> : <i>AL0700</i> , <i>AL0800</i> , <i>AL0900</i> , <i>AL1000</i> , <i>AL1100</i> , <i>AL1200</i>

Tabelle 7: Fortgeschriebene Vorwelleninformationen in Welle 11, Personenfragebogen (Fortsetzung)

Konstrukt	Fragenr.	Anmerkung	Fortschreibung in Var.
		in BIO-Spellldatensatz für Aufsatzspeils fortgeschrieben	<i>bio_spells:</i> AL0600, AL0601 <i>PENDDAT:</i> alg1abez
Zeiten von fortgeschriebenen Aktivitäten im Ein-Euro-Spellldatensatz			<i>ee_spells:</i> EE0800a, EE0800b
Angaben zum vorzeitigen Ende im Ein-Euro-Spellldatensatz			<i>ee_spells:</i> EE0900, EE1000a-EE1000e, EE1001a-EE1001e

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 Merkmalen 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, User Guide PASS Wave 6).

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 Vorwelleninformationen 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 nachfolgender Tabelle aufgeführt.

Tabelle 8: Typen der einfachen generierten Variablen in den Querschnittsdatensätzen (*HHENDDAT*; *PENDDAT*) bei Haushalten bzw. Personen, die schon in einer früheren Welle Angaben zum jeweiligen Thema gemacht haben

<i>Typ</i>	<i>Generierung auf Basis von Quelldaten aus: Welle der erstmaligen Erfassung des Themas bei HH/ Person</i>	<i>Generierung auf Basis von Quelldaten aus: akt. Welle</i>	<i>Beschreibung</i>
<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. Bsp: <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. Bsp: <i>schul1</i> (Höchster Schulabschluss)
<i>unabhängig (neu)</i>	nein	ja	Die Variable wird in jeder Welle unabhängig von Angaben aus der Vorwelle aus den Angaben der aktuellen Welle neu generiert. Bsp: <i>hhincome</i> (Haushaltsnettoeinkommen)

Zum Typ „unveränderlich (uv)“ muss mit Blick auf die einfachen Generierungen für *PENDDAT* eine vertiefende Erläuterung vorgenommen werden. Eine erstmalige Erfassung eines Themas findet bei einer Person u.U. nicht nur in der ersten Welle, in der die Person ein Personen-/Senioreninterview gibt, statt. Zwei Personengruppen werden erneut als erstbefragte Person behandelt, auch wenn sie zum wiederholten Mal ein Personen-/ Senioreneninterview geben.

Zum einen handelt es sich um Personen, die in einen Haushalt wiedereinziehen. Personen, die aus ihrem bisherigen Haushalt in einen Splithaushalt umziehen (siehe auch Kapitel 2.4), nehmen ihre Preloadinformationen mit. Somit können sie auch im Splithaushalt korrekt als erst- oder wiederzubefragende Personen gesteuert werden. Wenn eine Person hingegen aus einem Splithaushalt in einen in einer Vorwelle bewohnten Haushalt des Panels zurückzieht, findet keine Übernahme des Preloads für diese Person aus dem Splithaushalt in den Ursprungshaushalt statt. Die wiedereinziehenden Personen werden daher

wie neubefragte Personen behandelt. Dieser Sachverhalt tritt seit Welle 3 auf, da es in Welle 2 erstmals zu Auszügen aus wiederbefragten Haushalten und somit seit Welle 3 erstmals zu Wiedereinzügen von in einer Vorwelle ausgezogenen Personen kommen kann.

Zum anderen wird für eine Person nur dann ein personenbezogener Preload für das Dependent Interviewing (siehe Kapitel 4.3) erstellt, wenn sie in einer der beiden direkten Vorwellen ein Personen-/Senioreninterview gegeben hat. Hintergrund ist, dass eine Grenze gezogen werden soll, bis zu welchem Zeitpunkt sich eine Person an in Spellform erhobene Ereignisse zurückerinnern soll. Bei Personen, die ein Personen-/Senioreninterview letztmals in der dritten Vorgängerwelle oder früher gegeben haben, läge der Referenzzeitpunkt vor dem für erstbefragte Personen relevanten Datum. Um die Belastung der Zielperson zu begrenzen und in der Annahme, dass die Validität der erhobenen Informationen jenseits dieser Grenze zu stark bedroht ist, werden Personen, deren Referenzzeitpunkt für Auskünfte über Spellereignisse vor dem für erstbefragte Personen relevanten Datum liegt, wie erstbefragte Personen behandelt³⁰. Dieser Sachverhalt tritt seit Welle 4 auf, da hier erstmals ein früheres Personeninterview mehr als zwei Wellen zurückliegen kann.

Für diese beiden Personengruppen werden die den „unveränderlichen“ Generierungen zugrunde liegenden Informationen erneut erhoben (z.B. im Modul „Soziale Herkunft“), da diese Personen erneut als erstbefragte Personen gesteuert werden. Auch in der Datenaufbereitung werden die erhobenen Informationen genau wie die Angaben von tatsächlich zum ersten Mal im Rahmen des PASS befragten Personen weiterverarbeitet. Diese Generierungen, z.B. bei den Statusangaben zu Mutter und Vater, basieren also auf der aktuellen Welle. Eine Übernahme von Informationen aus der Vorwelle findet nicht statt, genauso wenig wie eine Plausibilitätsprüfung mit früheren Angaben. Zwar kann grundsätzlich angenommen werden, dass die Angaben der Zielpersonen, die zu den „unveränderlichen“ Generierungen weiterverarbeitet werden, im Falle der erneuten Erhebung konsistent gegenüber früheren Angaben sind. Inkonsistenzen und damit Abweichungen gegenüber Informationen aus den Vorwellen können jedoch nicht grundsätzlich ausgeschlossen werden. Personen, die zu einer der beiden beschriebenen Gruppen zählen, können im *PENDDAT* darüber identifiziert werden, dass sie in mehr als einer Welle in der Kennvariable *altbefr* als erstbefragt (Code „0“ bzw. für Welle 1 Code „-9“) geflaggt sind.

Die einfachen generierten Variablen sind in den folgenden sechs datensatzspezifischen Tabellen dargestellt. Die einzelnen Variablen werden dabei jeweils kurz beschrieben. Darüber hinaus werden die für die Generierung der Variable notwendigen Quellvariablen aufgeführt³¹. Für die Querschnittsdatsätze (*HHENDDAT*; *PENDDAT*) ist zudem jeweils an-

³⁰ Hiervon ausgenommen ist die Information, ob eine Person in einem früheren Interview bereits zu ihrer Zuspieldbereitschaft befragt worden ist. Diese Preloadinformation wird generiert, unabhängig davon, wie lange ein früheres Personeninterview zurückliegt. Es wird auf diese Weise vermieden, dass Personen, die in einer Vorwelle ihre Einwilligung gegeben haben, diese Frage (*RegP0100*) in einer folgenden Welle verneinen und ihre Einwilligung damit faktisch zurückziehen. Die Möglichkeit für die Zielperson, ihre Zuspieldbereitschaft zu widerrufen, bleibt von dieser Entscheidung unberührt.

³¹ Wie die Variablen in den Querschnittsdatsätzen (*HHENDDAT*; *PENDDAT*) für Beobachtungen der Vorwellen generiert wurden, ist in den jeweiligen Datenreports dokumentiert. Ebenfalls in der Dokumentation der jeweiligen Welle beschrieben ist die Generierung der wellenspezifischen Variablen in den Registerdatsätzen. Die generierten Variablen in den Spelldatsätzen wurden immer in den bereits fortgeschriebenen Datsätzen erzeugt. Wurde ein Spell nicht fortgeschrieben, blieben die entsprechenden generierten Va-

gegeben, um welchen Typ der in vorherigen Tabelle dargestellten einfachen 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 oder mit in der aktuellen Welle erhobenen Angaben fortgeschrieben wurde. Auch die Registerdatensätze folgen einer anderen Logik, so dass hier ebenfalls auf eine weitere Differenzierung verzichtet wurde.

riablen 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 zurückgegriffen, d.h. auf die mit den Informationen aus der aktuellen Welle gefüllten Variablen bzw. die für die aktuelle Welle relevanten Querschnittvariablen in den Spells.

Tabelle 9: Einfache Generierungen für Welle 11 im Haushaltsdatensatz (*HHEND-DAT*) und im *KINDER*-Datensatz (alphabetische Reihenfolge)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 10
<i>alg2abez</i>	<i>Aktueller ALG-II-Bezug, 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 (<i>AL22700</i>); <i>hintjahr</i> (<i>HHEND-DAT</i>)
<i>anzgeschw</i>	<i>Anzahl der Geschwister im Haushalt:</i> Gibt die Zahl der Geschwister einer Person an. Elternschaft und Geschwisterstatus werden unabhängig voneinander erhoben. Es kann Personen mit einem gemeinsamen Elternteil geben, die einander nicht als Geschwister bezeichnen. <i>anzgeschw</i> stimmt daher in seltenen Fällen nicht mit dem Geschwisterstatus überein, der im <i>p_register</i> über die Elternzeiger generiert werden kann (neu).	Informationen zu Beziehungen im Haushalt (<i>Haushaltsgrid</i>)
<i>bik</i>	<i>BIK-Regionsgrößenklassen (GKBIK10), generiert:</i> Die Information zur Regionsgrößenklasse wurde von infas generiert, indem von der in den Adressdaten vorliegenden Postleitzahl auf <i>GKBIK10</i> umgeschlüsselt wurde (neu).	vom Erhebungsinstitut geliefert
<i>blneuall</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 infas auf Basis der in den Adressdaten vorliegenden Postleitzahlen bestimmt (neu).	<i>bundesld</i> als die vom Erhebungsinstitut generierte und gelieferte Information zum Bundesland, in dem sich der Haushalt zum Befragungszeitpunkt befindet

Tabelle 9: Einfache Generierungen für Welle 11 im Haushaltsdatensatz (HHENDDAT) und im KINDER-Datensatz (alphabetische Reihenfolge) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>butaber</i>	<i>BuT-berechtigt zum Interviewzeitpunkt:</i> Ein Haushalt ist berechtigt zum Bezug von Leistungen des Bildungs- und Teilhabepakets (BuT), wenn er seit Januar des Jahres vor dem aktuellen Erhebungsjahr jemals mindestens eine der Leistungen ALG II, Kinderzuschlag, Wohngeld oder Sozialhilfe bezogen hat (neu).	AL20200; AL20400; AL20500 (<i>alg2_spells</i>); HA0250a-b; HW1800; HW1950; HEK0100; HEK0115; HEK1630; HEK1645 (HHENDDAT)
<i>hhinckat</i>	<i>Kategorisiertes Haushaltseinkommen pro Monat (in Euro), generiert:</i> Zusammenfassung der mehrstufig erhobenen kategorialen Angaben zum Haushaltseinkommen zu einer Variablen (neu).	HEK0700; HEK0800; HEK0900; HEK1000; HEK1100 (HHENDDAT)
<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).	HEK0600; HEK0700; HEK0800; HEK0900; HEK1000; HEK1100 (HHENDDAT)
<i>hintdat</i>	<i>Interviewdatum Haushaltsinterview:</i> Generierte Variable mit dem Datum, an dem das Haushaltsinterview geführt wurde im Format JJMMTT (neu).	<i>hintjahr; hintmon; hinttag</i> (HHENDDAT)
<i>hintnum</i>	<i>Interviewer Haushaltsinterview:</i> Die systemfreie Kennung kennzeichnet den Interviewer, der die Befragung durchgeführt hat. Die Information ist sowohl zwischen HHENDDAT und PEND-DAT als auch wellenübergreifend einheitlich vergeben worden. Eine bestimmte Ausprägung der Kennung bezeichnet also stets ein und den selben Interviewer (neu).	vom Erhebungsinstitut generierte und gelieferte Information
<i>kennungfbvers</i>	<i>Versionskennung des HH-Fragebogens Welle 11:</i> Im laufenden Feld der Welle 11 wurden etwa zwei bis drei Wochen nach Feldstart der Zugangsstichproben Änderungen im Fragebogen vorgenommen, die sich auf die Teilstichprobe der syrischen / irakischen Haushalte beziehen. Die Kennung markiert, ob ein Fall mit der ursprünglichen oder der überarbeiteten Fragebogenfassung befragt wurde. Eine genaue Darstellung der Änderungen im Fragebogen ist in Kapitel 1.3 enthalten (neu).	vom Erhebungsinstitut generierte und gelieferte Information

Tabelle 9: Einfache Generierungen für Welle 11 im Haushaltsdatensatz (HHENDDAT) und im KINDER-Datensatz (alphabetische Reihenfolge) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>kindu4</i>	<i>Steuervariable: Kind unter 4 Jahre im HH:</i> Die Variable zeigt an, dass in der Welle mind.eine Person im Haushalt unter vier Jahre alt ist. Da die Generierung nur auf den Altersangaben im Haushaltsdatensatz basiert, ist unerheblich, ob diese Person unter vier Jahren auch tatsächlich Kind einer anderen im Haushalt lebenden Person ist (neu).	<i>HD0200a - HD0200o (HHENDDAT)</i>
<i>kindu13</i>	<i>Steuervariable: Kind unter 13 Jahre im HH:</i> Die Variable zeigt an, dass in der Welle mind. eine Person im Haushalt unter 13 Jahre alt ist. Da die Generierung nur auf den Altersangaben im Haushaltsdatensatz basiert, ist unerheblich, ob diese Person unter 13 Jahren auch tatsächlich Kind einer anderen im Haushalt lebenden Person ist (neu).	<i>HD0200a - HD0200o (HHENDDAT)</i>
<i>kindu15</i>	<i>Steuervariable: Kind unter 15 Jahre im HH:</i> Die Variable zeigt an, dass in der Welle mind. eine Person im Haushalt unter 15 Jahre alt ist. Da die Generierung nur auf den Altersangaben im Haushaltsdatensatz basiert, ist unerheblich, ob diese Person unter 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; kategoriale Nachfrage zur Altersgruppe (bei KA in HD0200 (HHENDDAT))</i>
<i>kindu25</i>	<i>Steuervariable: Kinder unter 18 Jahre bzw. Schüler unter 25 Jahre im HH:</i> Die Variable zeigt an, dass in der Welle mind. eine Person im Haushalt unter 18 Jahre alt ist oder mind. eine Person zwischen 18 und 24 Jahren alt und Schüler ist. Da die Generierung nur auf den Altersangaben im Haushaltsdatensatz basiert, ist unerheblich, ob diese Person 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; kategoriale Nachfrage zur Altersgruppe (bei KA in HD0200); HD1100a-o (HHEND-DAT)</i>

Tabelle 9: Einfache Generierungen für Welle 11 im Haushaltsdatensatz (HHENDDAT) und im KINDER-Datensatz (alphabetische Reihenfolge) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<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 Welle 2 nur erhoben, wenn seit der Vorwelle ein Umzug oder eine Veränderung der Wohnung stattgefunden hatte (fs).	Bei Ersterfassung: HW1000 (HHEND-DAT) Bei wiederholter Erfassung: wohnfl aus Vorwelle; HW1000; (HHEND-DAT)

Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT) (alphabetische Reihenfolge)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>akt1euro</i>	<i>Akt. Teiln. an einem 1-Euro-Job, generiert:</i> <u>Indikator:</u> Befragungsperson nimmt zum Interviewzeitpunkt an einem 1-Euro-Job teil (neu).	<i>zensiert (ee_spells)</i>
<i>alakt</i>	<i>Aktuell arbeitslos gemeldet, generiert (ab Welle 2):</i> Zeigt an, dass die ZP zum Zeitpunkt des Personeninterviews der jeweiligen Welle arbeitslos gemeldet war (neu).	<i>zensiert; spintegr; BIO0101 (bio_spells)</i>
<i>alg1abez</i>	<i>Aktueller ALG I-Bezug, generiert:</i> <u>Indikator:</u> Befragungsperson bezieht zum Interviewzeitpunkt Arbeitslosengeld I. In der 11. Welle wurden die Zeiten seit Januar 2015 erhoben, in denen der Befragte arbeitslos gemeldet war. Für jede Episode wurde zudem erfragt, ob und in welchem Zeitraum Arbeitslosengeld I bezogen wurde (neu).	<i>AL0700; AL1000; AL1100; AL1200 (bio_spells)</i>
<i>apartner</i>	<i>Steuervariable: Unverh. Partner im HH:</i> <u>Indikator:</u> Befragungsperson hat einen unspezifischen oder unverheirateten Partner im HH (neu).	Informationen zu Beziehungen im Haushalt (Haushaltsgird); PD0500 - PD0800 (PENDDAT)
<i>azhpt1</i>	<i>Akt. vertragl. Arb.zeit Haupterwerbst. ohne Minijobs), gen.:</i> Wöchentliche, vertraglich vereinbarte Arbeitszeit der Haupterwerbstätigkeit, die der Befragte zum Interviewzeitpunkt ausübt, generiert aus offenen Angaben zur Arbeitszeit (neu).	<i>ET2009 (bio_spells)</i>

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>azhpt2</i>	<i>Akt. tatsächl. Arb.zeit Haupterwerbst. (ohne Minijobs, inkl. kat. Ang.), gen.:</i> Wöchentliche, tatsächliche Arbeitszeit der Haupterwerbstätigkeit, die der Befragte zum Interviewzeitpunkt ausübt, generiert aus offenen Angaben zur Arbeitszeit und kategorialer Nachfrage bei unregelmäßigen Arbeitszeiten (neu).	<i>ET2109; ET2209 (bio_spells)</i>
<i>azges1</i>	<i>Akt. vertragl. Arb.zeit insgesamt (ohne Minijobs), gen.:</i> Wöchentliche, vertraglich vereinbarte Arbeitszeit aller Erwerbstätigkeiten, die der Befragte zum Interviewzeitpunkt ausübt, generiert aus offenen Angaben zur Arbeitszeit (neu).	<i>ET2009 (bio_spells)</i>
<i>azges2</i>	<i>Akt. tatsächl. Arb.zeit insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.:</i> Wöchentliche, tatsächliche Arbeitszeit aller Erwerbstätigkeiten, die der Befragte zum Interviewzeitpunkt ausübt, generiert aus offenen Angaben zur Arbeitszeit und kategorialer Nachfrage bei unregelmäßigen Arbeitszeiten (neu).	<i>ET2109; ET2209 (bio_spells)</i>
<i>befrist</i>	<i>Jetzige Tätigkeit: Befristung? Generiert (alle Wellen):</i> <u>Indikator:</u> Die Erwerbstätigkeit, die der Befragte zum Interviewzeitpunkt ausübt, ist befristet (neu).	<i>PET2510a; PET2510b (PENDDAT)</i>
<i>begjeewt</i>	<i>Anfangsjahr der ersten Erwerbstätigkeit, generiert:</i> Jahr, in dem der Befragte erstmals hauptberuflich tätig war. Für die Generierung wurden Informationen zur ersten Erwerbstätigkeit mit Informationen aus den Erwerbstätigkeitsperioden kombiniert, wenn der Befragte die erste Erwerbstätigkeit bereits bei der Erhebung der Erwerbstätigkeitsperioden seit Januar 2015 genannt hat (uv).	<u>Bei Ersterfassung:</u> <i>bjahr (bio_spells); PET3200b (PENDDAT)</i> <u>Nach Ersterfassung:</u> <i>begjeewt</i> aus Vorwelle (PENDDAT)
<i>begminj</i>	<i>Anfangsjahr des aktuellen Minijobs, generiert:</i> Jahr, seit dem der Befragte in seinem aktuellen (Haupt-)Minijob tätig ist (neu).	<i>PMJ0800b</i>
<i>begmeewt</i>	<i>Anfangsmonat der ersten Erwerbstätigkeit, generiert:</i> Monat, in dem der Befragte erstmals hauptberuflich tätig war (Generierung siehe <i>begjeewt</i>) (uv).	<u>Bei Ersterfassung:</u> <i>bmonat (bio_spells); PET3200a (PENDDAT);</i> <u>Nach Ersterfassung:</u> <i>begmeewt</i> aus Vorwelle (PENDDAT)

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>begmminj</i>	<i>Anfangsmonat des aktuellen Minijobs, generiert: Monat, seit dem der Befragte in seinem aktuellen (Haupt-)Minijob tätig ist (neu).</i>	<i>PMJ0800a</i>
<i>berabj</i>	<i>Jahr des höchsten Berufsabschlusses: Jahr, in dem der Befragte seinen zum Interviewzeitpunkt höchsten Berufsabschluss erworben hat (fs) Anmerkung: Die Abschlussjahre der in der Welle 1 berichteten Berufsabschlüsse wurden in der Welle 2 nacherhoben.</i>	<u>Bei Ersterfassung:</u> <i>PB1310aj-kj (PENDDAT)</i> <u>Bei wiederholter Erfassung:</u> <i>berabj</i> aus Vorwelle; <i>PB1310aj-kj (PENDDAT)</i>
<i>beruf1</i>	<i>Höchster beruflicher Abschluss, ohne ausl. Abschl. und off. Angaben, generiert: Identifikation des zum jeweiligen Interviewzeitpunkt höchsten Berufsabschlusses durch Hierarchisierung der von den Befragungspersonen genannten beruflichen Abschlüsse, ohne offene Angaben (fs).</i>	<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: Wie beruf1 mit folgenden Unterschieden: 1. Einbeziehen von offenen Angaben; 2. Einbeziehen von Angaben zu ausländischen Abschlüssen; beruf2 aus Vorwelle; 3. Differenzierung der Studienabschlüsse nicht nach FH/Uni, sondern nach Abschlussniveau (Bachelor; Master/Diplom; Promotion) (fs).</i>	<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>

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>brges</i>	<i>Akt. Bruttoeinkommen insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.:</i> Enthält die über alle ETs (>450 Euro) kumulierte Angabe zum Bruttoeinkommen. Generiert aus offenen Angaben zum Bruttoeinkommen und kategorialer Nachfrage bei wn/ka- Angabe bei der offenen Abfrage (neu).	ET2806; ET2906; ET3006; ET3106; ET3206; ET3306 (<i>bio_spells</i>)
<i>brutto</i>	<i>Akt. Bruttoeinkommen Haupterwerbst (ohne Minijobs, inkl. kat. Ang.), gen.:</i> Generierung einer integrierten Variable aus kategorisierter und offener Angabe zum Bruttoeinkommen (neu).	ET2806; ET2906; ET3006; ET3106; ET3206; ET3306 (<i>bio_spells</i>)
<i>bruttokat</i>	<i>Akt. Bruttoeinkommen Haupterwerbst. (ohne Minijobs, kategorisiert), gen.:</i> Zusammenfassung der mehrstufig erhobenen kategorialen Bruttoeinkommensangaben zu einer Variablen (neu).	ET2806; ET2906; ET3006; ET3106; ET3206; ET3306 (<i>bio_spells</i>)
<i>emonlewt</i>	<i>Endzeitpunkt letzte Erwerbstätigkeit (Monat):</i> Monat, 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 2013 nicht erwerbstätig war (fs).	Bei Ersterfassung: PET1200b (PENDDAT); ejahr; emonat (<i>bio_spells</i>) Bei wiederholter Erfassung: ejhrlewt aus Vorwelle (PENDDAT); ejahr; emonat (<i>bio_spells</i>)
<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 2015 nicht erwerbstätig war (fs).	Bei Ersterfassung: PET1200b (PENDDAT); ejahr; emonat (<i>bio_spells</i>) Bei wiederholter Erfassung: ejhrlewt aus Vorwelle (PENDDAT) ejahr; emonat (<i>bio_spells</i>)

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>ekin1517</i>	<i>Steuervariable: eigenes Kind zwischen 15 und 17 Jahren im HH:</i> Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/ Adoptivkind im Alter zwischen 15 und 17 Jahre im Haushalt hat (neu).	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<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). In seltenen Haushaltskonstellationen kann es vorkommen, dass eine Person laut <i>ekind</i> im HH lebende Kinder hat, ihre <i>pnr</i> jedoch in den Zeigern <i>zmhh</i> und <i>zvhh</i> des <i>p_registers</i> nicht erscheint. Dies kann z.B. bei gleichgeschlechtlichen Beziehungen mit Kindern auftreten oder, wenn in einem HH neben dem aktuellen Partner auch noch der ehemalige Partner lebt.	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<i>ekin614</i>	<i>Steuervariable: eigenes Kind zwischen 6 und 14 Jahren im HH:</i> Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/ Adoptivkind im Alter zw. 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 Jahre 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 Jahre im Haushalt hat (neu).	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<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)
<i>etakt</i>	<i>Aktuell erwerbstaetig (>450 Euro pro Monat), generiert (ab Welle 2):</i> Zeigt an, dass die ZP zum Zeitpunkt des Personeninterviews der jeweiligen Welle einen andauernden ET-Spell (d.h. eine ET >450 Euro) hatte (neu).	<i>zensiert, spintegr, BIO0101 (bio_spells)</i>

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>famstand</i>	<i>Familienstand, generiert:</i> Generierung einer integrierten Familienstandsvariable aus den Angaben im Personenbogen und der aus dem Haushaltsdatensatz generierten Steuerungsvariable <i>epartner</i> (neu).	<i>epartner</i> ; PD0500; PD0700 (PENDDAT)
<i>gebhalbj</i>	<i>Geburtshalbjahr, generiert:</i> Indikator, ob Geburtstag im 1. oder 2. Halbjahr des Geburtsjahres liegt (neu).	Informationen zum Geburtsmonat
<i>kindzges</i>	<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).	Informationen zu Beziehungen im Haushalt (Haushaltsgrid) PD0900; PD1000; PD1100 (PENDDAT)
<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). <i>Anmerkung: Bei Verwendung der Variable sollte deren Personenbezogenheit beachtet werden. Dies bedeutet, dass ein Kind, das mit seinen Eltern in einem Haushalt lebt, sowohl beim Vater als auch bei der Mutter als „Kind im HH“ gezählt wird. Eine Aggregation dieser Variable über die Mitglieder des Haushalts ist also nicht sinnvoll.</i>	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<i>mberuf1</i>	<i>höchster ber. Abschluss der Mutter, inkl. Mut. im HH, ohne off. Ang., gen.:</i> Der Berufsabschluss der Mutter wurde in der Welle 1 nur dann erhoben, wenn diese nicht im Haushalt wohnte. Wohnte sie im Haushalt, wurde die Angabe zum Berufsabschluss aus ihrem Personeninterview ergänzt. Ab der Welle 2 wurde die Information zum Berufsabschluss der Mutter bei allen Neubefragten Personen erhoben, unabhängig davon, ob die Mutter im Haushalt lebte oder nicht. Ab der Welle 2 wurden für wiederholt Befragte die Werte aus der generierten Variable <i>mberuf1</i> aus der Vorwelle übernommen (uv).	Bei Ersterfassung: PSH0300a-i (PENDDAT) Nach Ersterfassung: <i>mberuf1</i> aus Vorwelle (PENDDAT)

Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>mberuf2</i>	<i>höchster ber. Abschluss der Mutter, inkl. Mut. im HH, off. 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).	Bei Ersterfassung: <i>PSH0301a-i</i> (PENDDAT) Nach Ersterfassung: <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)
<i>migration</i>	<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). <i>Anmerkung: Das Generierungskonzept der Variable wurde ab der Welle 2 überarbeitet. Statt nur die Information, ob die Befragungsperson in Deutschland geboren ist, und die Angaben, welche Vorfahren nach D. zugezogen sind, zur Generierung zu verwenden, wird nun auch die Information, ob und ggf. welcher Vorfahr außerhalb von D. geboren ist, einbezogen. Um eine konsistente Logik über die Wellen zu gewährleisten, wurde die Variable für Welle 1 ebenfalls neu generiert.</i>	Bei Ersterfassung: <i>PMI0100; PMI0700; PMI0800a-f; PMI0900a-f</i> (PENDDAT) Nach Ersterfassung: <i>migration</i> aus Vorwelle (PENDDAT)
<i>mschul1</i>	<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 wohnte. Wohnte sie im Haushalt, wurden die Angabe zum Schulabschluss aus ihrem Personeninterview ergänzt (uv). Ab der Welle 2 wurde die Information zum höchsten Schulabschluss der Mutter bei allen neubefragten Personen erhoben, unabhängig davon ob die Mutter im Haushalt lebte oder nicht.	Bei Ersterfassung: <i>PSH0200</i> (PENDDAT) Nach Ersterfassung: <i>mschul1</i> aus Vorwelle (PENDDAT)

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>mschul2</i>	<i>Höchster allg. Schulabschluss d. Mutter, inkl. Mut. im HH, inkl. off. Ang., gen.:</i> Wie <i>mschul1</i> mit dem Unterschied, dass bei der Generierung von <i>mschul2</i> auch die gemachten offenen Angaben berücksichtigt wurden (uv).	Bei Ersterfassung: <i>PSH0201 (PENDDAT)</i> Nach Ersterfassung: <i>mschul2</i> aus Vorwelle (<i>PENDDAT</i>)
<i>mstib</i>	<i>Stellung im Beruf der Mutter, Kennziffer, generiert:</i> Detaillierte Stellung der Mutter im Beruf, generiert aus den Einzelvariablen (uv).	Bei Ersterfassung: <i>PSH0320; PSH0330; PSH0340; PSH0360; PSH0370; PSH0380 (PENDDAT)</i> Nach Ersterfassung: <i>mstib (PENDDAT)</i>
<i>netges</i>	<i>Akt. Nettoeinkommen insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.:</i> Enthält die über alle ETs (>450 Euro) kumulierte Angabe zum Nettoeinkommen. Generiert aus offenen Angaben zum Nettoeinkommen und ab Welle 2 kategorialer Nachfrage bei wn/ka- Angabe bei der offenen Abfrage (neu).	<i>ET3406; ET3506; ET3606; ET3706; ET3806; ET3906 (bio_spells)</i>
<i>netto</i>	<i>Akt. Nettoeinkommen Haupterwerbst. (ohne Minijobs, inkl. kat. Ang.), gen.:</i> Generierung einer integrierten Variable aus kategorisierter und offener Angabe zum Nettoeinkommen (neu).	<i>ET3406; ET3506; ET3606; ET3706; ET3806; ET3906 (bio_spells)</i>
<i>nettokat</i>	<i>Akt. Nettoeinkommen Haupterwerbst. (ohne Minijobs, kategorisiert), gen.:</i> Zusammenfassung der mehrstufig erhobenen kategorialen Nettoeinkommensangaben zu einer Variablen (neu).	<i>ET3406; ET3506; ET3606; ET3706; ET3806; ET3906 (bio_spells)</i>
<i>ostaatansyr</i>	<i>Staatsangehörigkeit syr./irak. HH, inkl. off. Ang., Kategorien (anon.):</i> Anhand dieser Kennung kann ab Welle 10 für die Teilstichproben der syrischen und irakischen Haushalte betrachtet werden, ob eine Person die syrische oder eine andere Nationalität aufweist. Auf eine separate Ausweisung der Personen mit irakischer Nationalität musste wegen geringer Fallzahlen verzichtet werden .	Nicht anonymisierte Variante von <i>ostaatan</i> , <i>sample (PENDDAT)</i>

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>palter</i>	<i>Alter (Welle 1: gen. aus P1; ab Welle 2: beste Inf.), generiert:</i> Alter der Befragungsperson, generiert auf Basis der Angaben zum Geburtsdatum und des Datums des Personeninterviews in der der aktuellen Welle (neu).	<i>PD0100; pintjahr, pintmon, pinttag (PENDDAT)</i>
<i>panel</i>	<i>Panelbereitschaft:</i> (neu).	Vom Erhebungsinstitut gelieferte Information zur Panelbereitschaft der Haushalte
<i>pintdat</i>	<i>Interviewdatum Personeninterview:</i> Generierte Variable mit dem Datum, an dem das Personeninterview geführt wurde im Format JJMMTT (neu).	<i>pintjahr, pintmon, pinttag (PENDDAT)</i>
<i>pinnum</i>	<i>Interviewer Personeninterview:</i> Die systemfreie Kennung kennzeichnet den Interviewer, der die Befragung durchgeführt hat. Die Information ist sowohl zwischen <i>PENDDAT</i> und <i>HHENDDAT</i> als auch wellenübergreifend einheitlich vergeben worden. Eine bestimmte Ausprägung der Kennung bezeichnet also stets ein und den selben Interviewer (neu).	Vom Erhebungsinstitut generierte und gelieferte Information
<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).	Bei Ersterfassung: <i>PB0200; PB0220; PB0230; PB0300; PB0400 (PENDDAT)</i> Bei wiederholter Erfassung: <i>schul1</i> aus Vorwelle; <i>PB0200; PB0220; PB0230; PB0300; PB0400 (PENDDAT)</i>
<i>schul2</i>	<i>Höchster Schulabschluss, mit ausl. Abschl. und offene Angaben:</i> Wie <i>schul1</i> mit folgenden Unterschieden: 1. Einbeziehen von offenen Angaben; 2. Einbeziehen von Angaben zu ausländischen Abschlüssen (fs).	Bei Ersterfassung: <i>PB0200; PB0220; PB0231; PB0300; PB0401 (PENDDAT)</i> Bei wiederholter Erfassung: <i>schul2</i> aus Vorwelle; <i>PB0200; PB0220; PB0231; PB0300; PB0401 (PENDDAT)</i>

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>schulabj</i>	<i>Jahr des höchsten Schulabschlusses:</i> Jahr, in dem die Befragungsperson ihren höchsten Schulabschluss erworben hat (fs). <i>Anmerkung:</i> Wiederholt Befragte, für die bereits Informationen zum Schulabschluss aus einer Vorwelle vorliegen, wurden in der aktuellen Welle nicht zum Abschlussjahr befragt, 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 Welle 11 vor Mai 2017 geführt, wurde davon ausgegangen, dass der Abschluss in 2016 erworben wurde, bei später geführten Interviews wurde 2017 als Abschlussjahr angenommen.	Bei Ersterfassung: <i>PB0220; PB0230; PB0410; pintjahr; pintmon (PENDDAT)</i> Bei wiederholter Erfassung: <i>schulabj</i> aus Vorwelle; <i>PB0220; PB0230; PB0410; pintjahr; pintmon (PENDDAT)</i>
<i>statakt</i>	<i>Aktueller Hauptstatus, generiert (ab Welle 2):</i> Zeigt an, welchen Hauptstatus die ZP zum Zeitpunkt des Personeninterviews der jeweiligen Welle hatte (neu).	<i>zensiert; spintegr; BIO0101; azges2 (bio_spells)</i>
<i>stib</i>	<i>Stellung im Beruf, Kennziffer, aktuelle Erwerbstätigkeit, generiert:</i> Generierung der detaillierten Kennziffer der Stellung im Beruf aus den Einzelvariablen. Generierung der Variable mit Informationen aus dem Modul „Erwerbstätigkeit“ (<i>ET060*-ET120*</i>). Bei mehreren aktuell andauernden Erwerbstätigkeitsperioden wurde die Episode mit dem höheren Stundenumfang gewählt. Bei mehreren aktuell andauernden Episoden mit exakt gleichem Stundenumfang, diejenige die eher begonnen hat (neu).	<i>ET0609; ET0709; ET0809; ET0909; ET1009; ET1109; ET1209 (bio_spells)</i>
<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 Erwerbstätigkeitsperioden kombiniert, wenn der Befragte angab, die erste Erwerbstätigkeit bereits bei der Erhebung der Erwerbstätigkeitsperioden seit Januar 2015 genannt zu haben (uv).	Bei Ersterfassung: <i>PET3300; PET3400; PET3500; PET3600; PET3700; PET3800; PET3900 (PENDDAT)</i> <i>ET0609; ET0709; ET0809; ET0909; ET1009; ET1109; ET1209 (bio_spells)</i> Nach Ersterfassung: <i>stibewt</i> aus Vorwelle (PENDDAT)

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<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 2015 nicht erwerbstätig war (fs).	Bei Ersterfassung: <i>PET1210; PET1220; PET1230; PET1240; PET1250; PET1260; PET1270 (PENDDAT) ET0609; ET0709; ET0809; ET0909; ET1009; ET1109; ET1209 (bio_spells)</i> Bei wiederholter Erfassung: <i>stiblewt</i> aus Vorwelle (PENDDAT) <i>ET0609; ET0709; ET0809; ET0909; ET1009; ET1109; ET1209 (bio_spells)</i>
<i>vberuf1</i>	<i>Höchster ber. Abschluss des Vaters, inkl. Vater im HH, ohne off. Ang., gen.:</i> Generierung für den höchsten Berufsabschluss des Vaters analog zu <i>mberuf1</i> (uv).	Bei Ersterfassung: <i>PSH0600a-i (PENDDAT)</i> Nach Ersterfassung: <i>mberuf1</i> aus Vorwelle (PENDDAT)
<i>vberuf2</i>	<i>Höchster ber. Abschluss des Vaters, inkl. Vater im HH, mit off. Ang., gen.:</i> Generierung für den höchsten Berufsabschluss des Vaters (unter Berücksichtigung der offenen Angaben) analog zu <i>mberuf1</i> (uv).	Bei Ersterfassung: <i>PSH0601a-i (PENDDAT)</i> Nach Ersterfassung: <i>mberuf1</i> 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 <i>(Haushaltsgrid)</i>
<i>vschul1</i>	<i>Höchster allg. Schulabschluss d. Vaters, inkl. Vater im HH, o. off. Ang., gen.:</i> Generierung für den höchsten Schulabschluss des Vaters analog zu <i>mschul1</i> (uv).	Bei Ersterfassung: <i>PSH0500 (PENDDAT)</i> Nach Ersterfassung: <i>vschul1</i> aus Vorwelle (PENDDAT)

**Tabelle 10: Einfache Generierungen für Welle 11 im Personendatensatz (PENDDAT)
(alphabetische Reihenfolge) (Fortsetzung)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>vschul2</i>	<i>Höchster allg. Schulabschluss d. Vaters, inkl. Vater im HH, inkl. off. Ang., gen.:</i> Generierung für den höchsten Schulabschluss des Vaters (unter Berücksichtigung der offenen Angaben) analog zu <i>mschul2</i> (uv).	Bei Ersterfassung: <i>PSH0501</i> (PENDDAT) Nach Ersterfassung: <i>vschul2</i> aus Vorwelle (PENDDAT)
<i>vstib</i>	<i>Stellung im Beruf des Vaters, Kennziffer, generiert:</i> Detaillierte Stellung des Vaters im Beruf, generiert aus den Einzelvariablen (uv).	Bei Ersterfassung: <i>PSH0620; PSH0630; PSH0640; PSH0660; PSH0670; PSH0680</i> (PENDDAT) Nach Ersterfassung: <i>vstib</i> aus Vorwelle (PENDDAT)

**Tabelle 11: Einfache Generierungen für Welle 11 im Arbeitslosengeld-II-Datensatz
(*alg2_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz)**

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>bmonat</i>	<i>ALG2-Episode: Anfangsmonat, generiert:</i> Beginnmonat der ALG-II- 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 Welle 2 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 <u>Anmerkung:</u> siehe <i>bmonat</i>	<i>AL20200</i> <i>(alg2_spells)</i>
<i>emonat</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) hintmon</i> <i>(HHENDDAT)</i>
<i>ejahr</i>	<i>ALG2-Episode: Endjahr, generiert:</i> Endjahr der Arbeitslosengeld-II- Episode, in der bei rechtszensierten (d.h. zum Befragungszeitpunkt des Haushalts aktuell andauernden) Spells das Interviewjahr eingesetzt wurde. <u>Anmerkung:</u> siehe <i>bmonat</i>	<i>AL20400</i> <i>(alg2_spells) hintjahr</i> <i>(HHENDDAT)</i>

Tabelle 11: Einfache Generierungen für Welle 11 im Arbeitslosengeld-II-Datensatz (*alg2_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>alg2kbma- alg2kbmi</i>	ALG2: 1. Kürzung: Anfangsmonat, generiert bis ALG2: 9. Kürzung: Anfangsmonat, generiert: Beginnmonat der Arbeitslosengeld II- Kürzung, in der für die Generierung die Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden. <u>Anmerkung:</u> Die Arbeitslosengeld-II- Kürzungen sind in die Arbeitslosengeld-II- Bezugsspellis eingebettet. Über den Indikator am Ende der jeweiligen Variablen (a-i) können die Informationen zu den einzelnen Kürzungsepi-soden unterschieden werden. Die generierten Datums-variablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben des Befragten sind ab der Welle 2 in den Quellvariablen enthalten.	1. Kürzung: AL21000a (<i>alg2_spells</i>) bis 9. Kürzung: AL21000i (<i>alg2_spells</i>)
<i>alg2kbja- alg2kbji</i>	ALG2: 1. Kürzung: Anfangsjahr, generiert bis ALG2: 9. Kürzung: Anfangsjahr, generiert: Beginnjahr der Arbeitslosengeld II- Kürzung <u>Anmerkung:</u> siehe <i>alg2kma - alg2kbmi</i>	1. Kürzung: AL21100a (<i>alg2_spells</i>) bis 9. Kürzung: AL21100i (<i>alg2_spells</i>)
<i>alg2kema- alg2kemi</i>	ALG2: 1. Kürzung: Endmonat, generiert bis ALG2: 9. Kürzung: Endmonat, generiert: 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. <u>Anmerkung:</u> siehe <i>alg2kma - alg2kbmi</i>	1. Kürzung: <i>alg2kbma; alg2kbja; AL21200a; AL21201a; AL21202a</i> (<i>alg2_spells</i>) bis 9. Kürzung: <i>alg2kbmi; alg2kbji; AL21200i; AL21201i; AL21202i</i> (<i>alg2_spells</i>)
<i>alg2keja- alg2keji</i>	ALG2: 1. Kürzung: Endjahr, generiert bis ALG2: 9. Kürzung: Endjahr, generiert: 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. <u>Anmerkung:</u> siehe <i>alg2kma - alg2kbmi</i>	1. Kürzung: <i>alg2kbma; alg2kbja; AL21200a; AL21201a; AL21202a</i> (<i>alg2_spells</i>) bis 9. Kürzung: <i>alg2kbmi; alg2kbji; AL21200i; AL21201i; AL21202i</i> (<i>alg2_spells</i>)

Tabelle 11: Einfache Generierungen für Welle 11 im Arbeitslosengeld-II-Datensatz (*alg2_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
AL22150a- AL22150i	<i>ALG2: 1. Kürzung: Welchem HH-Mitglied wurde gekürzt, generiert bis ALG2: 9. 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 nur von Welle 2 bis Welle 4 erhoben wurden, ist die Frage in allen in der ersten Welle und ab der fünften Welle berichteten Kürzungen an allen 15 Stellen mit „I“ (Item in Welle nicht erhoben) gefüllt (s. u.). 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. <u>Codes:</u> 1 - dem HH-Mitglied wurde das ALG II gekürzt 2 - dem HH-Mitglied wurde das ALG II 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	Information, für welches Haushaltsmitglied in der jeweiligen Kürzungsepisode gekürzt wurde (nur bis Welle 4 erhoben)

Tabelle 11: Einfache Generierungen für Welle 11 im Arbeitslosengeld-II-Datensatz (*alg2_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>zensiert</i>	<i>ALG2-Episode: Spell dauert z. Zeitpkt. d. letzten HH-Int. an (Rechtszens.), gen.:</i> Der Zensierungsindikator gibt an, ob ein Spell zum Zeitpunkt des letzten Haushaltsinterviews noch andauerte. <u>Anmerkung:</u> Ein Spell gilt als zensiert, wenn eine der folgenden Bedingungen erfüllt ist: (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. (b) Ein wiederbefragter Haushalt gibt an, dass ein ALG-II-Spell auch zum Interviewdatum der Welle 11 noch andauert. Oder ein Enddatum wird genannt, das identisch mit dem Interviewdatum der Welle 11 ist und in der Nachfrage wird bestätigt, dass der Bezug aktuell noch andauert.	AL20300; AL20400, AL20500 (<i>alg2_spells</i>)

Tabelle 12: Einfache Generierungen für Welle 11 im BIO-Spelldatensatz (*bio_spells*)
(Reihenfolge analog zur Reihenfolge im Datensatz)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>bmonat</i>	<i>Alle Aktivitäten: Anfangsmonat, generiert</i> Beginnmonat der Tätigkeitsepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden. <i>Anmerkung: 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:</i> <i>21 Jahresanfang/Winter: Januar</i> <i>24 Frühjahr/Ostern: April</i> <i>27 Jahresmitte/Sommer: Juli</i> <i>30 Herbst: Oktober</i> <i>32 Jahresende: Dezember</i>	<i>BIO0200 (bio_spells)</i>
<i>bjahr</i>	<i>Alle Aktivitäten: Anfangsjahr, generiert</i> Beginnjahr der Tätigkeitsepisode <i>Anmerkung: siehe bmonat</i>	<i>BIO0300 (bio_spells)</i>
<i>emonat</i>	<i>Alle Aktivitäten: Endmonat, generiert</i> Endmonat der Tätigkeitsepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden und bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells der Intermonat eingesetzt wurde. <i>Anmerkung: siehe bmonat</i>	<i>BIO0400, BIO0600 (bio_spells); pintmon (PENDDAT)</i>
<i>ejahr</i>	<i>Alle Aktivitäten: Endjahr, generiert</i> Endjahr der Tätigkeitsepisode, in der bei rechtszensierten (d.h. zum Befragungszeitpunkt des Haushalts aktuell andauernden) Spells das Interviewjahr	<i>BIO0500, BIO0600 (bio_spells); pintjahr (PENDDAT)</i>

Tabelle 12: Einfache Generierungen für Welle 11 im BIO-Spellldatensatz (*bio_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
	eingesetzt wurde. <i>Anmerkung: siehe bmonat</i>	
<i>zensiert</i>	<i>Alle Aktivitäten: ggw. noch andauernder Spell (Rechtszensurierung)</i> Der Zensurierungsindikator gibt an, ob ein Spell zum Zeitpunkt des Personeninterviews in der letzten Be-frragungswelle noch andauerte, d.h. es handelt sich um einen rechtszensierten Spell. <i>Anmerkung: Ein Spell gilt als zensiert, wenn eine der beiden folgenden Bedingungen erfüllt ist:</i> <i>(a) Die Person gibt bezüglich des Enddatums des BIO-Spells an, zum Zeitpunkt des Interviews die Tätigkeit noch auszuüben.</i> <i>(b) Oder es wird ein Enddatum für die Tätigkeit genannt, welches identisch mit dem Interviewdatum ist und in der Nachfrage bestätigt, dass die Aktivität derzeit noch andauert.</i>	<i>BIO0400; BIO0500; BIO0600 (bio_spells)</i>
<i>stib</i>	<i>Stellung im Beruf, Kennziffer, generiert</i> Generierung der detaillierten Kennziffer der Stellung im Beruf aus den Einzelvariablen.	Erfassung von Spellinformationen in Welle 11 <i>ET0609; ET0709; ET0809; ET0909; ET1009; ET1109; ET1209 (bio_spells)</i> Ansonsten bleibt der Wert der Vorwelle bestehen.
<i>az1</i>	<i>Vertragl. Arbeitszeit, gen.</i>	Erfassung von Spellinformationen in Welle 11 <i>ET2009 (bio_spells)</i> Ansonsten bleibt der Wert der Vorwelle bestehen. Ausnahme: Wenn eine Tätigkeit bislang eine abhängige Beschäftigung war und die Stellung im Beruf nun in eine

Tabelle 12: Einfache Generierungen für Welle 11 im BIO-Spellldatensatz (*bio_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
		Selbständigkeit/Mithelfender Familienangehöriger, Keine Angabe oder Weiß nicht abgeändert wird, dann wird az1 mit Code -3 belegt.
<i>az2</i>	<i>Tatsächliche. 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.	Erfassung von Spellinformationen in Welle 11 <i>ET2109; ET2209 (bio_spells)</i> Ansonsten bleibt der Wert der Vorwelle bestehen.
<i>alg1bm</i>	<i>ALG1-Bezug: Anfangsmonat, generiert</i> Beginnmonat des Arbeitslosengeld-I-Bezugs, in der für die Generierung die Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden. <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. Umrechnung der Monatsangaben siehe <i>bmonat</i>.	<i>AL0800 (bio_spells)</i>
<i>alg1bj</i>	<i>ALG1-Bezug: Anfangsjahr, generiert</i> Beginnjahr des Arbeitslosengeld-I-Bezugs <u>Anmerkung:</u> siehe <i>alg1bm</i>	<i>AL0900 (bio_spells)</i>

Tabelle 12: Einfache Generierungen für Welle 11 im BIO-Spell Datensatz (*bio_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>alg1em</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>Anmerkung: siehe alg2kbma - alg2kbmi</i>	<i>AL1000; AL1200</i> <i>(bio_spells) pintmon</i> <i>(PENDDAT)</i>
<i>alg1ej</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>Anmerkung: siehe alg2kbma - alg2kbmi</i>	<i>AL1100; AL1200</i> <i>(bio_spells) pintjahr</i> <i>(PENDDAT)</i>
<i>alg1akt</i>	<i>ALG1-Bezug: ggw. noch andauernder Spell (Rechtszensierung)</i> Der Zensierungsindikator gibt an, ob der ALG-I-Bezug zum Zeitpunkt des Personeninterviews in der letzten Welle noch andauerte, d.h. es sich um einen rechtszensierten Spell handelt. <i>Anmerkung: Ein Spell gilt als zensiert, wenn eine der beiden folgenden Bedingungen erfüllt ist: Die Person gibt bezüglich des Enddatums des Arbeitslosengeld-I-Bezug-Spells an, zum Zeitpunkt des Interviews noch Arbeitslosengeld I zu erhalten. Oder es wird ein Enddatum genannt, welches identisch mit dem Interviewdatum ist und in der Nachfrage bestätigt, dass der Bezug derzeit noch andauert. Die Generierung basiert auf den plausibilitätsgeprüften, generierten Datumsvariablen.</i>	<i>emonat; ejahr;</i> <i>AL1000; AL1100;</i> <i>AL1200 (bio_spells)</i>

Tabelle 12: Einfache Generierungen für Welle 11 im BIO-Spelldatensatz (*bio_spells*) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>br</i>	<i>Bruttoeinkommen (inkl. kat. Ang.), generiert</i> Für andauernde Spells aus Welle 11 erfolgt die Berechnung aus den Befragungsdaten der aktuellen Welle. Für nicht andauernde oder nicht fortgeschriebene Spells wird in Welle 11 auf die Informationen aus Welle 10 zurückgegriffen.	<i>ET280*; ET290*; ET300*; ET310*; ET320*; ET330*</i> <i>(bio_spells)</i>
<i>net</i>	<i>Nettoeinkommen (inkl. kat. Ang.), generiert</i> Für andauernde Spells aus Welle 11 erfolgt die Berechnung aus den Befragungsdaten der aktuellen Welle. Für nicht andauernde oder nicht fortgeschriebene Spells wird in Welle 11 auf die Informationen aus Welle 10 zurückgegriffen.	<i>ET340*; ET350*; ET360*; ET370*; ET380*; ET390*</i> <i>(bio_spells)</i>

Tabelle 13: Einfache Generierungen für Welle 11 im Ein-Euro-Spellldatensatz (ee_ - spells) (Reihenfolge analog zur Reihenfolge im Datensatz)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>bmonat</i>	<i>Massnahme: Anfangsmonat, generiert</i> Beginnmonat der Massnahmenepisode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden. <i>Anmerkung: 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.</i> <i>Die Jahreszeitangaben wurden nach folgenden Regeln in Monatswerte recodiert:</i> <i>21 Jahresanfang/Winter: Januar</i> <i>24 Frühjahr/Ostern: April</i> <i>27 Jahresmitte/Sommer: Juli</i> <i>30 Herbst: Oktober</i> <i>32 Jahresende: Dezember</i>	<i>EE0600a (ee_spells)</i>
<i>bjahr</i>	<i>Massnahme: Anfangsjahr, generiert</i> Beginnjahr der Massnahmenepisode <i>Anmerkung: siehe bmonat</i>	<i>EE0600b (ee_spells)</i>
<i>emonat</i>	<i>Massnahme: Endmonat, generiert</i> Endmonat der Massnahmenepisode, 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. <i>Anmerkung: siehe bmonat</i>	<i>EE0600a; EE0600b; EE0700; EE0800a; EE0800b (ee_spells); pintmon, pintjahr (PENDDAT)</i>
<i>ejahr</i>	<i>Massnahme: Endjahr, generiert</i>	<i>EE0600a; EE0600b; EE0700; EE0800a;</i>

Tabelle 13: Einfache Generierungen für Welle 11 im Ein-Euro-Spell Datensatz (ee_spells) (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
	Endjahr der Massnahmenepisode, in der bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells das Interviewdatum eingesetzt wurde. <i>Anmerkung: siehe bmonat</i>	EE0800b (ee_spells)
zensiert	<i>Massnahme: ggw. 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 sich um einen rechtszensierten Spell handelt.	EE0700 (ee_spells)

Tabelle 14: Einfache Generierungen für Welle 11 im Personenregisterdatensatz (*p_register*) (alphabetische Reihenfolge)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
<i>alter11</i>	<i>Alter der Person in Welle 11 (2017)</i> Variable enthält die „beste“ verfügbare Altersinformation einer Person. Dabei handelt es sich entweder (a) um das aus dem Geburtsdatum, das in Welle 11 berichtet wurde, berechnete Alter oder (b), wenn kein Geburtsdatum aus Welle 11 vorliegt, um die Altersangabe aus dem Haushaltsinterview. Die Information aus <i>alter11</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 im Feld. 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 11 im Haushaltsdatensatz enthaltene „beste“ Altersinformation wurde in den Plausibilitätsprüfungen und bei der Generierung der Bedarfsgemeinschaften- und Haushaltstypen berücksichtigt.	<i>PD0100; pintjahr; pintmon; pinttag (PENDDAT); HD0200a bis HD0200o (HHENDDAT)</i>
<i>erwprox11</i>	<i>Erwerbsstatus laut HH-Interview in Welle 11 (2017)</i> Variable ist eine unveränderte Übernahme der <i>HD1101*</i> der aktuellen Welle aus <i>HHENDDAT</i>.	<i>HD1101*</i>
<i>kinddat11</i>	<i>Person im Kinderdatensatz enthalten für Welle 11 (2017)</i> Zeigt an, ob die Person in der aktuellen Welle im	<i>pnr (KINDER)</i>

Tabelle 14: Einfache Generierungen für Welle 11 im Personenregisterdatensatz (p_register) (alphabetische Reihenfolge) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
	KINDER-Datensatz enthalten ist. Alle Kinder im Alter von unter 15 Jahren sowie ab Welle 6 zusätzlich alle Haushaltsmitglieder im Alter von 16 bis unter 25 Jahren, für die Proxy-Angaben in den Modulen Soziale Teilhabe und Bildungs- und Teilhabepaket erhoben werden, sind im <i>KINDER</i> -Datensatz enthalten.	
<i>korrex</i>	<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.	HD0100a bis HD0100o aller Wellen (HENDDAT)
<i>lastint</i>	<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.	Personeninterviews aller Wellen <i>PENDDAT</i>
<i>neuj11</i>	<i>Jahr, seit dem P. neu im aktuellen HH ist, berichtet in Welle 11 (2017)</i> Variable zeigt an, seit welchem Jahr die Person neu zu dem Haushalt gehört, dessen Haushaltsmitglied sie in Welle 11 ist. <i>Anmerkung: Information zum Datum stammt aus dem Welle 11-Interview mit dem wiederbefragten Haushalt, in den die Person seit der Vorwelle zugezogen ist bzw. in dem sie geboren wurde</i>	Informationen zum Datum, seit dem eine Person neu im Haushalt ist. Erhoben im Haushaltsgrid
<i>neum11</i>	<i>Monat, seit dem P. neu im aktuellen HH ist, berichtet in Welle 11 (2017)</i> Variable zeigt an, seit welchem Monat die Person neu zu dem Haushalt gehört, dessen Haushaltsmitglied sie in Welle 11 ist. <i>Anmerkung: siehe neuj11</i>	Informationen zum Datum, seit dem eine Person neu im Haushalt ist. Erhoben im Haushaltsgrid
<i>wegj11</i>	<i>Jahr, seit dem P. nicht mehr im vorherigen HH ist, ber. in Welle 11 (2017)</i>	Informationen zum Datum, seit dem

Tabelle 14: Einfache Generierungen für Welle 11 im Personenregisterdatensatz (*p_register*) (alphabetische Reihenfolge) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
	Variable zeigt an, seit welchem Jahr die Person nicht mehr zum Haushalt der Vorwelle gehört. <i>Anmerkung: Information zum Datum stammt aus dem Welle 11 Interview mit dem Haushalt, in dem die Person in der Vorwelle gelebt hat.</i>	eine Person nicht mehr im Haushalt ist. Erhoben im Haushaltsgrid
<i>wegm11</i>	<i>Monat, seit dem P. nicht mehr im vorherigen HH ist, ber. in Welle 11 (2017)</i> Variable zeigt an, seit welchem Jahr die Person nicht mehr zum Haushalt der Vorwelle gehört. <i>Anmerkung: siehe wegj11</i>	Informationen zum Datum, seit dem eine Person nicht mehr im Haushalt ist. Erhoben im Haushaltsgrid
<i>zdub11</i>	<i>Zeiger: Personennr. der durch die ZP gedoppelten Person in Welle 11 (2017)</i> Zeigt an, dass eine aus einem Ur-HH bekannte Person aktuell in einem Split-HH dieses Ur-HH lebt, ohne dass der Ur-HH den Auszug dieser Person gemeldet hat. Für Matchvorgänge mit dem <i>p_register</i> über die Personennummer muss zunächst eine Matchvariable generiert werden, die <i>zdub*</i> entspricht, sofern diese größer 0 ist und ansonsten der <i>pnr</i>. <i>Anmerkung: Eine detaillierte Erläuterung zu den Gründen der Einführung dieser Variable findet sich in Kapitel 5.4.1.2 des Datenreports zur Welle 5 (vgl. Berg et.al, FDZ Datenreport 06/2012).</i>	Informationen zu allen Haushaltsmitgliedern eines Ur-HH und aller seiner Split-HH im Haushaltsgrid der aktuellen und der Vorwelle
<i>zmhh11</i>	<i>Zeiger: Personnummer der Mutter der Zielperson im HH in Welle 11 (2017)</i> Identifiziert die Personnummer der Mütter, sofern sie im Haushalt lebt. Als Mutter gelten: leibliche Mütter, Stiefmütter, Adoptivmütter oder Pflegemütter sowie Mütter mit nicht weiter spezifiziertem Status.	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
<i>zparth11</i>	<i>Zeiger: Personnummer des Partners der Zielperson im HH in Welle 11 (2017)</i> Identifiziert die Personnummer eines im Haushalt	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)

Tabelle 14: Einfache Generierungen für Welle 11 im Personenregisterdatensatz (p_register) (alphabetische Reihenfolge) (Fortsetzung)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 11
	lebenden Partners. Als Partner gelten: Ehepartner, eingetragene Lebenspartner, unverheiratete Partner und Partner mit nicht weiter spezifiziertem Status.	
<i>zupanel</i>	<i>Befragungswelle des Zugangs der Person zum Panel</i> Variable gibt an, in welcher Welle die Person erstmals Mitglied eines Befragungshaushaltes war.	Informationen zu den im Haushalt lebenden Personen aller Wellen (Haushaltsgrid)
<i>zvhh11</i>	<i>Zeiger: Personenummer des Vaters der Zielperson im HH in Welle 11 (2017)</i> Identifiziert die Personenummer des Vaters, sofern er im Haushalt lebt. Als Vater gelten: leibliche Väter, Stiefväter, Adoptivväter oder Pflegeväter sowie Väter mit nicht weiter spezifiziertem Status.	Informationen zu Beziehungen im Haushalt (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 11

	PENDDAT							BIO-Spells	EE_Spells
	Aktueller Status	Erwerbshistorie		Soziale Herkunft		Minijob	Ausbildungsaspiration	Erwerbs- und Arbeitslosigkeitsbiografie	Ein-Euro-Job-Teilnahme
		letzte ET	erste ET	Mutter	Vater				
Bildung	berabj								
	beruf1			mberuf1	vberuf1				
	beruf2			mberuf2	vberuf2				
	schulabj								
	schul1			mschul1	vschul1				
	schul2			mschul2	vschul2				
Bildungs-klassifikation	casmin			mcasmin	vcasmin				
	isc97			misc97	visc97				
	bilzeit			mbilzeit	vbilzeit				
Informationen zum aktuellen Status	akt1euro								
	alakt								
	etakt								
	statakt							spelltyp	
Sozio-ökonomische Position	egp	egplewt	egpeewt	megp	vegp			egp	
	esec	eseclewt	eseceewt	mesec	vesec			esec	
	isei1	iseilewt1	iseieewt1	misei1	visei1			isei1	
	isei2	iseilewt2	iseieewt2	misei2	visei2			isei2	
	mps	mpslewt	mpseewt	mmps	vmps			mps	
	siops1	siopslewt1	siopseewt1	msiops1	vsiops1			siops1	
	siops2	siopslewt2	siopseewt2	msiops2	vsiops2			siops2	
	stib	stiblewt	stibeewt	mstib	vstib			stib	
Stellung im Beruf	stibkz								
			begmeewt			begmminj		bmonat	bmonat
			begjeewt			begjminj		bjahr	bjahr
		emonlewt						emonat	emonat
Datierung der Erwerbstätigkeit		ejhrlewt						ejahr	ejahr
								alg1bm	
								alg1bj	
								alg1em	
Datierung der Arbeitslosigkeit								alg1ej	
Informationen zur Erwerbstätigkeit	befrist								
	azhpt1							az1	
	azhpt2							az2	
	azges1								
	azges2								
Berufliche Tätigkeit	isco88	isco88lewt	isco88eewt	misco88	visco88	isco88minj		isco88	
	isco08	isco08lewt	isco08eewt	misco08	visco08	isco08minj	isco08berufswunsch	isco08	
							isco08angberuf		
	kldb1992	kldb1992lewt	kldb1992eewt	mkldb1992	vkldb1992	kldb1992minj		kldb1992	
	kldb2010	kldb2010lewt	kldb2010eewt	mkldb2010	vkldb2010	kldb2010minj	kldb2010berufswunsch	kldb2010	
							kldb2010angberuf		
Tätig in Branche	branche1					brancheminj1		branche1	
	branche2					brancheminj2		branche2	

	PENDDAT						BIO-Spells	EE_Spells	
	Aktueller Status	Erwerbshistorie		Soziale Herkunft		Minijob	Ausbildungsaspiration	Erwerbs- und Arbeitslosigkeitsbiografie	Ein-Euro-Job-Teilnahme
		letzte ET	erste ET	Mutter	Vater				
Einkommen	netges								
	brges								
	netto								
	nettokat								
	brutto								
	bruttokat								
Leistungsbezug	alg1abez						alg1akt		
Haushaltskontext und Familienstand	hhgr								
	famstand								
	vhh								
	mhh								
	apartner								
	epartner								
	ekind								
	ekin614								
	ekinu15								
	ekinu18								
	ekin1517								
	kindzges								
	kindzihh								
	Migrationshintergrund	ogebland							
ostaatan									
ozulanda									
ozulandb									
ozulandc									
ozulandd									
ozulande									
ozulandf									
migration									
Informationen zur Person	gebhalbj								
	palter								
	zpalthh								
	zpsex								
Allgemein	altbefr								
	fb_vers								
	panel								
	pintdat								
	RegP0100								
	sample								
Freizeitverhalten	freiz1								
	freiz2								
	freiz3								
	frwunsch								

4.5 Konstruktvariablen

Unter Konstruktvariablen fallen diejenigen Variablen, für deren Generierungen umfangreiche 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

Tabelle 15: Bildung in Jahren

Variablenname:	<i>bilzeit</i>
Variablenlabel:	Schul- und Berufsbildung in Jahren, generiert
Ausgangsvariablen:	<i>schul2; beruf2</i>
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	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: 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 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.

Tabelle 15: Bildung in Jahren (Fortsetzung)

	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)). 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. 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
Literatur:	Helberger (1988)

Tabelle 16: Bildung in Jahren, Mutter

Variablenname:	<i>mbilzeit</i>
Variablenlabel:	Schul- und Berufsbildung der Mutter in Jahren, generiert
Ausgangsvariablen:	<i>mschul2; mberuf2</i>
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	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

Tabelle 16: Bildung in Jahren, Mutter (Fortsetzung)

	Universität: +5 Jahre Anderer Abschluss, D: +1,5 Jahre Anderer Abschluss, Ausl.: +1,5 Jahre
Literatur:	Helberger (1988)

Tabelle 17: Bildung in Jahren, Vater

Variablenname:	<i>vbilzeit</i>
Variablenlabel:	Schul- und Berufsbildung des Vaters in Jahren, generiert
AusgangsvARIABLEN:	<i>vschul2; vberuf2</i>
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	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
Literatur:	Helberger (1988)

Tabelle 18: CASMIN

Variablenname:	<i>casmin</i>
Variablenlabel:	Bildung klassifiziert nach CASMIN, aktual. Version, generiert
AusgangsvARIABLEN:	<i>schul2; beruf2</i>
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph

Tabelle 18: CASMIN (Fortsetzung)

Erläuterung:

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.

CASMIN (Befragte)		Schüler	n. gest.	TNZ	KA	WN	ohne Abschl.	Sonderschule	HS	RS	FHR	Abi	And. dt. Abschl.	And. aus. Abschl.
schulberuf	nicht erhob.	-10	-	-	-	-	-	-	-	-	-	-	-	-
unplan. Wert	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Schüler	-	-8	-	-	-	-	-	-	-	-	-	-	-	-
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
Änder. 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
Berufsfachsch.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Schul. d. Ges.wes.	-	-	-	-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
Disert.	-	-	-	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)

Tabelle 19: MCASMIN

Variablenname:	<i>mcasmin</i>
Variablenlabel:	Bildung der Mutter, klassifiziert nach CASMIN, aktual. Version, generiert
Ausgangsvariablen:	<i>mschul2; mberuf2</i>
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph

Tabelle 19: MCASMIN (Fortsetzung)

Erläuterung:

Generelle Darstellung: vgl. CASMIN Aufgrund der unterschiedli-

chen Erhebung der Bildungsvariablen für Befragte und deren El-

tern weicht das Vercodungsschema von *mcasmin* und *vcasmin*

leicht von dem bei *casmin* verwendeten ab. Zu Details vgl. die

folgende Abbildung.

CASMIN (Eltern)

Schul- Beruf	nicht erhob.	Platz fehlt	Eltern- 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
Platz fehlt	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Eltern- 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
PH	-	-	-	-	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)

Tabelle 20: VCASMIN

Variablenname:	<i>vcasmin</i>
Variablenlabel:	Bildung des Vaters, klassifiziert nach CASMIN, aktual. Version, generiert
Ausgangsvariablen:	<i>vschul2; vberuf2</i>
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph

Tabelle 20: VCASMIN (Fortsetzung)

Erläuterung:	Generelle Darstellung: vgl. CASMIN Aufgrund der unterschiedlichen Erhebung der Bildungsvariablen für Befragte und deren Eltern weicht das Vercodungsschema von <i>mcasmin</i> und <i>vcasmin</i> leicht von dem bei <i>casmin</i> verwendeten ab. Zu Details vgl. die folgende Abbildung. CASMIN (Eltern) <table> <tr> <th>Schul Beruf nicht erhob.</th><th>nicht erhob.</th><th>Plat fehlt</th><th>Eltern- unbek.</th><th>nicht gest.</th><th>TNZ</th><th>KA</th><th>WN</th><th>ohne Abschl.</th><th>Sonder- Schule</th><th>HS</th><th>RS</th><th>FHR</th><th>Abi</th><th>And. dt. Abschl.</th><th>And. au. Abschl.</th></tr> <tr> <td>unplan- bar</td><td>-10</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr> <td>Plat fehlt</td><td>-</td><td>-6</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr> <td>Eltern- unbek.</td><td>-</td><td>-</td><td>-5</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr> <td>nicht gest.</td><td>-</td><td>-</td><td>-</td><td>-4</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr> <tr> <td>TNZ</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-3</td><td>-3</td><td>-3</td><td>-3</td><td>-3</td><td>-3</td><td>-3</td><td>-3</td><td>-3</td><td>-3</td><td>-3</td></tr> <tr> <td>KA</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-3</td><td>-2</td><td>-2</td><td>-2</td><td>-2</td><td>-2</td><td>-2</td><td>-2</td><td>-2</td><td>-2</td><td>-2</td></tr> <tr> <td>WN</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-3</td><td>-2</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td><td>-1</td></tr> <tr> <td>ohne Abschl.</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-3</td><td>-2</td><td>-1</td><td>1a</td><td>1a</td><td>1b</td><td>2b</td><td>2c_gen</td><td>2c_gen</td><td>1b</td><td>1b</td></tr> <tr> <td>Akadem- ausbild.</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-3</td><td>-2</td><td>-1</td><td>1a</td><td>1a</td><td>1b</td><td>2b</td><td>2c_gen</td><td>2c_gen</td><td>1b</td><td>1b</td></tr> <tr> <td>Lehre</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-3</td><td>-2</td><td>-1</td><td>1c</td><td>1c</td><td>1c</td><td>2a</td><td>2c_voc</td><td>2c_voc</td><td>1c</td><td>1c</td></tr> <tr> <td>Meister</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-3</td><td>-2</td><td>-1</td><td>1c</td><td>1c</td><td>1c</td><td>2a</td><td>2c_voc</td><td>2c_voc</td><td>1c</td><td>1c</td></tr> <tr> <td>BA</td><td>-</td><td>-</td><td>-</td><td>-</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td></tr> <tr> <td>FH</td><td>-</td><td>-</td><td>-</td><td>-</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td><td>3a</td></tr> <tr> <td>Uni</td><td>-</td><td>-</td><td>-</td><td>-</td><td>3b</td><td>3b</td><td>3b</td><td>3b</td><td>3b</td><td>3b</td><td>3b</td><td>3b</td><td>3b</td><td>3b</td><td>3b</td></tr> <tr> <td>And. dt. Abschl.</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-3</td><td>-2</td><td>-1</td><td>1c</td><td>1c</td><td>1c</td><td>2a</td><td>2c_voc</td><td>2c_voc</td><td>1c</td><td>1c</td></tr> <tr> <td>And. au. Abschl.</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-3</td><td>-2</td><td>-1</td><td>1c</td><td>1c</td><td>1c</td><td>2a</td><td>2c_voc</td><td>2c_voc</td><td>1c</td><td>1c</td></tr> </table>	Schul Beruf nicht erhob.	nicht erhob.	Plat fehlt	Eltern- unbek.	nicht gest.	TNZ	KA	WN	ohne Abschl.	Sonder- Schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. au. Abschl.	unplan- bar	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Plat fehlt	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-	Eltern- 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	Akadem- 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. au. Abschl.	-	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
Schul Beruf nicht erhob.	nicht erhob.	Plat fehlt	Eltern- unbek.	nicht gest.	TNZ	KA	WN	ohne Abschl.	Sonder- Schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. au. Abschl.																																																																																																																																																																																																																																																																		
unplan- bar	-10	-	-	-	-	-	-	-	-	-	-	-	-	-	-																																																																																																																																																																																																																																																																		
Plat fehlt	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-																																																																																																																																																																																																																																																																		
Eltern- 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																																																																																																																																																																																																																																																																		
Akadem- 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. au. 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)																																																																																																																																																																																																																																																																																

Tabelle 21: ISCED 97

Variablenname:	<i>iscsed97</i>
Variablenlabel:	Bildung klassifiziert nach isced97, aktual. Version, generiert
Ausgangsvariablen:	<i>schul2; beruf2</i>
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	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)).

Tabelle 21: ISCED 97 (Fortsetzung)

	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. Zellen mit Kombinationen, die gültige ISCED-Werte ergeben, sind hellgrau, solche mit definierten missing values dunkelgrau unterlegt.														
	ISCED 97 (Befragte)														
	Schulberuf	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	-	-	-	-	-	-	-	-	-	-	-	-	-
	unplan. 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
	Anerk. 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-fachsch.	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
	Schul- u. Ges.-wes.	-	-	-	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b	5b
	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
	Uni/ 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)														

Tabelle 22: MISCED 97

Variablenname:	<i>miscd97</i>
Variablenlabel:	Bildung der Mutter klassifiziert nach <i>iscd97</i> , aktual. Version, generiert
Ausgangsvariablen:	<i>mschul2; mberuf2</i>
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	Zum theoretischen Hintergrund und zur Generierung vgl. ISCED-97.

Tabelle 22: MISCED 97 (Fortsetzung)

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.

ISCED 97 (Eltern)

Schul Beruf	nicht erhob.	Plat fehl	Eltern 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
Plat fehl	-	-4	-	-	-	-	-	-	-	-	-	-	-	-	-
Eltern 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
Anker- ausbild.	-	-	-	-	-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
PH	-	-	-	-	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. au. Abschl.	-	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2

Literatur: BMBF (2003); OECD (1999)

Tabelle 23: VISCED 97

	VISCED 97
Variablenname:	<i>visced97</i>
Variablenlabel:	Bildung des Vaters klassifiziert nach isced97, aktual. Version, generiert
Ausgangsvariablen:	<i>vschul2; vberuf2</i>
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	Zum theoretischen Hintergrund und zur Generierung vgl. ISCED-97.

Tabelle 23: VISCED 97 (Fortsetzung)

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.

ISCED 97 (Eltern)

Schul Beruf	nicht erhob.	Plat fehl	Ehrent. 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
Plat fehl	-	-6	-	-	-	-	-	-	-	-	-	-	-	-	-
Ehrent. 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
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
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)

Tabelle 24: International Standard Classification of Occupations 1988 (ISCO88)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	Aktuelle (<i>PENDDAT</i>) - <i>isco88</i> - <i>ET2500</i> Spelldaten (<i>bio_spells</i>) - <i>isco88</i> - <i>ET2500</i> erste (<i>PENDDAT</i>) - <i>isco88ewt</i> - <i>ET2500</i> , <i>PET1280</i> , <i>PET3950</i> letzte (<i>PENDDAT</i>) - <i>isco88lewt</i> - <i>ET2500</i> , <i>PET1280</i> des Vaters (<i>PENDDAT</i>) - <i>visco88</i> - <i>PSH0800</i> der Mutter (<i>PENDDAT</i>) - <i>misco88</i> - <i>PSH0700</i> Minijob - <i>isco88minj</i> - <i>PMJ0900</i>
Variablenlabel:	aktuelle Ewt.: Intern. Standard Classification of Occupations 88, aktuelle Erwerbstätigkeit, generiert Spelldaten (<i>bio_spells</i>): Intern. Standard Classification of Occupations 88, generiert erste Ewt.: ISCO 88, erste Erwerbstätigkeit, generiert letzte Ewt.: ISCO 88, letzte Erwerbstätigkeit, generiert Vater: ISCO 88 des Vaters, generiert Mutter: ISCO 88 der Mutter, generiert Minijob: ISCO 88, aktueller Minijob, generiert
Typ / Datensatz:	Beruf / Personendaten

Tabelle 24: Internat. Standard Class. of Occupations 1988 (ISCO88) (Fortsetzung)

Ansprechpartner:	Bernhard Christoph
Erläuterung:	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. Bei der Codierung der Angaben zum Minijob liegen keine Angaben zur Stellung im Beruf vor. Da es sich bei Minijobs in der weit überwiegenden Zahl der Fälle um geringqualifizierte Stellen handelt, wurde in allen Fällen, bei denen üblicherweise die Stellung im Beruf zum Entscheid zwischen verschiedenen in Frage kommenden Berufscodes genutzt wird, angenommen, dass bei der Tätigkeit keine Leitungsfunktion vorliegt. Es wurde dann immer der Beruf mit dem niedrigeren Prestige vercodet.
Literatur:	ILO (1990)

Tabelle 25: International Standard Classification of Occupations 2008 (ISCO08)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	aktuelle (<u>PENDDAT</u>) - <i>isco08</i> - <i>ET2500</i> Spelldaten (<i>bio_spells</i>) - <i>kldb</i> - <i>ET2500</i> erste (<u>PENDDAT</u>) - <i>isco08ewt</i> - <i>ET2500</i> , <i>PET1280</i> , <i>PET3950</i> letzte (<u>PENDDAT</u>) - <i>isco08ewt</i> - <i>ET2500</i> , <i>PET1280</i> des Vaters (<u>PENDDAT</u>) - <i>visco08</i> - <i>PSH0800</i> der Mutter (<u>PENDDAT</u>) - <i>misco08</i> - <i>PSH0700</i> Minijob - <i>isco08mini</i> - <i>PMJ0900</i> Ausbildungsaspiration, Wunschberuf - <i>isco08berufswunsch</i> - <i>PAA0100</i> Ausbildungsaspiration, angestrebter Beruf - <i>isco08angberuf</i> - <i>PAA1000</i>
Variablenlabel:	aktuelle Ewt.: Intern. Standard Classification of Occupations 08, aktuelle Erwerbstätigkeit, gen. Spelldaten (<i>bio_spells</i>):International Standard Classification of Occupations, gen. erste Ewt.: ISCO08, erste Erwerbstätigkeit, generiert letzte Ewt.: ISCO08, letzte Erwerbstätigkeit, generiert Vater: ISCO08 des Vaters, generiert Mutter: ISCO08 der Mutter, generiert Minijob: ISCO08, aktueller Minijob, generiert

Tabelle 25: Internat. Standard Class. of Occupations 2008 (ISCO08) (Fortsetzung)

	Ausbildungsaspiration, Wunschberuf: ISCO 08, Berufsvorstellungen: Berufswunsch, gen. Ausbildungsaspiration, angestrebter Beruf: ISCO 08, Berufsvorstellungen: Angestrebter Beruf, gen.
Typ / Datensatz:	Beruf / Personendaten
Ansprechpartner:	Bernhard Christoph
Erläuterung:	Die International Standard Classification of Occupations (ISCO) ist eine von der ILO entwickelte, international vergleichbare Berufsklassifikation. Die ISCO-08-Klassifizierung stellt eine Aktualisierung von ISCO-88 dar. Der Rahmen und die Konzepte, auf denen ISCO-08 basiert, sind im Wesentlichen unverändert gegenüber ISCO-88. Die Definition dieser Konzepte wurde aktualisiert und die Richtlinien für ihre Anwendung auf die Gestaltung der Einstufung wurde überarbeitet, um Mängeln in ISCO-88 zu begegnen. Berichtete Erwerbstätigkeiten sind in ISCO-08 codiert, sofern es sich um Berufsepisoden handelt, die ab der zehnten Befragungswelle aus der Vorwelle fortgeschrieben wurden oder sofern es sich um eine neue ab Welle 10 gemachte Angabe handelt. Vor Welle 10 angegebene und nicht in Welle 10ff. fortgeführte Berufsepisoden liegen ausschließlich in der ISCO-88-Codierung vor. Bei der Codierung der Angaben zum Minijob liegen keine Angaben zur Stellung im Beruf vor. Da es sich bei Minijobs in der weit überwiegenden Zahl der Fälle um geringqualifizierte Stellen handelt, wurde in allen Fällen, bei denen üblicherweise die Stellung im Beruf zum Entscheid zwischen verschiedenen in Frage kommenden Berufscodes genutzt wird, angenommen, dass bei der Tätigkeit keine Leitungsfunktion vorliegt. Es wurde dann immer der Beruf mit dem niedrigeren Prestige vercodet.

Tabelle 25: Internat. Standard Class. of Occupations 2008 (ISCO08) (Fortsetzung)

	Auch bei der Codierung der Angaben zur Ausbildungsaspiration liegen keine Angaben zur Stellung im Beruf vor, dort kann zudem nicht auf den Ausbildungsabschluss zurückgegriffen werden, da die Abfrage sich an Personen richtet, die noch keinen solchen Abschluss erworben haben. Es wird hier nach der Regel verfahren, dass Auszubildende demjenigen Beruf zugeordnet werden, für den sie bei zukünftiger Realisierung des Ausbildungswunschs ausgebildet würden. Wenn in den ISCO-08-Codes nach verschiedenen Qualifikationsniveaus unterschieden wird, werden jene Codes vermieden, die für eine ohne Ausbildung mögliche Tätigkeit oder für eine Anlern­tätigkeit stehen. Wenn in den Qualifikationsniveaus zwischen dem Absolvieren einer Berufsfachschule einerseits und dem Absolvieren einer Fachhochschule oder Universität andererseits unterschieden wird, wird der auf das Absolvieren einer Berufsfachschule ausgerichtete Code gewählt.
Literatur:	ILO (2012)

Tabelle 26: Klassifizierung der Berufe 1992 (KldB1992)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	aktuelle - <i>kldb1992</i> - ET2500 Spelldaten - (<i>bio_spells</i>) - <i>kldb1992</i> - ET2500 erste - <i>kldb1992ewt</i> - ET2500, PET1280, PET3950 letzte - <i>kldb1992lewt</i> - ET2500, PET1280 des Vaters - <i>vkldb1992</i> - PSH0800 der Mutter - <i>mkldb1991</i> - PSH0700 Minijob - <i>kdlb1992mini</i> - PMJ0900
Variablenlabel	aktuelle Ewt.: Klassifizierung der Berufe 1992, aktuelle Erwerbstätigkeit, gen. Spelldaten - (<i>bio_spells</i>): Klassifizierung der Berufe 1992, generiert erste Ewt.: Klassifizierung der Berufe 1992, erste Erwerbstätigkeit, gen. letzte Ewt.: Klassifizierung der Berufe 1992, letzte Erwerbstätigkeit, gen. Vater: Klassifizierung der Berufe 1992 des Vaters, generiert Mutter: Klassifizierung der Berufe 1992 der Mutter, generiert Minijob: Klassifizierung der Berufe 1992, aktueller Minijob, gen.
Typ/ Datensatz:	Beruf/ Personendaten
Ansprechpartner:	Bernhard Christoph

Tabelle 26: Klassifizierung der Berufe 1992 (KldB1992) (Fortsetzung)

Erläuterungen:	Die KldB92 ist die Version der Klassifizierung der Berufe in der Variante des Statistischen Bundesamtes aus dem Jahr 1992. Sie ist eine spezifisch auf Deutschland bezogene Klassifikation. Die Klassifikation basiert ausschließlich auf der beruflichen Tätigkeit.
Literatur:	StBA (1992)

Tabelle 27: Klassifizierung der Berufe 2010 (KldB2010)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	aktuelle - <i>kldb2010</i> - <i>ET2500</i> Spelldaten (<i>bio_spells</i>) - <i>kldb2010</i> - <i>ET2500</i> erste - <i>kldb2010ewt</i> - <i>ET2500</i> , <i>PET1280</i> , <i>PET3950</i> letzte - <i>kldb2010lewt</i> - <i>ET2500</i> , <i>PET1280</i> des Vaters - <i>vkldb2010</i> - <i>PSH0800</i> der Mutter - <i>mkldb2010</i> - <i>PSH0700</i> Minijob - <i>kldb2010mini</i> - <i>PMJ0900</i> Ausbildungsaspiration, Wunschberuf - <i>kldb2010berufswunsch</i> - <i>PAA0100</i> Ausbildungsaspiration, angestrebter Beruf - <i>kldb2010angberuf</i> - <i>PAA1000</i>
Variablenlabel	aktuelle Ewt.: Klassifizierung der Berufe 2010, aktuelle Erwerbstätigkeit, gen. Spelldaten <i>bio_spells</i>): Klassifizierung der Berufe 2010, generiert erste Ewt.: Klassifizierung der Berufe 2010, erste Erwerbstätigkeit, gen. letzte Ewt.: Klassifizierung der Berufe 2010, letzte Erwerbst., gen. Vater: Klassifizierung der Berufe 2010 des Vaters, generiert Mutter: Klassifizierung der Berufe 2010 der Mutter, generiert Minijob: Klassifizierung der Berufe 2010, aktueller Minijob, gen. Ausbildungsaspiration, Wunschberuf: Klassifizierung der Berufe 2010, Berufsvorstellungen: Berufswunsch, gen. Ausbildungsaspiration, angestrebter Beruf: Klassifizierung der Berufe 2010, Berufsvorstellungen: Angestrebter Beruf, gen.
Typ/ Datensatz	Beruf/ Personendaten
Ansprechpartner	Christian Dickmann

Tabelle 27: Klassifizierung der Berufe 2010 (KldB2010) (Fortsetzung)

Erläuterung	Bei der KldB2010 handelt es sich um eine vollständige Neuentwicklung, die die aktuelle Berufslandschaft in Deutschland realitätsnah abbildet. Damit können die in den letzten Jahrzehnten deutlich veränderten Berufsstrukturen in Statistiken und Analysen besser als bisher dargestellt werden. Ein weiterer Vorteil der KldB 2010 ist die hohe Kompatibilität zur internationalen Berufsklassifikation, der ISCO-08 (International Classification of Occupations 2008), da durch sie die internationale Vergleichbarkeit von Berufsinformationen in den amtlichen Statistiken und in der Forschung deutlich verbessert wird. Berichtete Erwerbstätigkeiten sind in KldB-2010 codiert, sofern es sich um Berufsepisoden handelt, die ab der zehnten Befragungswelle aus der Vorwelle fortgeschrieben wurden oder sofern es sich um neue ab Welle 10 gemachte Angaben handelt. Vor Welle 10 angegebene und nicht in Welle 10ff. fortgeführte Berufsepisoden liegen ausschließlich in der KldB-1992-Codierung vor. Bei der Codierung der Angaben zur Ausbildungsaspiration kann nicht auf den Ausbildungsabschluss zurückgegriffen werden, da die Abfrage sich an Personen richtet, die noch keinen solchen Abschluss erworben haben. Es wird hier nach der Regel verfahren, dass Auszubildende demjenigen Beruf zugeordnet werden, für den sie bei zukünftiger Realisierung des Ausbildungswunschs ausgebildet würden. Wenn in den KldB-2010-Codes nach verschiedenen Qualifikationsniveaus unterschieden wird, werden jene Codes vermieden, die für eine ohne Ausbildung mögliche Tätigkeit oder für eine Anlern Tätigkeit stehen. Wenn in den Qualifikationsniveaus zwischen dem Absolvieren einer Berufsfachschule einerseits und dem Absolvieren einer Fachhochschule oder Universität andererseits unterschieden wird, wird der auf das Absolvieren einer Berufsfachschule ausgerichtete Code gewählt.
Literatur	Bundesagentur für Arbeit (2011)

Tabelle 28: Klassenschema nach Erikson, Goldthorpe und Portocarrero (EGP)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	aktuelle - <i>egp - isco88, stib</i> Spelldaten - (<i>bio_spells</i>) - <i>egp - isco88, stib</i> erste - <i>egpeewt - isco88eewt, stibeewt</i> letzte - <i>egplewt - isco88lewt, stiblewt</i> des Vaters - <i>vegp - visco88, vstib</i> der Mutter - <i>megp - misco88, mstib</i>

Tabelle 28: Klassenschema (EGP) (Fortsetzung)

Variablenlabel:	aktuelle Ewt.: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), akt. Beruf, gen. Spelldaten (<i>bio_spells</i>): Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), gen. erste Ewt.: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), erster Beruf letzte Ewt.: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), erster Beruf Vater: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), Beruf des Vaters Mutter: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), Beruf der Mutter
Typ / Datensatz:	sozioökonomische Position / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	Das Klassenschema von Erikson, Goldthorpe und Portocarrero (Erikson et al. (1979, 1982); Erikson & Goldthorpe (1992)) ist eines der gebräuchlichsten Instrumente zur Operationalisierung der Klassenposition. 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 Classification (ESeC) – mithelfende Familienangehörige nicht als Selbständige, sondern als abhängig Erwerbstätige vercodet. 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. Für die Berufsangaben zum Minijob wurde EGP nicht generiert, da die üblicherweise erfragten Informationen zur Stellung im Beruf im Minijobmodul nicht erhoben worden sind
Literatur:	Christoph et al. (2005); Erikson & Goldthorpe (1992); Erikson et al. (1982); Erikson et al. (1979)

Tabelle 29: European Socio-economic Classification (ESeC)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	aktuelle - <i>esec - isco88, stib, PET2000, PET2700</i> Spelldaten (<i>bio_spells</i>) - <i>esec - isco88, stib, ET1100, ET1101, ET1102, ET1103, ET1104, ET1105, ET1300, ET1301, ET1302, ET1303, ET1304, ET1305</i> erste - <i>eseceewt - isco88eewt, stibeewt, PET1261</i> letzte - <i>eseclewt - isco88lewt, stiblewt, PET3801</i> des Vaters - <i>vesec - visco88, vstib, PSH0670</i> der Mutter - <i>mesec - misco88, mstib, PSH0370</i>
Variablenlabel:	aktuelle Ewt.: European Socio-economic Classification (ESeC), akt. Beruf, gen. Spelldaten (<i>bio_spells</i>): European Socio-economic Classification (ESeC), gen. erste Ewt.: European Socio-economic Classification (ESeC), erster Beruf, gen. letzte Ewt.: European Socio-economic Classification (ESeC), letzter Beruf, gen. Vater: European Socio-economic Classification (ESeC), Beruf des Vaters, gen. Mutter: European Socio-economic Classification (ESeC), akt. Beruf der Mutter, gen.
Typ / Datensatz:	sozioökonomische Position / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	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. Für die Berufsangaben zum Minijob wurde ESeC nicht generiert, da die üblicherweise erfragten Informationen zur Stellung im Beruf im Minijobmodul nicht erhoben worden sind.

Tabelle 29: European Socio-economic Classification (ESeC) (Fortsetzung)

Literatur:	Fischer & Wirth (2007); Harrison & Rose (2006); Müller et al. (2006, 2007); Rose & Harrison (2007)
------------	--

Tabelle 30: Magnitude-Prestigeskala (MPS)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	aktuelle - <i>mps</i> - <i>isco88</i> Spelldaten (<i>bio_spells</i>) - <i>mps</i> - <i>isco88</i> erste - <i>mpseewt</i> - <i>isco88eewt</i> letzte - <i>mpslewt</i> - <i>isco88lewt</i> des Vaters - <i>vmmps</i> - <i>visco88</i> der Mutter - <i>mmmps</i> - <i>misco88</i>
Variablenlabel:	aktuelle Ewt.: Magnitude-Prestigeskala, aktuelle Ewt. gen. Spelldaten (<i>bio_spells</i>): Magnitude-Prestigeskala, gen. erste Ewt.: Magnitude-Prestigeskala, erste Erwerbstätigkeit, gen. letzte Ewt.: Magnitude-Prestigeskala, letzte Erwerbst., gen. Vater: Magnitude-Prestigeskala, Beruf des Vaters, gen. Mutter: Magnitude-Prestigeskala, Beruf der Mutter, gen.
Typ / Datensatz:	sozioökonomische Position / Personendaten
Ansprechpartner:	Bernhard Christoph
Erläuterung:	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 KldB 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 infas. Für die Berufsangaben zum Minijob wurde MPS nicht generiert, da die üblicherweise erfragten Informationen zur Stellung im Beruf im Minijobmodul nicht erhoben worden sind.
Literatur:	Christoph (2005); Wegener (1985, 1988)

Tabelle 31: Standard International Occupational Prestige Scale (SIOPS/Treiman-Skala)- Basis ISCO-88

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	aktuelle - <i>siops1</i> - <i>isco88</i>

**Tabelle 31: Standard Internat. Occ. Prestige Scale (SIOPS/Treiman-Skala)
(Fortsetzung)**

	Spelldaten (<i>bio_spells</i>) - <i>siops1</i> - <i>isco88</i> erste - <i>siopseewt1</i> - <i>isco88eewt</i> letzte - <i>siopslewt1</i> - <i>isco88eewt</i> des Vaters - <i>vsiops1</i> - <i>visco88</i> der Mutter - <i>msiops1</i> - <i>misco88</i>
Variablenlabel:	aktuelle Ewt.: Standard International Occupational Prestige Scale (Basis ISCO-88), aktuelle Ewt., gen. Spelldaten (<i>bio_spells</i>): Standard International Occupational Prestige Scale (Basis ISCO-88), gen. erste Ewt.: SIOPS (Basis ISCO-88), erste Erwerbstät., gen. letzte Ewt.: SIOPS (Basis ISCO-88), letzte Erwerbstät., gen. Vater: SIOPS (Basis ISCO-88), Beruf des Vaters, gen. Mutter: SIOPS (Basis ISCO-88), Beruf der Mutter, gen.
Typ / Datensatz:	sozioökonomische Position / Personendaten
Ansprechpartner:	Bernhard Christoph
Erläuterung:	Die ursprünglich von Treiman (1977) für ISCO-68 entwickelte Treiman-Prestige-skala 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 Zuweisung erfolgte im Rahmen der Berufsvercodung durch infas. Für die Berufsangaben zum Minijob wurde SIOPS nicht generiert, da die üblicherweise erfragten Informationen zur Stellung im Beruf im Minijobmodul nicht erhoben worden sind.
Literatur:	Ganzeboom & Treiman (1996, 2003); Treiman (1977)

Tabelle 32: Standard International Occupational Prestige Scale (SIOPS/Treiman-Skala) – Basis ISCO-08

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen aktuelle - <i>siops2</i> - <i>isco08</i> Spelldaten (<i>bio_spells</i>) - <i>siops2</i> - <i>isco08</i> erste - <i>siopseewt2</i> - <i>isco08eewt</i> letzte - <i>siopslewt2</i> - <i>isco08lewt</i> des Vaters - <i>vsiops2</i> - <i>visco08</i> der Mutter - <i>msiops2</i> - <i>misco08</i>
Variablenlabel:	aktuelle Ewt.: Standard Intern. Occupational Prestige Scale (Basis ISCO08), aktuelle Ewt., gen. Spelldaten (<i>bio_spells</i>): Standard Intern. Occupational Prestige Scale (Basis ISCO08), gen.

Tabelle 32: Standard Internat. Occ. Prestige Scale (SIOPS/Treiman-Skala)
(Fortsetzung)

	erste Ewt.: SIOPS (Basis ISCO08), erste Erwerbstät., gen. letzte Ewt.: SIOPS (Basis ISCO08), letzte Erwerbstät., gen. Vater: SIOPS (Basis ISCO08), Beruf des Vaters, gen. Mutter: SIOPS (Basis ISCO08), Beruf der Mutter, gen.
Typ / Datensatz:	sozioökonomische Position / Personendaten
Ansprechpartner:	Christian Dickmann
Erläuterung:	Auch für ISCO-08 haben Ganzeboom und Treiman eine aktualisierte Variante von SIOPS entwickelt und eine Generierungssyntax zur Verfügung gestellt. Für berichtete Erwerbstätigkeiten wurde SIOPS auf der Basis von ISCO-08 generiert, sofern es sich um Berufsepisoden handelt, die ab der zehnten Befragungswelle aus der Vorwelle fortgeschrieben wurden oder sofern es sich um neue ab Welle 10 gemachte Angaben handelt. Für vor Welle 10 angegebene und nicht in Welle 10ff. fortgeführte Berufsepisoden liegt ausschließlich SIOPS auf der Basis von ISCO-88 vor. Für die Berufsangaben zum Minijob und zur Ausbildungsaspiration wurde SIOPS nicht generiert, da die üblicherweise erfragten Informationen zur Stellung im Beruf in diesen Modulen nicht erhoben worden sind.
Literatur:	Ganzeboom & Treiman (2010, 2011)

Tabelle 33: International Socio-Economic Index (ISEI) – Basis ISCO-88

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	aktuelle - <i>isei1 - isco88</i> Spelldaten (<i>bio_spells</i>) - <i>isei1 - isco88</i> erste - <i>iseieewt1 - isco88eewt</i> letzte - <i>iseilewt1 - isco88lewt</i> des Vaters - <i>visei1 - visco88</i> der Mutter - <i>misei1 - misco88</i>
Variablenlabel:	aktuelle Ewt.: International Socio-Economic Index (Basis ISCO88), aktuelle Ewt., gen. Spelldaten (<i>bio_spells</i>): Standard Intern. Occupational Prestige Scale (Basis ISCO88), gen. erste Ewt.: ISEI (Basis ISCO88), erste Erwerbstätigkeit., gen. letzte Ewt.: ISEI (Basis ISCO88), letzte Erwerbstätigkeit, gen. Vater: ISEI (Basis ISCO88), Beruf des Vaters, gen. Mutter: ISEI (Basis ISCO88), Beruf der Mutter, gen.
Typ / Datensatz:	sozioökonomische Position / Personendaten

**Tabelle 33: International-Socio Economic Index (ISEI) - Basis ISCO-88
(Fortsetzung)**

Ansprechpartner:	Bernhard Christoph
Erläuterung:	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. Der ISEI wurde zunächst für ISCO-68 entwickelt (Ganzeboom et al. (1992)), später jedoch auch für ISCO-88 umgesetzt (Ganzeboom & Treiman (1996, 2003)). Die Zuspiegelung erfolgte im Rahmen der Berufsvercodung durch infas. Für die Berufsangaben zum Minijob wurde ISEI nicht generiert, da die üblicherweise erfragten Informationen zur Stellung im Beruf im Minijobmodul nicht erhoben worden sind.
Literatur:	Ganzeboom et al. (1992); Ganzeboom & Treiman (1996,2003)

Tabelle 34: International Socio-Economic Index (ISEI) – Basis ISCO-08

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	aktuelle - <i>isei2 - isco08</i> Spelldaten (<i>bio_spells</i>) - <i>isei2 - isco08</i> erste - <i>iseieewt2 - isco08eewt</i> letzte - <i>iseilewt2 - isco08lewt</i> des Vaters - <i>visei2 - visco08</i> der Mutter - <i>misei2 - misco08</i>
Variablenlabel:	aktuelle Ewt.: International Socio-Economic Index (Basis ISCO08), aktuelle Ewt., gen. Spelldaten (<i>bio_spells</i>): Standard Intern. Occupational Prestige Scale (Basis ISCO08), gen. erste Ewt.: ISEI (Basis ISCO08), erste Erwerbstätigkeit., gen. letzte Ewt.: ISEI (Basis ISCO08), letzte Erwerbstätigkeit, gen. Vater: ISEI (Basis ISCO08), Beruf des Vaters, gen. Mutter: ISEI (Basis ISCO08), Beruf der Mutter, gen.
Typ / Datensatz:	sozioökonomische Position / Personendaten
Ansprechpartner:	Christian Dickmann
Erläuterung:	Basis des ISEI-08-Indizes sind die Datensätze des International Social Survey Programme (ISSP) der Jahre 2002 bis 2007. Die Zuspiegelung erfolgte im Rahmen der Berufsvercodung durch infas.

**Tabelle 34: International-Socio Economic Index (ISEI)- Basis ISCO-08
(Fortsetzung)**

	Für berichtete Erwerbstätigkeiten wurde ISEI auf der Basis von ISCO-08 generiert, sofern es sich um Berufsepisoden handelt, die ab der zehnten Befragungswelle aus der Vorwelle fortgeschrieben wurden oder sofern es sich um neue ab Welle 10 gemachte Angaben handelt. Für vor Welle 10 angegebene und nicht in Welle 10ff. fortgeführte Berufsepisoden liegt ausschließlich ISEI auf der Basis von ISCO-88 vor. Für die Berufsangaben zum Minijob und zur Ausbildungsaspiration wurde ISEI nicht generiert, da die üblicherweise erfragten Informationen zur Stellung im Beruf in diesen Modulen nicht erhoben worden sind.
Literatur:	Ganzeboom (2010)

Tabelle 35: Klassifikation der Wirtschaftszweige 2003 (WZ2003)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	aktuelle - <i>branche1</i> - ET2600 Spelldaten (<i>bio_spells</i>) - <i>branche1</i> - ET2600 Minijob - <i>brancheminj1</i> - PMJ1300
Variablenlabel:	aktuelle Ewt.: Jetzige Tätigkeit: Wirtschaftszweig/Branche, aktuelle Ewt. (WZ 2003) Spelldaten (<i>bio_spells</i>): Wirtschaftszweig/Branche (WZ 2003), generiert Minijob: Wirtschaftszweig/Branche, aktueller Minijob (WZ 2003)
Typ / Datensatz:	sozioökonomische Position / Personendaten
Ansprechpartner:	Bernhard Christoph
Erläuterung:	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.
Literatur:	StBA (2002); EG (2002)

Tabelle 36: Klassifikation der Wirtschaftszweige 2008 (WZ2008)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	aktuelle - <i>branche2</i> - ET2600 Spelldaten (<i>bio_spells</i>) - <i>branche2</i> - ET2600

Tabelle 36: Klassifikation der Wirtschaftszweige 2008 (WZ2008) (Fortsetzung)

	Minijob - <i>brancheminj2</i> - <i>PMJ1300</i>
Variablenlabel:	aktuelle Ewt.: Jetzige Tätigkeit: Wirtschaftszweig/Branche, aktuelle Ewt. (WZ 2008) Spelldaten (<i>bio_spells</i>): Wirtschaftszweig/Branche (WZ 2003), generiert Minijob: Wirtschaftszweig/Branche, aktueller Minijob (WZ 2003)
Typ / Datensatz:	sozioökonomische Position / Personendaten
Ansprechpartner:	Christian Dickmann
Erläuterung:	Die Vercodung der offenen Branchenangabe erfolgte anhand des Zweistellers der Klassifikation der Wirtschaftszweige des statistischen Bundesamts (WZ2008). Die Ebene des Zweistellers wird auch als Gliederungsebene der Abteilungen bezeichnet. Sie basiert auf der Internationalen Systematik der Wirtschaftszweige (ISIC Rev. 4) der Vereinten Nationen und der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft (NACE Rev. 2) auf. Auf der Ebene der Zweisteller sind diese beiden Grundlagen der Wirtschaftszweig-Codierung identisch. Berichtete Branchenangaben sind in WZ-2008 codiert, sofern es sich um Berufsepisoden handelt, die ab der zehnten Befragungswelle aus der Vorwelle fortgeschrieben wurden oder sofern es sich um neue ab Welle 10 gemachte Angaben handelt. Branchenangaben von vor der Welle 10 angegebenen und nicht in Welle 10ff. fortgeführten Berufsepisoden liegen ausschließlich in der WZ-2003-Codierung vor.
Literatur:	StBA (2002); EG (2002)

Tabelle 37: Körperliche Skala des SF12v2 (SOEP-Version, NBS)

Variablenname:	<i>pcs</i>
Variablenlabel:	körperliche Skala des SF12v2 (SOEP-Version, NBS), generiert
Ausgangsvariablen:	<i>PG1200; PG1205; PG1210; PG1215*</i>
Typ / Datensatz:	Gesundheit / Personendaten
Bearbeitung:	Christian Dickmann
Erläuterung:	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.

Tabelle 37: Körperliche Skala des SF12v2 (SOEP-Version, NBS) (Fortsetzung)

	Die Generierung von <i>pcs</i> in PASS basiert auf der bei Nübling et al. (2006) wiedergegebenen SPSS- Syntax. Die SF-12-Indikatoren wurden im PASS bislang in den Wellen 3, 6 und 9 erhoben.
Literatur:	Nübling et al. (2006); Andersen et al. (2007)

Tabelle 38: Psychische Skala des SF12v2 (SOEP-Version, NBS)

Variablenname:	<i>mcs</i>
Variablenlabel:	psychische Skala des SF12v2 (SOEP-Version, NBS), generiert
Ausgangsvariablen:	<i>PG1200; PG1205; PG1210; PG1215*</i>
Typ / Datensatz:	Gesundheit / Personendaten
Bearbeitung:	Christian Dickmann
Erläuterung:	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. Die Generierung von <i>mcs</i> in PASS basiert auf der bei Nübling et al. (2006) wiedergegebenen SPSS- Syntax. Die SF-12-Indikatoren wurden im PASS bislang in den Wellen 3, 6 und 9 erhoben.
Literatur:	Nübling et al. (2006); Andersen et al. (2007)

4.5.2 Haushalts- oder Bedarfsgemeinschaftsebene

Tabelle 39: Bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung

Variablenname:	<i>oecdinca</i>
Variablenlabel:	Bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung (gerundet)
Ausgangsvariablen:	<i>HD0200a-HD0200o; HA0100; hhincome</i>
Typ / Datensatz:	sozioökonomische Position / Haushaltsdaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	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). 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.
Literatur:	Hauser (1996); OECD (1982)

Tabelle 40: Bedarfsgewichtetes Haushaltseinkommen, neue OECD-Gewichtung

Variablenname:	<i>oecdincn</i>
Variablenlabel:	bedarfsgewichtetes Haushaltseinkommen, neue OECD-Gewichtung (gerundet)
Ausgangsvariablen:	<i>HD0200a-HD0200o; HA0100; hhincome</i>
Typ / Datensatz:	sozioökonomische Position / Haushaltsdaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	<u>Generelle Darstellung:</u> vgl. Bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung. 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.
Literatur:	Hagenaars et al. (1994)

Tabelle 41: Deprivationsindex, ungewichtet

Variablenname:	<i>depindug2</i>
Variablenlabel:	Alle Wellen: Deprivationsindex, ungewichtet (Itemsumme: 23)
Ausgangsvariablen:	<i>HLS0100a-HLS0400a; HLS0100b-HLS0400b; HLS0600a-HLS1200a;</i> <i>HLS0600b-HLS1200b; HLS1400a-HLS2500a; HLS1400b-HLS2500b;</i>
Typ / Datensatz:	materielle Situation / Haushaltsdaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	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. 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)). Als genereller Tenor der Forschung ist festzuhalten, dass der Ressourcen- und der Deprivationsansatz zum Teil unterschiedliche Personengruppen erfassen bzw. klassifizieren. Für eine Abgrenzung der Armutspopulation im engeren Sinne wird daher auch oft vorgeschlagen, Maße für Einkommensarmut und Deprivation zu kombinieren und nur diejenigen, die von beiden Ansätzen klassifiziert werden, als Armutspopulation im engeren Sinne zu betrachten (vgl. Halleröd (1995); Nolan & Whelan (1996); Andreß & Lipsmeier (2001)). Dem Index liegt eine Liste von 23 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.

Tabelle 41: Deprivationsindex, ungewichtet (Fortsetzung)

	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 ("listwise deletion"). Bei der Abfrage von 23 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. Für die Wellen 1 bis 4 liegt mit der Variable <i>depindug</i> eine Version des ungewichteten Deprivationsindex vor, der auf 26 statt 23 Items basiert, nämlich zusätzlich zu den oben genannten Items auch auf <i>HLS0500*</i>, <i>HLS1300*</i> und <i>HLS2600*</i>. Diese drei Items werden seit Welle 5 nicht mehr erhoben. Daher wurde <i>depindug2</i> neu in den Datensatz eingefügt und rückwirkend seit Welle 1 generiert.
Literatur:	Andreß & Lipsmeier (2001); Halleröd (1995); Nolan & Whelan (1996); Ringen (1988)

Tabelle 42: Deprivationsindex, gewichtet

Variablenname:	<i>depindg2</i>
Variablenlabel:	Alle Wellen: Deprivationsindex, gew. (Itemsumme bis W7: 11.08, ab W8: 10.59)
Ausgangsvariablen:	<i>HLS0100a-HLS0400a; HLS0100b-HLS0400b;</i> <i>HLS0600a-HLS1200a; HLS0600b-HLS1200b;</i> <i>HLS1400a-HLS2500a; HLS1400b-HLS2500b;</i> <i>PLS0100-PLS0400; PLS0600-PLS1200;</i> <i>PLS1400-PLS2500;</i>
Typ / Datensatz:	materielle Situation / Haushaltsdaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	Für eine generelle Darstellung: vgl. Deprivationsindex, ungewichtet

Tabelle 42: Deprivationsindex, gewichtet (Fortsetzung)

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)).

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.

In Welle 8 wurden die Fragen zur Notwendigkeit der Deprivationsitems erneut gestellt. Die Gewichtung des Deprivationsindex basiert für die Wellen 1 bis einschließlich 7 auf den Daten der 1. Welle, ab Welle 8 dann auf den Daten der 8. Welle.

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.).

Tabelle 42: Deprivationsindex, gewichtet (Fortsetzung)

	Für die Wellen 1 bis 4 liegt mit der Variable <i>depindg</i> eine Version des gewichteten Deprivationsindex vor, der auf 26 statt 23 Items basiert, nämlich zusätzlich zu den oben genannten Items auch auf <i>HLS0500*</i> , <i>HLS1300*</i> und <i>HLS2600*</i> sowie <i>PLS0500</i> , <i>PLS1300</i> und <i>PLS2600</i> . Die drei HLS-Items werden seit Welle 5 nicht mehr erhoben. Daher wurde <i>depindg2</i> neu in den Datensatz eingefügt und rückwirkend seit Welle 1 generiert.
Literatur:	Andreß & Lipsmeier (1995, 2001); Andreß et al. (1996); Halleröd (1995); Lipsmeier (1999); Nolan & Whelan (1996)

Tabelle 43: Haushaltstypologie

Variablenname:	<i>hhtyp</i>
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.

Tabelle 43: Haushaltstypologie (Fortsetzung)

	<u>Definition der Beziehungen bei d. Generierung d. Haushaltstyps:</u> ▪ <u>Paare:</u> Ehepaare; eingetragene Lebenspartnerschaften; unverheiratete Partnerschaften sowie Partnerschaften mit nicht weiter spezifiziertem Status (fehlender Wert in Nachfrage nach Art der Partnerschaft). ▪ <u>Kind einer Person:</u> leibliches Kind; Stief-, Adoptiv- oder Pflegekind; Kind mit nicht weiter spezifiziertem Status (fehlender Wert in Nachfrage nach Art der Beziehung zum Kind). ▪ <u>Elternteil einer Person:</u> leibliches Elternteil; Stief-, Adoptiv- und Pflegeelternteil; Eltern mit nicht weiter spezifiziertem Status (fehlender Wert in Nachfrage nach Art der Elternschaft). <u>Definition der Haushaltstypen:</u> ▪ <u>1-Personen Haushalt:</u> Haushalt mit nur einer Person im Haushalt. ▪ <u>Paar ohne Kinder:</u> Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen. ▪ <u>1-Eltern-Haushalt:</u> Haushalt besteht ausschließlich aus einem Elternteil und dessen Kindern. Hinsichtlich des Alters der Kinder werden keine Einschränkungen vorgenommen. ▪ <u>Paar mit Kindern < 16 Jahre:</u> Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen und deren jeweiligen bzw. gemeinsamen Kindern. Alle Kinder sind unter 16 Jahre alt. ▪ <u>Paar mit Kindern >= 16 Jahre:</u> Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen und deren jeweiligen bzw. gemeinsamen Kindern. Alle Kinder sind 16 Jahre oder älter. ▪ <u>Paar mit Kindern < 16 Jahre und >= 16 Jahre:</u> 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.
--	--

Tabelle 43: Haushaltstypologie (Fortsetzung)

	▪ <u>Mehrgenerationenhaushalt</u>: 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 Person im Haushalt. Ausgehend von den zu diesem Mehrgenerationenkern gehörenden Personen leben im Haushalt ausschließlich Eltern, Kinder, Geschwister, Partner sowie Geschwister der Partner. ▪ <u>Anderer Haushaltstyp</u>: Haushalt, der keinem der anderen definierten Haushaltstypen zugeordnet werden konnte. ▪ <u>Generierung nicht möglich (fehlende Werte)</u>: 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)

Tabelle 44: Bedarfsgemeinschaftsnummer, Welle 11

Variablenname:	<i>bgnr11</i>
Variablenlabel:	Bedarfsgemeinschaftsnummer in Welle 11 (2017)
Ausgangsvariablen:	Haushaltsinformationen zu Alter und Verwandtschaftsbeziehungen
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Gerrit Müller
Erläuterung:	Die Variable <i>bgnr11</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 Bedarfsgemeinschaft. Die Variable <i>bgnr11</i> setzt sich aus der bekannten Haushaltsnummer sowie einem zweistelligen Bedarfsgemeinschafts-Indikator zusammen.

Tabelle 44: Bedarfsgemeinschaftsnummer, Welle 11 (Fortsetzung)

	Zur Identifikation der Bedarfsgemeinschafts-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 Bedarfsgemeinschaften 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 und der entsprechenden Zugehörigkeit von Personen zu dieser im Rahmen einer Befragung nicht direkt zu erheben ist. Inhaltlich erfolgt die Zuordnung von Personen zu Bedarfsgemeinschaften auf Basis der aktuell gültigen Fassung des § 7 Abs. 3 SGB II (zuletzt geändert am 29. Dezember 2016). Danach bilden grundsätzlich alle Personen, die das 25. Lebensjahr vollendet und das 65. Lebensjahr noch nicht vollendet haben, jeweils eine eigene Bedarfsgemeinschaft; es sei denn, diese Personen leben in einer Partnerschaft und/oder haben ein Kind(ern), 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 Bedarfsgemeinschaft. Sollten zwei Personen mit einem gemeinsamen Kind im gleichen Haushalt leben, aber im HH-Grid nicht angeben, Partner voneinander zu sein, wird dies im Sinne des §7 Abs. (3a) trotzdem unterstellt und die entsprechenden Personen gemeinsam mit dem Kind(ern) zu einer Bedarfsgemeinschaft 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 Bedarfsgemeinschaft.
--	--

Tabelle 44: Bedarfsgemeinschaftsnummer, Welle 11 (Fortsetzung)

	Personen, die das 65. Lebensjahr vollendet haben, fallen nicht in den Geltungsbereich des SGB II und gehören daher keiner Bedarfsgemeinschaft 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 Bedarfsgemeinschaft an (Code 0). Sie fallen in den Geltungsbereich des SGB XII. Für Haushalte mit fehlenden Beziehungs- bzw. Altersangaben zu einzelnen Personen wurde keine Bedarfsgemeinschaft -Zuordnung vorgenommen, sondern für alle Haushaltsmitglieder Code 99 vergeben. Diese Haushalte können approximativ als Ein-Bedarfsgemeinschafts-Haushalte interpretiert werden.
Literatur:	Sozialgesetzbuch, Zweites Buch - Grundsicherung für Arbeitssuchende (SGB II)

Tabelle 45: Bedarfsgemeinschaftstypologie, Welle 11

Variablenname:	<i>bgtyp11</i>
Variablenlabel:	Typ der Bedarfsgemeinschaft in Welle 11 (2017)
Ausgangsvariablen:	Haushaltsinformationen zu Alter und Verwandtschaftsbeziehungen
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Gerrit Müller

Tabelle 45: Bedarfsgemeinschaftstypologie, Welle 11 (Fortsetzung)

Erläuterung:	Der Bedarfsgemeinschaftstypologie liegt das gleiche Konzept der synthetischen Bedarfsgemeinschaft wie bei der Variable <i>bgnr11</i> zugrunde. Kinder gehören bis zur Vollendung des 25. Lebensjahres der Bedarfsgemeinschaft ihrer Eltern an, sofern sie nicht selbst Partner oder Kinder 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. Code 0, keine Bedarfsgemeinschaft, wurde vergeben, wenn eine oder mehrere Personen nicht in den Geltungsbereich des SGB II fallen (vgl. Code 0 bei <i>bgnr11</i>). Code -5, Generierung nicht möglich (fehlende Werte), wurde für Haushalte mit fehlenden Beziehungs- bzw. Altersangaben zu einzelnen Personen (vgl. Code 99 bei <i>bgnr11</i>) vergeben.
Literatur:	-

Tabelle 46: Arbeitslosengeld II-Bezug der Bedarfsgemeinschaft zum Zeitpunkt der Stichprobenziehung, Welle 11

Variablenname:	<i>bgbez11</i>
Variablenlabel:	ALG2-Bezug der Bedarfsgemeinschaft bei Stichprobenziehung in Welle 11 (2017)
Ausgangsvariablen:	<i>HA0250*</i> , <i>HA0300</i> , <i>AL20100</i> , <i>AL20200</i> , <i>AL20300</i> , <i>AL20400</i> , <i>AL20610</i> , <i>AL20710*</i> , <i>HA0400</i> , <i>sample</i> , <i>hnr</i> , <i>bgnr11</i> , <i>hhgr</i>
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Mark Trappmann
Erläuterung:	Diese Variable gibt für jede Bedarfsgemeinschaft, die entsprechend dem bei Variable <i>bgnr11</i> beschriebenen Vorgehen gebildet wurde, an, ob diese Bedarfsgemeinschaft zum Zeitpunkt der Stichprobenziehung für Welle 11 Arbeitslosengeld II bezogen hat oder nicht.
Literatur:	-

Tabelle 47: Arbeitslosengeld II-Bezug der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 11

Variablenname:	<i>bgbezb11</i>
----------------	-----------------

Tabelle 47: Arbeitslosengeld II-Bezug der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 11 (Fortsetzung)

Variablenlabel:	ALG2-Bezug d. Bedarfsgemeinschaft zum Befragungszeitpkt. in Welle 11 (2017)
Ausgangsvariablen:	<i>AL20610, AL20710*, zensiert (alg2_spells), sample, hhgr, bgnr11</i>
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Daniel Gebhardt
Erläuterung:	Diese Variable gibt für jede Bedarfsgemeinschaft, die entsprechend dem bei Variable <i>bgnr10</i> beschriebenen Vorgehen gebildet wurde, an, ob diese Bedarfsgemeinschaft zum Befragungszeitpunkt in Welle 10 Arbeitslosengeld II bezogen hat oder nicht.
Literatur:	-

Tabelle 48: Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 10

Variablenname:	<i>bgbezbkorr10</i>
Variablenlabel:	Korrektur des ALG2-Bezug d. BG zum Befragungszeitpkt. in W10 (2016)
Ausgangsvariablen:	<i>hintmon, hintjahr, kennungfbvers (HHENDDAT), AL20100-AL20400, AL20610, AL20710*, zensiert (alg2_spells), sample, hhgr, bgnr11, bgbezb10</i>
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Christian Dickmann
Erläuterung:	Siehe Abschnitt unten
Literatur:	-

Tabelle 49: Flag zur Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 10

Variablenname:	<i>bgbezbkorrflag10</i>
Variablenlabel:	Flag z. Korrektur des ALG2-Bezug d. BG zum Befragungszeitpkt. in W10 (2016)
Ausgangsvariablen:	<i>kennungfbvers (HHENDDAT), sample, bgbezb10, bgbezbkorr10</i>
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Christian Dickmann
Erläuterung:	Siehe Abschnitt unten
Literatur:	-

Aufgrund der Panelstruktur eignen sich die PASS-Daten insbesondere für Analysen von Übergängen im SGB II. Mit den generierten Variablen *bgbezs** und *bgbezb** stehen dazu im Personenregister zwei Variablen zur Verfügung, die den ALG-II-Bezugsstatus auf Personenebene zu unterschiedlichen Zeitpunkten wiedergeben. Dabei enthält *bgbezs** den Bezugsstatus zum Stichprobenziehungszeitpunkt, *bgbezb** den zum Befragungszeitpunkt. Die Variable *bgbezb** wird in allen Wellen für alle Teilstichproben aus den Befragungsangaben generiert und ist damit über die gesamte Zeit hinweg vergleichbar erhoben. Auch die Variable *bgbezs** wird in allen Wellen für alle Teilstichproben aus den Befragungsangaben gebildet. Für alle Zugangsstichproben, die aus Grundsicherungsregistern der Bundesagentur für Arbeit gezogen werden (alle Teilstichproben außer den beiden Bevölkerungstichproben, *sample=2* und *sample=6*), wird jedoch in der jeweils ersten Befragungswelle eines Haushalts die Registerinformation als Korrekturfaktor herangezogen. Das heißt, sie ist in der Erstbefragung bei jedem Haushalt dieser Stichproben mindestens für eine Bedarfsgemeinschaft auf eins (Bedarfsgemeinschaft bezieht Grundsicherungsleistungen) gesetzt, auch falls die Angabe in der Befragung davon abweicht. In den folgenden Wellen wird dann auch diese Variable ausschließlich auf der Basis von Befragungsangaben gebildet. Aufgrund der unterschiedlichen Herkunft der Variablen, empfiehlt es sich Dynamik in der Grundsicherung entweder direkt mit den Spelldaten zum Grundsicherungsbezug zu untersuchen oder mittels der Variable *bgbezb**. Will man die Variable *bgbezs** mit einbeziehen, so sollte die erste Befragungswelle eines jeden Haushalts nicht verwendet werden, da sonst die Gefahr besteht, dass mögliche Messunterschiede zwischen administrativen Daten und Befragungsdaten mit der wahren Veränderung konfundiert sind. Zu diesen Messunterschieden gibt es auf Basis der PASS-Daten inzwischen eine umfangreiche Literatur (siehe Bruckmeier et al. (2014); Bruckmeier et al. (2015); Eggs (2016); Kreuter et al. (2010); Kreuter et al. (2014)).

Während der Feldzeit der Welle 11 ließen Auswertungen der bis dahin verfügbaren Daten der Welle 10 und Rückmeldungen der Interviewer aus dem Feld vermuten, dass die Frage zum ALG-II-Leistungsbezug im Haushaltsfragebogen von Personen in der Teilstichprobe der syrischen und irakischen Haushalte teilweise missverständlich aufgefasst wurde. Im Vergleich zu den anderen BA-Zugangsstichproben (aus vorherigen Wellen oder derselben Welle ohne die syrischen und irakischen Haushalte) liegt der Anteil an Haushalten, die angeben niemals ALG II erhalten zu haben, auffällig hoch.

Um diesem Problem entgegenzukommen, wurden im laufenden Feld (13 Wochen nach Feldstart und 3 Wochen nach Start des Fremdsprachenfelds sowie der neuen BA-Zugangsstichproben) Änderungen im ALG-II-Modul vorgenommen. Die Änderungen betrafen ausschließlich die Teilstichprobe der syrischen und irakischen Haushalte (*sample = 14* oder *17*). Für diese Gruppe wurde der Einführungstext am Anfang des ALG-II-Moduls (HABLK01) um eine zusätzliche Erläuterung erweitert sowie ein Interviewerhinweis in der Frage HA0300 aufgenommen. Die genauen Änderungen lassen sich dem Haushaltsfragebogen der Welle 11 entnehmen. An entsprechender Stelle existieren zwei Versionen. Ver-

sion 1 beinhaltet das Frageprogramm vor der Änderung (während der laufenden Feldzeit), Version 2 das Frageprogramm danach. Anhand der Variable *kennungfbvers* im Haushaltssenddatensatz (HHENDDAT) lässt sich nachvollziehen, welche Frageversion im Haushaltinterview gestellt wurde.

Aus dieser Änderung im Fragebogen ergeben sich Besonderheiten für die Datenaufbereitung der Angaben zum ALG-II-Bezug. Die bestehenden Aufbereitungsregeln für die Angaben der Panel-Haushalte der syrischen und irakischen Teilstichprobe zum Beginn des ALG-II-Bezugs aus der Welle 11 werden beibehalten. In den generierten Variablen *bmonat* und *bjahr* im ALG-II-Spellldatensatz (*alg2_spells*) wird das Beginndatum des ALG-II-Bezugs weiterhin auf das Datum des vorherigen Interviews gesetzt, falls die Angabe in der Befragung zeitlich davor liegt. Die tatsächlichen Angaben zum Bezugszeitraum bleiben für den Nutzer in den Variablen *AL20100* und *AL20200* erkennbar. Die bereits im Scientific-Use-File der Welle 10 im Personenregister (*p_register*) ausgelieferte Variable *bgbezb10* wird nicht korrigiert. Stattdessen wird im Scientific-Use-File der Welle 11 eine neue Variable *bgbezb10_korr* generiert, in der für die Teilstichprobe zusätzlich zu den Angaben aus Welle 10 auch die in Welle 11 gemachten Angaben zur Bestimmung des ALG-II-Bezugs zum Befragungszeitpunkt verwendet werden. Falls im Haushaltsinterview der Welle 11 ein Bezug für den Zeitpunkt des Haushaltsinterviews der Welle 10 berichtet wurde, wird dieser in der Variable *bgbezb10_korr* aufgenommen. Ob eine solche Korrektur vorliegt, lässt sich der zusätzlichen Variable *bgbezbkorrflag10* entnehmen. Für Haushalte, die in Welle 11 die Teilnahme aussetzen oder die in Welle 11 noch die Frageversion 1 erhalten haben, werden die zukünftigen Angaben aus Welle 12 zusätzlich berücksichtigt, um diese im Scientific-Use-File der Welle 12 in den Variablen *bgbezb10_korr* und *bgbezb11_korr* zu berücksichtigen.

Tabelle 50: Anzahl der Bedarfsgemeinschaften im Haushalt

Variablenname:	<i>anzbg</i>
Variablenlabel:	Anzahl der synthetischen Bedarfsgemeinschaften im HH, generiert
Ausgangsvariablen:	<i>bgnr11</i> , <i>hnr</i>
Typ / Datensatz:	Bedarfsgemeinschaft / Haushaltsdatensatz
Bearbeitung:	Daniel Gebhardt
Erläuterung:	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>bgnr11</i> beschriebenen Vorgehensweise.
Literatur:	-

Tabelle 51: Anzahl der zum Stichprobenzeitpunkt Leistungen beziehenden Bedarfsgemeinschaften im Haushalt

Variablenname:	<i>nbgbezug</i>
Variablenlabel:	Anzahl zum Stichprobenzeitpunkt beziehender BGs im HH
Ausgangsvariablen:	<i>bgbzs11, bgnr11, hnr</i>
Typ / Datensatz:	Bedarfsgemeinschaft / Haushaltsdatensatz
Bearbeitung:	Daniel Gebhardt
Erläuterung:	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>bgbzs11</i> aus dem Personenregister beziehenden Bedarfsgemeinschaften über die Haushaltsnummer gebildet.
Literatur:	-

5 Datenaufbereitung

Die Datenaufbereitung wird seit Welle 3 nicht vom IAB selbst, sondern von infas durchgeführt. Um 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 Entwicklung des in Welle 4 erstmals erstellten *bio_spells*-Datensatzes, 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 Welle 11 erhobenen Informationen liegen bei infas zunächst als ASCII-Daten vor. In einer ersten Stufe erstellt infas aus diesen Rohdaten die folgenden Datensätze³²:

- Haushaltsdatensatz für im Querschnitt erhobene Fragen inklusive der später in Spellform umgebrochenen Fragen des Moduls „Kinderbetreuung“
- Haushaltsdatensatz für im Längsschnitt erhobene Daten (Modul „Arbeitslosengeld II“)
- Datensatz über die Fortschreibung der Haushaltszusammensetzung (Matrix)
- Datensatz über die Fortschreibung der Verwandtschaftsbeziehungen im Haushalt (Beziehungsmatrix)
- Personen-/Seniorentatsatz für im Querschnitt erhobene Fragen
- Personendatsatz für im Längsschnitt erhobene Daten I (Modul „Erwerbsbiografie [Spells]“)
- Personendatsatz für im Längsschnitt erhobene Daten II (Modul „Maßnahmen“)
- Datensatz für offene Texte (übergreifend über alle Haushalts-, Personen- und Seniorenbefragungen)

In einer zweiten Stufe werden die Daten einer ausführlichen formalen und inhaltlichen Prüfung unterzogen und zum Scientific Use File aufbereitet. Darüber hinaus liefert infas einen Bruttodatsatz sowie weitere Spezialdatensätze, die sich jedoch nicht direkt aus den eigentlichen Erhebungsinstrumenten ergeben.

Die sich anschließende Datenprüfung bei infas unterteilte 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

³² Im Rahmen der Datenaufbereitung kamen die Softwarepakete Stata Version 13 sowie SPSS Version 25 zum Einsatz.

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

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

Tabelle 52: Übersicht über die Schritte der Datenaufbereitung in der Welle 11 des PASS

Nr.	Arbeitsschritt
1	Einlesen der erhobenen Rohdaten in Arbeitsdatensätze
2	Prüfung der Haushaltsstruktur (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 eine Strukturprüfung statt. Dabei sollten zunächst Interviews identifiziert werden, die im Sinne des PASS als nicht realisiert gelten und die 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 Welle 11 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. Die entsprechenden Fälle wurden in einem formalisierten Verfahren zwischen infas und dem IAB diskutiert. Abschließend wurde dann eine endgültige Entscheidung im IAB getroffen, wie mit diesen Fällen zu verfahren ist. Es sollte beachtet werden, dass die folgende Darstellung den Umfang der durchgeführten Prüfungen angibt. Es führen nicht alle Prüfungen in allen Wellen zur Identifikation von Problemen, sondern das Ergebnis einer Prüfung besteht üblicherweise darin, dass ein geprüfter Sachverhalt in geringer Fallzahl oder gar nicht auftritt. Zudem werden bekannte Fehlerquellen möglichst schon während des Interviews abgefangen. So sieht das Erhebungsinstrument beispielsweise vor, dass aus einem Panel-HH nicht alle bisher dort bekannten Zielpersonen gleichzeitig ausziehen können und dass unter den nach Auszügen verbleibenden Personen mindestens eine Person 15 Jahre oder älter sein muss.

- 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.
- Weiterhin wurde 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 ak-

tuellen 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.

- Generell gilt, dass das Geburtsdatum aus dem Personen-/Senioreninterview der aktuellen Welle alle anderen Altersangaben zu dieser Person z.B. aus dem Haushaltsgrid sticht und Basis aller Generierungen ist, die u.a. auf dem Alter basieren. In einer speziellen Konstellation wird die Angabe des Geburtsdatums in *PD0100* jedoch korrigiert. Wenn sich das Geburtsjahr einer Person gemäß *PD0100* deutlich ändert, die Angaben zum Tag und Monat jedoch unverändert bleiben, das bisher laut *PD0100* bekannte Geburtsdatum sich noch nie geändert hat und mindestens zwei Angaben zum Geburtsdatum aus *PD0100* aus Vorwellen vorliegen, dann wird das Geburtsjahr unter Berücksichtigung der gesamten Haushaltskonstellation auf den aus den Vorwellen bekannten Wert zurückgesetzt. Ein theoretisches Beispiel ist eine Person, deren Geburtsdatum aus mindestens zwei Vorwellen als 01.02.1972 bekannt ist, deren Geburtsdatum nun als 01.02.1992 erfasst wird und die damit jünger wäre als die ebenfalls im Haushalt lebenden Kinder. Eine solche Konstellation würde ohne Korrektur zu einer Unplausibilität in der Beziehungsstruktur führen und in der Konsequenz dazu, dass z.B. die synthetischen Bedarfsgemeinschaften nicht generiert werden können. Daher wird die Angabe im Beispielfall in der aktuellen Welle auf den 01.02.1972 korrigiert.
- 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 vorlag³³.
- Breiten Raum nehmen auch die Prüfungen von Ein- und Auszügen ein. Panelhaushalte, in denen Auszüge berichtet wurden, wurden grundsätzlich auf ihren Haushaltskontext sichtgeprüft 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 Jahren, 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 Ausziehenden 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 Nichtrealisierung des Splithaushalts entschieden, die Auszüge als plausibel anzuerkennen, alle Ausgezogenen jedoch nachträglich in einem

³³ Bei neubefragten Haushalten, für die zwar ein Haushalts- aber kein gültiges Personeninterview vorlag, wurden die Haushaltsinterviews analog zum Vorgehen seit Welle 1 entfernt. Bei wiederbefragten und abgespaltenen Haushalten wurden die Haushaltsinterviews dagegen beibehalten.

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 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 Splithaushalten wird bei Personen, die laut Feld nicht aus dem Panelhaushalt bekannt sind, sondern über den Splithaushalt neu zum PASS hinzukommen, geprüft, ob sie nicht doch aus dem Panelhaushalt stammen. Es bestehen zwei Konstellationen, die solche Fälle begünstigen. Zum einen kommt es vor, dass im Panelhaushalt bei mehreren ausziehenden Personen berichtet wird, dass die Splitpersonen mehr als einen Splithaushalt bilden. In diesem Fall werden dynamisch für das laufende Feld Preload für alle durch den Panelhaushalt benannten Splithaushalte erzeugt. Wenn jedoch tatsächlich Personen, die laut Panelhaushalt in verschiedenen Splithaushalten leben, in einem gemeinsamen Splithaushalt angetroffen werden, haben jene Personen, die durch den Panelhaushalt nicht diesem, sondern einem anderen Splithaushalt zugerechnet wurden, in diesem Splithaushalt keinen Preload und werden als neue Person erfasst.
- Zum anderen ist es möglich, dass Personen aus einem Panelhaushalt aus- und in einen Haushalt einziehen, der in einer Vorwelle als Splithaushalt entstand und seinerzeit bereits realisiert wurde. Es findet also nach der damaligen Abspaltung des Splithaushalts ein weiterer Umzug aus dem ursprünglichen Panelhaushalt in diesen Splithaushalt statt. Unabhängig davon, ob der Panelhaushalt, aus dem der betroffene Splithaushalt hervorgegangen ist, in der Welle des erneuten Umzugs vom Panel in den Splithaushalt realisiert wurde, ist es feldseitig nicht vorgesehen, solche Fälle steuern zu können. Hierzu müssten dem Splithaushalt die Personeninformationen aller Personen des Panelhaushalts (und ggf. aller Personen in weiteren Splithaushalten dieses Panelhaushalts) als Preload hinterlegt werden. Die wenigen zu erwartenden Fälle, in denen eine solche Konstellation eintritt, rechtfertigen einen solchen Aufwand im Feld nicht. Stattdessen müssen solche Fälle in den Strukturprüfungen entdeckt werden. In diesem Zusammenhang ist anzumerken, dass Splithaushalte in Bezug auf die Strukturprüfungen auch in den Folgewellen ihrer Erstrealisierung in ihrer Eigenschaft als Splithaushalt betrachtet werden müssen, auch wenn sie nach der Erstrealisierung in der Feldsteuerung als Panelhaushalte gelten.

In beiden Fällen wird die *pnr* der betroffenen Personen im Splithaushalt nachträglich korrigiert. Zu beachten ist auch, dass solche Personen in ihren Personen-/Senioren-

interviews als Neubefragt gesteuert werden, obwohl sie u.U. altbefragte Personen sind. Diese Abweichung wird grundsätzlich nicht korrigiert, siehe hierzu auch Kapitel 4.4.

- 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 Personen 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 ALG-II-Kürzung auf die im Rahmen der Strukturprüfung als richtig definierte Position. Umgekehrt wird auch geprüft, ob Personen, die im Feld als wiedereinziehend markiert werden, tatsächlich mit der in einer Vorwelle ausgezogenen Person identisch sind. Falls nicht, handelt es sich um den Einzug einer neu zum PASS hinzukommenden Person. Auch dann werden die beschriebenen Änderungen in der Haushaltsstruktur vorgenommen.
- Bei Wiedereinzügen wird geprüft, ob der Splithaushalt, in dem die Person vor dem Wiedereinzug in den Panelhaushalt lebte, in der aktuellen Welle realisiert wurde und ob in diesem Fall die wiedereinziehende Person im Splithaushalt als ausgezogen bezeichnet wird. Auch Personen, die in einer Vorwelle in ihren Panelhaushalt zurückgezogen sind, müssen weiterhin mit Blick auf ihren Status im Splithaushalt überprüft werden, solange der Splithaushalt zur aktuellen Panelstichprobe zählt. Wenn eine wiedereinziehende Person weiterhin in ihrem Splithaushalt als aktuelles Haushaltsmitglied geführt wird, wurde in der Datenaufbereitung für die betroffenen Fälle entschieden, ob dies als plausibel anerkannt wurde oder ob die Haushaltsstruktur des Panel- oder des Splithaushalts bereinigt werden musste.
- Nicht nur Wiedereinzüge können dazu führen, dass eine Person als aktuelles Haushaltsmitglied mehrerer Haushalte gilt. Es kann auch vorkommen, dass eine Person als Mitglied eines Splithaushalts benannt wird, obwohl sie im Panelhaushalt nicht als ausgezogen benannt wurde. Unter Prüfung der Haushaltskonstellationen der beteiligten Haushalte kann dies im Einzelfall als plausibel anerkannt werden. Solche Fälle werden in den *zdub**-Variablen im Personenregister dokumentiert. Weitere Erläuterungen hierzu sind in Kapitel 4.4 dieses Dokuments und Kapitel 5.4.1.2 des Datenreports zur Welle 5 (vgl. Berg et al. (2012)) enthalten.
- Im Verhältnis zwischen einem Panelhaushalt und seinen Splithaushalten können weitere Punkte auftreten. Es ist z.B. denkbar, dass Personen, die über einen Splithaushalt neu zum PASS hinzugekommen sind, in den Panelhaushalt umziehen.

Oder es ist denkbar, dass Personen von einem Splithaushalt in einen anderen umziehen. Grundsätzlich müssen alle Personen in einem Panelhaushalt und alle mit ihm verbundenen Splithaushalte als Netzwerk betrachtet werden. Die Strukturprüfungen sind so ausgelegt, dass Personenbewegungen zwischen den Haushalten eines solchen Netzwerks entdeckt werden, unabhängig von der Richtung, in der sich eine Person im Netzwerk bewegt.

- 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 in geringem Umfang geprüft werden. Bei erstbefragten Haushalten wird anhand der Angaben zu Vorname, Alter und Geschlecht betrachtet, ob einzelne Haushaltsmitglieder mehrfach aufgelistet werden. In diesem Fall wird von den doppelt berichteten Personen nur die zuerst berichtete Haushaltsposition behalten, die weiteren Haushaltspositionen werden 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 Haushaltsmitgliedern 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 auch ein Verschieben der personenbezogenen Angaben im Haushaltsinterview.
- Durch die Rückmeldung eines Feldinterviewers wurde in Welle 4 ein Haushalt entdeckt, der doppelt in der Panelstichprobe enthalten war. Der Haushalt 10015439 ist als identischer Haushalt 15044862 bereits seit Welle 1 doppelt in der Stichprobe. Beide Haushalte wurden jeweils in Welle 1 und in Welle 3 realisiert und jeweils in Welle 2 nicht realisiert. In Welle 4 wurde der Haushalt 10015439 realisiert. Da dem CAPI-Interviewer dieses Points „beide“ Haushalte zugeteilt worden waren, ist die Dopplung aufgefallen. Die Haushaltszusammensetzung blieb in beiden Haushalten über die Wellen hinweg unverändert. Der in Welle 4 nicht realisierte Haushalt 15044862 wird aus der Welle 5-Stichprobe gelöscht. Es findet keine rückwirkende Entfernung der Dublette in den Wellen 1 bis 3 statt, da bspw. Auswirkungen auf die Gewichtung die Folge wären. Im *hh_register* ist der Dublettenhaushalt in Variable *hnettod4* mit dem Code 26 „HH nicht realisiert, Dublette“ gekennzeichnet, der den Grund für die Nichtrealisierung transparent macht. Im *p_register* sind die Haushaltsmitglieder des Dublettenhaushalts in der Variable *pnettod4* entsprechend mit Code 56 „HH nicht realisiert, Dublette“ gekennzeichnet.
- 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 Welle 11 unklar war, wurden alle Interviews der Welle 11 entfernt. Sie werden in der Welle 12 wie Haushalte behandelt, die in der Welle 11 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 Welle 11 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 Vorwellenwert + 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 Auffrischungstichprobe unterschieden. Die als nicht realisiert geltenden Haushalte aus der Auffrischungstichprobe wurden analog zum Vorgehen in den Vorwellen entfernt. Bei wiederbefragten Haushalten ohne Interviews auf Personenebene wurde das Haushaltsinterview dagegen nicht gelöscht.

Einen Hinweis auf entfernte Interviews geben die Nettovariablen (*hnettok11*, *hnettod11*, *pnettok11*, *pnettod11*) in den Haushalts- und Personenregisterdatensätzen. Über die entsprechenden Variablen im Haushaltsregister ist nachvollziehbar, für welche wiederbefragten Haushalte die Haushaltsinterviews nachträglich entfernt wurden. Mittels 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 Auffrischungstichprobe der Welle 11, für die nicht mindestens ein gültiges Haushalts- und Personeninterview vorliegt, sind gelöschte Interviews nicht in den Registerdatensätzen 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 hätten³⁴. 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 Haushaltssstrukturprüfungen sein. Wenn der Auszug einer Person nachträglich als unplausibel verworfen 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 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 nachfolgender Tabelle 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.

³⁴ Die Bereinigung erfolgte dabei, wie in solchen Fällen üblich, immer von den zuerst erhobenen zu den später erfragten Items.

Tabelle 53: 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“ ³⁵

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 Welle 2 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 Welle 2 in keiner Fragebogenversion mehr erfragt wurden, für die Beobachtungen dieser Welle den Wert „-9“. Ebenso verhält es sich mit Beobachtungen der Welle 1. Erstmals nach der Welle 1 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. bis Welle 3 auch 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 Vorwellen durchgeführten Prüfungen für die aktuelle Welle angepasst und sinnvoll erweitert. Daneben wurden auch die Haushaltsstruktur sowie die Spelldaten plausibilitätsgeprüft – insbesondere auf unzulässige Überschneidungen innerhalb der einzelnen Spelltypen. Dabei wurden grundsätzlich nur die im Querschnitt der Welle 11 erhobenen Daten für die

³⁵ Ab Welle 4 dient Code „-10“ ausschließlich der Unterscheidung zwischen Personen- und Seniorenfragebogen. Bis einschließlich Welle 3 wurde zusätzlich auch auf der Haushaltsebene zwischen erst- und wiederbefragten Haushalten unterschieden. Die Unterscheidung auf Haushaltsebene entfällt ab Welle 4 infolge der Zusammenfassung der bisher getrennten Fragebogenversionen zu einem einzigen gemeinsamen Haushaltsfragebogen.

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:

- Widerspruchsprüfung: Generell wurden Widersprüche nur dann korrigiert, wenn entweder die fehlende Plausibilität als besonders gravierend bzw. der Eingriff als vergleichsweise mild einzustufen war. Letzteres galt z.B., wenn nur eine geringe Zahl von Fällen betroffen war oder wenn nur ein Missing Code (z.B. „-3“) durch einen neuen („-8“) ersetzt wurde. Für die Bereinigung unplausibler Angaben wurden zwei Strategien verwendet: 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.
- Plausibilitätsprüfung der Haushaltsstruktur: Auf Grundlage der im Haushaltsinterview erhobenen Information zu den Verwandtschaftsbeziehungen der Haushaltsmitglieder untereinander sowie der Angaben zu Alter, Geschlecht und Vorname 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 ähnlich unplausible Kombination von Beziehungen identifiziert, so wurde auch hier versucht, die Beziehung im Rahmen des Haushaltskontexts 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 Bedarfsgemeinschaft ist.
- 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 Wellen 1 bis 10

Im Rahmen der Datenaufbereitung für den Scientific Use File (SUF) der Welle 11 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 Files der Welle 10 festgestellt wurden. Die korrigierten Daten können stets in den SUF-Datensätzen der aktuellen Welle, also derzeit Welle 11, genutzt werden. Die folgenden fünf Tabellen geben einen Überblick über die rückwirkenden Veränderungen an bereits ausgelieferten Wellen des PASS³⁶.

Tabelle 54: Übersicht über rückwirkenden Veränderungen im Haushaltsdatensatz (HHENDDAT, KINDER)

geänderte Variable	betroffener Datensatz	geänderte Welle	Art der Änderung	Beschreibung der Änderung
-	-	-	-	-

³⁶ Anpassungen an Werte- oder Variablenlabels sind hier nur berücksichtigt, wenn sie dadurch die Interpretation von Variablen oder Werten verändert.

**Tabelle 55: Übersicht über rückwirkenden Veränderungen im Personendatensatz
(PENDDAT)**

geänderte Variable	betroffener Datensatz	geänderte Welle	Art der Änderung	Beschreibung der Änderung
<i>PP1600</i>	<i>PENDDAT</i>	10	Korrektur	Die Interviews mit den Seniorenfragebogen (fb_vers=3) werden auf Sondercode -10 statt auf Sondercode -3 gesetzt.
<i>brancheminj2</i>	<i>PENDDAT</i>	10	Korrektur	Die Branchenangabe wird bei den Minijobs nicht erhoben, wenn der Befragte laut eigener Angabe in einem Privathaushalt beschäftigt ist. Wenn sich aus der Codierung der Branche heraus dennoch ergibt, dass es sich um einen Privathaushalt gehandelt haben soll, wird dies als unplausibel gewertet und die Angabe auf den Sondercode -8 umgesetzt. Der Fehler bestand darin, dass Tätigkeiten in Privathaushalten in der WZ2003-Codierung (<i>brancheminj1</i>) auf Code 95 abgelegt werden, in der WZ2008-Codierung (<i>brancheminj2</i>) hingegen auf Code 97. Dieser Unterschied wurde nicht beachtet und so wurde auch in <i>brancheminj2</i> der Code 95 (Reparatur von Datenverarbeitungsgeräten und Gebrauchsgütern) auf Sondercode -8 umgesetzt, während Code 97 unbereinigt blieb. In der Korrektur wurden zwei Fälle vom fälschlich vergebenen Sondercode -8 auf Code 95 sowie acht fälschlich unbereinigte Fälle von Code 97 auf Sondercode -8 umgesetzt.

Tabelle 55: Übersicht über rückwirkenden Veränderungen im Personendatensatz (PENDDAT) (Fortsetzung)

<i>PET1000*, alakt, alg1abez, PET0920, etakt, statakt</i>	<i>PENDDAT</i>	10	Korrektur	In den bio_spells wurden in Welle 10 in 63 Fällen doppelt angegebene Arbeitslosigkeits-Spells zwar zunächst korrekt identifiziert, dann jedoch beide Spells statt nur des überzähligen doppelten entfernt. Dies hat auch Auswirkungen auf einige generierte Variable in PENDDAT. In <i>PET1000*</i> werden 63 Fälle vom Sondercode -3 auf inhaltliche Angaben umgesetzt. In <i>alakt</i> werden 17 Fälle von Sondercode -3 und 46 Fälle von Code 2 auf den Code 1 umgesetzt. In <i>alg1abez</i> werden vier Fälle von Code 0 auf Sondercode -5, weitere fünf Fälle von Sondercode -5 auf Code 0 und ein Fall von 0 auf 1 umgesetzt. In <i>PET0920</i> werden 61 Fälle vom Sondercode -3 auf inhaltliche Angaben umgesetzt. In <i>etakt</i> werden 17 Fälle von Sondercode -3 auf Code 2 umgesetzt. In <i>statakt</i> werden 63 Fälle von anderen inhaltlichen Codes auf Code 2 umgesetzt.
<i>zpalthh, palter</i>	<i>PENDDAT</i>	5-6	Korrektur	Die Verweigerung der Altersangabe wurde in Welle 5 in einem Fall als Alter von 99 Jahren fehlinterpretiert und in Welle 6 auf das Alter 100 fortgeschrieben. Beide Werte wurden für diese Person auf den Sondercode -2 umgesetzt.

Tabelle 56: Übersicht über rückwirkenden Veränderungen in Spelldatensätzen
(*bio_spells, alg2_spells, ee_spells*)

geänderte Variable	betroffener Datensatz	geänderte Welle	Art der Änderung	Beschreibung der Änderung
Alle AL-spezifischen Spellvariablen, <i>spellnr</i> , <i>spellnral</i>	<i>bio_spells</i>	10	Korrektur und Ergänzung	Durch einen Fehler in der Plausibilitätsprüfung der Arbeitslosigkeits-Spells wurden in der Aufbereitung der Welle 10 versehentlich 63 Arbeitslosigkeits-Spells zu viel als unplausibel markiert und in der Folge aus dem Datensatz gelöscht bzw. nicht korrekt fortgeschrieben. Dies hatte zur Folge, dass im Bio-Spelldatensatz der Welle 10 ursprünglich 51 Arbeitslosigkeits-Spells zu wenig enthalten waren sowie 12 weitere Spells nicht richtig fortgeschrieben wurden. Die fehlenden Spells wurden ergänzt. Die nicht richtig fortgeschriebenen Spells wurden korrigiert. Die Spellnummerierung wurde anschließend neu generiert. Von der Korrektur sind daher neben den AL-spezifischen Variablen in den 63 korrigierten Spells auch die <i>spellnr</i> und <i>spellnral</i> anderer Spells betroffen.

Tabelle 57: Übersicht über rückwirkenden Veränderungen in Registerdatensätzen
(*hh_register; p_register*)

geänderte Variable	betroffener Datensatz	geänderte Welle	Art der Änderung	Beschreibung der Änderung
-	-	-	-	-

Tabelle 58: Übersicht über rückwirkenden Veränderungen in den Gewichtungsdatensätzen (*hweights*; *pweights*)

geänderte Variable	betroffener Datensatz	geänderte Welle	Art der Änderung	Beschreibung der Änderung
-	-	-	-	-

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*³⁷, *bgtyp*³⁸), Indikatorvariablen zu Partnern im Haushalt (*apartner*; *epartner*³⁹), Zeigervariablen auf Eltern und Partner im Haushalt (*zmhh*; *zvhh*; *zparthh*⁴⁰) sowie verschiedene Indikatorvariablen, die anzeigen ob Eltern (*mhh*; *vhh*⁴¹) oder Kinder der Zielperson (z.B. *ekind*⁴²) im Haushalt leben. Nachfolgende Tabelle gibt einen Überblick über die betroffenen Variablen und die vorgenommenen Anonymisierungsschritte⁴³ im Personendatensatz. In den beiden darauf folgenden Tabellen sind die anonymisierten Variablen des Erwerbstätigkeitsspelldatensatzes und des KINDER-Datensatzes dargestellt.

³⁷ Enthalten im Haushaltsdatensatz (*HHENDDAT*), siehe Kapitel 4.5.2.

³⁸ Wellenspezifische Variablen enthalten im Personenregister (*p_register*), siehe Kapitel 4.4.

³⁹ Enthalten im Personendatensatz (*PENDDAT*), siehe Kapitel 4.4.

⁴⁰ Wellenspezifische Variablen enthalten im Personenregister (*p_register*), siehe Kapitel 4.4.

⁴¹ Enthalten im Personendatensatz (*PENDDAT*), siehe Kapitel 4.4.

⁴² Enthalten im Personendatensatz (*PENDDAT*), siehe Kapitel 4.4.

⁴³ Falls für Ihr Forschungsvorhaben die Verwendung der nicht anonymisierten Versionen einer oder mehrerer 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 59: Übersicht über die in Welle 11 anonymisierten Variablen im Personen-
datensatz (*PENDDAT*)**

Varname	Variablenlabel	Vorgehen
<i>PD0100</i>	Geburtsdatum (anon.)	Das genaue Geburtsdatum wurde zum Geburtsjahr verkürzt.
<i>gebhalbj</i>	Geburtshalbjahr, generiert	Das genaue Geburtsdatum wurde zu einem Indikator für das 1. bzw. 2. Halbjahr verkürzt.
<i>PET1210</i>	Letzte Stellung im Beruf, einfache Einteilung (ohne Spellinfo) (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>	Letzte StiB Beamter: Det. Erfassung, inkl. Soldaten (ohne Spellinfo) (anon.)	Diese Variable enthält zusätzliche Fälle. Die Berufs- und Zeitsoldaten aus <i>PET1240</i> wurden der entsprechenden Beamtenkategorie zugeschlagen. Die Variable für Berufs- und Zeitsoldaten <i>PET1240</i> wird nicht ausgeliefert.
<i>PET1211</i>	Letzte Stellung im Beruf, einfache Einteilung (inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu <i>PET1210</i> .
<i>PET1251</i>	Letzte StiB Beamter: Det. Erf., inkl. Soldaten (inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten <i>PET1240</i> wird nicht ausgeliefert.
<i>stiblewt</i>	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>	Jetzige Stellung im Beruf, einf. Eint., Erh. ab W2 (anon.)	Vorgehen analog zu <i>PET1210</i> .
<i>PET1900</i>	Jetzige StiB Beamter: Det. Erfassung, inkl. Soldaten (anon.)	Vorgehen analog zu <i>PET1250</i> . Die in den Senioreninterviews erhobene Variable für Berufs- und Zeitsoldaten <i>PET1800</i> wird nicht ausgeliefert. Für die Personeninterviews wird aus den Erwerbstätigkeitsspiels <i>ET090*</i> keine generierte Variable für Berufs- und Zeitsoldaten in den Personen- datensatz übernommen.

Tabelle 59: Übersicht über die in Welle 11 anonymisierten Variablen im Personendatensatz (PENDDAT) (Fortsetzung)

Varname	Variablenlabel	Vorgehen
<i>Stibkz</i>	Jetzige Stellung im Beruf, einf. Eint., harmonisiert (anon.)	Bei der Generierung der Stellung im Beruf werden die Berufs- und Zeitsoldaten der entsprechenden Beamtenkategorie zugeordnet.
<i>Stib</i>	Stellung im Beruf, Kennziffer, aktuelle Erwerbstätigkeit, generiert	Vorgehen analog zu <i>stiblewt</i> .
<i>PET3300</i>	Erste Stellung im Beruf, einfache Einteilung (anon.)	Vorgehen analog zu <i>PET1210</i> .
<i>PET3700</i>	Erste StiB Beamter: Det. Erfassung, inkl. Soldaten	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten <i>PET3600</i> wird nicht ausgeliefert.
<i>PET3301</i>	Erste Stellung im Beruf, einf. Eint.(zus.gef., inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu <i>PET1210</i> .
<i>PET3701</i>	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 <i>PET3600</i> wird nicht ausgeliefert.
<i>stibeewt</i>	Stellung im Beruf, erste Erwerbstätigkeit, Kennziffer, generiert	Vorgehen analog zu <i>stiblewt</i> .
<i>PSH0320</i>	Damalige Stellung im Beruf der Mutter, einfache Einteilung (anon.)	Vorgehen analog zu <i>PET1210</i> .
<i>PSH0360</i>	Damalige StiB d. Mutter Beamte, inkl. Soldaten: Det. Erfassung (anon.)	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten <i>PSH0350</i> wird nicht ausgeliefert.
<i>mstib</i>	Stellung im Beruf der Mutter, Kennziffer, generiert	Vorgehen analog zu <i>stiblewt</i> .
<i>PSH0620</i>	Damalige Stellung im Beruf des Vaters, einfache Einteilung (anon.)	Vorgehen analog zu <i>PET1210</i> .
<i>PSH0660</i>	Damalige StiB d. Vaters Beamter inkl., Soldaten: Det. Erfassung (anon.)	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten <i>PSH0650</i> wird nicht ausgeliefert.
<i>vstib</i>	Stellung im Beruf des Vaters, Kennziffer, generiert	Vorgehen analog zu <i>stiblewt</i> .
<i>PMI0200</i>	Nicht in D geboren: Geburtsland, Kategorien (anon.)	Länder mit zu geringen Fallzahlen wurden zu größeren Kategorien zusammengefasst.
<i>ogebland</i>	Geburtsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu <i>PMI0200</i> .
<i>PMI0500</i>	Keine d. Staatsangehörigkeit: Welche Staatsangehörigkeit? (anon.)	Staatsangehörigkeiten von Ländern mit zu geringen Fallzahlen wurden zu größeren Kategorien zusammengefasst.

Tabelle 59: Übersicht über die in Welle 11 anonymisierten Variablen im Personendatensatz (PENDDAT) (Fortsetzung)

Varname	Variablenlabel	Vorgehen
<i>ostaatan</i>	Staatsangehörigkeit, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu <i>PMI0500</i> .
<i>ostaatansyr</i>	Staatsangehörigkeitsyr./irak. HH, inkl. off. Ang., Kategorien (anon.)	Für die Teilstichproben der syrischen und irakischen Haushalte wird die syrische Nationalität separat ausgewiesen.
<i>PMI1000a</i>	Vater: Zuzugsland (anon.)	Zuzugsländer mit zu geringen Fallzahlen wurden zu größeren Kategorien zusammengefasst.
<i>PMI1000b</i>	Mutter: Zuzugsland (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>PMI1000c</i>	Vater des Vaters: Zuzugsland (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>PMI1000d</i>	Mutter des Vaters: Zuzugsland (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>PMI1000e</i>	Vater der Mutter: Zuzugsland (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>PMI1000f</i>	Mutter der Mutter: Zuzugsland (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulanda</i>	Vater: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulandb</i>	Mutter: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulandc</i>	Vater des Vaters: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulandd</i>	Mutter des Vaters: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulande</i>	Vater der Mutter: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu <i>PMI1000a</i> .
<i>ozulandf</i>	Mutter der Mutter: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu <i>PMI1000a</i> .

Tabelle 60: Übersicht über die in Welle 11 anonymisierten Variablen im BIO-Spelldatensatz (*bio_spells*)

Varname	Variablenlabel	Vorgehen
<i>ET0609</i>	Welle 11, Stellung im Beruf: Einfache Einteilung (anon.)	Vorgehen analog zu <i>PET1210</i> .
<i>ET1009</i>	Welle 11, Stellung im Beruf: Beamter/Richter/Soldat, detaillierte Erfassung (anon.)	Vorgehen analog zu <i>PET1250</i> . Die Variable für Berufs- und Zeitsoldaten wird nicht ausgeliefert.
<i>stib</i>	Stellung im Beruf, Kennziffer, generiert	Vorgehen analog zu <i>stiblewt</i> .

Tabelle 61: Übersicht über die in Welle 11 anonymisierten Variablen im Kinder-Datensatz (*KINDER*)

Varname	Variablenlabel	Vorgehen
<i>alter12u14m</i>	Kinderalter 12 Monate und jünger als 14 Monate	Seit Welle 10 wird für Kinder unter 7 Jahren das Alter einmalig monatsgenau erfragt. Die Information über Geburtsmonat und -jahr wurde auf einen Indikator reduziert, ob das Kind zum Zeitpunkt der Befragung mindestens 12 und weniger als 14 Monate alt war. Basierend auf diesen Angaben wurde der Indikator auch für die vorherigen Befragungswellen gefüllt.

5.6 Arbeitslosengeld-II-Bezüge

Bereits in den Wellen 1 bis 10 wurden die Arbeitslosengeld-II-Bezüge auf Haushaltsebene in Episodenform erfasst. Dieses Konzept wurde in Welle 11 mit einem leicht überarbeiteten Fragenprogramm fortgeführt.

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 Dependent-Interviewing-Fragen enthalten. Hierüber setzen Haushalte mit andauernden Spells zur Vorwelle bei diesen andauernden Spells wieder in die Befragung ein.

Die in Welle 11 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 2015 oder wurde dazu keine Angabe gemacht, wurden die Arbeitslosengeld-II-Bezüge des Haushalts seit Januar 2015 erhoben.

5.6.2 Struktur des Arbeitslosengeld-II-Spelldatensatzes

Durch die Integration der in Welle 11 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 11 nicht mehr erhobenen Variablen.

1. Auch in Welle 11 wurden mit *AL20610*, *AL20710a* bis *AL20710o*, *AL20810* und *AL 20910* neue, wellenspezifische Querschnittvariablen in den Arbeitslosengeld-II-Spelldatensatz eingefügt. Diese Variablen sind auf den Interviewzeitpunkt der Welle 11 bezogen. Für die Befragungszeitpunkte der Vorwellen existieren ebenfalls Querschnittvariablen, in denen die auf die jeweilige Welle bezogenen analogen Informationen abgelegt sind. Folgende Tabelle gibt einen Überblick über die im Arbeitslosengeld-II-Spelldatensatz enthaltenen Querschnittinformationen.

Tabelle 62: Querschnittvariablen im ALG II-Spelldatensatz (*alg2_spells*)

	Welle 1	Welle 2	Welle 3	...	Welle 11
Erhält der HH ALG II für alle HH-Mitglieder?	<i>AL20600</i>	<i>AL20601</i>	<i>AL20602</i>	...	<i>AL20610</i>
Erhält der HH ALG II für die Personen 1 bis 15	<i>AL20700a-AL20700o</i>	<i>AL20701a-AL20701o</i>	<i>AL20702a-AL20702o</i>	...	<i>AL20710a-AL20710o</i>
Höhe des monatlichen ALG II-Bezugs?	<i>AL20800</i>	<i>AL20801</i>	<i>AL20802</i>	...	<i>AL20810</i>
Hat eine Kürzung des ALG II begonnen?	<i>AL20900</i>	<i>AL20901</i>	<i>AL20902</i>	...	<i>AL20910</i>

2. Liegen in Welle 11 im Vergleich zu Welle 10 nicht vor.
3. Liegen in Welle 11 im Vergleich zu Welle 10 nicht vor.

5.6.3 Plausibilitätsprüfung und –bereinigung des Arbeitslosengeld-II Spelldatensatzes

Wie in den Wellen 1 bis 10 wurden die Informationen zu Arbeitslosengeld-II- Bezügen auch in der Welle 11 einer Reihe von Plausibilitätsprüfungen unterzogen. Unzulässige Überschneidungen und Datierungen der Arbeitslosengeld-II-Spells bzw. der Kürzungs-episoden wurden dabei ggf. bereinigt. Veränderungen wurden dabei grundsätzlich nur in den generierten Datumsvariablen (*bmonat*; *bjahr*; *emonat*; *ejahr*) des Arbeitslosengeld-II-Spells, der Kürzungs-episoden (*alg2kbm**; *alg2kbj**; *alg2kem**; *alg2kej**) sowie im Zensierungsindikator des Arbeitslosengeld-II-Spells (*zensiert*) vorgenommen. Konnten fehlende Plausibilitäten nicht durch Korrektur der Datumsvariablen 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 11 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 11 erhobenen Informationen fortgeschrieben. Dabei sind zwei 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 - 11 - 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 11 berichtete Spells zugespielt (3). Diese drei Varianten sollen im Folgenden kurz dargestellt werden:

1. *Fälle, in denen der Haushalt in Welle 11 einem andauernden Arbeitslosengeld-II-Bezug zum Befragungszeitpunkt der Vorwelle widerspricht.*

Hat der Haushalt der Information, dass zum Zeitpunkt der Vorwelle ein aktuell andauernder Arbeitslosengeld-II-Bezug vorlag, in der Fortschreibungsfrage 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 unverändert⁴⁴. Das generierte Enddatum des Arbeitslosengeld-II-Spells (*emonat*; *ejahr*) wurde bereits in der Vorwelle auf das Interviewdatum der Vorwelle gesetzt.

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

Liegen in Welle 11 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

⁴⁴ Hier gilt dasselbe. Verändert wird nur der Zensierungsindikator. Das berichtete Enddatum, die Nachfrage für andauernde Spells und das generierte Enddatum bleiben unverändert.

(AL20300; AL20400), das generierte Enddatum (*emonat; ejahr*), die Nachfrage, ob der Arbeits-

losengeld-II-Bezug aktuell noch andauert (AL20500) sowie der Zensierungsindikator (*zensiert*) mit den in der Vorwelle erhobenen Informationen überschrieben. Weiterhin wurden die in der 11. Welle berichteten Kürzungsepisoden sowie die auf die Welle 11 bezogenen Querschnittinformationen (AL20610; AL20710a bis AL20710o, AL20810, AL20910) übernommen.

3. *In Welle 11 neu berichtete Arbeitslosengeld-II-Spells, die keine in der Vorwelle zensierte Episode fortschreiben.*

In Welle 11 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 den Wellen 2 und 3 wurden die Erwerbstätigkeits-, Arbeitslosigkeits- und Lückenphasen auf Personenebene in Episodenform erfasst. Dieses Konzept der modularen Spellerhebung wurde in Welle 4 zu einer integrierten Erhebung der Erwerbsbiografie umgewandelt. Für Personen, die in Welle 11 erstmals nach ihrer Erwerbsbiografie gefragt wurden, wurde das Referenzdatum für den Beginn des Retrospektivintervalls angepasst. In Welle 11 sollten hier alle Erwerbstätigkeits- und Arbeitslosigkeitsepisoden seit Januar 2015 berichtet werden. 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 Variablen zum (Nicht-)Erwerbsstatus im PENDDAT

Das Konzept zur Erhebung der Erwerbstätigkeiten wurde über die Wellen hinweg mehrfach überarbeitet:

- Welle 1: Panelkonzept, d.h. nur Erhebung am aktuellen Rand
- Wellen 2 und 3: Modulare Erhebung von ET-/AL-Episoden + Auffüllen von Lücken von > 3 Monaten und am aktuellen Rand
- Ab Welle 4: Integrierte Erhebung von ET-/AL-/LU-Episoden

Aufgrund der Änderungen des Erhebungskonzeptes variieren die für die einzelnen Wellen vorliegenden Informationen hinsichtlich:

- Form der vorliegenden Informationen (Panel vs. Episoden)
- Detailgrad der vorliegenden Informationen (Hauptstatus vs. parallele Zustände)

- Konsistenz der vorliegenden Parallelitäten (Lückenfüllung vs. vollständige Erhebung paralleler Zustände)

Das in den Wellen 2 und 3 angewendete Konzept der generierten Variablen zum (Nicht-)Erwerbsstatus orientiert sich sehr stark an der Erhebungslogik der 1. Welle. Diese Logik sah – vereinfacht dargestellt – wie folgt aus:

- Liegt eine ET von min. 1/h pro Woche vor?
- Wenn ET: Eine oder mehrere ET?
- Wenn ET (Angabe für Haupt-ET): Sukzessive Bestimmung, ob ET ein Minijob, Ein-Euro-Job o.ä., oder Teil einer Ausbildung ist
- Wenn keine ET (oder Haupt-ET=Minijob): Bestimmung des Status der Nichterwerbstätigkeit (AL oder anderer Status)

Das Konzept der generierten Variablen (*erwerb*, *erwerb2*, *nichterw*, *nichtew2*) folgt im weitesten Sinne dieser Erhebungslogik aus Welle 1. Während in Welle 1 durch die Abfrage-logik keine konkurrierenden Zustände möglich waren (ZP mit ET, die kein Minijob ist, wurden nicht nach weiteren Aktivitäten gefragt) ergab sich ab Welle 2 die Notwendigkeit, Entscheidungen zu treffen, falls mehr als eine andauernde Episode vorlag. Bei der Generierung der Variablen zum (Nicht-)Erwerbsstatus wurde dabei in den Wellen 1 bis 3 die folgende Logik angelegt:

Tabelle 63: Generierungslogik *erwerb*, *erwerb2*, *nichterw*, *nichterw2*

Variable	Generierungslogik Welle 1	Generierungslogik Wellen 2 und 3
<i>erwerb</i>	<i>(1) Ausdifferenzierung Status Haupt-ET</i> - Keine Haupt-ET - Haupt-ET: nicht Lehre/ ABM/ Minijob - Haupt-ET: Teil einer Lehre - Haupt-ET: ABM etc. - Haupt-ET: Minijob <i>(2) Ausdifferenzierung Status Haupt-ET ist Basis für weitere Generierung</i> - Haupt-ET: nicht Lehre/ ABM/ Minijob → ET als Erwerbsstatus (Ausnahmen: Auszubildende (aus PB0100) mit arbeitszeit <21 → Auszubildende; Schüler (aus PB0100) mit arbeitszeit >0 & arbeitszeit <24 → Schüler; Studenten (aus PB0100) mit arbeitszeit >0 & arbeitszeit <21 → Studenten; Erwerbstätige mit arbeitszeit >0 & arbeitszeit <16 → Sonstige) - keine Haupt-ET oder Haupt-ET: Minijob → Erwerbsstatus übernehmen aus PET0801 (d.h. Einsetzen des Status der Nichterwerbstätigkeit) - keine Haupt-ET + laut PB0100 Schüler/ Studenten → Erwerbsstatus übernehmen aus PB0100 - Haupt-ET: ABM etc. → Übernahme als Erwerbsstatus (ABM, 1-Euro-Job, etc.) <i>(3) Widersprüchliche Angaben entscheiden</i> - <i>erwerb</i>: ABM etc. + PB0100: Schüler/ Studenten/ Azubi → -8 - <i>erwerb</i>: Schüler + PB0100: Student → -8 - <i>erwerb</i>: Rentner + PB0100: Azubi → -8 - <i>erwerb</i>: Schüler + PB0100: Azubi → -8 - <i>erwerb</i>: Sonstiges + PB0100: Schüler/ Student/ Azubi → Erwerbsstatus aus PB0100	nicht generiert (-9)

Tabelle 63: Generierungslogik *erwerb*, *erwerb2*, *nichterw*, *nichterw2*
(Fortsetzung)

Variable	Generierungslogik Welle 1	Generierungslogik Wellen 2 und 3
erwerb2	(1) Recode von <i>erwerb</i> - Zusammenlegung von Kategorien: - Arbeitslos + ABM/ 1-Euro-Job/ u.ä. → Arbeitslos - Lehre/Ausbildung/Weiterbildung - Umschulung + Student → Lehre/ Ausbildung/ Studium	(1) Recode von <i>nichtew2</i> (2) ETs integrieren - Ersetzen der Werte, wenn aktuelle ET (>400 Euro aus ET-Spells) vorliegt (3) Anpassungen vornehmen . <i>erwerb2</i>: ET + PB0100: Student + Arbeitszeit ≤20h → Student - <i>erwerb2</i>: AL + PB0100: Student → Student - <i>erwerb2</i>: Schüler + PB0100: Student → Unklarer Status
nichterw	(1) Recode von <i>PET0800</i>	(1) Recode von <i>LU0100</i> (Lückenstatus ohne offene Angaben) + aktuelle AL aus AL-Spells) - Zusammenfassung von Kategorien: - Arbeitslos gemeldet + nicht gemeldet → Arbeitslos - Ausbildung/ Lehre/ Studium + Sonstiges → Sonstiges - Bestimmung MV aus <i>PET0151/ PET0911</i> + Indikator für fehlerhaft nicht ins Lückenmodul gefilterte Fälle
nichterw	(1) Recode von <i>PET0801</i> - Zusammenlegung von Kategorien: - Arbeitslos + ABM/ 1-Euro-Job/ u.ä. → Arbeitslos - Lehre/ Ausbildung/ Weiterbildung - Umschulung + Student → Lehre/ Ausbildung/ Studium	(1) Recode von <i>LU0101</i> (Lückenstatus mit offenen Angaben) - Zusammenfassung von Kategorien - Arbeitslos gemeldet + nicht gemeldet → Arbeitslos - Etwas anderes/Hauptstatus unklar → Sonstiges/ Hauptstatus unklar (2) Schüler/Studenten/Azubis aus <i>PB0100</i> berücksichtigen - Wenn bisher kein gültiger Status vorhanden → Übernahme der Information aus <i>PB0100</i>

Die generierten Variablen führen also die Logik des Erhebungskonzepts von Welle 1 fort,

die sich auch als Grundlogik in der generierten Variable findet: ET schlägt grundsätzlich alle anderen Zustände (bis auf Ausnahmen); AL schlägt alle Zustände außer ET (bis auf Ausnahmen)

In Welle 1 wäre eine andere Logik (bspw. AL schlägt ET) nicht umsetzbar gewesen, da die Erhebungslogik die ET der ZP priorisiert hat und andere Zustände nur ersatzweise erhoben wurden. Das Vorgehen bei der Variablengenerierung folgt also dem Vorgehen bei der Erhebung.

Allerdings ist dieses Vorgehen für die Bestimmung des Hauptstatus der Person nicht wirklich sinnvoll und übergeht darüber hinaus Grundgedanken, die sich bspw. in der Definition von Arbeitslosigkeit (§§16, 119 SGB III; gilt auch für SGB II nach §53a SGB II) finden.

Arbeitslosigkeit hat (laut der Definition des SGB III) bestimmte Voraussetzungen:

- Beschäftigungslosigkeit (d.h. keine bezahlte Beschäftigung bzw. Beschäftigung nur bis zur Grenze von 15h/Woche; Schwankungen sind möglich) (§119 SGB III)
- Verfügbarkeit (d.h. steht Vermittlungsbemühungen der BA zur Verfügung; Suche nach und Bereitschaft zu Beschäftigung $\geq$ 15h/Woche; Kann Eingliederungsvorschlägen zeitnah Folge leisten; Bereitschaft zu Teilnahme an beruflichen Eingliederungsmaßnahmen) (§119 SGB III)
- Eigenbemühungen (d.h. Bemühungen, die Beschäftigungslosigkeit zu beenden) (§119 SGB III)
- Meldung (d.h. persönliche Arbeitslosmeldung bei BA) (§16 SGB III)
- Kein aktueller Maßnahmeteilnehmer (§16 SGB III)

Die bisherige Logik, in der AL von ET unabhängig vom Stundenumfang geschlagen wird, ist demnach eher durch die Erhebungslogik der Welle 1 getrieben als durch eine inhaltliche Abwägung, was tatsächlich als Hauptstatus anzusehen ist.

Weitere Kritik an den (Nicht-)Erwerbsstatusvariablen betrifft das grundlegende Ziel dieser Variablen. Was sollen sie anzeigen? Den Hauptstatus der Person? Den Erwerbsstatus (wenn ja, was soll das genau sein)? Das Ziel erscheint bei genauerer Betrachtung nicht konsistent, da zwei Konzepte vermischt werden:

- Die Darstellung des Hauptstatus der ZP (d.h. bei konkurrierenden Zuständen wird eine Entscheidung getroffen, welcher Status unter welchen Bedingungen einen anderen „schlägt“)
- Die Darstellung, ob die ZP aktuell in einem bestimmten Zustand ist (auch wenn dieser mglw. im Hauptstatus von einem anderen „geschlagen“ wird)

Für die Generierung der (Nicht-)Erwerbsstatusvariablen ab Welle 4 gibt es grundsätzlich zwei Möglichkeiten:

- Weiterführen der bisherigen Generierungslogik auf neuer Datenbasis

- Überarbeitung der Generierungslogik, mit dem Ziel:
 - Konzeptionelle Straffung (Was genau bilden die Variablen ab?)
 - Verbesserung von Entscheidungen, die in der Vergangenheit vor dem Hintergrund der vorhandenen Daten getroffen wurden aber inhaltlich suboptimal sind (d.h. nicht nur Weiterführung auf neuer Datenbasis, sondern inhaltliche Nutzung der detaillierteren Datenbasis)
 - Entschlackung (d.h. Entfernung von Variablen mit äußerst begrenztem Zusatznutzen)

Es wurde entschieden, die Generierungslogik grundlegend zu überarbeiten. Mit den bisherigen Variablen wird wie folgt verfahren:

Tabelle 64: Überarbeitung *erwerb*, *erwerb2*, *nichterw*, *nichterw2*

Variable	Entscheidung	Begründung
erwerb	beibehalten (Welle1: inhaltlich generiert) (Welle 2ff: -9)	Variable bildet Erhebungskonzept Welle 1 optimal ab. Der Schwerpunkt liegt auf ET (diese schlagen vereinfacht gesprochen Arbeitslosigkeit, die wiederum alles andere schlägt.). Gegen eine Fortführung sprechen inhaltliche Bedenken, denen dank der detaillierteren Datenbasis durch eine Neukonzeption begegnet werden kann. Für Welle 1 bleibt die Variable erhalten, da sie auf das Erhebungskonzept zugeschnitten ist. Die Besonderheiten (keine Parallelitäten; Konzentration auf ET; keine Differenzierung der Arbeitslosigkeit in gemeldete und nicht gemeldete) bleiben so auf Welle 1 begrenzt.
erwerb2	entfällt aus SUF	Mit dieser Variablen wird die Logik des Erhebungskonzepts aus Welle 1 in harmonisierter Form fortgeführt. Dabei treten aber verschiedene Probleme auf: (1) Welche ETs erhoben werden, ändert sich (Welle 1: 1h/Woche vs. Welle 2ff.: >400 Euro) (2) Änderung der Schwerpunktsetzung (Welle 1: Wenn ET [nicht Minijob] vorhanden → keine Erhebung paralleler ALs/LU-Zustände; Welle 2ff.: ET/AL/(teilweise auch LU) parallel möglich) (3) Durch Festhalten an Logik der Welle 1 können die Möglichkeiten der neuen Datenbasis nicht sinnvoll genutzt werden (um bspw. inhaltlich sinnvollere Entscheidungen zu treffen) Fazit: Harmonisierte Variable mit Schwerpunkt auf ET (wie bisher in <i>erwerb2</i>) ist einzige Möglichkeit für wellenübergreifend harmonisierte Variable. Deren Erstellung wäre zwar möglich, aber nur auf Basis inhaltlicher/konzeptioneller Entscheidungen die nicht sinnvoll sind. Da das Konzept der Welle 1 als problematisch angesehen wird, wird auf die Auslieferung einer harmonisierten Variable verzichtet.

Tabelle 64: Überarbeitung *erwerb*, *erwerb2*, *nichterw*, *nichterw2*
(Fortsetzung)

Variable	Überarbeitung	Begründung
nichterw	entfällt aus SUF	Bisherige Aufteilung in Erwerbs- und Nichterwerbsstatus wird aufgegeben und durch Hauptstatus + Indikator für aktuelle (sozialversicherungspflichtige) Erwerbstätigkeit + Indikator für aktuelle Arbeitslosmeldung ersetzt. Welle 1: Variable bietet im Vergleich zu neuer Hauptstatusvariable keine Zusatzinformation Welle 2ff.: Zusatzinformation im Vergleich zu neuer Hauptstatusvariable nur gering Fazit: Generell eher zusätzliche Komplexität mit sehr begrenztem Nutzen (bspw. Studenten > 20h/Arbeitszeit pro Woche). Für Auswertungen wahrscheinlich eine eigene Bestimmung der Nebenzustände eher sinnvoll als bisher ausgelieferte Variablen.
nichterw2	entfällt aus SUF	(siehe nichterw)

Ab Welle 2 werden die folgenden Variablen generiert:

- *etakt*: Aktuell erwerbstätig (>400 Euro pro Monat), generiert (ab Welle 2)
- *alakt*: Aktuell arbeitslos gemeldet, generiert (ab Welle 2)
- *statakt*: Aktueller Hauptstatus, generiert (ab Welle 2)

Mit der Überarbeitung werden die folgenden Ziele verfolgt:

- Trennung der Information zum Hauptstatus (*statakt*) von der Information zu aktuell andauernden Episodentypen (*etakt*, *alakt*)
- Klarere Dokumentation der Regeln bei der Erstellung des Hauptstatus
- Differenzierung zwischen gemeldeter und nicht gemeldeter AL (wo möglich)

etakt (Aktuell erwerbstätig (>400 Euro pro Monat), generiert (ab Welle 2))

Die Variable zeigt an, dass die ZP zum Zeitpunkt des Personeninterviews der jeweiligen Welle einen andauernden ET-Spell (d.h. eine ET > 400 Euro) hatte. Für Welle 1 kann die Variable entsprechend nicht generiert werden, da sich das Erhebungskonzept zwischen Welle 1 und den Folgewellen unterscheidet (Welle 1: min 1h/Woche; Welle 2ff. > 400 Euro/Monat). Eine aktuelle ET liegt vor, wenn im Spelldatensatz der jeweiligen Welle ein zensierter ET-Spell vorhanden ist.

Ausprägungen der zu generierenden Variable:

- -10 Item fuer Fragebogenversion nicht relevant

- -5 Generierung nicht mgl. (fehlende Werte)
- -3 Trifft nicht zu (Filter)
- 1 ZP akt. erwerbstätig (> 400 Euro)
- 2 ZP akt. nicht erwerbstätig (> 400 Euro)

alakt (Aktuell arbeitslos gemeldet, generiert (ab Welle 2))

Die Variable zeigt an, dass die ZP zum Zeitpunkt des Personeninterviews der jeweiligen Welle arbeitslos gemeldet war. Für Welle 1 kann die Variable entsprechend nicht generiert werden, da sich das Erhebungskonzept zwischen Welle 1 und den Folgewellen unterscheidet (Welle 1: AL nur erhoben, wenn keine ET vorliegt; Welle 1: arbeitslos; Welle 2ff.: arbeitslos gemeldet). Eine aktuelle, gemeldete AL liegt vor, wenn im Spelldatensatz der jeweiligen Welle ein zensierter AL-Spell (gemeldete AL) vorhanden ist.

Ausprägungen der zu generierenden Variable:

- -10 Item fuer Fragebogenversion nicht relevant
- -5 Generierung nicht mgl. (fehlende Werte)
- -3 Trifft nicht zu (Filter)
- 1 ZP akt. arbeitslos gemeldet
- 2 ZP akt. nicht arbeitslos gemeldet

statakt (Aktueller Hauptstatus, generiert (ab Welle 2))

Die Variable zeigt an, welchen Hauptstatus die ZP zum Zeitpunkt des Personeninterviews der jeweiligen Welle hatte.

Basis für die Generierung sind die Spelldatensätze (Wellen 2 und 3: *ET-/AL-/LU-Spells*; Welle 4ff.: *BIO-Spells*) sowie der Status als Schüler/Student/Azubi in PB0100.

Wenn in der jeweiligen Welle ein Spell eines bestimmten Typs aktuell andauert, dann liegt für die Person der entsprechende Zustand vor. Der Spelltyp wird in den Wellen 2 und 3 über den jeweiligen Spelldatensatz (*ET-/AL-Spells*) bzw. den Lückenzustand (*LU0101* in *LU-Spells*) bestimmt. Ab Welle 4 kann die Variable *spelltyp* genutzt werden. In allen Wellen werden nur zum Interviewdatum andauernde Spells (d.h. *zensiert=1* im SUF der jeweiligen Welle) berücksichtigt. Der aktuelle Status als Schüler bzw. als Student/Azubi aus PB0100 wird berücksichtigt, als ob in der jeweiligen Welle ein aktuell andauernder Spell vorliegen würde.

Ausprägungen der zu generierenden Variable:

- -10 Item fuer Fragebogenversion nicht relevant
- -5 Generierung nicht mgl. (fehlende Werte)
- -3 Trifft nicht zu (Filter)

- 1 Erwerbstätig m. Eink.> 400 Euro/Monat
- 2 Arbeitslos gemeldet
- 3 Schüler (bis W3 in Spells n. kat. erh.)
- 4 Ausbildung/Lehre/Studium
- 5 Wehr-/Zivildienst/etc.
- 6 Hausfrau/Hausmann
- 7 Mutterschutz/Erzieh.urlaub/Elternzeit
- 8 Rentner/Pensionär/Vorruhestand
- 9 Sonstiges/Hauptstatus unklar
- 10 Arb.los, n. gemeldet (ab W4 aus off. A.)
- 11 Krank/beruf.unf./erw.unf./beh.(off. A.)
- 12 Selbst./mithelf. Fam.angeh. (off. A.)

Die Zuweisung der Codes sollte dabei schrittweise erfolgen:

Tabelle 65: Grobzuweisung - Spell höherer Priorität schlägt Spell niedrigerer Priorität

Priorität eines aktuellen Spells (bzw. analoger Zustand aus PB0100)	Code in stat (analog zu Variable spell-typ)	Bedeutung
1	2	Arbeitslos gemeldet/ Massnahmenteiln.
2	1	Erwerbstätig m. Eink.> 400 Euro/Monat
3	8	Rentner/ Pensionär/ Vorruhestand
4	7	Mutterschutz/ Erzieh.urlaub/ Elternzeit
5	5	Wehr-/ Zivildienst/ etc.
6	4	Ausbildung/ Lehre/ Studium
7	3	Schüler (bis Welle 3 aus off. A.)
8	12	Selbst./mithelf. Fam.angeh. (off. A.)
9	11	Krank/ beruf.unf./ erw.unf./ beh.(off. A.)
10	10	Arb.los, n. gemeldet (ab Welle 4 aus off. A.)
11	6	Hausfrau/ Hausmann
12	9	Sonstiges/ Hauptstatus unklar

Liegen für die Zusatzinformation keine gültigen Werte vor, bleibt die Grobzuweisung erhalten.

Tabelle 66: Feinzuweisung für Sonderfälle

Grobzuweisung	Zusatzinformation	Entscheidung
Arbeitslos gemeldet	Erwerbstätig m. Eink.>400 Euro/Monat + Arbeitszeit (az2ges; tatsächliche Arbeitszeit, Summe über zensierte ETs)>=15h	Erwerbstätig m. Eink.>400 Euro/Monat
Erwerbstätig m. Eink. > 400 Euro/Monat	Ausbildung/ Lehre/ Studium + Arbeitszeit (az2ges; tatsächliche Arbeitszeit, Summe über zensierte ETs)<=20h	Ausbildung/ Lehre/ Studium

Eine aktuelle, gemeldete AL liegt vor, wenn im Spelldatensatz der jeweiligen Welle (Wellen 2 und 3: *AL-Spells*; Welle 4ff.: *BIO-Spells*) ein zensierter AL-Spell (gemeldete AL) vorhanden ist.

5.7.2 Einkommensvariablen und Arbeitszeit im PENDDAT und in BIO-Spells

In den Wellen 1 bis 4 beziehen sich die Variablen zur aktuellen Erwerbstätigkeit auf die Haupterwerbstätigkeit⁴⁵. Ausgenommen sind hier die Angaben zum Brutto-/Nettoeinkommen in den Wellen 2 bis 4 – diese beziehen sich auf alle aktuell andauernden Erwerbstätigkeiten > 400 Euro (Unschärfe hinsichtlich Minijoblöhnen). Spellspezifische Angaben liegen nicht vor und werden erst ab Welle 5 erhoben. Die Angaben werden nur als Gesamtwert über alle Erwerbstätigkeiten erhoben. Das führt zu zwei Teilproblemen:

1. Die generierten Variablen zu Arbeitsstunden und Brutto/Netto beziehen sich ab Welle 2 auf verschiedene Erwerbstätigkeiten (Haupt-ET bzw. alle ET). Werden auf dieser Basis Stundenlöhne berechnet, kommt es zu Fehlern bei ZP mit mehreren ETs.
2. Die unterschiedlichen Bezüge sind aus den Variablenlabels nicht ersichtlich.

Die generierten Variablen zu Einkommen und Arbeitszeit werden deshalb in Welle 4 entsprechend überarbeitet.

Einkommensvariablen

Das Erhebungskonzept der Einkommensvariablen hat sich zwischen Welle 1 und 2 stark verändert, ohne dass dies zur Bildung neuer Variablen geführt hat: *brutto* (*bruttokat*) und *netto* (*nettokat*) geben in Welle 1 das Einkommen der Haupterwerbstätigkeit wieder, ab Welle 2 das Einkommen aller nicht geringfügigen Erwerbstätigkeiten. Dies ist inkonsistent und führt potenziell zu Fehlern in der Auswertung. Dieses Problem soll mit der Überarbeitung korrigiert werden:

⁴⁵ Wellen 2 und 3; Es handelt sich um die zensierte ET im ET-Spelldatensatz. Bei mehreren zensierten Spells wurde der Spell mit dem größten Stundenumfang ausgewählt. Bei mehreren zensierten Spells mit gleichem Stundenumfang wurde der am längsten andauernde Spell ausgewählt. Für Senioren wurde nur eine Erwerbstätigkeit erhoben.

Tabelle 67: Überarbeitung Einkommensvariable

Variable - Inhalt - Datensatz	Generiert für W1 - W2 - W3 - W4 - W5ff.	Basis OFFA - KatA
bruttokat - Haupt-ET, Brutto - PENDDAT	1 - 0 - 0 - 0 - 1	0 - 1
brutto - Haupt-ET, Brutto - PENDDAT	1 - 0 - 0 - 0 - 1	1 - 1
nettokat - Haupt-ET, Netto - PENDDAT	1 - 0 - 0 - 0 - 1	0 - 1
netto - Haupt-ET, Netto - PENDDAT	1 - 0 - 0 - 0 - 1	1 - 1
brges - GesET, Brutto - PENDDAT	0 - 1 - 1 - 1 - 1	1 - 1
netges - GesET, Netto - PENDDAT	1 - 1 - 1 - 1 - 1	1 - 1*
br - SpellET, Brutto - BIO-Spells	0 - 0 - 0 - 0 - 1	1 - 1
net - SpellET, Netto - BIO-Spells	0 - 0 - 0 - 0 - 1	1 - 1

Anmerkung: In Welle 1 existiert nur für den Nettolohn der Haupterwerbstätigkeit eine kategoriale Nachfrage, nicht aber für die weiteren Tätigkeiten. Das wird bei der Generierung von *netges* akzeptiert. Fehlen die Angaben (MV) zum Nettoeinkommen der weiteren Tätigkeiten, kann die Variable *netges* nicht generiert werden.

Überarbeitete Variablen (bereits im Datensatz in den Wellen 1 bis 3):

- **bruttokat** (Akt. Bruttoeinkommen Haupterwerb. (ohne Minijobs, kategorisiert), gen.)
- **brutto** (Akt. Bruttoeinkommen Haupterwerb. (ohne Minijobs, inkl. kat. Ang.), gen.)
- **nettokat** (Akt. Nettoeinkommen Haupterwerb. (ohne Minijobs, kategorisiert), gen.)
- **netto** (Akt. Nettoeinkommen Haupterwerb. (ohne Minijobs, inkl. kat. Ang.), gen.)

Diese Variablen beziehen sich in Welle 1 auf die jeweilige Haupt-ET. Ab Welle 2 wurden sie aber mit den kumulierten Angaben für alle ETs (>400 Euro) gefüllt, da nur diese erhoben wurden. Die Variablenlabel werden ab Welle 4 entsprechend angepasst. Für die Wellen 2 bis 4 werden die Variablen mit -9 gefüllt, da eine zu Welle 1 analoge Generierung nicht möglich ist.

Neue Variablen in Welle 4:

brges (Akt. Bruttoeinkommen insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.)

Diese Variable enthält die über alle ETs (>400 Euro) kumulierte Angabe zum Bruttoeinkommen. Für Welle 1 ist die Variable in dieser Form nicht generierbar, da nur für die Haupt-ET das Bruttoeinkommen erhoben wurde. Für Welle 2 und Welle 3 ist die Variable inhaltlich identisch zur Variable *brutto*, die im SUF der Welle 3 ausgeliefert wurde (d.h. vor der oben beschriebenen Überarbeitung). In den Wellen 2 bis 4 wurde nur das kumulierte Bruttoeinkommen erhoben – die in den Wellen 2 und 3 genutzten Quellvariablen enthalten also schon die entsprechenden Angaben zum Gesamteinkommen aus ETs (> 400 Euro).

Für Welle 4 soll die Variable analog zu den Wellen 2 und 3 erstellt werden. Ab Welle 5 erfolgt die Generierung auf Basis von spellspezifischen Einkommensangaben.

netges (*Akt. Nettoeinkommen insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.*)

Diese Variable enthält die über alle ETs (>400 Euro) kumulierte Angabe zum Nettoeinkommen. Für Welle 1 ist die Variable generierbar, in dem die offene und kategoriale Nettoeinkommensangabe der Haupterwerbstätigkeit mit der Angabe für die weiteren Tätigkeiten kombiniert wird (hier fehlt allerdings die kategoriale Nachfrage). Für die Wellen 2 und 3 ist die Variable identisch zur Variable *netto*, die im SUF der Welle 3 ausgeliefert wurde. In den Wellen 2 bis 4 wurde nur das kumulierte Nettoeinkommen erhoben – die in den Wellen 2 und 3 genutzten Quellvariablen enthalten also schon die entsprechenden Angaben zum Gesamteinkommen aus ETs (> 400 Euro). Für Welle 4 wurde die Variable analog zu den Wellen 2 und 3 erstellt. Ab Welle 5 erfolgt die Generierung auf Basis von spellspezifischen Einkommensangaben.

Arbeitszeit

Durch die Korrektur der Variablen zum (Brutto-/Netto)Einkommen (s.o im gleichen Abschnitt) können im Personendatensatz keine Stundenlöhne mehr generiert werden, da als Arbeitszeit nur die tatsächliche Arbeitszeit der Haupt-ET vorliegt (Variable *arbzeit* im *PENDDAT* des SUF der Welle 3). Analog zur Überarbeitung der Einkommensvariablen ist demnach auch eine Überarbeitung der Arbeitszeitvariablen im *PENDDAT* als auch in den *BIO-Spells* nötig.

Tabelle 68: Überarbeitung Arbeitszeitvariablen

Variable - Inhalt - Datensatz	Generiert für W1 - W2 - W3	Basis OFFA - KatA	Anmerkung
az1 - SpellET, vertrag. - Bio-Spells	0 - 1 - 1	1 - 0	Kat. Welle 2ff.
azhpt1 - Haupt-ET, vertrag. - PENDDAT	0 - 1 - 1	1 - 0	Kat. Welle 2ff.
azges1 - Gesamt, vertrag. - PENDDAT	0 - 1 - 1	1 - 0	Kat. Welle 2ff.
az2 - SpellET, vertrag. - Bio-Spells	0 - 1 - 1	1 - 1	entspricht bisheriger Variable arbeitszeit (BIO-Spells); Kat. Welle 2ff., ET mit max(az2) = Haupt-ET (wenn zwei identische: ET mit frühestem Beginn)
azhpt2 - Haupt-ET, vertrag. - PENDDAT	1 - 1 - 1	1 - 1	entspricht bisher Variable arbeitszeit (PENDDAT); Kat. Welle 1 != Kat Welle 2ff.
azges2 - Gesamt, vertrag. - PENDDAT	1 - 1 - 1	1 - 1*	Kat Welle 1 != Kat Welle 2ff.; in Welle 1 keine Kat für NebenET

Überarbeitete Variablen (bereits im Datensatz in den Wellen 1 bis 3):

arbeitszeit (*Wöchentliche Arbeitszeit inkl. Angaben bei unregelmässiger Arbeitszeit, gen.*)

Variable entfällt aus *PENDDAT* und *BIO-Spells*. Sie wird inhaltlich durch *azhpt2* (*PENDDAT*) und *az2* (*BIO-Spells*) ersetzt.

Neue Variablen in Welle 4:

az1 (*Vertragl. Arbeitszeit, gen.*)

Die Variable wird für alle Episoden in den *BIO-Spells* generiert. Sie enthält die jeweils für den Spell aktuellste Angabe zur vertraglichen Arbeitszeit (ET > 400 Euro). Basis für die Generierung sind dabei jeweils die Querschnittsvariablen, für die im jeweiligen Spell zuletzt Angaben erhoben wurden.

Bsp.:

- Spell angelegt in Welle 2, beendet in Welle 2: Querschnittsvariablen Welle 2
- Spell angelegt in Welle 2, fortgeschrieben in den Wellen 3 und 4: Querschnittsvariablen Welle 4

azhpt1 (*Vertragl. akt. Arb.zeit Haupterwerbst. (ohne Minijobs), gen.*)

Die Variable wird für den *PENDDAT* generiert. Sie enthält die vertragliche Arbeitszeit der in der jeweiligen Welle aktuell andauernden Haupt-ET aus den Spelldaten (ET > 400

Euro). Für Welle 1 ist die Variable nicht generierbar (-9), da die entsprechenden Angaben erst ab Welle 2 erhoben werden. Ab Welle 2 wird aus den jeweiligen Spelldaten die generierte Variable zur vertraglichen Arbeitszeit der Haupt-ET (*az1*) in den *PENDDAT* übernommen. Bei welchem aktuell andauernden Spell es sich um die Haupt-ET handelt, wird auch hier auf Basis der tatsächlichen Arbeitszeit bestimmt (generierte Variable *az2* in den Spelldaten; analog zum Vorgehen in den Wellen 2 und 3 wo die Variable *arbzeit* zur Bestimmung verwendet wurde).

azges1 (*Vertragl. akt. Arb.zeit insgesamt (ohne Minijobs), gen.*)

Die Variable wird für den *PENDDAT* generiert. Sie enthält die kumulierte vertragliche Arbeitszeit aller in der jeweiligen Welle aktuell andauernden Erwerbstätigkeiten aus den Spelldaten (ET > 400 Euro). Für Welle 1 ist die Variable nicht generierbar (-9), da die entsprechenden Angaben erst ab Welle 2 erhoben werden. Ab Welle 2 wird die Variable auf Basis der generierten Variable zur vertraglichen Arbeitszeit (*az1*) aus den Spelldaten generiert. Bei der Generierung werden die Angaben in der generierten Variable zur vertraglichen Arbeitszeit (*az1*) über alle Spells hinweg kumuliert, die zum Befragungszeitpunkt aktuell andauern. Diese Angabe wird in den *PENDDAT* übernommen.

az2 (*Tatsaechliche Arbeitszeit inkl. Angaben bei unregelmaessiger Arbeitszeit, gen.*)

Die Variable wird für alle Episoden in den *BIO-Spells* generiert. Sie enthält die jeweils für den Spell aktuellste Angabe zur tatsächlichen Arbeitszeit und integriert dabei auch die kategorialen Angaben bei unregelmäßiger Arbeitszeit. Basis für die Generierung sind dabei jeweils die Querschnittsvariablen, für die im jeweiligen Spell zuletzt Angaben erhoben wurden.

Bsp.:

- Spell angelegt in Welle 2, beendet in Welle 2: Querschnittsvariablen Welle 2
- Spell angelegt in Welle 2, fortgeschrieben in Welle 3 und Welle 4: Querschnittsvariablen Welle 4

Die Variable ersetzt die bisher in den ET-Spells generierte Variable *arbzeit* (die entsprechend entfällt). Die Generierung erfolgt analog zur Generierung von *arbzeit* in der Datenaufbereitung der Wellen 2 und 3.

Definition der Haupterwerbstätigkeit (Haupt-ET):

Die Variable *az2* dient der Bestimmung der Haupterwerbstätigkeit (Haupt-ET) in einer Welle, für die verschiedene Angaben in den *PENDDAT* übernommen werden. Die Haupt-ET ist die in der jeweiligen Welle andauernde ET mit dem größten Stundenumfang. Existieren mehrere Erwerbstätigkeiten mit gleichem Stundenumfang, so wird die zuerst beginnende ausgewählt. Bei mehreren andauernden ET mit gleichem Stundenanfang und identischem Beginndatum wird die durch die befragte Person zuerst genannte ET ausgewählt. Diese hat unter den in Frage kommenden Erwerbstätigkeiten die niedrigste Spell-Nummer.

azhpt2 (*Tatsaechl. akt. Arb.zeit Haupterwerbst. (ohne Minijobs, inkl. kat. Ang.), gen.*)

Die Variable wird für den *PENDDAT* generiert. Sie enthält die tatsächliche Arbeitszeit der aktuell andauernden Haupterwerbstätigkeit und integriert dabei auch die kategoriale Nachfrage bei unregelmäßiger Arbeitszeit. Die Variable ersetzt inhaltlich die im *PENDDAT* entfallene Variable *arbzeit*. Die Generierung erfolgt für die Wellen 1 bis 4 analog zur Generierung der entfallenen Variablen.

In Welle 1 wird die Variable auf Basis der Querschnittsdaten generiert. Sie kombiniert also sowohl die offenen Angaben zur tatsächlichen Arbeitszeit als auch die kategoriale Nachfrage. 1-Euro-Jobs, ABM, Minijobs und Tätigkeiten, die Teil einer Lehre sind, werden dabei nicht berücksichtigt – für diese Fälle kann die Variable nicht generiert werden (-3), da in den Wellen 2 bis 4 analoge Informationen nicht erhoben werden.

Ab Welle 2 wird aus den jeweiligen Spelldaten die generierte Variable zur tatsächlichen Arbeitszeit der Haupt-ET (*az2*) in den *PENDDAT* übernommen. Bei welchem aktuell andauernden Spell es sich um die Haupt-ET handelt, wird auch hier auf Basis der tatsächlichen Arbeitszeit bestimmt (generierte Variable *az2* in den Spelldaten; analog zum Vorgehen in den Wellen 2 und 3 wo die Variable *arbzeit* zur Bestimmung verwendet wurde). Die kategoriale Nachfrage bei unregelmäßiger Arbeitszeit unterscheidet sich zwischen Welle 1 und Welle 2ff. Trotzdem findet eine Integration über die Wellen hinweg statt.

azges2 (*Tatsaechl. akt. Arb.zeit insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.*)

Die Variable wird für den *PENDDAT* generiert. Sie enthält die kumulierte tatsächliche Arbeitszeit aller in der jeweiligen Welle aktuell andauernden Erwerbstätigkeiten.

In Welle 1 werden dazu die Stunden der Haupt-ET (nach Integration der kategorialen Angaben bei unregelmäßiger Arbeitszeit) mit den Angaben zur tatsächlichen Arbeitszeit der weiteren ETs kombiniert. 1-Euro-Jobs, ABM, Minijobs und Tätigkeiten, die Teil einer Lehre sind, werden dabei nicht berücksichtigt – für diese Fälle kann die Variable nicht generiert werden (-3), da in den Wellen 2 bis 4 analoge Informationen nicht erhoben werden.

Ab Welle 2 wird die Variable auf Basis der generierten Variable zur tatsächlichen Arbeitszeit (*az2*) aus den Spelldaten generiert. Bei der Generierung werden die Angaben in der generierten Variable zur tatsächlichen Arbeitszeit (*az2*) über alle Spells hinweg kumuliert, die zum Befragungszeitpunkt aktuell andauern. Diese Angabe wird in den *PENDDAT* übernommen.

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

In der Welle 11 wurden andauernde ET-, AL- und Lückenspells fortgeschrieben. Für die Fortschreibung der in der Vorwelle andauernden und damit im Spelldatensatz rechtszenzierten Spells sind in den Personenfragebögen Dependent- Interviewing- Fragen enthalten.

5.7.4 Struktur des BIO-Spelldatensatzes

Der BIO-Spelldatensatz orientiert sich seit Welle 4 hinsichtlich seiner Struktur an den modularen Spelldatensätzen ET, AL und LU der Wellen 2 bis 3. ET-spezifische Variablen haben dabei im BIO-Spelldatensatz ihren Variablennamen im Vergleich zum ET-SUF der Welle 3 behalten, analog auch AL- bzw. LU-spezifische Variablen. Zwischen ET, AL und LU gleiche Variablen wurden ab Welle 4 vereinheitlicht (*BIO0100*, *BIO0101*, *BIO0200*, *BIO0300*, *BIO0400*, *BIO0500*, *BIO0600*) bzw. lagen bereits einheitlich in den ursprünglichen Datensätzen des SUF Welle 3 vor (*bmonat*, *bjahr*, *emonat*, *ejahr*, *zensiert*). Darüber hinaus liegen Variablen für den Aktivitätstyp (*spelltyp*), für die Spellintegration (*spintegr*) und eine übergreifende Spellnummer (*spellnr*) vor.

Durch die Integration der in Welle 11 berichteten Erwerbstätigkeits- und Arbeitslosigkeitsphasen in den BIO-Spelldatensatz werden neue ET- bzw. AL-spezifische Variablen ergänzt. 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 11 nicht mehr erhobenen Variablen.

1. Im BIO-Spelldatensatz werden die ET-spezifischen Variablen *ET0600* bis *ET2200* als wellenspezifische, auf Welle 2 bezogene Querschnittinformationen verstanden. Die Variablen *ET0601* bis *ET2201* sind auf Welle 3, *ET0552* bis *ET2202* auf Welle 4, *ET0553* bis *ET2203* auf Welle 5, *ET0554* bis *ET2204* auf Welle 6, *ET0555* bis *ET2205* auf Welle 7, *ET0556* bis *ET2206* auf Welle 8, *ET0557* bis *ET2207* auf Welle 9, *ET0558* bis *ET2208* auf Welle 10 und *ET0559* bis *ET2209* auf Welle 11 bezogene Querschnittinformationen. Folgende Tabelle gibt einen Überblick über die im BIO-Spelldatensatz enthaltenen ET-spezifischen Querschnittinformationen.

Tabelle 69: ET-spezifische Querschnittvariablen im BIO-Spelldatensatz (*bio_* - *spells*)

	Welle 2	Welle 3	Welle 4	Welle 5	...	Welle 9	...	Welle 11
Stellung im Beruf (einfache und detail- lierte Einteilung)	<i>ET0600</i>	<i>ET0601</i>	<i>ET0552</i>	<i>ET0553</i>	...	<i>ET0557</i>	...	<i>ET0559</i>
	<i>ET0700</i>	<i>ET0701</i>	<i>ET0602</i>	<i>ET0603</i>		<i>ET0607</i>	...	<i>ET0609</i>
	<i>ET0800</i>	<i>ET0801</i>	<i>ET0702</i>	<i>ET0703</i>		<i>ET0707</i>	...	<i>ET0709</i>
	<i>ET0800</i>	<i>ET0801</i>	<i>ET0802</i>	<i>ET0803</i>		<i>ET0807</i>	...	<i>ET0809</i>
	<i>ET1000</i>	<i>ET1001</i>	<i>ET1002</i>	<i>ET1003</i>		<i>ET1007</i>	...	<i>ET1009</i>
	<i>ET1100</i>	<i>ET1101</i>	<i>ET1102</i>	<i>ET1103</i>		<i>ET1107</i>	...	<i>ET1109</i>
Aufsichtungsfunktion; Zahl der beaufsichtig- ten Arbeitskräfte	<i>ET1200</i>	<i>ET1201</i>	<i>ET1202</i>	<i>ET1203</i>		<i>ET1207</i>	...	<i>ET1209</i>
	<i>ET1300</i>	<i>ET1301</i>	<i>ET1302</i>	<i>ET1303</i>	...	<i>ET1307</i>	...	<i>ET1309</i>
Entfristung einer an- fangs unbefristeten Tätigkeit	<i>ET1400</i>	<i>ET1401</i>	<i>ET1402</i>	<i>ET1403</i>		<i>ET1407</i>	...	<i>ET1409</i>
	<i>ET1700</i>	<i>ET1701</i>	<i>ET1702</i>	<i>ET1703</i>	...	<i>ET1707</i>	...	<i>ET1709</i>
				<i>ET1753a</i> <i>ET1753b</i>		<i>ET1757a</i> <i>ET1757b</i>	...	<i>ET1759a</i> <i>ET1759b</i>

Tabelle 69: ET-spezifische Querschnittvariablen im BIO-Spellldatensatz (*bio_spells*)
(Fortsetzung)

	Welle 2	Welle 3	Welle 4	Welle 5	...	Welle 9	...	Welle 11
Arbeitszeit (vertraglich; tatsächlich; durchschnittlich bei unregelmäßiger	<i>ET2000</i> <i>ET2100</i> <i>ET2200</i>	<i>ET2001</i> <i>ET2101</i> <i>ET2201</i>	<i>ET1952</i> <i>ET2002</i> <i>ET2102</i> <i>ET2202</i>	<i>ET1953</i> <i>ET2003</i> <i>ET2103</i> <i>ET2203</i>	...	<i>ET1957</i> <i>ET2007</i> <i>ET2107</i> <i>ET2207</i>	...	<i>ET1959</i> <i>ET2009</i> <i>ET2109</i> <i>ET2209</i>
Einkommen für gegenwärtig andauernde Spells				<i>ET2800-ET3900</i>	...	<i>ET2804-ET3904</i>	...	<i>ET2806-ET3906</i>
Überstunden						<i>ET4100</i> <i>ET4200</i>	...	<i>ET4102</i> <i>ET4202</i>

Im BIO-Spellldatensatz gibt es darüber hinaus eine AL-spezifische Variable, die als wellenspezifische Querschnittinformation zu verstehen ist (*AL1300* für Welle 2; *AL1301* für Welle 3, *AL1302* für Welle 4, *AL1303* für Welle 5, *AL1304* für Welle 6, *AL1305* für Welle 7, *AL1306* für Welle 8, *AL1307* für Welle 9, *AL1308* für Welle 10 und *AL1309* für Welle 11). Nachfolgende Tabelle gibt einen Überblick über die im Spellldatensatz enthaltenen Querschnittinformationen.

Tabelle 70: AL-spezifische Querschnittvariablen im BIO-Spellldatensatz (*bio_spells*)

	Welle 2	Welle 3	Welle 4	Welle 5	...	Welle 11
Höhe des monatlichen ALG I-Bezugs?	<i>AL1300</i>	<i>AL1301</i>	<i>AL1302</i>	<i>AL1303</i>	...	<i>AL1309</i>

- Liegen in Welle 11 im Vergleich zu Welle 10 nicht vor.
- Frage ET4300 zu den Hauptkunden von Selbständigen, die zuvor abhängig beschäftigt waren, ist entfallen.

5.7.5 Plausibilitätsprüfung und –bereinigung der Spellldatensätze

Auf Personenebene orientierten sich die Plausibilitätsprüfungen und Bereinigungsregeln an Welle 2 bis Welle 4. Die Prüfungen fanden analog zu Welle 4 nur innerhalb eines Spelltyps statt. Spellentypübergreifende Prüfungen wurden nicht durchgeführt. Bereinigungen und Umcodierungen wurden, wie bei den Arbeitslosengeld-II-Spellldaten, nur in den generierten Datumsvariablen vorgenommen. Auch hier wurden Jahreszeitangaben in konkrete Monatsangaben umcodiert, „8“ als Werte für unplausible Angaben gesetzt oder Datumsangaben ersetzt oder plausibilisiert. Da lediglich die generierten Datumsvariablen editiert wurden, stehen Nutzerinnen und Nutzern mit den Datumsvariablen *BIO0200-BIO0500* und

AL0800-AL1100 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. So wurden bspw. offensichtlich doppelt erfasste Spells entfernt. Ebenfalls gelöscht wurden Episoden, die komplett außerhalb des abgefragten Zeitraums liegen und trotzdem erhoben wurden.

5.7.6 Fortschreibung der Spelldatensätze

Im Anschluss an die Umstrukturierung der in Welle 11 berichteten Episoden ins 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 Spells mit den in Welle 11 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 11-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 11 berichtete Spells zugespielt (3). Diese drei Varianten sollen im Folgenden kurz dargestellt werden:

1. *Fälle, in denen die Person in Welle 11 einem andauernden Spell zum Befragungszeitpunkt der Vorwelle widerspricht.*

Hat die Person der Information, dass zum Zeitpunkt der Vorwelle ein aktuell andauernder Spell vorlag, in der Fortschreibungsfrage explizit bzw. durch Angabe eines vor dem Interviewdatum der Vorwelle 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 Vorwelle hinaus keine verlässlichen Aussagen über die weitere Dauer des Spells möglich sind, wird angenommen, dass dieser im Interviewmonat der Vorwelle geendet hat. Die berichteten und generierten Variablen zum Enddatum des Spells (*BIO0400*, *BIO0500* sowie *emonat*, *ejahr*) sowie die Nachfrage, ob ein Spell noch andauert (*BIO0600*), bleiben dabei unverändert⁴⁶. Das generierte Enddatum des Spells (*emonat*; *ejahr*) wurde bereits in der Vorwelle auf das Interviewdatum der Vorwelle gesetzt.

2. *Fälle, in denen die Person Angaben zum Enddatum eines in der Vorwelle aktuell an-*

⁴⁶ 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 Datumsvariablen enthalten weiterhin die letzte gültige Information, die hier dem Interviewdatum der Welle entspricht, in welcher der Spell zensiert war.

dauernden Spells macht.

Liegen in Welle 11 Angaben zum Enddatum eines in der Vorwelle zensierten Spells vor, so wurde der in der Vorwelle zensierte Spell mit den aktuellen Angaben fortgeschrieben. Bei ET-Spells wurden das erhobene Enddatum (*BIO0400*; *BIO0500*), das generierte Enddatum (*emonat*; *ejahr*), die Nachfrage, ob der Spell aktuell noch andauert (*BIO0600*), der Grund für die Auflösung des Arbeitsvertrags (*ET2300*), die generierten Variablen zur Stellung im Beruf und der wöchentlichen Arbeitszeit (*stib*, *az1*, *az2*) sowie der Zensierungsindikator (*zensiert*) mit den in Welle 11 erhobenen Informationen überschrieben. Weiterhin wurden die auf die Welle 11 bezogenen Querschnittinformationen (*ET0559* bis *ET2209*) übernommen.

Bei AL-Spells wurden das erhobene Enddatum (*BIO0400*; *BIO0500*), das generierte Enddatum (*emonat*; *ejahr*), die Nachfrage, ob der Spell aktuell noch andauert (*BIO0600*), der Grund für die Beendigung der Arbeitslosigkeit (*AL0600*, *AL0601*) sowie der Zensierungsindikator (*zensiert*) mit den in Welle 11 erhobenen Informationen überschrieben. Weiterhin wurden die auf die Welle 11 bezogenen Querschnittinformationen (*AL1309*) übernommen. Bei den AL-Spellaten gibt es zudem noch die Besonderheit, dass innerhalb des AL-Spells die Arbeitslosengeld-I-Bezugsepisode (ALG I-Bezug) erhoben wird. Welche Informationen fortgeschrieben werden, ist abhängig davon, ob in diesem AL-Spell zu einem früheren Zeitpunkt schon einmal ALG I bezogen wurde und ob dieser Bezug in der Vorwelle noch andauert hat:

Lag in der Vorwelle in dem fortzuschreibenden AL-Spell auch ein andauernder Bezug von ALG I vor, wurde 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 11 erhobenen Informationen überschrieben.

Wurde in dem fortzuschreibenden AL-Spell in den Vorwellen noch nie ALG-I bezogen, wurden alle Informationen zum ALG-I-Bezug mit den in Welle 11 erhobenen Informationen überschrieben. Neben dem Indikator, ob im AL-Spell schon einmal ALG I 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 ALG-I-Bezug stattgefunden, der aber bereits in der Vorwelle beendet war, wurden in diesen Spells keine Veränderungen vorgenommen.

3. *In Welle 11 neu berichtete Spells, die keine in der Vorwelle zensierte Episode fortschreiben.*

In Welle 11 neu berichtete Spells wurden dem BIO-Spellatensatz angehängt. Danach wurde der Spellzähler neu generiert, um eine lückenlose Variable *spellnr* zu schaffen. Das Fortschreiben der Spellatensätze hat keine Auswirkung auf die Spellnummern des Vorwells-SUF. Bereits im SUF zu Welle 10 enthaltene Spells (*spellnret*, *spellnral*, *spellnrlu*, *spellnrl*) behalten ihre Spellnummer. Die neuen Spells aus Welle 11 werden an den entsprechenden Datensatz angehängt und die

Spellnummern fortgeschrieben.

5.8 Ein-Euro-Spell Datensatz (*ee_spells*)

In der 4. Welle wurde das Konzept zur Erhebung der Maßnahmeteilnahme grundlegend überarbeitet. Anstelle des MN-Spell Datensatzes trat ab Welle 4 der Ein-Euro-Spell Datensatz (*ee_spells*). Dieser wurde in Welle 11 erneut fortgeschrieben. Als Referenzzeitpunkt, ab dem Ein-Euro-Jobs erfragt wurden, diente für Welle 11 Januar 2016.

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

In der 11. Welle wurden andauernde Ein-Euro-Spells fortgeschrieben. Für die Fortschreibung der in der Vorwelle andauernden und damit im Spell Datensatz rechtszensierten Spells sind in den Personenfragebögen Dependent-Interviewing-Fragen enthalten.

5.8.2 Struktur des Ein-Euro-Spell Datensatzes

Durch die Integration der in Welle 11 berichteten Ein-Euro-Jobs in den Ein-Euro-Spell Datensatz werden neue Variablen ergänzt, die auf eine bestimmte Welle bezogen sind. Folgende Tabelle gibt einen Überblick über die im Ein-Euro-Spell Datensatz enthaltenen Querschnittinformationen.

Tabelle 71: Querschnittvariablen im Ein-Euro-Spell Datensatz (*ee_spells*)

	Welle 4	Welle 4	...	Welle 11
Arbeitszeit pro Woche im EEJ	<i>EE1100</i>	<i>EE1101</i>	...	<i>EE1107</i>
EEJ das gleiche was festangestellte Kollegen machen	<i>EE1200</i>	<i>EE1201</i>	...	<i>EE1207</i>
Welche Art der Ausbildung für EEJ erforderlich	<i>EE1300</i>	<i>EE1301</i>	...	<i>EE1307</i>
Ganze Zeit Arbeit oder auch Schulung/Kurs?	<i>EE1400</i>	<i>EE1401</i>	...	<i>EE1407</i>
Beurteilung EEJ	<i>EE1500a- EE1500h</i>	<i>EE1501a- EE1501h</i>	...	<i>EE1507a- EE1507h</i>

Im EEJ-Spell Datensatz ist zu beachten, dass Spells auch vorhanden sind, wenn der Ein-Euro-Job nicht durchgeführt wurde, d.h. es nicht zur Teilnahme kam.

5.8.3 Plausibilitätsprüfung und –bereinigung des EEJ-Spelldatensatzes

Der EEJ-Spelldatensatz zur Teilnahme an Ein-Euro-Jobs wurde plausibilitätsgeprüft und bereinigt. Die Plausibilitätsprüfungen beinhalteten Prüfungen der Zeitangaben, des Referenzzeitpunkts für in Welle 11 neu aufgenommene Spells (Januar 2016) sowie Prüfungen auf logische Inkonsistenzen bei Befragten mit mehreren Ein-Euro-Job-Spells.

Bereinigungen und Umcodierungen wurden nur in den generierten Datumsvariablen (*bmonat*, *bjahr*, *emonat*, *ejahr*) vorgenommen. Jahreszeitangaben wurden dort in konkrete Monatsangaben umcodiert, „-8“ als Werte für unplausible Angaben gesetzt oder Datumsangaben ersetzt oder plausibilisiert. Danach wurde ein Spellzähler *spellnr* generiert. Die Generierung wurde analog zu den chronologischen Zählern in den BIO-Spelldatensätzen vorgenommen, wobei nicht teilgenommene Spells bei der Sortierung miteinbezogen wurden und somit ihre ursprüngliche Erhebungsposition innerhalb der Befragungswelle beibehalten haben. Bereits im SUF zu Welle 10 enthaltene Spells haben ihre *spellnr* auch im SUF Welle 11 beibehalten.

6 Gewichtung Welle 11

Die Konstruktion der Gewichte für die elfte Welle erfolgte im Wesentlichen analog zur Welle 10 (vgl. Berg et al. (2017)). Ausgangspunkt für die Gewichtung der elften Welle und des Längsschnitts von Welle 10 auf Welle 11 waren nunmehr die Querschnittsgewichte aus Welle 10 für Haushalte bzw. Personen. Die zwei Gewichte jedes Haushalts und die zwei Gewichte jeder Person wurden erneut fortgeschrieben. Dieses Kapitel des Datenreports stellt insbesondere die technischen Details und Modelle bei der Erzeugung der Gewichte für Welle 11 dar. Ein Überblick über das Gewichtungskonzept von PASS findet sich in Kapitel 8 (Trappmann (2013a)) des User Guides von PASS (Bethmann, Fuchs und Wurdack (Hg.)(2013)). Anwendungsbeispiele sind dort in Kapitel 12 zu finden (Trappmann (2013b)).

6.1 Designgewichte für die Bestandsauffrischung (EWO) der Welle 11

In Welle 11 wurde PASS durch eine Bestandsauffrischung (Aufstockung) der Bevölkerungsstichprobe (EWO-Aufstockung) ergänzt. Nähere Informationen zur Auswahl der Primary Sampling Units (PSUs), zur Auswahl der Gemeinden für die Aufstockung der Bevölkerungsstichprobe und zur Auswahl der Haushalte für die Aufstockung der Bevölkerungsstichprobe finden sich im FDZ-Datenreport 06/2012 der Welle 5.

Die Designgewichte für die Bestandsauffrischung (EWO) der Bevölkerungsstichprobe (sample=15) sind als Kehrwert der Auswahlwahrscheinlichkeiten auf den verschiedenen Stufen des Ziehungsdesigns definiert. Die Auswahlwahrscheinlichkeiten bestimmen sich dabei über drei Auswahlstufen. Die Auswahlwahrscheinlichkeit des Points (korrigiert um ausgefallene Points in Welle 11), die Auswahlwahrscheinlichkeit der Gemeinde in den Fällen, in denen sich die PLZ über mehrere Gemeinden erstreckt (in allen anderen Fällen = 1) und die Auswahlwahrscheinlichkeit der Person im Point. Die Multiplikation dieser drei Auswahlwahrscheinlichkeiten ergibt die Auswahlwahrscheinlichkeit der ausgewählten Person in der Bruttostichprobe.

Die Transformation der Personenstichprobe in eine Haushaltsstichprobe stellt einen zusätzlichen Schritt bei der Aufstockung der Bevölkerungsstichprobe dar, der nur für die realisierten Fälle durchgeführt werden kann. Dieser zusätzliche Gewichtungsschritt, der die unterschiedlichen Auswahlwahrscheinlichkeiten auf Grund der unterschiedlichen (reduzierten) Haushaltgröße korrigiert, wurde nach der Berechnung der Teilnahmepropensities durchgeführt, d.h. nach dem Übergang von der Bruttostichprobe zur Nettostichprobe, indem die Auswahlwahrscheinlichkeit der Personen mit der geschätzten Teilnahmepropensity und der Zahl der Zielpersonen im Haushalt multipliziert wurde.

Eine detaillierte Beschreibung der Auswahlstufen und Berechnung der Auswahlwahrscheinlichkeiten für die strukturgleiche Bestandsauffrischung (EWO) in Welle 5, findet sich auch im Datenreport der Welle 5.

6.2 Integration der Designgewichte Bestandsauffrischung (EWO) mit den bestehenden Gewichten der Bevölkerungsstichprobe

Die Integration der Designgewichte für die Bestandsauffrischung (EWO) mit den bestehenden Gewichten der Bevölkerungsstichprobe (Microm, EWO Auffrischung Welle 5) erfolgte wie bisher nach den Propensity-Modellen, aber vor der Kalibrierung.

Die Gewichte der vereinigten Bevölkerungsstichproben sollen die Microm-Stichprobe aus Welle 1, die EWO-Auffrischung aus Welle 5 und die neue EWO-Auffrischung aus Welle 11 auf die Haushalte der Bundesrepublik hochrechnen. Deshalb wurden zunächst für die Bestandsstichprobe und die EWO-Auffrischung jeweils nach dem bisherigen Konzept separate Gewichte berechnet. Für das Bevölkerungsgewicht vor Kalibrierung wurde dann der Bestand Microm plus EWO-Auffrischung aus Welle 5 mit der EWO-Auffrischung Welle 11 (sample=15) über eine Konvexkombination integriert (Spieß & Rendtel 2000).

Im Anschluss daran erfolgte wie bisher die Integration von Bevölkerungsgewichten und BA-Gewichten zu Gesamtgewichten.

6.3 Designgewichte für die Bestandshaushalte der Welle 11

Unter Berücksichtigung von Zuzügen aus dem Inland wurden für die elfte Welle nunmehr aus den Haushaltsquerschnittsgewichten der zehnten Welle neue „Haushaltsdesigngewichte“ für die Welle 11 generiert. Dies erfolgte erneut über das für Welle 2 beschriebene „Weight-Share“-Verfahren (vgl. Gebhardt et.al (2009)). 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, von Welle 1 bis Welle 10 ins Sample zu gelangen. Das neue Designgewicht für Teilstichprobe i $dw_{i}hh_{11}$ errechnet sich also aus dem alten Querschnittsgewicht $wq_{i}hh_{10}$:

$$1/dw_{i}hh_{11} = 1/wq_{i}hh_{10} + (n_{sample_i}/n_{population_i})$$

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

6.4 Designgewichte für die Zugangsstichprobe der Welle 11

In der Welle 11 wurde das Panel erneut durch Nachziehung von Neuzugängen zum Leistungsbezug aufgefrischt. Eine Ziehungschance hatten alle Haushalte, die im Juli 2016 Leistungen bezogen haben, im selben Monat 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007 und 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 Jahre Leistungen bezogen hatte. Nachgezogen wurde wiederum in den 300 Points der ersten Welle und den 100 Aufstockungspoints der Welle 5. 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 Zugangsstichprobe für Welle 11 befindet sich für Fälle mit *sample* = 16 (normale Zugangsstichprobe) bzw. *sample* = 17 (Zugangsstichprobe syrische/ irakische Haushalte) in der Variable *dw_ba*.

6.5 Wiederteilnahme-Propensity Haushalte

In diesem Schritt wurde wiederum analog zum Vorgehen für Welle 10 für jeden Haushalt, der in Welle 10 teilgenommen hatte, die Wahrscheinlichkeit zur Wiederteilnahme an Welle 11 über Logit-Modelle für Panelbereitschaft, Erreichbarkeit und Teilnahme geschätzt. Auch Haushalte mit letzter Teilnahme in Welle 9, ohne Teilnahme in Welle 10 (temporäre Ausfälle), wurden in der Modellierung für Welle 11 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. die Zahl der Kontaktversuche. Die vorhergesagten Wahrscheinlichkeiten aller drei Modelle wurden multipliziert. Der Kehrwert dieses Produkts ist für jede Welle in der Variable *hpleib* zu finden. Das Längsschnittgewicht eines Haushalts aus einer der Stichproben der Welle 1 für den gesamten möglichen Zeitraum [t1 bis t11] zwischen allen elf Wellen lässt sich dann als Produkt aus Querschnittgewicht zu t1, *hpleib* (Welle 1 zu Welle 2) und *hpleib* (Welle 2 zu Welle 3 usw.) berechnen (vgl. hierzu auch Kapitel 12 des User Guides Trappmann (2013b)).

Die Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Wiederholerhaushalte sind → **Tabelle A9** zu entnehmen.

Die Logit-Modelle zur Wiederteilnahme für Panelbereitschaft, Erreichbarkeit und Teilnahme sind → **Tabelle A10** zu entnehmen.

6.6 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 vorausgegangenen Wellen noch nicht befragt wurden. Modelliert wurden also die Teilnahmepropensities für die erstbefragten Split-Haushalte, getrennt nach dem Kriterium in Welle 10 entstandener Split-Haushalt (Split-W10-Haushalte) und in Welle 11 entstandener Split-Haushalt (Split-W11-Haushalte). Die Wahrscheinlichkeit zur Wiederteilnahme wurde jeweils über Logit-Modelle für Erreichbarkeit und Teilnahme geschätzt. Bei den Split-W11-Haushalten wurde auf eine getrennte Modellierung der Erreichbarkeit und Teilnahmewahrscheinlichkeit verzichtet, da die Erreichbarkeit nahezu bei allen Haushalten gegeben war. Hier wurde für alle Haushalte nur ein Gesamtmodell 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 *hpsplit* zu finden.

Die Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der erstbefragten Split-Haushalte (Welle 10 und Welle 11) sind → [Tabelle A11](#) zu entnehmen.

Die Logit-Modelle zur Erstteilnahme von Split-W10-Haushalten sind → [Tabelle A12](#) zu entnehmen.

Die Logit-Modelle zur Erstteilnahme von Split-W11-Haushalten für Erreichbarkeit und Teilnahme sind → [Tabelle A13](#) zu entnehmen.

6.7 Ausfallgewichtung für Haushalte aus der Auffrischungsstichprobe der BA-Neuzugänge Welle 11

Für die Haushalte aus der Auffrischungsstichprobe der BA-Neuzugänge zum ALG II-Bezug in Welle 11 (*sample* = 16, normale Zugangsstichprobe bzw. *sample* = 17, Zugangsstichprobe syrische/ irakische Haushalte) wurde erneut eine Ausfallmodellierung (Teilnahme) analog zur Auffrischungsstichprobe der Welle 10 vorgenommen, jeweils für Erreichbarkeit und Teilnahme. Es wurde ein integriertes Modell für beide Auffrischungsstichproben geschätzt, die Variable (*samaufftyp_2*) unterscheidet die Zugehörigkeit zur Teilstichprobe der Syrer und Iraker und der normalen Zugangsstichprobe. Die aus dem Modell abgeleitete Teilnahmewahrscheinlichkeit befindet sich in der Variable *prop_t0*.

Die Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der BA-Auffrischungsstichprobe Welle 11 sind → [Tabelle A14](#) zu entnehmen.

Die Logit-Modelle zur Erstteilnahme der BA-Auffrischungsstichprobe Welle 11 sind → [Tabelle A15](#) zu entnehmen.

6.8 Ausfallgewichtung für Haushalte aus der Bestandsauffrischung (EWO) Welle 11

Für die Haushalte der erstmalig gezogenen Bestandsauffrischungsstichprobe (EWO) für die allgemeine Bevölkerung (*sample*=15) wurde eine zweistufige Ausfallmodellierung (Erreichbarkeit und Teilnahme) vorgenommen. Die daraus abgeleitete Teilnahmewahrscheinlichkeit befindet sich in der Variable *prop_t0*.

Die Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Bestandsauffrischungsstichprobe (EWO) Welle 11 sind → [Tabelle A16](#) zu entnehmen.

Die Logit-Modelle zur Erstteilnahme für Erreichbarkeit und Teilnahme der Bestandsauffrischungsstichprobe (EWO) Welle 11 sind → [Tabelle A17](#) zu entnehmen.

6.9 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. Die Teilnahmepropensities für Personen in Welle 11 wurden unter Einschluss von zusätzlichen Personenmerkmalen über Logit-Modelle für Panelbereitschaft, 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 [t1 bis t11] über alle elf Wellen lässt sich dann als Produkt aus Querschnittgewicht zu t1, *ppbleib* (Welle 1 zu Welle 2) und *ppbleib* (Welle 2 zu Welle 3 usw.) berechnen.

Die Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der Personen mit wiederholter Teilnahme sind → **Tabelle A18** zu entnehmen.

Die Logit-Modelle zur Wiederteilnahme Personen für Panelbereitschaft, Erreichbarkeit und Teilnahme sind → **Tabelle A19** zu entnehmen.

6.10 Integration der Gewichte zum Gesamtgewicht vor Kalibrierung

In diesem Schritt ging es erneut darum, die durch die Ausfallmodellierung modifizierten Haushaltsgewichte der neuen Auffrischungsstichprobe der BA-Neuzugänge Welle 11 und der Bestandsstichprobe (einschließlich der Auffrischer aus Welle 2 bis 9) zu kombinieren. Die mehrfache Auswahlwahrscheinlichkeit, die ein nun gezogener Leistungsempfänger, der in den vorausgegangenen Jahren 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 im Juli 2006, im Juli 2007, im Juli 2008, im Juli 2009, im Juli 2010, im Juli 2011, im Juli 2012, im Juli 2013, im Juli 2014, im Juli 2015 oder im Juli 2016 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 im Juli 2016. Dazu wurde die Inklusionswahrscheinlichkeit in der jeweils anderen Stichprobe für Fälle aus der Microm-Stichprobe (Welle 1), der EWO-Auffrischungen (Welle 5 und Welle 11) und der neuen Auffrischungsstichprobe der BA-Neuzugänge (Welle 11) geschätzt. Für Fälle aus der Auffrischungsstichprobe der BA-Neuzugänge wurde die mittlere Welle 1-Auswahlwahrscheinlichkeit in der Microm-Stichprobe, die mittlere Welle 5-Auswahlwahrscheinlichkeit in der EWO-Auffrischung (Welle 5), bzw. die mittlere Welle 11-Auswahlwahrscheinlichkeit in der EWO-Auffrischung (Welle 11) in der jeweiligen Postleitzahl und jeweils die durchschnittliche Teilnahmewahrscheinlichkeit (für Welle 1 bis Welle 11) in dieser Stichprobe unterstellt. Für Fälle aus der Microm- oder EWO-Auffrischungsstichprobe (Welle 5) wurde, falls es sich (laut Befragungsdaten) um Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug zwischen den zehn letzten

Stichprobenziehungszeitpunkten (Welle 2 bis Welle 10) handelt, die durchschnittliche Auswahlwahrscheinlichkeit eines Auffrischungshaushalts (BA-Neuzugänge Welle 11) in der jeweiligen PLZ und die durchschnittliche Teilnahmewahrscheinlichkeit in dieser Stichprobe unterstellt. Für Fälle aus der neuen EWO-Auffrischungsstichprobe (Welle 11) wurde, sofern dieser Haushalt jemals im Bezug war (Frage PA0980), die durchschnittliche Auswahlwahrscheinlichkeit aller BA-Stichproben unterstellt. Danach wurden die beiden Gewichte zu einem neuen Gesamtgewicht integriert.

6.11 Integration temporärer Ausfälle (Haushalte)

In Welle 11 konnten, wie auch schon in den Wellen zuvor, 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 Welle 9 (die Welle 10 ist der temporäre Ausfall).

Für die temporären Ausfälle wurde dann die Wahrscheinlichkeit der Nichtteilnahme in Welle 10 gegeben die Teilnahme in Welle 9 (Nichtteilnahme-Propensities Welle 10) und die Wahrscheinlichkeit der Teilnahme in Welle 11 gegeben die Nichtteilnahme in Welle 10 (Teilnahme-Propensities Welle 11) bestimmt. Die Wahrscheinlichkeit der Nichtteilnahme in Welle 10 ergibt sich über 1-Teilnahmewahrscheinlichkeit in Welle 10.

Die beschriebenen Teilnahme- bzw. Nichtteilnahme-Propensities wurden wiederum über Logit-Modelle geschätzt. Die vorhergesagten Wahrscheinlichkeiten der jeweiligen Modelle wurden multipliziert. Das modifizierte Haushaltsgewicht der temporären Ausfälle ergibt sich dann über die Multiplikation der kalibrierten Haushaltsgewichte der Welle 9 mit dem Kehrwert dieses Produkts.

Die Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der temporären Ausfälle sind → **Tabelle A20** zu entnehmen.

Die Logit-Modelle temporärer Ausfälle sind → **Tabelle A21** zu entnehmen.

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:

$dw_{ihh_{temp.Ausfall}} * (n_{temp.Ausfall_i} / (n_{temp.Ausfall_i} + n_{Bestand_i}))$ für die temporären Ausfälle

$dw_{ihh_{Bestand}} * (n_{Bestand_i} / (n_{temp.Ausfall_i} + n_{Bestand_i}))$ für die Bestandsstichprobe.

6.12 Kalibrierung zum Haushaltsgewicht Welle 11 Querschnitt

Anschließend fand wieder eine Kalibrierung der modifizierten Designgewichte inklusive der Ausfallgewichtung auf Haushaltsebene mittels GREG-Verfahren an die Eckwerte des Statistischen Bundesamts für 2016 und für Haushalte mit Leistungsbezug an die Eckwerte der Statistik der Bundesagentur für Arbeit für Juli 2016 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 (357.551) verwendet. Dabei werden diejenigen Fälle der bisherigen Stichproben aus den Wellen 1 bis 11 auf die Eckwerte zum Arbeitslosengeld-II-Bezug der Bundesagentur für Arbeit hochgerechnet, die laut Welle 11 des Surveys im Juli 2016 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 getrimmten Eingangsgewichten zu den kalibrierten Gewichten. Zu beachten war dabei, dass engere Grenzen für die Gewichte in geringerer Streuung und damit geringerer Varianz der Schätzungen resul-

tieren, 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.

6.13 Kalibrierung der BA-Stichprobe

Die Grundgesamtheit der kumulierten BA-Stichprobe aus allen elf Wellen bilden alle Haushalte in Deutschland mit mindestens einer Bedarfsgemeinschaft mit Leistungen nach dem SGB II zu einem der bisher zehn Ziehungszeitpunkte (im Juli 2006, Juli 2007, Juli 2008, Juli 2009, Juli 2010, Juli 2011, Juli 2012, Juli 2013, Juli 2014, Juli 2015 oder Juli 2016). Eine Kalibrierung findet in Welle 10 dagegen nur an die Eckwerte der BA-Statistik vom Juli 2016 statt. Damit beeinflusst die Kalibrierung nur die Gewichte der Haushalte aus der BA-Stichprobe, in denen auch im Juli 2016 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 147,46 bis 5007,61. Das Verhältnis der Gesamthochrechnungsfaktoren nach Kalibrierung zu getrimmten Designgewichten wurde nach unten auf minimal 0,1 und nach oben auf maximal 3,0 begrenzt. Danach bewegen sich die Gesamthochrechnungsfaktoren nach Kalibrierung im Bereich von minimal 14,75 bis maximal 7965,07.

Für folgende Merkmale wurde kalibriert:

Bedarfsgemeinschaften (BG) 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 Jahren in der Bedarfsgemeinschaft, nach West/Ost
- Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Kinder unter 15 Jahren in der Bedarfsgemeinschaft, nach West/Ost
- Anzahl BG im Leistungsbezug nach SGB II, die aus einer/einem Alleinerziehenden mit Kind(ern) bestehen, 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 (357.551).

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 Modalwert der jeweiligen Variablen.

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 Tabellenanhang dargestellten Werten.

Die Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Haushalte) sind → **Tabelle A22** zu entnehmen.

Die Kenngrößen der Verteilung der Gewichte (BA-Stichprobe, Haushalte) sind → **Tabelle A23** zu entnehmen.

6.14 Kalibrierung der Bevölkerungsstichprobe

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 2524,1 bis 32876,1. Das Verhältnis der Gesamthochrechnungsfaktoren nach Kalibrierung zu getrimmten Designgewichten wurde nach unten auf minimal 0,3 und nach oben auf maximal 2,5 begrenzt. Danach bewegen sich die Gesamthochrechnungsfaktoren nach Kalibrierung im Bereich von minimal 757,2 bis maximal 77904,1.

Für folgende Merkmale wurde kalibriert:

1. Bedarfsgemeinschaften: Grundlage BA-Statistik:
 - Anzahl BG im Leistungsbezug nach SGB II nach West/Ost
 - Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Personen unter 65 Jahre in der Bedarfsgemeinschaft
 - Anzahl BG im Leistungsbezug nach SGB II nach Zahl der Kinder unter 15 Jahre in der Bedarfsgemeinschaft
 - Anzahl BG im Leistungsbezug nach SGB II, die aus einer/einem Alleinerziehenden mit Kind(ern) besteht
2. Haushalte: Grundlage Mikrozensus 2016:
 - 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 Modalwert der jeweiligen Variablen.

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 Tabellenanhang dargestellten Werten.

Die Soll-Verteilungen und Verteilungen nach Kalibrierung (Bevölkerungsstichprobe, Haushalte) sind → **Tabelle A24** zu entnehmen.

Die Kenngrößen der Verteilung der Gewichte (Bevölkerungsstichprobe, Haushalte) sind → **Tabelle A25** zu entnehmen.

6.15 Kalibrierung der 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 149,4 bis 24178,7. Das Verhältnis der Gesamthochrechnungsfaktoren nach Kalibrierung zu getrimmten Designgewichten wurde nach unten auf minimal 0,1 und nach oben auf maximal 3,0 begrenzt. Danach bewegen sich die Gesamthochrechnungsfaktoren nach Kalibrierung im Bereich von minimal 14,9 bis maximal 32772,5.

Für folgende Merkmale wurde kalibriert:

1. 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 Kind(ern) bestehen, nach West/Ost
2. Haushalte Grundlage Mikrozensus 2016:

- 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 (357.551) 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 Modalwert der jeweiligen Variablen.

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 Tabellenanhang dargestellten Werten.

Die Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Haushalte) sind → **Tabelle A26** zu entnehmen.

Die Kenngrößen der Verteilung der Gewichte (Gesamtstichprobe, Haushalte) sind → **Tabelle A27** zu entnehmen.

6.16 Kalibrierung zum Personengewicht Welle 11 Querschnitt

Die Kalibrierung der Personengewichte wurde – wie in den Vorgängerwellen – unter der Restriktion möglichst geringer Abweichung von den kalibrierten Haushaltsgewichten vorgenommen. Die kalibrierten Haushaltsgewichte wurden 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 (463.063) verwendet. Es wurden erneut diejenigen Fälle der Stichproben aus allen Wellen auf die Eckwerte zum Arbeitslosengeld-II-Bezug der Bundesagentur für Arbeit hochgerechnet, die laut Welle 11 des Surveys im Juli 2016 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.

6.17 BA-Stichprobe

Die Grundgesamtheit der kumulierten BA-Stichprobe aller elf Wellen bilden alle Personen im Alter ab 15 Jahre, die in einem Haushalt leben, in dem es zu einem der bisher elf Ziehungszeitpunkte (im Juli 2006, Juli 2007, Juli 2008, Juli 2009, Juli 2010, Juli 2011, Juli 2012, Juli 2013, Juli 2014, Juli 2015 oder Juli 2016) 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 2016 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 55,6 bis 5072,9. Das Verhältnis der Gesamthochrechnungsfaktoren nach Kalibrierung zu getrimmten Designgewichten wurde nach unten auf minimal 0,1 und nach oben auf maximal 4,0 begrenzt. Danach bewegen sich die Gesamthochrechnungsfaktoren nach Kalibrierung im Bereich von minimal 13,7 bis maximal 14476,6.

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 (463.063) 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 Modalwert der

jeweiligen Variablen.

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 Tabellenanhang dargestellten Werten.

Die Soll-Verteilungen und Verteilungen nach Kalibrierung (BA-Stichprobe, Personen) sind → **Tabelle A28** zu entnehmen.

Die Kenngrößen der Verteilung der Gewichte (BA-Stichprobe, Personen) sind → **Tabelle A29** zu entnehmen.

6.18 Bevölkerungsstichprobe

Die Grundgesamtheit bilden alle Personen ab 15 Jahren in Privathaushalten in Deutschland. Ausgangspunkt für die Kalibrierung waren die kalibrierten Haushaltsgewichte der Bevölkerungsstichprobe. 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 2884,3 bis 37687,6. Das Verhältnis der Gesamthochrechnungsfaktoren nach Kalibrierung zu getrimmten Designgewichten wurde nach unten auf minimal 0,1 und nach oben auf maximal 5,0 begrenzt. Danach bewegen sich die Gesamthochrechnungsfaktoren nach Kalibrierung im Bereich von minimal 288,4 bis maximal 143976,4.

Für folgende Merkmale wurde kalibriert:

1. Leistungsempfänger Grundlage BA-Statistik:

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

2. Bevölkerung Grundlage Mikrozensus 2016:

- 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

3. 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 im 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 Modalwert der jeweiligen Variablen.

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 Tabellenanhang dargestellten Werten.

Die Soll-Verteilungen und Verteilungen nach Kalibrierung (Bevölkerungsstichprobe, Personen) sind → **Tabelle A30** zu entnehmen.

Die Kenngrößen der Verteilung der Gewichte (Bevölkerungsstichprobe, Personen) sind → **Tabelle A31** zu entnehmen.

6.19 Gesamtstichprobe

Die Grundgesamtheit bilden wie bei der Bevölkerungsstichprobe alle 15-jährigen und älteren Personen in Privathaushalten in Deutschland. Ausgangspunkt für die Kalibrierung waren die kalibrierten Haushaltsgewichte der Gesamtstichprobe. 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 166,7 bis 27811,3. Das Verhältnis der Gesamthochrechnungsfaktoren nach Kalibrierung zu getrimmten Designgewichten wurde nach unten auf minimal 0,1 und nach oben auf maximal 5,0 begrenzt.

Danach bewegen sich die Gesamthochrechnungsfaktoren nach Kalibrierung im Bereich von minimal 16,7 bis maximal 111636,9.

Für folgende Merkmale wurde kalibriert:

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

2. Bevölkerung Grundlage Mikrozensus 2016:

- 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 und 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

3. 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 im 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 (463.063) 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 Modalwert der jeweiligen Variablen.

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 Tabellenanhang dargestellten Werten.

Die Soll-Verteilungen und Verteilungen nach Kalibrierung (Gesamtstichprobe, Personen) sind → **Tabelle A32** zu entnehmen.

Die Kenngrößen der Verteilung der Gewichte (Gesamtstichprobe, Personen) sind → **Tabelle A33** zu entnehmen.

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

Schließlich verblieben auch in der Welle 11 noch Haushalte und Personen, denen durch Kalibrierung kein BA-Haushaltsquerschnittsgewicht oder BA-Personenquerschnittsgewicht zugewiesen werden konnte. Die Anzahl dieser Haushalte ist nun in Welle 11 erneut größer als in den Wellen zuvor, da ein größerer Teil der BA-Stichprobe der Wellen 1 bis 10 mittlerweile aus dem Leistungsbezug ausgeschieden ist. Es handelt sich um die folgenden drei Gruppen ohne Leistungsbezug im Juli 2016, die aber zur Grundgesamtheit der BA-Stichprobe (Haushalte bzw. Personen in Haushalten mit Arbeitslosengeld-II-Bezug im Juli 2006, Juli 2007, Juli 2008, Juli 2009, Juli 2010, Juli 2011, Juli 2012, Juli 2013, Juli 2014, Juli 2015 oder Juli 2016) gehören:

- Aus der Auffrischungsstichprobe: Nicht-BG-Mitglieder im Haushalt: Das Personengewicht wurde hier aus dem BA-Haushaltsgewicht in Welle 11 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 2016 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 10 erhalten als neues BA-Personengewicht das BA-Haushaltsgewicht ihres Haushalts für Welle 11 dividiert durch den Anteil solcher Personen, die teilnehmen, vorausgesetzt ihr Haushalt nimmt teil.
- Nicht-BG-Mitglieder in Bestandshaushalten, die im Juli 2016 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.

- Diese Personen und Haushalte wurden zusätzlich an einen Eckwert für die Personen bzw. Bedarfsgemeinschaften, die den Grundsicherungsbezug verlassen haben, angepasst. Die exakte Populationsgröße dieser Gruppe ist dabei unbekannt, sie lässt sich aber aus der Summe aller kumulierten BA-Teilstichproben abzüglich der aktuell im Bezug befindlichen Personen bzw. Bedarfsgemeinschaften näherungsweise bestimmen. Erstmals ab Welle 11 wurde dieser Eckwert auch um die geschätzte kumulierte Sterberate dieser Personengruppe seit 2005 um 1,6% reduziert. Die entsprechende Zahl der aus dem Grundsicherungsbezug abgegangenen Personen bemisst sich nun für Welle 11 auf 5.930.093, die Zahl der aus dem Grundsicherungsbezug abgegangenen Bedarfsgemeinschaften auf 4.367.707.

7 Anhang: Kurzbeschreibung des Datensatzes

Die Kurzbeschreibung des Datensatzes ist in den → [Tabellenanhang](#) dargestellt.

Literatur

AAPOR (2011). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 7. Auflage. Lanexa: AAPOR.

Achatz, J., Hirseland, A. & Promberger, M. (2007). Rahmenkonzept für das IAB-Panel „Arbeitsmarkt und Soziale Sicherung“. In M. Promberger (Hrsg.), *Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung „Arbeitsmarkt und Soziale Sicherung“*, *IAB-Forschungsbericht 12/2007* (S. 11-32), Nürnberg.

Andreß H.-J., Burkatzki, E., Lipsmeier, G., Salentin, K., Schulte, K. & Strengmann-Kuhn, W. (1996). *Leben in Armut. Analysen der Verhaltensweisen armer Haushalte mit Umfragedaten*. Endbericht des DFG-Projekts „Versorgungsstrategien privater Haushalte im unteren Einkommensbereich (VuE)“. Bielefeld.

Andreß, H.-J. & Lipsmeier, G. (1995). Was gehört zum notwendigen Lebensstandard und wer kann ihn sich leisten? Ein neues Konzept zur Armutsmessung. *Aus Politik und Zeitgeschichte* 31-32, S. 35-49.

Andreß, H.-J. & Lipsmeier, G. (2001). *Armut und Lebensstandard. Gutachten im Rahmen des Armuts- und Reichtumsberichts der Bundesregierung*. BMAS. Bonn.

Beckmann, P. & Trometer, R. (1991). Neue Dienstleistungen des ALLBUS: Haushalts- und Familientypologien, Klassenschema nach Goldthorpe. *ZUMA-Nachrichten*, 28, S. 7-17.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Marwinski, K., Gebhardt, D., Wenzig, C. & Wetzels, M. (2010). Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS). Bd. 1: Datenreport Welle 3. *FDZ Datenreport 06/2010*. Nürnberg.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, M., Bethmann, A., Fuchs, B., Huber, M. & Trappmann, M. (2014). Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS). Bd. 1: Datenreport Welle 7. *FDZ Datenreport 02/2014*. Nürnberg.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, M., Bethmann, A., Fuchs, B., Huber, M., Schwarz, S., Trappmann, M. & Reindl, A. (2016). Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS) Bd. 1: Datenreport Welle 9. *FDZ-Datenreport 07/2016*. Nürnberg.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, Beste, J., Dumert, S., Frodermann, C., Fuchs, B., Schwarz, S., Trappmann, M. & Trenkle, S. (2017). Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS) Bd. 1: Datenreport Welle 10. *FDZ-Datenreport 07/2017*. Nürnberg.

Bethmann, A., Fuchs, B. & Wurdack, A. (Hrsg.)(2013). User Guide "Panel Labour Market and Social Security"(PASS). Wave 6. *FDZ Datenreport 07/2013*. Nürnberg.

Brauns, H. & Steinmann, S. (1999). Educational Reform in France, West-Germany and the United Kingdom. Updating the Casmin Classification. *ZUMA-Nachrichten*, 44. S. 7-45.

Bruckmeier, K., Müller, G. & Riphahn, R. (2014). Who misreports welfare receipt in surveys? *Applied Economics Letters*, Vol. 21, No. 12. S. 812-816.

Bruckmeier, K., Müller, G. & Riphahn, R. (2015). Survey misreporting of welfare receipt - respondent, interviewer, and interview characteristics. *Economics Letters*, Vol. 129, S. 103 - 107.

Bundesagentur für Arbeit (2011). Klassifikation der Berufe 2010. Nürnberg.

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

Büngeler, K., Gensicke, M., Hartmann, J., Jäckle, R. & Tschersich, N. (2009). IAB-Haushaltspanel im Niedrigeinkommensbereich Welle 2 (2007/2008). Methoden- und Feldbericht. *FDZ-Methodenreport 08/2009*. Nürnberg.

Büngeler, K., Gensicke, M., Hartmann, J., Jäckle, R. & Tschersich, N. (2010). IAB-Haushaltspanel im Niedrigeinkommensbereich Welle 3 (2008/2009). Methoden- und Feldbericht. *FDZ-Methodenreport 10/2010*. Nürnberg.

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

Eggs, J. (2016). Patterns and impact of longitudinal measurement error for welfare receipt. Bielefeld: Bertelsmann 136 S. Reihe / Serie: IAB-Bibliothek Nr. 362.

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.

Europäische Gemeinschaften [EG] (2006). Verordnung (EG) Nr. 1893/2006 des europäischen Parlaments und des Rates vom 20. Dezember 2006 zur Aufstellung der statistischen Systematik der Wirtschaftszweige NACE Revision 2 und zur Änderung der Verordnung (EWG) Nr. 3037/90 des Rates sowie einiger Verordnungen der EG über bestimmte Bereiche der Statistik. Amtsblatt der Europäischen Gemeinschaften L393/1-L393-39. Brüssel.

Erikson, R. & Goldthorpe, J. (1992). *The Constant Flux. A Study of Class Mobility in*

Industrial Society. Oxford: Clarendon Press.

Erikson, R., Goldthorpe, J. & Portocarero, L. (1979). Intergenerational Class Mobility in Three Western Societies: England, France and Sweden. *British Journal of Sociology*, 30, S. 415-441.

Erikson, R., Goldthorpe, J. & Portocarero, L. (1982). Social Fluidity in Industrial Nations: England, France and Sweden. *British Journal of Sociology*, 33, S. 1-34.

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

Frick, J., Göbel, J. & Krause, P. (o.J.). \$HGEN: Generated Household Level Variables. [<http://www.diw.de/documents/dokumentenarchiv/17/60053/hgen.pdf> (8.11.2007)].

Fuchs, B. (2013). Structure of the scientific use file and its datasets. In: Bethmann, A., Fuchs, B. & Wurdack, A. (Hrsg.)(2013). User Guide "Panel Labour Market and Social Security"(PASS). Wave 6. *FDZ Datenreport 07/2013* (S. 27-47). Nürnberg.

Ganzeboom, H. & Treiman, D. (1996). Internationally Comparable Measures for Occupational Status for the 1988 International Standard Classification of Occupations. *Social Science Research*, 25, S. 201-239.

Ganzeboom, H. & Treiman, D. (2003). Three Internationally Standardised Measures for Comparative Research on Occupational Status. In H. Jürgen, P. Hoffmeyer-Zlotnik & C. Wolf (Hrsg.), *Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables* (S. 159-193), New York: Kluwer Academic / Plenum Publishers.

Ganzeboom, H. (2010). A new International Socio-Economic Index (ISEI) of Occupational Status for the International Standard Classification of Occupation 2008 (ISCO-08) constructed with Data from the ISSP 2002-2007; With an Analysis of Quality of Occupational Measurement in ISSP. Paper presented at Annual Conference of International Social Survey Programme, Lisbon, May 1 2010. [[http://www.harryganzeboom.nl/pdf/2010-ganzeboom-isei08-issp-lisbon-\(paper\).pdf](http://www.harryganzeboom.nl/pdf/2010-ganzeboom-isei08-issp-lisbon-(paper).pdf) (01.05.2010)]

Ganzeboom, H. & Treiman, D. (2010). Occupational Status Measures for the new International Standard Classification of Occupations ISCO-08; with a Discussion of the new Classification [www.harryganzeboom.nl/isol/isol2010c2-ganzeboom.pdf (24.05.2010)]

Ganzeboom, H. & Treiman, D. (2011). International Stratification and Mobility File: Conversion Tools [<http://www.harryganzeboom.nl/ismf/index.htm> (ohne Datum)]

Ganzeboom, H., De Graaf, P. & Treiman, D. (1992). A Standard International Socio-

Economic Index of Occupational Status. *Social Science Research*, 21, S. 1-56.

Gebhardt, D., Müller, G., Bethmann, A., Trappmann, M., Christoph, B., Gayer, C., Müller, B., Tisch, A., Siflinger, B., Kiesel, H., Huyer-May, B., Achatz, J., Wenzig, C., Rudolph, H., Graf, T. & Biedermann, A. (2009). Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS). Datenreport Welle 2 (2007/2008). *FDZ Datenreport 06/2009*. Nürnberg.

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

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

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

Harrison, E. & Rose, R. (2006). ESeC User Guide, Appendix 6 (SPSS-Syntax: Esec Full) [<http://www.iser.essex.ac.uk/esec/guide/docs/Appendix6.sps> (31.11.2007)]

Hartmann, J., Brink, K., Jäckle, R. & Tschersich, N. (2008). IAB-Haushaltspanel im Niedrigeinkommensbereich. Methoden- und Feldbericht. *FDZ Methodenreport 07/2008*. Nürnberg.

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

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

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

International Labour Office [ILO] (2012). *International Standard Classification of Occupations. ISCO-08*. Geneva: International Labour Office.

Jäckle, A. (2008). The Causes of Seam Effects in Panel Surveys. *ISEP Working Paper Series 2008-14*. Essex.

Jesske, B. & Quandt, S. (2011). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 4. Erhebungswelle – 2010 (Haupterhebung). *FDZ-Methodenreport 08/2011*.

Nürnberg.

Jesske, B. & Schulz, S. (2012). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 5. Erhebungswelle – 2011 (Haupterhebung), *FDZ Methodenreport 11/2012*. Nürnberg.

Jesske, B. & Schulz, S. (2013). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 6. Erhebungswelle – 2012 (Haupterhebung), *FDZ Methodenreport 10/2013*. Nürnberg.

Jesske, B. & Schulz, S. (2014). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 7. Erhebungswelle – 2013 (Haupterhebung), *FDZ Methodenreport 11/2014*. Nürnberg.

Jesske, B. & Schulz, S. (2015). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 8. Erhebungswelle – 2014 (Haupterhebung), *FDZ Methodenreport 11/2015*. Nürnberg.

Jesske, B., Knerr, P. & Schulz, S. (2016). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 9. Erhebungswelle – 2015 (Haupterhebung), *FDZ Methodenreport 04/2016*. Nürnberg.

Jesske, B., Knerr, P. & Kraft, L. (2017). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 10. Erhebungswelle – 2016 (Haupterhebung), *FDZ Methodenreport 09/2017*. Nürnberg.

Jesske, B. & Schulz, S. (2018). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 11. Erhebungswelle – 2017 (Haupterhebung), *FDZ Methodenreport im Erscheinen*. Nürnberg.

König, W., Lüttinger, P. & Müller, W. (1987). Eine vergleichende Analyse der Entwicklung und Struktur von Bildungssystemen. Methodologische Grundlagen und Konstruktion einer vergleich-baren Bildungsskala. *CASMIN-Projekt. Arbeitspapier Nr. 12*. Mannheim.

Kreuter, F., Müller, G. & Trappmann, M. (2010). Nonresponse and measurement error in employment research. Making use of administrative data. *Public Opinion Quarterly*, Vol. 74, No. 5, S. 880–906.

Kreuter, F., Müller, G. & Trappmann, M. (2014). A note on mechanisms leading to lower data quality of late or reluctant respondents. In: *Sociological Methods and Research*, Vol. 43, No. 3, S. 452-464.

Lechert, Y., Schroedter, J. & Lüttinger, P. (2006). Die Umsetzung der Bildungsklassifikation CASMIN für die Volkszählung 1970, die Mikrozensus- Zusatzerhebung 1971 und die

Mikrozensus 1976-2004. *ZUMA-Methodenbericht 2006/12*. Mannheim.

Lengerer, A., Bohr, J. & Janßen, A. (2005). Haushalte, Familien und Lebensformen im Mikrozensus – Konzepte und Typisierungen. *ZUMA-Arbeitsbericht 2005/05*. Mannheim.

Lipsmeier, G. (1999). Die Bestimmung des notwendigen Lebensstandards – Einschätzungsunterschiede und Entscheidungsprobleme. *Zeitschrift für Soziologie*, 28, S. 281-300.

Müller, W., Wirth, H., Bauer, G., Pollak, R. & Weiss, F. (2006). ESeC – Kurzbericht zur Validierung und Operationalisierung einer europäischen sozioökonomischen Klassifikation. *ZUMA-Nachrichten*, 59, S. 111–119.

Müller, W., Wirth, H., Bauer, G., Pollak, R. & Weiss, F. (2007): Entwicklung einer europäischen so-zioökonomischen Klassifikation. *Wirtschaft und Statistik*, 5, S. 527-530.

Nolan, B. & Whelan, C. (1996). Measuring Poverty Using Income and Deprivation Indicators: Alternative Approaches. *Journal of European Social Policy*, 6, S. 225-240.

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

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

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

Rammstedt, B. & John, O. (2005). Kurzversion des Big Five Inventory (BIF-K). *Diagnostica*, 51, S. 195-206.

Rendtel, U. & Harms, T. (2009). Weighting and calibration for household panels. In P. Lynn (Hrsg.), *Methodology of Longitudinal Surveys* (S. 265-286), Chichester: Wiley.

Ringen, S. (1988). Direct and Indirect Measurement of Poverty. *Journal of Social Policy*, 17, S. 351-365.

Rose, R. & Harrison, E. (2007). The European Socio-Economic Classification: A New Social Class Schema for Comparative European Research. *European Societies*, 9, S. 459-490.

Rudolph, H. & Trappmann, M. (2007). Design und Stichprobe des Panels „Arbeitsmarkt und Soziale Sicherung“ (PASS). In M. Promberger (Hrsg.), *Neue Daten für die Sozialstaatsforschung: Zur Konzeption der IAB-Panelerhebung „Arbeitsmarkt und Soziale*

Sicherung“, *IAB-Forschungsbericht 12/2007* (S. 60-101), Nürnberg.

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

Spieß, M. & Rendtel, U. (2000). Combining an ongoing panel with a new cross-sectional sample. *DIW-Discussion Papers 198*. Berlin.

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

Statistisches Bundesamt [StBA] (2008). Klassifikation der Wirtschaftszweige 2008 (WZ 2008) mit Erläuterungen. Wiesbaden: Statistisches Bundesamt.

Trappmann, M., Christoph, B., Achatz, J., Wenzig, C., Müller, G. & Gebhardt, D. (2009). Design and stratification of PASS. A New Panel Study for Research on Long Term Unemployment. *IAB-Discussion Paper 5/2009*. Nürnberg.

Trappmann, M. (2013a). Weighting. In: Bethmann, A., Fuchs, B. & Wurdack, A. (Hrsg.)(2013). User Guide "Panel Labour Market and Social Security"(PASS). Wave 6. *FDZ Datenreport 07/2013* (S. 56-66). Nürnberg.

Trappmann, M. (2014b). Weights. In: Bethmann, A., Fuchs, B. & Wurdack, A. (Hrsg.)(2013). User Guide "Panel Labour Market and Social Security"(PASS). Wave 6. *FDZ Datenreport 07/2013* (S. 81-99). Nürnberg.

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

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

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

Impressum

FDZ-Datenreport 06/2018

Herausgeber

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

Redaktion

Dana Müller, Dagmar Theune

Technische Herstellung

Dagmar Theune

Rechte

Nachdruck - auch auszugsweise - nur mit
Genehmigung des FDZ gestattet

Bezugsmöglichkeit

http://doku.iab.de/fdz/reporte/2018/DR_06-18.pdf

Internet

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

Rückfragen zum Inhalt an:

Forschungsdatenzentrum (FDZ) der Bundesagentur
für Arbeit im Institut für Arbeitsmarkt- und
Berufsforschung (IAB), Regensburger Str. 104,
90478 Nürnberg,
E-Mail: iab.fdz@iab.de

Jonas Beste, Institut für Arbeitsmarkt- und
Berufsforschung (IAB),
Regensburger Str. 104, 90478 Nürnberg,
Tel.: +49 (0) 911/179-2279
E-Mail: Jonas.Beste@iab.de