FORSCHUNGSDATENZENTRUM
der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung

FDZ-DATENREPORT

Dokumentation zu Arbeitsmarktdaten

10|2026 DE Codebuch und Dokumentation des Panel „Arbeitsmarkt und soziale Sicherung“ (PASS) Datenreport Welle 19

Fritz Anker et al.



Bundesagentur für Arbeit

Codebuch und Dokumentation des Panel „Arbeitsmarkt und soziale Sicherung“ (PASS) Datenreport Welle 19

Fritz Anker, Marco Berg, Ralph Cramer, Christian Dickmann, Vincent Gerber, Reiner Gilberg, Birgit Jesske, Martin Kleudgen (alle infas Institut für angewandte Sozialwissenschaft GmbH),

Jonas Beste, Nadja Bömmel, Madhu Chauhan, Sandra Dummert, Corinna Frodermann, Sonja Malich, Valentina Prospero, Claudia Wenzig, Mark Trappmann, Sebastian Bähr, Matthias Collischon, Anja Fehn, Stefanie Gundert, Jan Mackeben, Marcel Müller, Jens Stegmaier (alle Institut für Arbeitsmarkt- und Berufsforschung (IAB))

Die FDZ-Datenreporte beschreiben die Daten des FDZ im Detail. Diese Reihe hat somit eine doppelte Funktion: zum einen stellen Nutzerinnen und Nutzer fest, ob die angebotenen Daten für das Forschungsvorhaben geeignet sind, zum anderen dienen sie zur Vorbereitung der Auswertungen.

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.

Inhaltsverzeichnis

1 Einleitung 9
1.1 Ziele und Fragestellungen des Panel „Arbeitsmarkt und soziale Sicherung “ 9
1.2 Instrumente und Befragungsprogramm 10
1.2.1 Besonderheiten und Neuerungen in Welle 19 - Personenfragebogen 12
1.2.2 Besonderheiten und Neuerungen in Welle 19 - Seniorenfragebogen 14
1.2.3 Besonderheiten und Neuerungen in Welle 19 - Haushaltsfragebogen 15
1.2.4 Eingeschränkte Datenqualität bei einzelnen Variablen aus den fremdsprachigen Instrumenten 15
1.2.5 Allgemeine Hinweise zu früheren Wellen 17
1.3 Stichprobe und Datenaufbereitung 18
1.3.1 Qualitätskontrolle realisierter Interviews 20

2 Kennwerte 21
2.1 Stichprobenumfang 24
2.2 Response-Raten 26
2.3 Panelbereitschaft, Zuspielbereitschaft, Verknüpfung mit Prozessdaten 27
2.4 Abgespaltene Haushalte 28

3 Datensatzstruktur 30

4 Generierte Variablen 34
4.1 Vercodung offener Angaben 34
4.1.1 Offene Restkategorien und offene Items 34
4.1.2 Berufs- und Branchencodierung 39
4.2 Harmonisierung 43
4.3 Dependent Interviewing 46
4.4 Einfache Generierungen 49
4.5 Konstruktvariablen 79
4.5.1 Personenebene 79
4.5.2 Haushalts- oder Bedarfsgemeinschaftsebene 103

5 Datenaufbereitung 118
5.1 Strukturprüfungen und entfernte Interviews 120
5.2 Filterbereinigung und vergebene Missing Codes 126
5.3 Plausibilitätsprüfungen 128
5.4 Rückwirkende Veränderungen der Wellen 1 bis 18 130
5.5 Anonymisierung 132
5.6 Arbeitslosengeld-II-Bezüge/ Bürgergeld-Bezüge 136
5.6.1 Konzept für die Fortschreibung der in der Vorwelle aktuell andauernden Arbeitslosengeld-II-Spells/Bürgergeld-Spells 136

5.6.2 Struktur des Arbeitslosengeld-II Spelldatensatzes 137
5.6.3 Plausibilitätsprüfung und -bereinigung des Arbeitslosengeld-II Spelldatensatzes 138
5.6.4 Fortschreibung des Arbeitslosengeld-II-Spelldatensatzes 138
5.7 Erwerbsbiografien 139
5.7.1 Variablen zum (Nicht-)Erwerbsstatus im PENDDAT 140
5.7.2 Einkommensvariablen und Arbeitszeit im PENDDAT und in BIO-Spells 149
5.7.3 Konzept für die Fortschreibung der in der Vorwelle aktuell andauernden Spells 154
5.7.4 Struktur des BIO-Spelldatensatzes 154
5.7.5 Plausibilitätsprüfung und -bereinigung der Spelldatensätze 157
5.7.6 Fortschreibung der Spelldatensätze 157

6 Gewichtung Welle 19 160
6.1 Designgewichte für die Bestandshaushalte der Welle 19 160
6.2 Designgewichte für die Zugangsstichprobe und die Bestandsauffrischung der Welle 19 160
6.3 Wiederteilnahme-Propensity Haushalte 161
6.4 Teilnahme-Propensity erstbefragte Split-Haushalte 162
6.5 Ausfallgewichtung für Haushalte aus der Auffrischungsstichprobe der SGB-II-Neuzugänge Welle 19 162
6.6 Wiederteilnahme-Propensity Personen 163
6.7 Integration der Gewichte zum Gesamtgewicht vor Kalibrierung 164
6.8 Integration temporärer Ausfälle (Haushalte) 164
6.9 Kalibrierung zum Haushaltsgewicht Welle 19 Querschnitt 166
6.10 Kalibrierung der SGB-II-Stichprobe 167
6.11 Kalibrierung der Bevölkerungsstichprobe 168
6.12 Kalibrierung der Gesamtstichprobe 169
6.13 Kalibrierung zum Personengewicht Welle 19 Querschnitt 170
6.14 SGB-II-Stichprobe 171
6.15 Bevölkerungsstichprobe 172
6.16 Gesamtstichprobe 174
6.17 Schätzung der BA-Querschnittsgewichte für Haushalte und Personen ohne Arbeitslosengeld-II-Bezug 175

7 Anhang: Kurzbeschreibung des Datensatzes 177

Abbildungsverzeichnis

Abbildung 1: Realisierte Panelstichprobe Haushalte und Personen nach Erhebungswellen 26
Abbildung 2: Datensatzstruktur des PASS in Welle 19 32
Abbildung 3: Übersicht über generierte Variablen auf Personenebene in Welle 19 77

Tabellenverzeichnis

Tabelle 1: Änderungen im Personenfragebogen in W19 13
Tabelle 2: Änderungen im Seniorenfragebogen in W19 14
Tabelle 3: Änderungen im Haushaltsfragebogen in W19 15
Tabelle 4: Übersicht über Übersetzungsfehler bei Variablen mit und ohne Datenkorrekturen 16
Tabelle 5: Vercodung offener Angaben auf Haushaltsebene in Welle 19 35
Tabelle 6: Vercodung offener Angaben auf Personenebene in Welle 19 36
Tabelle 7: Codeschema der im PASS genutzten Zusatzvariablen 40
Tabelle 8: Harmonisierte Variablen im Personendatensatz (*PENDDAT*) 44
Tabelle 9: Wellenübergreifend generierte, aber nicht vollständig harmonisierte Variablen im Personendatensatz (*PENDDAT*) 45
Tabelle 10: Fortgeschriebene Vorwelleninformationen in Welle 19, Haushaltsfragebogen 47
Tabelle 11: Fortgeschriebene Vorwelleninformationen in Welle 19, Personenfragebogen . 48
Tabelle 12: 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 50
Tabelle 13: Einfache Generierungen für Welle 19 im Haushaltsdatensatz (*HHENDDAT*) und im *KINDER*-Datensatz (alphabetische Reihenfolge) 53
Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge) 55
Tabelle 15: Einfache Generierungen für Welle 19 im Arbeitslosengeld-II-Datensatz *(alg2_spells)* (Reihenfolge analog zur Reihenfolge im Datensatz) 67
Tabelle 16: Einfache Generierungen für Welle 19 im BIO-Spelldatensatz *(bio_spells)* (Reihenfolge analog zur Reihenfolge im Datensatz) 70
Tabelle 17: Einfache Generierungen für Welle 19 im Personenregisterdatensatz *(p_register* (alphabetische Reihenfolge) 74
Tabelle 18: Bildung in Jahren 79

Tabelle 19: Bildung in Jahren, Mutter 80
Tabelle 20: Bildung in Jahren, Vater 81
Tabelle 21: CASMIN 82
Tabelle 22: MCASMIN 83
Tabelle 23: VCASMIN 84
Tabelle 24: ISCED-97 85
Tabelle 25: MISCED-97 86
Tabelle 26: VISCED-97 86
Tabelle 27: International Standard Classification of Occupations 1988 (ISCO88) 87
Tabelle 28: International Standard Classification of Occupations 2008 (ISCO08) 88
Tabelle 29: Klassifizierung der Berufe 1992 (KldB1992) 90
Tabelle 30: Klassifizierung der Berufe 2010 (KldB2010) 91
Tabelle 31: Klassenschema nach Erikson, Goldthorpe und Portocarrero (EGP) 93
Tabelle 32: European Socio-economic Classification (ESeC) 94
Tabelle 33: Magnitude-Prestigeskala (MPS) 95
Tabelle 34: Standard International Occupational Prestige Scale (SIOPS/Treiman-Skala)-Basis ISCO-88 96
Tabelle 35: Standard International Occupational Prestige Scale (SIOPS/Treiman-Skala) – Basis ISCO-08 97
Tabelle 36: International Socio-Economic Index (ISEI) – Basis ISCO-88 98
Tabelle 37: International Socio-Economic Index (ISEI) – Basis ISCO-08 99
Tabelle 38: Klassifikation der Wirtschaftszweige 2003 (WZ2003) 100
Tabelle 39: Klassifikation der Wirtschaftszweige 2008 (WZ2008) 101
Tabelle 40: Körperliche Skala des SF12v2 (SOEP-Version, NBS) 101
Tabelle 41: Psychische Skala des SF12v2 (SOEP-Version, NBS) 102
Tabelle 42: Bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung 103
Tabelle 43: Bedarfsgewichtetes Haushaltseinkommen, neue OECD-Gewichtung 104
Tabelle 44: Deprivationsindex, ungewichtet 104
Tabelle 45: Deprivationsindex, gewichtet 106
Tabelle 46: Haushaltstypologie 108
Tabelle 47: Bedarfsgemeinschaftsnummer, Welle 19 110
Tabelle 48: Bedarfsgemeinschaftstypologie, Welle 19 111
Tabelle 49: Bürgergeld-Bezug der Bedarfsgemeinschaft zum Zeitpunkt der Stichprobenziehung, Welle 19 112
Tabelle 50: Bürgergeld-Bezug der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 19 113
Tabelle 51: Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 10 113
Tabelle 52: Flag zur Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 10 113

Tabelle 53: Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 11 114
Tabelle 54: Flag zur Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 11 114
Tabelle 55: Anzahl der Bedarfsgemeinschaften im Haushalt 116
Tabelle 56: Anzahl der zum Stichprobenzeitpunkt Leistungen beziehenden Bedarfsgemeinschaften im Haushalt 117
Tabelle 57: Übersicht über die Schritte der Datenaufbereitung in der Welle 19 des PASS . 119
Tabelle 58: Übersicht über die vergebenen Missing Codes 127
Tabelle 59: Übersicht über rückwirkende Veränderungen im Haushaltsdatensatz *(HHENDDAT, KINDER)* 130
Tabelle 60: Übersicht über rückwirkende Veränderungen im Personendatensatz *(PENDDAT)* 130
Tabelle 61: Übersicht über rückwirkende Veränderungen in Spelldatensätzen *(bio_spells, alg2_spells, ee_spells)* 131
Tabelle 62: Übersicht über rückwirkende Veränderungen in Registerdatensätzen *(hh_register; p_register)* 131
Tabelle 63: Übersicht über rückwirkende Veränderungen in den Gewichtungsdatensätzen *(hweights; pweights)* 131
Tabelle 64: Übersicht über die in Welle 19 anonymisierten Variablen im Personendatensatz *(PENDDAT)* 133
Tabelle 65: Übersicht über die in Welle 19 anonymisierten Variablen im BIO-Spelldatensatz *(bio_spells)* 136
Tabelle 66: Übersicht über die in Welle 19 anonymisierten Variablen im Kinder-Datensatz *(KINDER)* 136
Tabelle 67: Querschnittsvariablen im ALG II-Spelldatensatz *(alg2_spells)* 137
Tabelle 68: Generierungslogik *erwerb, erwerb2, nichterw, nichtew2* 141
Tabelle 69: Überarbeitung *erwerb, erwerb2, nichterw, nichtew2* 145
Tabelle 70: Grobzuweisung - Spell höherer Priorität schlägt Spell niedrigerer Priorität ... 148
Tabelle 71: Feinzuweisung für Sonderfälle 148
Tabelle 72: Überarbeitung Einkommensvariablen 150
Tabelle 73: Überarbeitung Arbeitszeitvariablen 151
Tabelle 74: ET-spezifische Querschnittsvariablen im BIO-Spelldatensatz (*bio_spells*) 155
Tabelle 75: AL-spezifische Querschnittsvariablen im BIO-Spelldatensatz (*bio_spells*) 156

Abstract

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 19 and is based upon the eighteenth wave's data report: Fritz Anker, Marco Berg, Ralph Cramer, Christian Dickmann, Vincent Gerber, Reiner Gilberg, Birgit Jesske, Martin Kleudgen (all infas Institute for Applied Social Sciences), Jonas Beste, Sandra Dummert, Corinna Frodermann, Sonja Malich, Valentins Prospero, Claudia Wenzig, Mark Trappmann, Sebastian Bähr, Matthias Collischon, Stefanie Gundert, Jan Mackeben, Marcel Müller, Jens Stegmaier, Anja Wunder (all Institute for Employment Research (IAB)): Codebook and documentation of the panel Study 'Labour Market and Social Security' (PASS), Datenreport wave 18, FDZ Datenreport, 10/2025 (en), Nürnberg.

Zusammenfassung

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 19 von PASS. Das Dokument basiert auf dem Datenreport der Welle 18: Fritz Anker, Marco Berg, Ralph Cramer, Christian Dickmann, Vincent Gerber, Reiner Gilberg, Birgit Jesske, Martin Kleudgen (alle infas Institut für angewandte Sozialwissenschaft GmbH), Jonas Beste, Sandra Dummert, Corinna Frodermann, Sonja Malich, Valentina Prospero, Claudia Wenzig, Mark Trappmann, Sebastian Bähr, Matthias Collischon, Stefanie Gundert, Jan Mackeben, Marcel Müller, Jens Stegmaier, Anja Wunder (alle Institut für Arbeitsmarkt- und Berufsforschung (IAB)): Codebuch und Dokumentation des Panel „Arbeitsmarkt und soziale Sicherung" (PASS): Datenreport Welle 18, FDZ Datenreport, 10/2025 (de), Nürnberg.

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-Forschung[1] 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[2]:

1. Welche Wege führen aus dem Bezug von Grundsicherungsleistungen (Arbeitslosengeld II/Bürgergeld)?
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 Grundsicherungsleistungen?

Dieser Datenreport gibt einen Überblick über die Befragungswelle 19, für die zwischen Februar 2025 und September 2025 11.391 Personen in 8.835 Haushalten[3] befragt wurden. Darunter befanden sich 8.882 Personen und 6.876 Haushalte, mit denen bereits zum wiederholten Mal ein Interview im Rahmen von PASS geführt wurde.

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

[2] Die Forschungsfragen wurden vor dem Hintergrund der Reform des SGB II angepasst und berücksichtigen seit Befragungswelle 17 das Bürgergeld.

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

Der vorliegende wellenspezifische Datenreport[4] der Welle 19 dokumentiert die wellenbezogenen Aspekte der Studie. Im Kapitel 1 findet sich ein Überblick über die Ziele und Fragestellungen der Studie mit einer Kurzbeschreibung der Instrumente und des Befragungsprogramms im Kapitel 1.2, sowie den Besonderheiten und Neuerungen der Welle 19 in den Kapiteln 1.2.1 bis 1.2.3. Die zentralen Kennzahlen zu Stichprobe und Ausschöpfungsquoten der Welle 19 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 Gewichtungsverfahren vorgestellt. Im Anhang findet sich ein Gesamtüberblick über die Datensätze im PASS über alle Wellen.

Bis Welle 16 wurden in gesonderten Tabellenbänden die Häufigkeiten aller im Scientific Use File enthaltenen Variablen ausgewiesen. Diese Bände werden ab Welle 17 nicht mehr veröffentlicht. Auf Anfrage sind Häufigkeitsauswertungen zu einzelnen Variablen möglich.

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 Haushaltsinterview[5] 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 66 Jahre erhalten dabei eine Kurzfassung des Personenfragebogens (Seniorenfragebogen)[6] , in dem für diese

[4] 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.

[5] 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); Jesske et al. (2019); Jesske et al. (2020); Jesske et al. (2021); Jesske et al. (2022); Jesske et al. (2023); Jesske et al. (2024)).

[6] Seit Welle 10 wird im Erhebungsinstrument berücksichtigt, dass das gesetzliche Rentenalter in den nächsten Jahren schrittweise auf 67 Jahre angehoben wird. Die Altersgrenze für die Vorlage des

Altersgruppe irrelevante Fragen nicht enthalten sind.

Erhebungsinstrumente und Befragungsprogramm der Welle 19 basieren auf der Welle 18 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 ermöglichen[7]. Um bei der erneuten Befragung Seam-Effekte[8] zu vermeiden und die Datenqualität zu erhöhen, wurden bereits seit der Welle 3 bei bestimmten Fragen per Dependent Interviewing Informationen, die die Befragungspersonen beim letzten Interview gegeben hatten, 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)[9] Im Panelverlauf ist es möglich, dass Haushalte oder Personen aus Panelhaushalten wieder teilnehmen, die eine Welle (oder auch mehrere Wellen) ausgesetzt, d.h. nicht teilgenommen hatten (sog. temporäre Ausfälle). Temporäre Aussetzer, die in genau einer Welle, also der Vorwelle, nicht teilgenommen haben, werden bei der Interviewsteuerung wie Panelhaushalte bzw. -befragte behandelt. Nimmt ein gesamter Haushalt an zwei aufeinander folgenden Wellen nicht teil, wird er nicht mehr kontaktiert. Auf Personenebene ist es jedoch möglich, dass Befragte in Panelhaushalten zwei oder mehr Wellen hintereinander nicht befragt werden und danach wieder an der Panelstudie teilnehmen. Im Personeninterview wird diese besondere Befragtengruppe nach der Befragungspause wie Neubefragte behandelt, d.h. unveränderliche Informationen (z.B. soziale Herkunft) werden erneut erhoben und es findet kein Dependent Interviewing statt.

Die einzelnen Instrumente und das Befragungsprogramm werden ausführlich im wellenübergreifenden User Guide (Chauhan et al. (2026), Kapitel 4) dargestellt. Zudem werden auch die wellenspezifischen Fragebögen bereitgestellt (zu den verwendeten Skalen und Instrumenten siehe auch Müller et al. (2020)). Der nachfolgende Abschnitt liefert einen Überblick zu den Besonderheiten und Neuerungen der Welle 19. Die Besonderheiten und Neuerungen in Welle 19 betreffen das Frageprogramm für den Haushalts- und

Seniorenfragebogens ist daher nicht mehr starr, sondern richtet sich monatsgenau nach dem gesetzlichen Rentenalter der Zielperson.

[7] 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)).

[8] 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)).

[9] Dabei werden Split-Haushalte wie neue Haushalte für die Abfrage behandelt.

Personenfragebogen (Änderung der Referenzzeiträume, Modifikation von Einzelfragen und neue Fragenmodule)[10], die Stichprobe und die Aufbereitung der Daten.

1.2.1 Besonderheiten und Neuerungen in Welle 19 - Personenfragebogen

Im Personenfragebogen werden die seit Welle 2 erhobenen erwerbsbiografischen Informationen fortgeschrieben[11]. Ab Welle 4 ist eine geänderte Abfragelogik der chronologischen Retrospektiverhebung enthalten (s.a. Kapitel 1.3.1 in Berg et al. (2011), FDZ Datenreport 08/2011).

Im Personenbogen wurden in Welle 19 kleinere Änderungen bzw. Ergänzungen in vorhandenen Modulen vorgenommen. Zusätzlich wurde ein Modul zu Einstellungen gegenüber sozialpolitischen Leistungen neu aufgenommen, sowie ein Modul zur Internetnutzung erneut eingesetzt. Zudem wurde die Obergrenze für Minijobs in den entsprechenden Nachfragen und Filterungen angepasst, von 538 Euro auf 556 Euro.

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

[11] 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.

Tabelle 1: Änderungen im Personenfragebogen in W19

Modul	Variable	Änderung	Erläuterung
Einstellung (sozialpolitische Leistungen)	PSE0100-PSE0700	eingefügt	Jeweils zwei von sechs Fragen zu sozialpolitischen Leistungen werden randomisiert vorgelegt. Es geht um Fragen zum BAföG, Bürgergeld, Kindergeld, Deutschlandticket, Arbeitslosengeld und Elterngeld.
Trägerkontakte	PTK3300	eingefügt	Frage zu Kontaktarten mit dem Jobcenter wurde ergänzt.
Politische Teilhabe	PPT0500	eingefügt	Die Sonntagsfrage wurde ergänzt.
Internetnutzung	PTS100*, PTS010-PTS2000	eingefügt	Modul zur Ausstattung und Nutzungshäufigkeit aus Vorwelle wieder aufgenommen.
Stigmabewusstsein	PSV0400	eingefügt	Eine neue Frage an alle Befragten zum Stigmabewusstsein.
Gesundheit	PG0750, PG0755	eingefügt	Zwei neue Fragen im Modul, ob der Arbeitgeber über eine Behinderung informiert wurde und wenn nein, warum nicht.
Zuspielerlaubnis	RegP0100	modifiziert	Die Frage wird nun mit zwei unterschiedlichen Textvarianten zufällig vorgelegt.
Politische Teilhabe	PPT0200*	modifiziert	Item G „dem öffentlichen Rundfunk“ ergänzt.
Zukunftssorgen	PA1600*	modifiziert	Variable wurde umbenannt (ehemals: PA1600a-c) und Item D „dass Sie Ihre aktuelle Wohnung verlieren“ ergänzt.
Zuspielerlaubnis 2	PReP0200	modifiziert	Die Nachfrage zur Zuspielungsbereitschaft wurde weiter hinten im Fragebogen platziert.
Weiterbildung	PWBLK01-PWB0700	modifiziert	Variablenbenennung wurde geändert (ehemals WBBLK01-WBH0700).
Social Media	PSM0300-PSM0500	gestrichen	Das Modul wurde gestrichen.

Tabelle 1: Änderungen im Personenfragebogen in W19 (Fortsetzung)

Qualität der Beschäftigung	PQB0200, PQB0300, PQB0400, PQB0500, PQB0600*, PQB0900*, PQB1000	gestrichen	Einzelne Fragen im Modul entfallen.
Einstellung (Verteilung)	PVG0100*	gestrichen	Das Modul wurde gestrichen.

1.2.2 Besonderheiten und Neuerungen in Welle 19 - Seniorenfragebogen

Aufgrund der schrittweisen Erhöhung des Rentenalters erfolgt - um den Seniorenbogen als Kurzversion weiterhin bereits für 66 Jährige und älter vorzusehen - die Filterung für Befragte mit validen Angaben zum Geburtstag ab Welle 10 monatsgenau. Gemäß des Renteneintrittsalters wurde die Altersgrenze für den Übergang vom Senioren bzw. Personenbogen in Welle 19 wie folgt angepasst: bei 66 Jahre und 2 Monate (Jahrgang 1959).

Wie im Personenfragebogen wurde auch im Seniorenfragebogen die Obergrenze für Minijobs in den entsprechenden Nachfragen und Filterungen angepasst, von 538 Euro auf 556 Euro. Aus der Liste der Änderungen für den Personenfragebogen wurden nachfolgende Änderungen auch im Seniorenfragebogen vorgenommen.

Tabelle 2: Änderungen im Seniorenfragebogen in W19

Modul	Variable	Änderung	Erläuterung
Einstellung (sozialpolitische Leistungen)	PSE0100-PSE0700	eingefügt	Jeweils zwei von sechs Fragen zu sozialpolitischen Leistungen werden randomisiert vorgelegt. Es geht um Fragen zum BAföG, Bürgergeld, Kindergeld, Deutschlandticket, Arbeitslosengeld und Elterngeld.
Politische Teilhabe	PPT0500	eingefügt	Die Sonntagsfrage wurde ergänzt.
Stigmabewusstsein	PSV0400	eingefügt	Eine neue Frage an alle Befragten zum Stigmabewusstsein.
Politische Teilhabe	PPT0200*	modifiziert	Item G „dem öffentlichen Rundfunk“ ergänzt.

1.2.3 Besonderheiten und Neuerungen in Welle 19 - Haushaltsfragebogen

Im Haushaltsfragebogen gab es zu den üblichen Anpassungen des Zeitbezugs und Filtersteuerungen, Anpassungen zur einheitlichen Darstellung von Fragen. Ansonsten gab es nur eine Änderung in Welle 19, die in der folgenden Übersicht aufgelistet ist.

Tabelle 3: Änderungen im Haushaltsfragebogen in W19

Modul	Variable	Änderung	Erläuterung
Kosten der Unterkunft	HKU0300	modifiziert	Item Code 4 ergänzt: „Ihre Wohnkosten entsprechen der Obergrenze.“

1.2.4 Eingeschränkte Datenqualität bei einzelnen Variablen aus den fremdsprachigen Instrumenten

Das Vorgehen zur Übersetzung der Erhebungsinstrumente im PASS orientiert sich an den Translation Guidelines des European Social Surveys (ESS), mit dem Ziel eine inhaltlich gleichwertige, funktional äquivalente Übersetzung zu gewährleisten, die auch kulturelle Unterschiede der Zielgruppe berücksichtigt. Für Details zum Vorgehen siehe Anker et al. (2025).

Trotz der intensiven Abstimmung und Korrekturschleifen der Übersetzer sowie den Rückmeldungen der fremdsprachigen Interviewenden, treten einzelne Übersetzungsfehler und Ungenauigkeiten auf, die sich auf die Datenqualität von PASS auswirken.

In der ukrainischen Sprachvariante der Welle 19 wurden einzelne Variablen korrigiert, wenn die erhobenen Inhalte erheblich von der deutschen Originalfrage abweichen. In diesen Fällen wurde der Sondercode -4 „Frage irrtümlich nicht gestellt“ vergeben [12]. Bei geringfügigen Abweichungen in der Übersetzung und eingeschränkter Vergleichbarkeit wurde auf eine Datenkorrektur verzichtet.

[12] Weitere Hinweise zu den Datenkorrekturen finden sich im Kapitel 5.4

Tabelle 4: Übersicht über Übersetzungsfehler bei Variablen mit und ohne Datenkorrekturen

Modul	Variable	Welle	Erläuterung	Korrektur
Trägerkontakte	PTK2500a	Welle 19	Bevormundung durch Jobcentermit-arbeitende übersetzt mit: „die Jobcenter-mitarbeitenden haben es mir verboten.“	Sondercode -4 „Frage irrtümlich nicht gestellt“.
Migration	PMI3000	Welle 17, 18, 19	Gründe für nicht Anerkennung von Abschlüssen: Wurde erhoben: Wird nicht anerkannt. Hätte erhoben werden sollen: Warum hat man nicht anerkennen lassen.	Sondercode -4 „Frage irrtümlich nicht gestellt“.
Erwerbsbiografie (spells) - Enddatum Arbeitslosigkeit, Monat	BIO0400	Welle 17, 18, 19	Wurde erhoben: „Ab wann“. Hätte erhoben werden sollen: „Bis wann“.	Keine Korrektur.
Erwerbsbiografie (spells) - Enddatum Arbeitslosigkeit, Jahr	BIO0500	Welle 17, 18, 19	Wurde erhoben: „Ab wann“. Hätte erhoben werden sollen: „Bis wann“.	Keine Korrektur.

Im Spellmodul zur Erwerbsbiografie bei den Variablen BIO0400/BIO0400 „Und bis wann waren Sie arbeitslos gemeldet? Falls Sie während Ihrer Arbeitslosigkeit an einer Maßnahme der Arbeitsagentur oder des Jobcenters teilgenommen haben, zählen Sie diese bitte nicht als Unterbrechung, sondern nennen Sie mir bitte das Ende des Gesamtzeitraums.“ wurde „bis wann“ in der ukrainischen Übersetzung fälschlicherweise mit „seit wann“ übersetzt. Aufgrund des Fragezusammenhangs (auch mit BIO0200/BIO0300) und nach Prüfung der erhobenen Daten, konnten jedoch keine Auffälligkeiten beim Enddatum der berichteten Arbeitslosigkeits-Spells festgestellt werden. Weder die durchschnittliche Dauern der berichteten Spells weichen für die ukrainische Befragtengruppe ab, noch zeigen sich vermehrt kurze Ein-Monats-Spells. Eine Prüfung anhand der Audiomitschnitte zeigt, dass die fremdsprachigen Interviewenden den Fehler in vielen Fällen direkt beim Vorlesen der Frage korrigieren. Auf dieser Grundlage wurde die Entscheidung gegen eine Datenkorrektur getroffen. Nutzende, die dennoch an einer Identifikation der betroffenen Spells (aufgrund der laufenden Fortschreibung kann sich dies für die einzelnen Spells über die Zeit ändern) interessiert sind, können auf Anfrage eine Liste mit Personen- und Spellnummer erhalten.

1.2.5 Allgemeine Hinweise zu früheren Wellen

Der Scientific Use File (SUF) enthält auch Variablen von Einzelfragen/Modulen aus früheren Wellen, die in der aktuellen Welle nicht mehr erhoben werden. In diesem Kapitel sollen daher wichtige Hinweise zu Modulen/Fragen aus früheren Datenreporten gesammelt werden, die im aktuellen Datenreport im Rahmen des Kapitels Neuerungen und Besonderheiten nicht sinnvoll thematisiert werden können. Häufig wird bei der kurzen Beschreibung auch auf die entsprechende Stelle im damaligen Datenreport verwiesen.

- In Welle 14 mussten Änderungen im bereits laufenden Feld vorgenommen werden. Dabei ging es um neu aufgenommene Fragen, die die besondere Situation der Covid-19 Pandemie aufgegriffen haben. Die Variable kennungfbversW14 kennzeichnet im Datensatz die Fälle bzw. Haushalte, die vor (Code 1) und nach (Code 2) der Änderung befragt wurden. Die Änderungen im Fragebogen waren ab dem 8. April 2020 aktiv. Für weitere Informationen siehe Berg et al. (2021, S. 14). Grundsätzlich sollte die Nutzung der Welle 14 vor dem Hintergrund der Covid-19-Pandemie betrachtet und intepretiert werden. Denn durch die eingeschränkten Erhebungsmöglichkeiten können bei einzelnen Kennwerten geringere Fallzahlen oder sinkende Raten auftreten (Berg et al., 2021, S. 16).
- Das Modul Vertrauensexperiment (PVX0100-PVX0600) konnte in Welle 14 durch die mit der Covid-19 Pandemie verbundenen Einschränkungen im CAPI-Feld ab Mitte März nicht mehr erhoben werden. Für weitere Informationen siehe Berg et al. (2021, S. 13).
- Das Modul Crowdworking (PCW0100-PCW0400) wurde bei PASS neu entwickelt und in den Wellen 12 und 13 erhoben. Analysen mit diesen Variablen haben erhebliche Inkonsistenzen gezeigt und geben Anlass zum Zweifel an der Validität des Messinstruments. Zu weiteren Informationen siehe Berg et al. (2019, S. 14).
- In Welle 11 mussten Änderungen im laufenden Feld vorgenommen werden. Diese Änderungen im Fragebogen bezogen sich auf die Teilstichprobe der syrischen / irakischen Haushalte. Die Variable kennungfbversW11 markiert, ob ein Fall mit der ursprünglichen oder der überarbeiteten Fragebogenfassung befragt wurde. Für weitere Informationen siehe Berg et al. (2018, S. 44).
- Bis zur Welle 15 wurde im Trägerkontakte-Modul eine Plausibilitätsprüfung zwischen PTK0500 und PTK0200 durchgeführt. Bei unplausibel hohen Abweichungen (mehr als doppelt so hoch) zwischen den Angaben wurde bislang die Angabe in PTK0500 auf den Sondercode -8 „unplausible Angabe“ gesetzt.
  Aufgrund von Änderungen im Fragebogen (Interviewerhinweis) und tatsächlichen Änderungen in der Jobcenter-Praxis infolge der Corona-Pandemie wurde entschieden, die Plausibilitätsprüfung rückwirkend ab Welle 14 nicht mehr durchzuführen und die Variable PTK0500 mit den ursprünglichen Befragtenangaben (und ohne Sondercode) auszuliefern (Berg et al., 2023, S. 159).

- Im Feldverlauf der Welle 17 wurden bei zwei Fragen Fehler in der Übersetzung festgestellt, weshalb alle Fälle, die bis zu diesem Zeitpunkt die Frage durchlaufen haben, auf Code -4 „Frage irrtümlich nicht gestellt“ gesetzt wurden. Zum einen betrifft es Frage PMI3600 in der arabischen Übersetzung des Fragebogens, bei der ein Label falsch übersetzt wurde. Darüber hinaus ist die Itembatterie zur Selbstwirksamkeit (PEO0100a-e) im russischen Instrument (Personen- und Seniorenbogen) betroffen. Dieser Fehler lag auch bereits in Welle 11 vor. Die erhobenen Werte sind für Welle 17 und für die zurückliegende Welle 11 bereits im SUF-Welle 16 auf den Sondercode -4 umgesetzt worden (siehe Berg et al., 2024, S. 16).
- In Welle 18 wurden Übersetzungsfehler festgestellt, die dazu geführt haben, dass die betreffenden Fälle – teilweise auch in den Vorwellen- auf Code -4 „Frage irrtümlich nicht gestellt“ gesetzt werden mussten. In der ukrainischen Sprachvariante sind die Module Deprivation (HLS0200a-b), Migration (PMI1800), Trägerkontakte (PTK2500a) und Gesundheit (PG0950h, PG0951h) betroffen, im Russischen die Module Einstellungen (Selbstwirksamkeit, PEO0100) und Trägerkontakte (PTK1700m, PTK2500a) sowie im Arabischen das Modul zur Erwerbsbiografie (ET3706-13).

Weitere Ungenauigkeiten in der ukrainischen Übersetzung wurden zwar nicht korrigiert, bedürfen aber einer kritischen Einzelfallentscheidung durch die Datennutzenden. Betroffen sind die Module Deprivation (HLS-Batterie), Minijob (PET0510), Arbeitssuche (PAS0930) und Netzwerke (PSK0100). Hier findet sich im Datensatz ein Hinweis bei der Variablenbeschreibung (stata notes). Für ausführliche Informationen siehe Anker et al. (2025).

1.3 Stichprobe und Datenaufbereitung

Wie in den Vorwellen, wurde auch in Welle 19 eine sogenannte Zugangsstichprobe für die SGB-II-Teilstichprobe[13] gezogen. Ziel ist es, die Repräsentativität der SGB-II-Stichprobe im Querschnitt sicherzustellen. Für die Auffrischung wurden Neuzugänge in den Bedarfsgemeinschaften gezogen, d.h. solche die im Juli 2024 Bürgergeld bezogen haben, nicht jedoch zum Zeitpunkt der Stichprobenziehungen der Wellen 1 bis 18 (vgl. Kapitel 2.1 und zum Auffrischungskonzept Trappmann et al., 2009, S. 11 ff.)). Zusätzlich wurde eine Auffrischung des SGB-II-Bestands (inkl. der Zugänge) aus den neuen PSUs der EWO-Stichprobe Welle 17 gezogen (vgl. Methodenberichte Welle 17 und Welle 19).

Alle in Welle 19 neubefragten Haushalte können wiederum über den Stichprobenindikator (*sample*) identifiziert werden.Bereits in Erhebungswelle 10 hat PASS auf die

[13] 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.

Herausforderung reagiert, dass sich die Zusammensetzung der Zielgruppe der Personen mit SGB-II-Bezug verändert hat. Ausgelöst durch die verstärkte Flüchtlingszuwanderung seit 2014 in Deutschland finden sich unter den Personen mit Leistungsbezug größere Anteile aus arabischsprachigen Herkunftsländern, die über die nächsten Jahre im PASS befragt werden sollen. Ab der Welle 10 wurde somit arabisch als weitere Erhebungssprache eingesetzt. Mit der Erweiterung der Erhebungssprache ist sichergestellt, dass anerkannte Flüchtlinge aus den häufigsten Herkunftsländern (Syrien und Irak) im Rahmen der jährlichen Auffrischungsstichproben der Grundsicherungsempfangenden erreicht und dann auch im Panel weiterverfolgt werden können. Während in den Wellen 11 bis 14 Auffrischungsstichproben von Bedarfsgemeinschaften innerhalb der Sampling Points von PASS nach dem üblichen Verfahren gezogen wurden (weitere Details s. Methodenbericht der Welle 12[14]), erfolgte in den Wellen 15 und 16 ein Oversampling von neu zugegangenen Bedarfsgemeinschaften mit Personen syrischer und irakischer Nationalität ins SGB II, damit Flüchtlinge in ausreichender Zahl befragt werden konnten. Vor dem Hintergrund, dass sich die Personen mit Leistungsbezug 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 Staatsangehörigkeit 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.

Durch den russischen Angriff auf die Ukraine im Februar 2022 vermerkt Deutschland einen vermehrten Zuzug von ukrainischen Staatsbürgern und Staatsbürgerinnen. Dieser Zuzug von Personen aus der Ukraine spiegelt sich auch in einem größeren Anteil im Leistungsbezug wider. Daher soll diese Personengruppe auch in den nächsten Jahren im PASS berücksichtigt und befragt werden. Seit Welle 17 wurde daher zusätzlich zu Russisch und Arabisch auch Ukrainisch als weitere Erhebungssprache eingesetzt, um eine Befragung dieser Personengruppe zu gewährleisten. In Welle 17 erfolgte ein Sampling von neu zugegangenen Bedarfsgemeinschaften mit Personen ukrainischer Nationalität ins SGB II, um diese Gruppe in ausreichender Zahl befragen zu können. In Welle 19 wurden somit im Rahmen der üblichen Zugangsstichproben auch ukrainische Gelüchtete berücksichtigt, die im Juni 2024 - aber noch nicht im Juni 2023 - zu den Leitungsbeziehenden im SGB II gehörten. Die Gruppen wurden sowohl bei den Neuzugängen als auch bei der Bestandsauffrischung berücksichtigt. Ab dem Scientific Use File der Welle 17 wird die

[14] Jesske & Schulz (2018)

zusätzliche Variable ostaatanukr ausgeliefert. Als Kategorien werden „ukrainische Staatsangehörigkeit Ja“ oder „ukrainische Staatsangehörigkeit Nein“ ausgewiesen.

1.3.1 Qualitätskontrolle realisierter Interviews

In der PASS-Studie sind umfangreiche Maßnahmen zur Qualitätssicherung auf allen Ebenen des Erhebungsprozesses implementiert. So greifen bereits am Anfang Revisions- und Testschleifen bei der Fragebogenprogrammierung, mit denen ein fehlerfreies Instrument für den Feldeinsatz sichergestellt wird und stehen am Ende umfangreiche Prüfschritte bei der Datenaufbereitung, die die Qualität der Auswertungsdaten garantieren. Mit verschiedenen methodischen Experimenten leistet PASS einen Beitrag zur Optimierung der Erhebungen im Panel und zur Verbesserung der Datenqualität [15]. In Welle 14 wurden verbesserte Qualitätskontrollen entwickelt, mit denen abweichendes Verhalten der Interviewenden aufgedeckt werden soll. Im Fokus der Kontrollen standen dabei insbesondere statistische Kontrollmethoden wie die Analyse von Zeitmarken, längsschnittlichen Korrelationsanalysen sowie weitere statistische Indikatoren (siehe dazu Beste et al., 2021). Während die Erweiterung in Welle 14 noch zu einer nachträglichen Entfernung von Interviews geführt hat (siehe dazu Berg et al., 2021, S. 16 ff.), werden die realisierten Interviews seit der Welle 15 bereits im laufenden Feld kontrolliert. In Welle 18 gab es keine Auffälligkeiten, so dass auch keine Datensätze nachträglich entfernt werden mussten. Der Umfang und das Vorgehen werden im Methodenbericht der Welle 18 ausführlich dargelegt (Jesske et al., 2025).

[15] Das Vorgehen und die verschiedenen Experimente werden in den Methodenberichten der jeweiligen Wellen dokumentiert.

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 1: Teilstichprobe 1 (SGB-II-Stichprobe W1[16]) bezeichnet dabei im Folgenden die Leistungsempfangendenstichprobe aus den Prozessdaten der Bundesagentur für Arbeit.
- Sample 2: Teilstichprobe 2 (MICROM-Stichprobe W1) verweist auf die geschichtete Bevölkerungsstichprobe.
- Sample 3: Zugangsstichprobe (SGB-II-Zugang W2) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 1 und Welle 2.
- Sample 4: Zugangsstichprobe (SGB-II-Zugang W3) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 2 und Welle 3.
- Sample 5: Zugangsstichprobe (SGB-II-Zugang W4) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 3 und Welle 4.
- Sample 6: Bestandsauffrischung/Aufstockung (Aufstockung EWO W5) benennt die Stichprobe aus den Zugängen der Einwohnermeldeamtsregister in 100 neuen PLZ-Bereichen (Welle 5).
- Sample 7: Bestandsauffrischung/Aufstockung (Aufstockung SGB II W5) benennt die Stichprobe aus den SGB II-Zugängen in 100 neuen PLZ-Bereichen (Welle 5).
- Sample 8: Zugangsstichprobe (SGB-II-Zugang W5) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 4 und Welle 5.
- Sample 9: Zugangsstichprobe (SGB-II-Zugang W6) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 5 und Welle 6.
- Sample 10: Zugangsstichprobe (SGB-II-Zugang W7) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 6 und Welle 7.
- Sample 11: Zugangsstichprobe (SGB-II-Zugang W8) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 7 und Welle 8.
- Sample 12: Zugangsstichprobe (SGB-II-Zugang W9) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 8 und Welle 9.
- Sample 13: Zugangsstichprobe (SGB-II-Zugang W10) benennt die Stichprobe aus den SGB II- Zugängen zwischen Welle 9 und Welle 10.

[16] In Vorgängerversionen des PASS Datenreports war von BA-Stichprobe die Rede. Ab Welle 19 ändert sich diese Bezeichnung zu SGB-II-Stichprobe. Mit dieser sprachlichen Anpassung geht aber keine inhaltliche Änderung einher.

- Sample 14: Zugangsstichprobe (SGB-II-Zugang W10 syrische/irakische Haushalte) benennt die Stichprobe aus dem Oversampling syrischer/irakischer Haushalte.
- Sample 15: Bestandsauffrischung/Aufstockung (Aufstockung EWO W11) benennt die Stichprobe aus den Zugängen der Einwohnermeldeamtsregister in den PLZ-Bereichen der Welle 5 (Welle 11).
- Sample 16: Zugangsstichprobe (SGB-II-Zugang W11) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 10 und Welle 11.
- Sample 17: Zugangsstichprobe (SGB-II-Zugang W11 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/ irakischer Haushalte zwischen Welle 10 und Welle 11.
- Sample 18: Zugangsstichprobe (SGB-II-Zugang W12) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 11 und Welle 12.
- Sample 19: Zugangsstichprobe (SGB-II-Zugang W12 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/irakischer Haushalte zwischen Welle 11 und Welle 12.
- Sample 20: Zugangsstichprobe (SGB-II-Zugang W13) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 12 und Welle 13.
- Sample 21: Zugangsstichprobe (SGB-II-Zugang W13 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/irakischer Haushalte zwischen Welle 12 und Welle 13.
- Sample 22: Zugangsstichprobe (SGB-II-Zugang W14) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 13 und Welle 14.
- Sample 23: Zugangsstichprobe (SGB-II-Zugang W14 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/irakischer Haushalte zwischen Welle 13 und Welle 14.
- Sample 24: Bestandsauffrischung/Aufstockung (Aufstockung SGB II W14 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/irakischer Nationalität (Welle 14).
- Sample 25: Zugangsstichprobe (SGB-II-Zugang W15) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 14 und Welle 15.
- Sample 26: Zugangsstichprobe (SGB-II-Zugang W15 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/irakischer Haushalte zwischen Welle 14 und Welle 15.
- Sample 27: Bestandsauffrischung/Aufstockung (Aufstockung SGB II W15 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/irakischer Nationalität (Welle 10).
- Sample 28: Zugangsstichprobe (SGB-II-Zugang W16) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 15 und Welle 16.
- Sample 29: Zugangsstichprobe (SGB-II-Zugang W16 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/ irakischer Haushalte zwischen Welle 15 und Welle 16.

- Sample 30: Bestandsauffrischung/Aufstockung (Aufstockung EWO W17) benennt die Stichprobe aus den Zugängen der Einwohnermeldeamtsregister in 100 PLZ-Bereichen (Welle 17).
- Sample 31: Zugangsstichprobe (SGB-II-Zugang W17) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 16 und Welle 17.
- Sample 32: Zugangsstichprobe (SGB-II-Zugang W17 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/ irakischer Haushalte zwischen Welle 16 und Welle 17.
- Sample 33: Zugangsstichprobe (SGB-II-Zugang W17 ukrainische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen ukrainischer Haushalte.
- Sample 34: Zugangsstichprobe (SGB-II-Zugang W18) benennt die Stichprobe aus den SGB II-Zugängen zwischen Welle 17 und Welle 18.
- Sample 35: Zugangsstichprobe (SGB-II-Zugang W18 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen syrischer/ irakischer Haushalte zwischen Welle 17 und Welle 18.
- Sample 36: Zugangsstichprobe (SGB-II-Zugang W18 ukrainische Haushalte) benennt die Stichprobe aus den SGB II-Zugängen ukrainischer Haushalte zwischen Welle 17 und Welle 18.
- Sample 37: Bestandsauffrischung/Aufstockung (Aufstockung EWO W17/2) benennt die Stichprobe aus den Zugängen der Einwohnermeldeamtsregister in 100 PLZ-Bereichen (Nachbearbeitung aus Welle 17).
- Sample 38: Zugangsstichprobe (SGB-II-Zugang W19) benennt die Stichprobe aus den SGB-II-Zugängen zwischen Welle 18 und Welle 19.
- Sample 39: Zugangsstichprobe (SGB-II-Zugang W19 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB-II-Zugängen syrischer/ irakischer Haushalte zwischen Welle 18 und Welle 19.
- Sample 40: Zugangsstichprobe (SGB-II-Zugang W19 ukrainische Haushalte) benennt die Stichprobe aus den SGB-II-Zugängen ukrainischer Haushalte zwischen Welle 18 und Welle 19.
- Sample 41: Bestandsauffrischung/Aufstockung (Aufstockung SGB II W19) benennt die Stichprobe aus den SGB-II-Zugängen in 100 neuen PLZ-Bereichen (Welle 19).
- Sample 42: Bestandsauffrischung/Aufstockung (Aufstockung SGB II W19 syrische/irakische Haushalte) benennt die Stichprobe aus den SGB-II-Zugängen in 100 neuen PLZ-Bereichen (Welle 19).
- Sample 43: Bestandsauffrischung/Aufstockung (Aufstockung SGB II W19 ukrainische Haushalte) benennt die Stichprobe aus den SGB-II-Zugängen in 100 neuen PLZ-Bereichen (Welle 19).

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 18[17]. Dabei muss berücksichtigt werden, dass nur solche erstbefragten Haushalte auch bei den Wiederholungsbefragungen zur Verfügung stehen, für welche die Panelbereitschaft[18] 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 19 beschrieben[19].

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

Dieser Fall konnte erstmalig am Ende der Welle 3 eintreten und wirkt sich auf das

[17] 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))

[18] Die Panelbereitschaft wird durch den oder die Haushaltsbevollmächtigte erteilt und zählt damit für alle Haushaltsmitglieder. Panelbereite Haushalte haben der Speicherung ihrer Adresse zum Zweck von Wiederholungsbefragungen im Rahmen der Studie zugestimmt.

[19] 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); Jesske et al. (2019); Jesske et al. (2020); Jesske & Schulz (2021); Jesske et al. (2022); Jesske et al. (2023); Jesske et al. (2024); Jesske et al. (2025); Jesske et al. (2026).

Panelbrutto in den Wellen 4[20] bis 19 aus. Die Bruttoeinsatzstichprobe, die im Feld für die Welle 19 von PASS bearbeitet wurde, umfasste insgesamt 9.648 Panelhaushalte. Dazu kommen erstbefragte Haushalte aus den SGB-II-Zugangsstichproben (n=6.015, davon 1.105 syrische/irakische Haushalte und 1.501 ukrainische Haushalte) und der SGB-II-Bestandsauffrischung (n=7.478 Haushalte, davon 541 syrische/irakische Haushalte und 699 ukrainische Haushalte) sowie den Splithaushalten, die sich in den Wellen 18[21] (n=171) oder 19 (n=236) neu ergeben haben[22].

Die Fallzahlen für den Bruttostichprobenumfang der Panelhaushalte in den jeweiligen Erhebungswellen und Teilstichproben[23] werden in → **Tabelle A1**[24] ausgewiesen. In der Welle 19 konnte in 6.922 Haushalten der Panelstichprobe mindestens ein Interview geführt werden. Dazu kommen 950 erstbefragte Haushalte aus der üblichen SGB-II-Zugangsstichprobe, von denen 878 ihre Panelbereitschaft erklärt haben, sowie 963 Haushalte aus der Bestandsauffrischung, von denen 913 ihre Panelbereitschaft erklärt haben. Aus den Stichproben Zugang- und Bestandsauffrischung entfallen 224 Haushalte auf syrische/irakische Haushalte, von denen 211 panelbereit waren; 377 entfallen auf ukrainische Haushalte, von denen 357 panelbereit waren. Ebenfalls erstbefragte Haushalte der Welle 19 sind 36 Splithaushalte, die sich aus den Teilstichproben der bisherigen Wellen 1 bis 18 ergeben und in den Haushalten der Panelstichprobe enthalten sind.

Den 8.835 realisierten Haushaltsinterviews in der Welle 19 entsprechen 11.391 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. Ab Welle 17 wurde zudem Ukrainisch als weitere Fremdsprache aufgenommen, damit ukrainische Haushalte ebenfalls befragt werden können. Die→ **Tabelle A3** gibt an, wie viele Haushalte bzw. Personen in den drei Erhebungssprachen

[20] Zusätzlich auf das Panelbrutto in der Welle 4 wirkte sich der Wechsel des Erhebungsinstituts aus. Damit die Adressen der Panelteilnehmenden 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).

[21] 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.

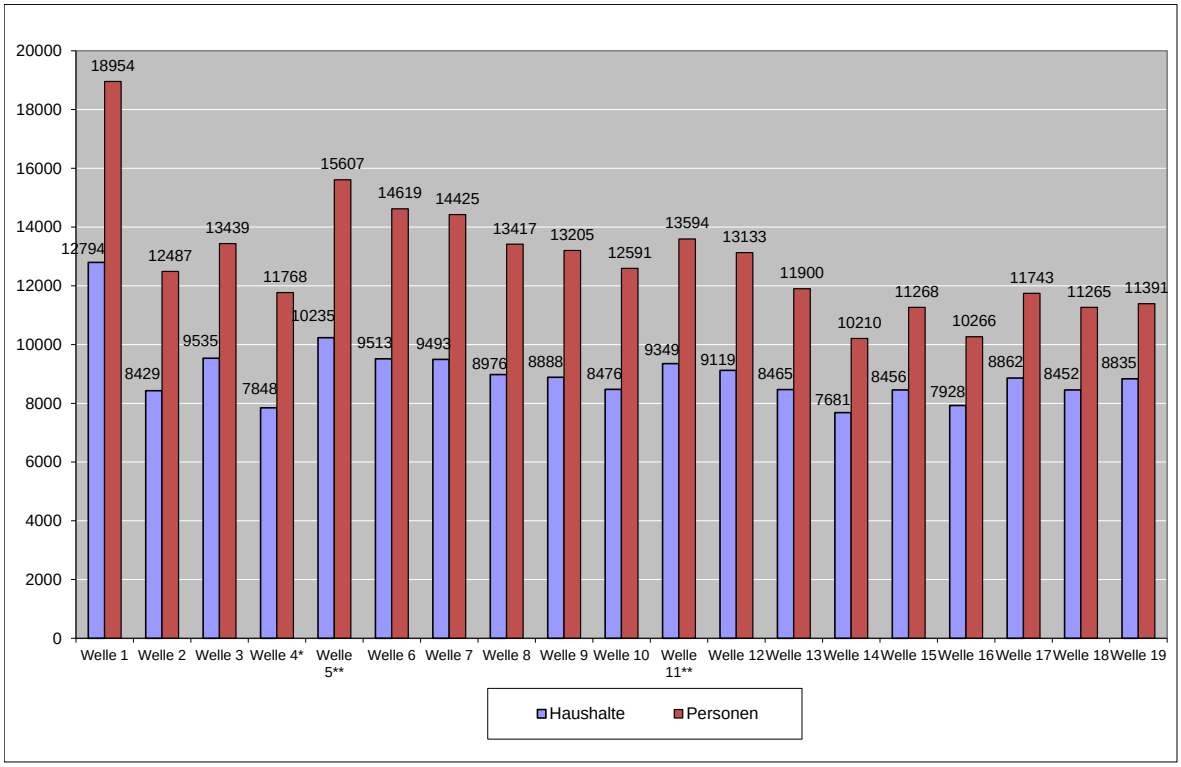
[22] Zahlen der Einsatzstichprobe s. Methodenbericht Welle 19 (Jesske et al., 2026)

[23] Die Zahlen umfassen alle im Registerfile enthaltenen Fälle. Abweichungen zum Methodendatensatz sind aufgrund der nachtäglichen Datenprüfungen und Bereinigungen möglich.

[24] Die Registerfiles des Scientific Use Files umfassen immer die Nettostichprobe realisierter Interviews der jeweiligen Wellen. Im Falle von Splithaushalten ist es möglich, dass es nachträglich zur Erweiterung des Panel-HH-bruttos der Vorwelle kommt, wenn der Splithaushalt in der Vorwelle zwar identifiziert wurde, aber noch nicht realisiert werden konnte.

befragt wurden.

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



** Erweiterung der Bruttostichprobe durch Aufstockung
* Reduktion der Bruttostichprobe durch Widerrufsverfahren

Abbildung 1: Realisierte Panelstichprobe Haushalte und Personen nach Erhebungswellen

2.2 Response-Raten

Die Berechnung der Response-Rate erfolgt nach den Standards der American Association for Public Opinion Research (AAPOR) (AAPOR, 2023). Berichtet wird die Ausschöpfungsquote RR1, in der auch alle Fälle ungeklärter Eligibilität in den Nenner eingehen und die somit den geringsten Wert aller Ausschöpfungsquoten annimmt[25]. Als

[25] 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

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 19 auf Haushaltsebene dargestellt.

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

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

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 Haushalte[26] im PASS ist in →

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.

[26] Bei allen Haushalten der Welle 1 handelt es sich um erstmals befragte Haushalte. Ab Welle 2 handelt es sich nur bei Haushalten der jeweiligen Auffrischungsstichproben (in Welle 5 Haushalte der Aufstockungsstichprobe) um erstmals befragte Haushalte sowie bei den erstmals teilnehmenden

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 übertragen. 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 Kapitel 2.1)[27]. Im Gegensatz zur Panelbereitschaft wurde die Erlaubnis zur Zuspielung von Prozessdaten der Bundesagentur für Arbeit an die Befragungsdaten für jede befragte Person erhoben, die mit dem Personenfragebogen interviewt wurde. Für Personen ab 66 Jahre entfällt die Nachfrage, da sie im Seniorenfragebogen nicht enthalten ist. Die Zuspielbereitschaft wird dabei nicht in jeder Welle neu erhoben[28].

→ **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.

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.

[27] 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. Widersprach der oder die Befragte dieser Ankündigung nicht explizit, wurde der Haushalt als weiterhin panelbereit gewertet und die Variable panel im Personendatensatz (*PENDDAT*) entsprechend fortgeschrieben.

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

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 & Harms, 2009, S. 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 19 liegen insgesamt 2.054 abgespaltene Haushalte vor, von denen in Welle 19 genau 569 befragt werden konnten. Darunter befanden sich 36 neue Splithaushalte aus Welle 19 sowie 42 erstbefragte Splithaushalte, die bereits in Welle 18 identifiziert werden konnten. Weitere Informationen zu den Splithaushalten finden sich im Methodenbericht der Erhebungswelle 19 (Jesske et al., 2026).

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

3 Datensatzstruktur

Die übliche Struktur für die Aufbereitung eines Paneldatensatzes, wie sie z. B. in Erhebungen wie dem Sozio-oekonomischen Panel (SOEP) oder der britischen Längsschnittstudie „The UK-Household Longitudinal Study“ (Understanding Society) 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 Haus-halts- 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) 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 SGB-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 für die Wellen 4 bis 17 ein Ein-Euro-Spelldatensatz vor (*ee_spells*). Informationen zu *ee_spells* können dem PASS-Datenreport der Welle 17 (Jesske et al., 2024) entnommen werden. 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_SUB*) 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. In den Wellen 12 und 15 wurden erneut Vignetten-Module erhoben. Der Datensatz wurde in Welle 12 *VIGDAT_KON* benannt und in Welle 15 *VIGDAT_MUK*. Die in den Wellen 5, 12 und 15 erhobenen Vignetten-Module unterscheiden sich inhaltlich voneinander. Die *VIGDAT*-Datensätze aus den verschiedenen Erhebungswellen sollten daher nicht zusammengeführt werden. Detaillierte Informationen zur Struktur der einzelnen Datensätze finden sich im User-Guide von PASS (Chauhan et al., 2026, Kapitel 5).

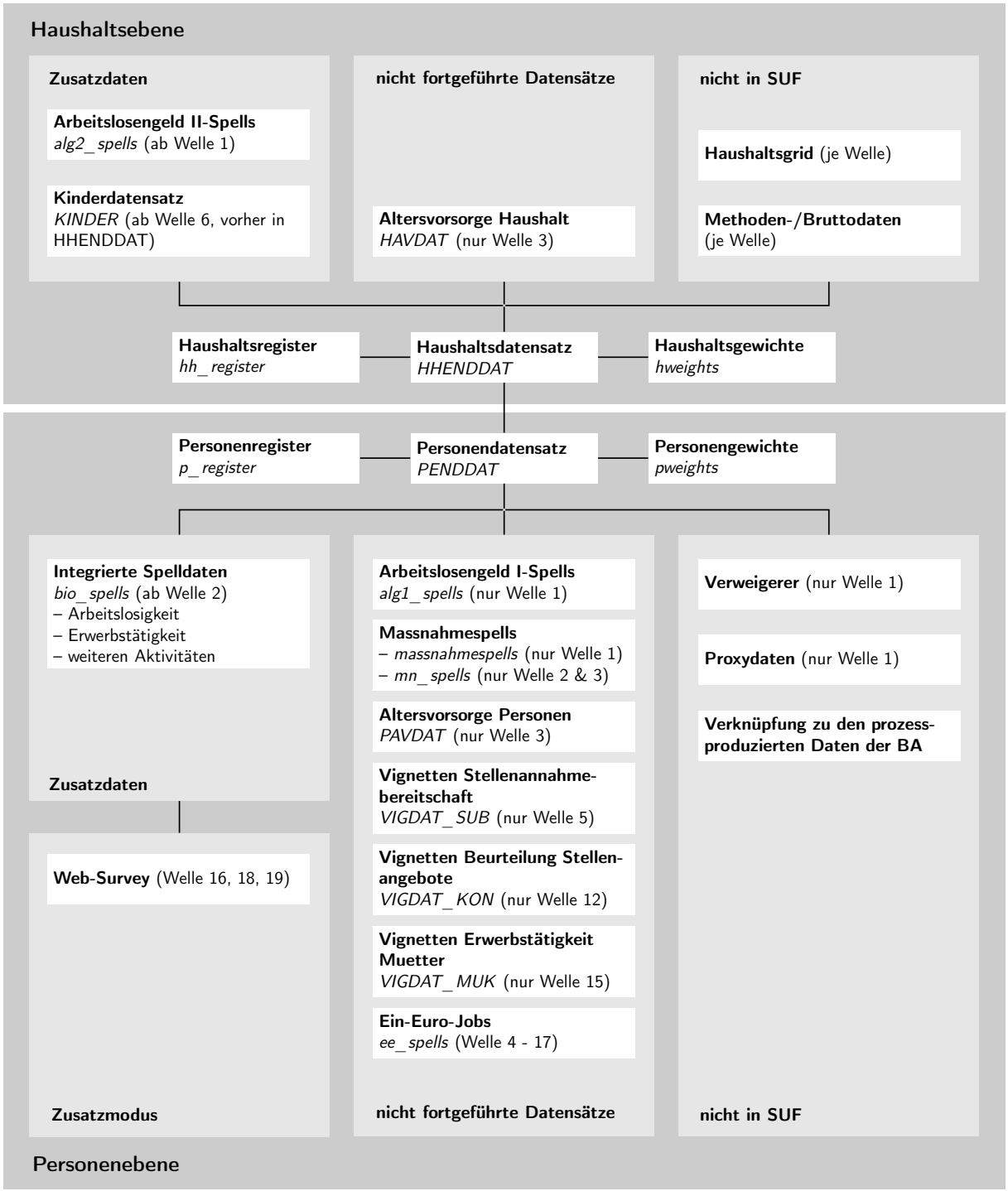


Abbildung 2: Datensatzstruktur des PASS in Welle 19

Um einen Einstieg in die selbstadministrierte Erhebung im Webmodus zu vollziehen, wurde im Anschluss an Welle 16 im Jahr 2022 erstmalig ein Websurvey im Rahmen des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS) durchgeführt. Das Ziel war dabei insbesondere die Erprobung optimaler Abläufe bei der Durchführung von Webbefragungen in der PASS-Stichprobe (Dickmann et al. 2023). Im Anschluss an die Wellen 18 und 19 wurden erneut Websurvey-Befragungen durchgeführt.

In der Labelung aller Datensätze des Scientific Use File werden Umlaute und Sonderzeichen aufgelöst, um eine einheitliche Anzeige der Labeltexte unabhängig vom bei den einzelnen Nutzerinnen und Nutzern eingestellten Zeichensatz zu gewährleisten.

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 Variablennamen[29]. Die nachfolgenden beiden Tabellen geben einen Überblick über die offenen Angaben, die in der Welle 19 vercodet wurden[30].

[29] *ogebland* (Geburtsland); *ostaatan* (Staatsangehörigkeit); *ostaatansyr* (syrische Staatsangehörigkeit); *ostaatanukr* (ukrainische Staatsangehörigkeit); *ozulanda* bis *ozulandf* (Zuzugsländer Eltern/Großeltern)

[30] 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 5: Vercodung offener Angaben auf Haushaltsebene in Welle 19

Fester Variablenname	zugecodet zu Variable	Datensatz	Bezeichnung
HD1100a-o	*HD1101a-o*	*HHENDDAT*	Anderer Erwerbsstatus der HH-Mitglieder, ggf. Proxy-Angabe
HW0880a-i	*HW0881a-j*	*HHENDDAT*	Anderer, nicht in Liste enthaltener Auszugsgrund
HKU0300	*HKU0301*	*HHENDDAT*	Wohnkosten liegen unterhalb / oberhalb der Obergrenze des Jobcenters
HKU0400a-f	*HKU0401a-f*	*HHENDDAT*	Anderer Grund des Jobcenters, über der Obergrenze liegende Wohnkosten anzuerkennen
HKU0800a-g	*HKU0801a-g*	*HHENDDAT*	Andere Maßnahme zur Senkung der Wohnkosten
AL20550a-h	*AL20551a-i*	*alg2_spells*	Anderer Grund für den Beginn des ALG2-Bezugs
AL21300a-i bis *AL22100a-i*	*AL21301a-i* *AL21401a-i* *AL21501a-i* *AL21601a-i* *AL21701a-i* *AL21801a-i* *AL21851a-i* *AL21901a-i* *AL22001a-i* *AL22101a-i* *AL22102a-i* *AL22103a-i*	*alg2_spells*	Anderer, nicht in Liste enthaltener Grund für Kürzung
AL22200a-h	*AL22201a-h*	*alg2_spells*	Anderer, nicht in Liste enthaltener Grund, warum kein ALG II mehr bezogen wurde

Tabelle 6: Vercodung offener Angaben auf Personenebene in Welle 19

Fester Variablenname	zugecodet zu Variable	Datensatz	Bezeichnung
PB0230 (Code 6)	*PB0231*	*PENDDAT*	Anderer, nicht in Liste enthaltener deutscher Schulabschluss (Fortschreibung)
PB0230 (Code 7)	*PB0231*	*PENDDAT*	Anderer, nicht in Liste enthaltener ausländischer Schulabschluss (Fortschreibung)
PB0400 (Code 9)	*PB0401*	*PENDDAT*	Anderer, nicht in Liste enthaltener deutscher Schulabschluss (Ersterhebung o. keine Angabe in VW)
PB0400 (Code 10)	*PB0401*	*PENDDAT*	Anderer, nicht in Liste enthaltener ausländischer Schulabschluss (Ersterhebung o. keine Angabe in VW)
PB1000	*PB1001*	*PENDDAT*	Anderer, nicht in Liste enthaltener ausländischer Schulabschluss (Ersterhebung o. keine Angabe in VW)
PB1300a-j (Item I)	*PB1301a-j*	*PENDDAT*	Anderer, nicht in Liste enthaltener deutscher Ausbildungsabschluss (Ersterhebung o. keine Angabe in VW)
PB1300a-j (Item J)	*PB1301a-j*	*PENDDAT*	Anderer, nicht in Liste enthaltener ausländischer Ausbildungsabschluss (Ersterhebung o. keine Angabe in VW)
PB1600	*PB1601*	*PENDDAT*	Anderer, nicht in Liste enthaltener Ausbildungsabschluss, dem der ausländische Abschluss entspricht
AL0600	*AL0601*	*bio_spells*	Anderer, nicht in Liste enthaltener Grund, weshalb nicht mehr arbeitslos gemeldet
BIO0100	*BIO0101*	*bio_spells*	Anderer, nicht in Liste enthaltener Aktivitätstyp
ET2400	*ET2401*	*bio_spells*	Andere Art, von der Stelle zu erfahren
ET2420	*ET2421*	*bio_spells*	Über anderes soziales Netzwerk von Stelle erfahren
PTK0320b-g	*PTK0321b-g*	*PENDDAT*	Anderer, nicht in Liste enthaltener Grund, weshalb keine Arbeit gesucht werden muss
PTK1700a-lm	*PTK1701a-im*	*PENDDAT*	Anderes Förderangebot des Jobcenters

Tabelle 6: Vercodung offener Angaben auf Personenebene in Welle 19 (Fortsetzung)

PTK3200a-j	*PTK3201a-j*	*PENDDAT*	Andere Gesprächsthemen mit dem Jobcenter
PKO1700a-g	*PKO1701a-g*	*PENDDAT*	Anderer Grund für frühere Bürgergeldbeziehende, noch keinen Kooperationsplan vollständig erstellt zu haben
PKO1705a-g	*PKO1706a-g*	*PENDDAT*	Anderer Grund für aktuelle Bürgergeldbeziehende, noch keinen Kooperationsplan vollständig erstellt zu haben
PAS0900a-g	*PAS0901a-g* *PAS0901i*	*PENDDAT*	Andere, nicht in Liste enthaltene Stellen, bei denen sich ZP über offene Stellen informiert hat
PAS0920a-l	*PAS0921a-l*	*PENDDAT*	Anderes für die Stellensuche genutztes soziales Netzwerk
PER0200a-e	*PER0201a-e*	*PENDDAT*	Andere Gründe für Erwerbstätigkeit im Ruhestand
PER0400	*PER0401*	*PENDDAT*	Andere Art, von der Stelle zu erfahren
PER1000a-e	*PER1001a-e*	*PENDDAT*	Andere Gründe für die geplante Aufnahme einer Erwerbstätigkeit im Ruhestand
PER1300a-g	*PER1301a-g*	*PENDDAT*	Andere Gründe gegen die Aufnahme einer Erwerbstätigkeit im Ruhestand
PSK0700a-e	*PSK0701a-e*	*PENDDAT*	Andere Gründe für freiwilliges Engagement
PSK1100a-e	*PSK1101a-e*	*PENDDAT*	Andere Gründe für die geplante Aufnahme eines freiwilligen Engagements
PSK1200a-f	*PSK1201a-f*	*PENDDAT*	Andere Gründe gegen die Aufnahme eines freiwilligen Engagements
PG0950a-i	*PG0951a-i*	*PENDDAT*	Andere Art der Behinderung/ Einschränkung
PG1310	*PG1311*	*PENDDAT*	Gesetzliche Krankenversicherung
PG1320	*PG1321*	*PENDDAT*	Private Krankenversicherung
PG1330	*PG1331*	*PENDDAT*	Private Krankenversicherung, Beihilfeberechtigung
PG1340	*PG1341*	*PENDDAT*	Heilfürsorge

Tabelle 6: Vercodung offener Angaben auf Personenebene in Welle 19 (Fortsetzung)

PG1350	*PG1351*	*PENDDAT*	Andere Krankenversicherung
PMI0200	*ogebland*	*PENDDAT*	Anderes, nicht in Liste enthaltenes Geburtsland
PMI0500	*ostaatan*	*PENDDAT*	Andere, nicht in Liste enthaltene Staatsangehörigkeit
PMI1000a-f	*ozulanda-f*	*PENDDAT*	Anderes, nicht in der Liste enthaltenes Zuzugsland des Elternteils/Großelternteils
PMI1700	*PMI1701*	*PENDDAT*	Rechtliche Grundlage des Zuzugs
PMI3000	*PMI3001*	*PENDDAT*	Anderer Grund, die Anerkennung eines im Ausland erworbenen beruflichen Abschlusses in Deutschland nicht zu beantragen
PSH0200 (Code 9)	*PSH0201*	*PENDDAT*	Anderer, nicht in Liste enthaltener deutscher Schulabschluss der Mutter
PSH0200 (Code 10)	*PSH0201*	*PENDDAT*	Anderer, nicht in Liste enthaltener ausländischer Schulabschluss der Mutter
PSH0300a-i (Code 7)	*PSH0301a-i*	*PENDDAT*	Anderer, nicht in Liste enthaltener deutscher Berufsabschluss der Mutter
PSH0300a-i (Code 8)	*PSH0301a-i*	*PENDDAT*	Anderer, nicht in Liste enthaltener ausländischer Berufsabschluss der Mutter
PSH0500 (Code 9)	*PSH0501*	*PENDDAT*	Anderer, nicht in Liste enthaltener deutscher Schulabschluss des Vaters
PSH0500 (Code 10)	*PSH0501*	*PENDDAT*	Anderer, nicht in Liste enthaltener ausländischer Schulabschluss des Vaters
PSH0600a-i (Code 7)	*PSH0601a-i*	*PENDDAT*	Anderer, nicht in Liste enthaltener deutscher Berufsabschluss des Vaters
PSH0600a-i (Code 8)	*PSH0601a-i*	*PENDDAT*	Anderer, nicht in Liste enthaltener ausländischer Berufsabschluss 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 Codierenden in Schulungen vermittelt wird. Die Schulungen finden unter Einsatz standardisierter Schulungsunterlagen statt. Die erste Schulung neuer Codierender beinhaltet eine Präsentation, in der die Grundregeln des Codierens und der ISCO/KldB-Codierung vermittelt werden, sowie die Codierung und Besprechung von ausgewählten Testfällen unterschiedlichen Schwierigkeitsgrades. Die Schulung umfasst eineinhalb Tage.

Sofern Codierende länger als ein halbes Jahr keine Berufscodierungen durchgeführt haben, erfolgt zu Beginn eines neuen Projekts eine Auffrischung der Codierregeln sowie ein einmaliger Abgleich aller Codierender untereinander. Hierfür werden mindestens 500 randomisierte Fälle von allen beteiligten Personen codiert und die Abweichungen analysiert. Anhand dieses Vorgehens können systematische Fehler einzelner Codierender 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 Codierenden ein individuelles Feedback über eventuell auftretende Abweichungen. Hierfür werden für alle Codierenden Fälle aufgelistet, bei welchen ein vorgeschlagener Code abgelehnt wurde. Zeichnen sich systematische Fehler ab, werden diese mit dem jeweiligen Codierenden 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 Mitarbeitenden der Codierung mit den Mitarbeitenden 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 7: Codeschema der im PASS genutzten Zusatzvariablen

Abkürzung	Titel
StiB_g	**Grobe Einteilung der Stellung im Beruf**
ang	Angestellte/Angestellter
arb	Arbeiterin/Arbeiter
bea	Beamtin/Beamter oder Richterin/Richter
selbst_f	Sebstständig im freien Beruf
selbst_H/DL	Sebstständig Handel/Gew./Ind./Dienstl.
landw	Selbstständige Landwirtin/Selbstständiger Landwirt
mith_f	Mithelfende Familienangehörige/Mithelfender Familienangehöriger
sol	Berufssoldatin/Berufssoldat
k.A.	keine Angabe
wn	weiß nicht
StiB_f	**Feine Einteilung der Stellung im Beruf**
xxHektar	Landwirtin/Landwirt mit xx Hektar
xxMitarbeiter	Selbstständiger oder Akademischer freier Beruf mit xx Mitarbeitenden
40	Beamtin/Beamter im einfachen Dienst
41	Beamtin/Beamter im mittleren Dienst
42	Beamtin/Beamter im gehobenen Dienst
43	Beamtin/Beamter im höheren Dienst
45	Mannschaftsgrad
46	Unteroffiziersgrad
47	Offiziersgrad bis einschl. Hauptmann
48	Offiziersgrad ab Major
51	Angestellte/Angestellter mit einfacher Tätigkeit
52	Angestellte/Angestellter mit qualifizierter Tätigkeit
53	Angestellte/Angestellter mit hochqualifizierter Tätigkeit/Leitungsfunktion
54	Angestellte/Angestellter mit höheren leitungsaufgaben/umfassenden Entscheidungsbefugnissen
60	Ungelernte Arbeiterin/Ungelernter Arbeiter
61	Angelernte Arbeiterin/Angelernter Arbeiter
62	Gelernte Arbeiterin/Gelernter Arbeiter/Facharb./Geselle

Tabelle 7: Codeschema der im PASS genutzten Zusatzvariablen (Fortsetzung)

63	Vorarbeiterin/Kolonnenführerin/Vorarbeiter/Kolonnenführer
64	Meisterin/Meister/Polier
k.A.	keine Angabe
wn	weiß nicht
Aufs,x	**Aufsichtsfunktion, Zahl der beaufsichtigten Mitarbeitenden**
Aufs,x	Aufsichtsfunktion, Zahl der beaufsichtigten Mitarbeitenden
k.Aufs	keine Aufsichtsfunktion
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 Schülerin/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; Teil-facharbeiterabschluss; 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

Tabelle 7: Codeschema der im PASS genutzten Zusatzvariablen (Fortsetzung)

Prom/Hab	Promotion, Habilitation
Schüler	noch Schülerin/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 umgesetzten Berufscodierungen fortlaufend gefüllt. Das Internet kann ebenfalls für die Recherche von Tätigkeiten genutzt werden (z. B. berufenet der Bundesagentur 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 Vorwellen des Projektes. Diese angepassten Codierregeln werden dokumentiert und den Codierenden 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 Codierenden auch die ggf. abgegebenen Kommentare zu Schwierigkeiten im Codierprozess.

2. **Erstcodierung**

   Die Erstcodierung erfolgt mit Hilfe einer automatischen Vorcodierungssoftware (supervised learning). Die Daten werden in das elektronische Codier-System importiert und dort vorcodiert. Mit Hilfe der elektronischen Vorcodierungssoftware können bereits bis zu 40 Prozent der offenen Angaben eindeutig einem Code zugeordnet werden, d.h. in 40 Prozent der Fälle vergibt die automatische Vorcodierung genau einen Code. Alle übrigen Daten werden mit zwei vorgeschlagenen Codes vorcodiert.

   Voraussetzung für eine hohe Trefferquote sind eine gute Qualität der Daten und die Erhebung aller relevanter Zusatzinformationen.

3. **Zweitcodierung**
   Für alle Nennungen erfolgt im Rahmen einer Zweitcodierung eine kritische Prüfung der Fälle der automatischen Vorcodierung auf inhaltliche Plausibilität. Dabei sehen die Zweitcodierenden das Ergebnis/die Ergebnisvorschläge der Erstcodierung und können dieses annehmen (ein Vorschlag) bzw. auswählen (zwei Vorschläge) sowie korrigieren. Mit Hilfe der automatischen Vorcodierungssoftware kann mittlerweile eine Trefferquote von 70 Prozent richtiger Codes erzielt werden, die in der Zweitcodierung nicht mehr durch einen Codierer korrigiert werden.
4. **Drittcodierung**
   Codeabweichungen zwischen Erst- und Zweitcodierung werden von dritten Codierenden entschieden. Problemfälle werden in Diskussionsrunden besprochen und entschieden. Stimmen die Drittcodierenden der Erst- oder der Zweitcodierung eindeutig zu, da die andere Codierung eindeutig falsch ist, wird der richtige Code in die Entscheidungsspalte eingetragen. Können sich die Drittcodierenden nicht für einen der beiden Codes entscheiden oder schlagen einen weiteren Code vor, so wird dies gekennzeichnet. Damit ist der Fall in der Problemfallbesprechung zu thematisieren. Zudem kann eine Kommentarspalte verwendet werden, um die Entscheidung zu begründen.
5. **Problemfallbesprechung**
   Die Codierenden treffen sich regelmäßig für die Besprechung und Entscheidung der Problemfälle.
6. **Letzte Überprüfung**
   Abschließend werden die Richtigkeit der Codes und die Einhaltung der wichtigsten Codevergaberegeln durch die hauptverantwortlichen Mitarbeitenden 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.

Tabelle 8: Harmonisierte Variablen im Personendatensatz (*PENDDAT*)

Variable	thematischer Bereich	Bezeichnung
stibkz	Erwerbstätigkeit	Jetzige Stellung im Beruf, einf. Eint., harmonisiert (anonymisiert)

Während die ausdrücklich harmonisierten Variablen – neben Veränderungen am Erhebungskonzept – auch Veränderungen an Kategorien und den befragten Gruppen über die Wellen berücksichtigen, werden von einem zweiten Variablentyp Veränderungen an den befragten Gruppen nicht explizit berücksichtigt. Diese Variablen werden zwar für alle 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 Typus 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 ist[31].

[31] Bspw. wurden in Welle 1 andere Befragtengruppen nach einer Erwerbstätigkeit gefragt als in den Folgewellen. Entsprechend haben auch jeweils andere Gruppen Angaben zur StiB, zur beruflichen Tätigkeit,

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

Tabelle 9: Wellenübergreifend generierte, aber nicht vollständig harmonisierte Variablen im Personendatensatz (*PENDDAT*)

Variable	thematischer Bereich	Bezeichnung
isco88	Erwerbstätigkeit	Intern. Standard Classification of Occupations 88, aktuelle Erwerbstätigkeit, gen.
kldb1992	Erwerbstätigkeit	Klassifizierung der Berufe 1992, aktuelle Erwerbstätigkeit, gen.
azhpt2	Erwerbstätigkeit	Akt. tatsächl. Arb.zeit Haupterwerbst. (ohne Minijobs, inkl. kat. Ang.), gen.
azges2	Erwerbstätigkeit	Akt. tatsächl. Arb.zeit insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.
befrist	Erwerbstätigkeit	jetzige Tätigkeit: Befristung? generiert (alle Wellen)
mps	Erwerbstätigkeit	Magnitude-Prestigeskala, aktuelle Ewt., gen.
siops1	Erwerbstätigkeit	Standard Intern. Occupational Prestige Scale (Basis ISCO88), aktuelle Ewt., gen.
isei1	Erwerbstätigkeit	International Socio-Economic Index (Basis ISCO88), aktuelle Ewt., gen.
egp	Erwerbstätigkeit	Klassenschema nach Erikson, Goldthorpe und Portocarrero (EGP), akt. Beruf, gen.
esec	Erwerbstätigkeit	European Socio-economic Classification (ESeC), akt. Beruf, gen.
stib	Erwerbstätigkeit	Stellung im Beruf, Kennziffer, aktuelle Erwerbstätigkeit, generiert
netges	Erwerbstätigkeit	Akt. Nettoeinkommen insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.
alg1abez	Leistungsbezug	aktueller ALG 1-Bezug, generiert
aktmassn	Maßnahmeteilnahme	Akt. Teiln. an einem v. Arbeitsamt finanz./ geförd. Programm, generiert
isco08minj	Minijob	ISCO 08, aktueller Minijob, generiert
kldb2010minj	Minijob	Klassifizierung der Berufe 2010, aktueller Minijob, gen.

zur Arbeitszeit, zur Befristung etc. gemacht.

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 bestimmten Fragekomplex vorlagen[32]. An diesen Stellen wurden Vorwelleninformationen zur Filtersteuerung genutzt. Zweitens sollten den 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 abzugrenzen[33]. An anderen Stellen, insbesondere bei der Fortschreibung von Episodeninformationen[34], wurden auch Antworten der Befragten aus der Vorwelle in die Fragetexte integriert, die dazu dienten, die 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 sind[35].

Im Zuge der Datenaufbereitung werden die erhobenen Veränderungen mit Informationen

[32] 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.

[33] 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? “

[34] Beispiele sind die Fortschreibung des Arbeitslosengeld-II-Bezugs/Bürgergeld-Bezugs aus der Vorwelle im Haushaltsinterview der jeweils aktuellen Welle bzw. die Fortschreibung von Erwerbstätigkeiten oder Arbeitslosigkeiten im Personeninterview.

[35] 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 erworben, 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.

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 Querschnittsdatensätzen (*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.

Tabelle 10: Fortgeschriebene Vorwelleninformationen in Welle 19, Haushaltsfragebogen

Konstrukt	Fragenr.	Anmerkung	Fortschreibung in Var.
Wohnsituation		Wohnform, Art des Mietverhältnisses und die Art des Wohnheims während des Interviews fortgeschrieben	*HHENDDAT*: *HW0200a* bis *HW0400*
Haushaltsstruktur		Haushaltsgröße während des Interviews fortgeschrieben	*HHENDDAT*: *HA0100*
		Geschlecht der Personen im HH während des Interviews ggf. korrigiert	*HHENDDAT*: *HD0100a* bis *HD0100o*
		Alter der Personen im HH während des Interviews fortgeschrieben	*HHENDDAT*: *HD0200a* bis *HD0200o*
		Verwandtschaftsbeziehungen während des Interviews fortgeschrieben	nicht im SUF ausgeliefert
Wohnfläche der Wohnung in qm	HW1000	in generierter Variable fortgeschrieben	*HHENDDAT*: *wohnfl*
Arbeitslosengeld-II-Bezug / Bürgergeld-Bezug	Modul „Bürgergeld / Arbeitslosen-geld II“	in Arbeitslosengeld-II-Spelldatensatz fortgeschrieben	*alg2_spells*: Variablen des Arbeitslosengeld-II-Spelldatensatzes
		Information zum aktuellen Arbeitslosengeld-II-Bezug / Bürgergeld-Bezug des HH	*HHENDDAT*: *alg2abez*
		Information zum Arbeitslosengeld II Bezug der Bedarfsgemeinschaft	*p_register*: *bgbezs19; bgbezb19*

Tabelle 11: Fortgeschriebene Vorwelleninformationen in Welle 19, Personenfragebogen

Konstrukt	Fragenr.	Anmerkung	Fortschreibung in Var.
höchster allgemeinbildender Schulabschluss	PB0220-PB1100	in generierter Variable fortgeschrieben	*PENDDAT*: *schul1* (ohne offene Ang.) *schul2* (mit offenen Ang.)
Jahr des höchsten Schulabschlusses	PB0410	in generierter Variable fortgeschrieben	*PENDDAT*: *schulabj*
Berufsabschluss	PB1200-PB1600	höchster Berufsabschluss in generierter Variable fortgeschrieben	*PENDDAT*: *beruf1* (ohne offene Ang.) *beruf2* (mit offenen Ang.)
Jahr des Berufsabschlusses	PB1310a-k	in generierter Variable fortgeschrieben	*berabj*
Zeiten von fortgeschriebenen Aktivitäten im BIO-Spelldatensatz	BIO0600z1, BIO0600z2, BIO0400z, BIO0500z	in BIO-Spelldatensatz für Aufsatzspells fortgeschrieben	*bio_spells*: *BIO0400, BIO0500, BIO0600*
		Informationen zu beendeten Tätigkeiten	*bio_spells*: *ET2300, ET2700*
		Informationen zur aktuellen Erwerbstätigkeit in generierten Variablen fortgeschrieben	*PENDDAT*: *isco88; isco08; kldb1992; kdlb2010; stib; stibkz; azhpt1; azhpt2; azges1; azges2; befrist; mps; siops1; siops2; isei1; isei2; egp; esec; branche1; branche2*
		Informationen zum aktuellen (Nicht-) Erwerbstätigkeitsstatus in generierten Variablen fortgeschrieben	*PENDDAT*: *etakt; alakt; statakt*
Zeiten von Arbeitslosengeld I Bezug bei fortgeschriebenen Arbeitslosigkeits- Spells		Informationen zum aktuellen Arbeitslosengeld I Bezug	*bio_spells*: *AL0700, AL0800, AL0900, AL1000, AL1100, AL1200*

Tabelle 11: Fortgeschriebene Vorwelleninformationen in Welle 19, Personenfragebogen (Fortsetzung)

		Informationen zu beendeten Arbeitslosigkeitsperioden	*bio_spells*: *AL0600, AL0601*
		Informationen zum aktuellen Arbeitslosengeld-I-Bezug	*PENDDAT*: *alg1abez*

Von diesen Merkmalen, bei denen in der Vergangenheit erhobene Informationen mit Angaben zu Veränderungen zwischen den Befragungszeitpunkten fortgeschrieben bzw. aktualisiert werden, sind die sogenannten „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).

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 (wie z. B. beim höchsten Bildungsabschluss, siehe Kapitel 4.3) oder wichtige Informationen aus anderen Teildatensätzen zugespielt wurden (z.B. Indikatoren für den aktuellen Bezug von Arbeitslosengeld I oder SGB-II-Leistungen).

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 Querschnittsdatensätzen (*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.

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-/Seniorenінterview gibt, statt. Zwei Personengruppen werden erneut als erstbefragte Person behandelt, auch wenn sie zum wiederholten Mal ein

Tabelle 12: 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

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

Personen-/Senioreninterview 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 behandelt[36]. 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. 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 fünf datensatzspezifischen Tabellen dargestellt. Die einzelnen Variablen werden dabei jeweils kurz beschrieben. Darüber hinaus werden die für die Generierung der Variable notwendigen Quellvariablen aufgeführt[37]. Für die Querschnittsdatensätze (*HHENDDAT; PENDDAT*) ist zudem jeweils

[36] Hiervon ausgenommen ist die Information, ob eine Person in einem früheren Interview bereits zu ihrer Zuspielbereitschaft 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 in einer folgenden Welle verneinen und ihre Einwilligung damit faktisch zurückziehen. Die Möglichkeit für die Zielperson, ihre Zuspielbereitschaft zu widerrufen, bleibt von dieser Entscheidung unberührt.

[37] Wie die Variablen in den Querschnittsdatensätzen (*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 Registerdatensätzen. Die generierten Variablen in den Spelldatensätzen wurden immer in den bereits fortgeschriebenen Datensätzen erzeugt. Wurde ein Spell nicht fortgeschrieben, blieben die entsprechenden generierten Variablen unverändert (ggf. mit Ausnahme, dass im Zensierungsindikator ein Sondercode gesetzt wurde, falls der Spell aus technischen Gründen nicht fortgeführt werden konnte). Wurde ein Spell fortgeschrieben, wurde immer auf die aktuellsten Informationen zurückgegriffen, d.h. auf die mit den

angegeben, um welchen Typ es sich in Anlehnung an die einfache Generierung aus vorheriger Tabelle 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.

Informationen aus der aktuellen Welle gefüllten Variablen bzw. die für die aktuelle Welle relevanten Querschnittsvariablen in den Spells.

Tabelle 13: Einfache Generierungen für Welle 19 im Haushaltsdatensatz (*HHENDDAT*) und im *KINDER*-Datensatz (alphabetische Reihenfolge)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 19
alg2abez	*Aktueller ALG-II-Bezug, generiert*: Indikator für aktuellen Arbeitslosengeld-II-Bezug / Bürgergeld-Bezug des Haushalts (neu).	*zensiert; AL20300; AL20400; AL20500 (alg2_spells)* Information zu weiteren Arbeitslosengeld-II-Bezügen / Bürgergeld-Bezügen *(AL22700); hintjahr (HHENDDAT)*
anzgeschw	*Anzahl der Geschwister im Haushalt*: 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. *anzgeschw* stimmt daher in seltenen Fällen nicht mit dem Geschwisterstatus überein, der im *p_register* über die Elternzeiger generiert werden kann (neu).	Informationen zu Beziehungen im Haushalt *(Haushaltsgrid)*
bik	*BIK-Regionsgrößenklassen (GKBIK10), generiert*: Die Information zur Regionsgrößenklasse wurde von infas generiert, indem von der in den Adressdaten vorliegenden Postleitzahl auf *GKBIK10* umgeschlüsselt wurde (neu).	vom Erhebungsinstitut geliefert
blneualt	*Alte oder neue Bundesländer, generiert*: 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).	*bundesld* als die vom Erhebungsinstitut generierte und gelieferte Information zum Bundesland, in dem sich der Haushalt zum Befragungszeitpunkt befindet
butaber	*BuT-berechtigt zum Interviewzeitpunkt*: 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 (alg2_spells); HA0250a-b; HW1800; HW1950; HEK0100; HEK0115; HEK1630; HEK1645 (HHENDDAT)*
hhinckat	*Kategorisiertes Haushaltseinkommen pro Monat (in Euro), generiert*: Zusammenfassung der mehrstufig erhobenen kategorialen Angaben zum Haushaltseinkommen zu einer Variablen (neu).	*HEK0700; HEK0800; HEK0900; HEK1000; HEK1100 (HHENDDAT)*

Tabelle 13: Einfache Generierungen für Welle 19 im Haushaltsdatensatz (*HHENDDAT*) und im *KINDER*-Datensatz (alphabetische Reihenfolge) (Fortsetzung)

hhincome	*Haushaltseinkommen pro Monat (in Euro) inkl. kategorisierter Angaben, generiert*: Generierung einer integrierten Variable aus kategorisierter und offener Angabe zum Haushaltsnettoeinkommen (neu).	*HEK0600; HEK0700; HEK0800; HEK0900; HEK1000; HEK1100 (HHENDDAT)*
hintdat	*Interviewdatum Haushaltsinterview*: Generierte Variable mit dem Datum, an dem das Haushaltsinterview geführt wurde im Format JJMMTT (neu).	*hintjahr; hintmon; hinttag (HHENDDAT)*
hintnum	*Interviewer Haushaltsinterview*: Die systemfreie Kennung kennzeichnet den Interviewer bzw. die Interviewerin, der bzw. die die Befragung durchgeführt hat. Die Information ist sowohl zwischen *HHENDDAT* und *PENDDAT* als auch wellenübergreifend einheitlich vergeben worden. Eine bestimmte Ausprägung der Kennung bezeichnet also stets ein und den selben Interviewer bzw. dieselbe Interviewerin (neu).	vom Erhebungsinstitut generierte und gelieferte Information
kindu4	*Steuervariable: Kind unter 4 Jahre im HH*: 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).	*HD0200a - HD0200o (HHENDDAT)*
kindu13	*Steuervariable: Kind unter 13 Jahre im HH*: 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).	*HD0200a - HD0200o (HHENDDAT)*
kindu15	*Steuervariable: Kind unter 15 Jahre im HH*: 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).	*HD0200a - HD0200o;* kategoriale Nachfrage zur Altersgruppe (bei KA in *HD0200 (HHENDDAT)*)

Tabelle 13: Einfache Generierungen für Welle 19 im Haushaltsdatensatz (*HHENDDAT*) und im *KINDER*-Datensatz (alphabetische Reihenfolge) (Fortsetzung)

kindu18	*Steuervariable: Kind unter 18 Jahre im HH*: Die Variable zeigt an, dass in der Welle mind. eine Person im Haushalt unter 18 Jahre alt ist. Da die Generierung nur auf den Altersangaben im Haushaltsdatensatz basiert, ist unerheblich, ob diese Person unter 18 Jahren auch tatsächlich Kind einer anderen im Haushalt lebenden Person ist (neu).	*HD0200a - HD0200o;* kategoriale Nachfrage zur Altersgruppe (bei KA in HD0200) *(HHENDDAT)*)
kindu25	*Steuervariable: Kinder unter 18 Jahre bzw. Schüler unter 25 Jahre im HH*: 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 oder Schülerin 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).	*HD0200a - HD0200o;* kategoriale Nachfrage zur Altersgruppe (bei KA in *HD0200*); *HD1100a-o (HHENDDAT)*
kind5u15	*Steuervariable: Kind von 5 bis unter 15 Jahre im HH*: Die Variable zeigt an, dass in der Welle mind. eine Person im Haushalt 5 bis unter 15 Jahre alt ist. Da die Generierung nur auf den Altersangaben im Haushaltsdatensatz basiert, ist unerheblich, ob diese Person im Alter von 5 bis unter 15 Jahren auch tatsächlich Kind einer anderen im Haushalt lebenden Person ist (neu).	*HD0200a - HD0200o;* *(HHENDDAT)*)
wohnfl	*Wohnfläche in qm, generiert*: 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 (HHENDDAT)* Bei wiederholter Erfassung: wohnfl aus Vorwelle; *HW1000; (HHENDDAT)*

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 19
alakt	*Aktuell arbeitslos gemeldet, generiert (ab Welle 2)*: Zeigt an, dass die ZP zum Zeitpunkt des Personeninterviews der jeweiligen Welle arbeitslos gemeldet war (neu).	*zensiert; spintegr; BIO0101 (bio_spells)*

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge)

(Fortsetzung)

alg1abez	*Aktueller ALG I-Bezug, generiert*: Indikator: Befragungsperson bezieht zum Interviewzeitpunkt Arbeitslosengeld I. In der 19. Welle wurden die Zeiten seit Januar 2023 erhoben, in denen die Befragten arbeitslos gemeldet waren. Für jede Episode wurde zudem erfragt, ob und in welchem Zeitraum Arbeitslosengeld I bezogen wurde (neu).	*AL0700; AL1000; AL1100; AL1200 (bio_spells)*
apartner	*Steuervariable: Unverh. Partner im HH*: Indikator: Befragungsperson hat einen bzw. eine unspezifische/n oder unverheiratete/n Partner bzw. Partnerin im HH (neu).	Informationen zu Beziehungen im Haushalt *(Haushaltsgrid); PD0500 - PD0800 (PENDDAT)*
azhpt1	*Akt. vertragl. Arb.zeit Haupterwerbst. ohne Minijobs), gen.*: Wöchentliche, vertraglich vereinbarte Arbeitszeit der Haupterwerbstätigkeit, die die Befragten zum Interviewzeitpunkt ausüben, generiert aus offenen Angaben zur Arbeitszeit (neu).	*ET2017 (bio_spells)*
azhpt2	*Akt. tatsächl. Arb.zeit Haupterwerbst. (ohne Minijobs, inkl. kat. Ang.), gen.*: Wöchentliche, tatsächliche Arbeitszeit der Haupterwerbstätigkeit, die die Befragten zum Interviewzeitpunkt ausüben, generiert aus offenen Angaben zur Arbeitszeit und kategorialer Nachfrage bei unregelmäßigen Arbeitszeiten (neu).	*ET2117; ET2217 (bio_spells)*
azges1	*Akt. vertragl. Arb.zeit insgesamt (ohne Minijobs), gen.*: Wöchentliche, vertraglich vereinbarte Arbeitszeit aller Erwerbstätigkeiten, die die Befragten zum Interviewzeitpunkt ausüben, generiert aus offenen Angaben zur Arbeitszeit (neu).	*ET2017 (bio_spells)*
azges2	*Akt. tatsächl. Arb.zeit insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.*: Wöchentliche, tatsächliche Arbeitszeit aller Erwerbstätigkeiten, die die Befragten zum Interviewzeitpunkt ausüben, generiert aus offenen Angaben zur Arbeitszeit und kategorialer Nachfrage bei unregelmäßigen Arbeitszeiten (neu).	*ET2117; ET2217 (bio_spells)*
befrist	*Jetzige Tätigkeit: Befristung? Generiert (alle Wellen)*: Indikator: Die Erwerbstätigkeit, die die Befragten zum Interviewzeitpunkt ausüben, ist befristet (neu).	*PET2510a; PET2510b (PENDDAT)*

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge) (Fortsetzung)

begjeewt	*Anfangsjahr der ersten Erwerbstätigkeit, generiert*: Jahr, in dem die Befragten erstmals hauptberuflich tätig waren. Für die Generierung wurden Informationen zur ersten Erwerbstätigkeit mit Informationen aus den Erwerbstätigkeitsepisoden kombiniert, wenn die Befragten die erste Erwerbstätigkeit bereits bei der Erhebung der Erwerbstätigkeitsepisoden seit Januar 2023 genannt hatten (uv).	Bei Ersterfassung: *bjahr (bio_spells); PET3200b (PENDDAT)* Nach Ersterfassung: *begjeewt* aus Vorwelle *(PENDDAT)*
begjminj	*Anfangsjahr des aktuellen Minijobs, generiert*: Jahr, seit dem die Befragten im aktuellen (Haupt-)Minijob tätig sind (neu).	*PMJ0800b*
begmeewt	*Anfangsmonat der ersten Erwerbstätigkeit, generiert*: Monat, in dem die Befragten erstmals hauptberuflich tätig waren (Generierung siehe *begjeewt*) (uv).	Bei Ersterfassung: *bmonat (bio_spells); PET3200a (PENDDAT)*; Nach Ersterfassung: *begmeewt* aus Vorwelle *(PENDDAT)*
begmminj	*Anfangsmonat des aktuellen Minijobs, generiert*: Monat, seit dem die Befragten im aktuellen (Haupt-)Minijob tätig sind (neu).	*PMJ0800a*
berabj	*Jahr des höchsten Berufsabschlusses*: Jahr, in dem die Befragten den zum Interviewzeitpunkt höchsten Berufsabschluss erworben haben (fs) *Anmerkung: Die Abschlussjahre der in der Welle 1 berichteten Berufsabschlüsse wurden in der Welle 2 nacherhoben.*	Bei Ersterfassung: *PB1310aj-kj (PENDDAT)* Bei wiederholter Erfassung: *berabj* aus Vorwelle; *PB1310aj-kj (PENDDAT)*
beruf1	*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).	Bei Ersterfassung: *PB0100; PB0200; PB0300; PB1200b; PB1200c; PB1300a-j (PENDDAT)* Bei wiederholter Erfassung: *beruf1* aus Vorwelle; *PB0100; PB0200; PB1200a; PB1300a-j (PENDDAT)*

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge)

(Fortsetzung)

beruf2	*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).	Bei Ersterfassung: *PB0200; PB1301a-j; PB1500a; PB1500b; PB1500c; PB1601 (PENDDAT)* Bei wiederholter Erfassung: *PB0200; PB1301a-j; PB1500a; PB1500b; PB1500c; PB1601 (PENDDAT)*
brges	*Akt. Bruttoeinkommen insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.*: Enthält die über alle ETs (oberhalb der Minijobverdienstgrenze) kumulierte Angabe zum Bruttoeinkommen. Generiert aus offenen Angaben zum Bruttoeinkommen und kategorialer Nachfrage bei wn/ka-Angabe bei der offenen Abfrage (neu).	*ET2814; ET2914; ET3014; ET3114; ET3214; ET3314 (bio_spells)*
brutto	*Akt. Bruttoeinkommen Haupterwerbst (ohne Minijobs, inkl. kat. Ang.), gen.*: Generierung einer integrierten Variable aus kategorisierter und offener Angabe zum Bruttoeinkommen (neu).	*ET2814; ET2914; ET3014; ET3114; ET3214; ET3314 (bio_spells)*
bruttokat	*Akt. Bruttoeinkommen Haupterwerbst. (ohne Minijobs, kategorisiert), gen.*: Zusammenfassung der mehrstufig erhobenen kategorialen Bruttoeinkommensangaben zu einer Variable (neu).	*ET2814; ET2914; ET3014; ET3114; ET3214; ET3314 (bio_spells)*
emonlewt	*Endzeitpunkt letzte Erwerbstätigkeit (Monat)*: Monat, in dem die Befragten zuletzt erwerbstätig waren. Generierung siehe ejhrlewt (fs).	Bei Ersterfassung: *PET1200a (PENDDAT); ejahr; emonat (bio_spells)* Bei wiederholter Erfassung: *ejhrlewt* aus Vorwelle *(PENDDAT); ejahr; emonat (bio_spells)*

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge)

(Fortsetzung)

ejhrlewt	*Endzeitpunkt letzte Erwerbstätigkeit (Jahr)*: Jahr, in dem die Befragten zuletzt erwerbstätig waren. Für die Generierung wurden Informationen aus den Erwerbstätigkeitsepisoden mit den Angaben zur letzten Erwerbstätigkeit kombiniert, wenn die Befragungsperson seit Januar 2023 nicht erwerbstätig war (fs).	Bei Ersterfassung: *PET1200b (PENDDAT); ejahr; emonat (bio_spells)* Bei wiederholter Erfassung: *ejhrlewt* aus Vorwelle *(PENDDAT); ejahr; emonat (bio_spells)*
ekin1517	*Steuervariable: eigenes Kind zwischen 15 und 17 Jahren im HH*: Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/ Adoptivkind im Alter zwischen 15 und 17 Jahren im Haushalt hat (neu).	Informationen zu Beziehungen im Haushalt (Haushaltsgrid); Information zum Alter: *HD0200a - HD0200o* (HHENDDAT)
ekind	*Steuervariable: eigenes Kind im HH*: 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 *ekind* im HH lebende Kinder hat, ihre *pnr* jedoch in den Zeigern *zmhh* und *zvhh* des *p_registers* nicht erscheint. Dies kann z.B. bei gleichgeschlechtlichen Beziehungen mit Kindern auftreten oder, wenn in einem HH neben dem aktuellen Partner oder der aktuellen Partnerin auch noch der ehemalige Partner beziehungsweise die ehemalige Partnerin lebt.	Informationen zu Beziehungen im Haushalt *(Haushaltsgrid)*
ekin614	*Steuervariable: eigenes Kind zwischen 6 und 14 Jahren im HH*: 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); Information zum Alter: *HD0200a - HD0200o* (HHENDDAT)
ekinu15	*Steuervariable: eigenes Kind unter 15 Jahre im HH*: Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/ Adoptivkind im Alter von unter 15 Jahren im Haushalt hat (neu).	Informationen zu Beziehungen im Haushalt (Haushaltsgrid); Information zum Alter: *HD0200a - HD0200o* (HHENDDAT)

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge)

(Fortsetzung)

ekinu18	*Steuervariable: eigenes Kind unter 18 Jahre im HH*: Indikator, dass Befragungsperson ein unspezifisches, leibliches oder Stief-/ Adoptivkind im Alter von unter 18 Jahren im Haushalt hat (neu).	Informationen zu Beziehungen im Haushalt (Haushaltsgrid); Information zum Alter: *HD0200a - HD0200o* (HHENDDAT)
epartner	*Steuervariable: Ehepartner oder eingetr. Lebenspartner im HH*: Indikator, dass Befragungsperson einen Ehepartner oder einen gleichgeschlechtlichen, eingetragenen Lebenspartner bzw. eine Ehepartnerin oder eine gleichgeschlechtliche, eingetragene Lebenspartnerin im Haushalt hat (neu).	Informationen zu Beziehungen im Haushalt *(Haushaltsgrid)*
etakt	*Aktuell erwerbstätig (oberhalb Minijobverdienstgrenze), generiert (ab Welle 2)*: Zeigt an, dass die ZP zum Zeitpunkt des Personeninterviews der jeweiligen Welle einen andauernden ET-Spell (d.h. eine ET oberhalb der Minijobverdienstgrenze) hatte (neu).	*zensiert, spintegr, BIO0101 (bio_spells)*
famstand	*Familienstand, generiert*: Generierung einer integrierten Familienstandsvariable aus den Angaben im Personenbogen und der aus dem Haushaltsdatensatz generierten Steuerungsvariable *epartner* (neu).	*epartner; PD0500; PD0700 (PENDDAT)*
gebhalbj	*Geburtshalbjahr, generiert*: Indikator, ob Geburtstag im 1. oder 2. Halbjahr des Geburtsjahres liegt (neu).	Informationen zum Geburtsmonat
kindzges	*Anzahl eigene Kinder gesamt (inner- und außerhalb HH), generiert*: 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)*
kindzihh	*Anzahl eigene Kinder innerhalb des HH, generiert*: 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).	

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge) (Fortsetzung)

	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.	Informationen zu Beziehungen im Haushalt *(Haushaltsgrid)*
mberuf1	*Höchster ber. Abschluss der Mutter, inkl. Mut. im HH, ohne off. Ang., gen.*: 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 *mberuf1* aus der Vorwelle übernommen (uv).	Bei Ersterfassung: *PSH0300a-i (PENDDAT)* Nach Ersterfassung: *mberuf1* aus Vorwelle *(PENDDAT)*
mberuf2	*Höchster ber. Abschluss der Mutter, inkl. Mut. im HH, off. Ang., gen.*: Wie *mberuf1* mit dem Unterschied, dass bei der Generierung von *mberuf2* auch die gemachten offenen Angaben berücksichtigt wurden (uv).	Bei Ersterfassung: *PSH0301a-i (PENDDAT)* Nach Ersterfassung: *mberuf2* aus Vorwelle *(PENDDAT)*
mhh	*Steuervariable: Mutter lebt im HH*: Indikator, der anzeigt, dass die unspezifische, leibliche, Stief-/Adoptivmutter der Befragungsperson im Haushalt lebt (neu).	Informationen zu Beziehungen im Haushalt *(Haushaltsgrid)*
migration	*Migrationshintergrund d. Befragten, generiert*: 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).	
	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.	Bei Ersterfassung: *PMI0100; PMI0700; PMI0800a-f; PMI0900a-f (PENDDAT)* Nach Ersterfassung: *migration* aus Vorwelle *(PENDDAT)*

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge) (Fortsetzung)

mschul1	*Höchster allg. Schulabschluss d. Mutter, inkl. Mutter im HH, o. offene Ang., gen.*: Der höchste Schulabschluss der Mutter wurde in der Welle 1 nur dann erhoben, wenn diese nicht im Haushalt wohnte. Wohnte sie im Haushalt, wurde 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: *PSH0200 (PENDDAT)* Nach Ersterfassung: *mschul1* aus Vorwelle *(PENDDAT)*
mschul2	*Höchster allg. Schulabschluss d. Mutter, inkl. Mut. im HH, inkl. off. Ang., gen.*: Wie *mschul1* mit dem Unterschied, dass bei der Generierung von *mschul2* auch die gemachten offenen Angaben berücksichtigt wurden (uv).	Bei Ersterfassung: *PSH0201 (PENDDAT)* Nach Ersterfassung: *mschul2* aus Vorwelle *(PENDDAT)*
mstib	*Stellung im Beruf der Mutter, Kennziffer, generiert*: Detaillierte Stellung der Mutter im Beruf, generiert aus den Einzelvariablen (uv).	Bei Ersterfassung: *PSH0320; PSH0330; PSH0340; PSH0360; PSH0370; PSH0380 (PENDDAT)* Nach Ersterfassung: *mstib (PENDDAT)*
netges	*Akt. Nettoeinkommen insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.*: Enthält die über alle ETs (oberhalb der Minijobverdienstgrenze) 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).	*ET3414; ET3514; ET3614; ET3714; ET3814; ET3914 (bio_spells)*
netto	*Akt. Nettoeinkommen Haupterwerbst. (ohne Minijobs, inkl. kat. Ang.), gen.*: Generierung einer integrierten Variable aus kategorisierter und offener Angabe zum Nettoeinkommen (neu).	*ET3414; ET3514; ET3614; ET3714; ET3814; ET3914 (bio_spells)*
nettokat	*Akt. Nettoeinkommen Haupterwerbst. (ohne Minijobs, kategorisiert), gen.*: Zusammenfassung der mehrstufig erhobenen kategorialen Nettoeinkommensangaben zu einer Variablen (neu).	*ET3414; ET3514; ET3614; ET3714; ET3814; ET3914 (bio_spells)*

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge)

(Fortsetzung)

ostaatansyr	*Staatsangehörigkeit syr./irak. HH, inkl. off. Ang., Kategorien (anon.)*: 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 *ostaatan, sample (PENDDAT)*
ostaatanukr	*Staatsangehörigkeit ukrain. HH, inkl. off. Ang., Kategorien (anon.)*: Anhand dieser Kennung kann ab Welle 17 für die Teilstichproben der ukrainischen Haushalte betrachtet werden, ob eine Person die ukrainische oder eine andere Nationalität aufweist (neu).	Nicht anonymisierte Variante von *ostaatan, sample (PENDDAT)*
palter	*Alter (Welle 1: gen. aus P1; ab Welle 2: beste Inf.), generiert*: Alter der Befragungsperson, generiert auf Basis der Angaben zum Geburtsdatum und des Datums des Personeninterviews in der der aktuellen Welle (neu).	*PD0100; pintjahr, pintmon, pinttag (PENDDAT)*
panel	*Panelbereitschaft*: (neu).	Vom Erhebungsinstitut gelieferte Information zur Panelbereitschaft der Haushalte
pintdat	*Interviewdatum Personeninterview*: Generierte Variable mit dem Datum, an dem das Personeninterview geführt wurde im Format JJMMTT (neu).	*pintjahr, pintmon, pinttag (PENDDAT)*
pintnum	*Interviewer Personeninterview*: Die systemfreie Kennung kennzeichnet den Interviewer, der die Befragung durchgeführt hat. Die Information ist sowohl zwischen *PENDDAT* und *HHENDDAT* als auch wellenübergreifend einheitlich vergeben worden. Eine bestimmte Ausprägung der Kennung bezeichnet stets denselben Interviewer bzw. dieselbe Interviewerin (neu).	Vom Erhebungsinstitut generierte und gelieferte Information

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge) (Fortsetzung)

schul1	*Höchster Schulabschluss, ohne ausl. Abschl. und offene Angaben*: Variable für den höchsten Schulabschluss; äquivalente west- und ostdeutsche Abschlüsse wurden zusammengefasst (z.B. EOS und Abitur); ohne offene Angaben (fs).	Bei Ersterfassung: *PB0200; PB0220; PB0230; PB0300; PB0400 (PENDDAT)* Bei wiederholter Erfassung: *schul1* aus Vorwelle; *PB0200; PB0220; PB0230; PB0300; PB0400 (PENDDAT)*
schul2	*Höchster Schulabschluss, mit ausl. Abschl. und offene Angaben*: Wie *schul1* mit folgenden Unterschieden: 1. Einbeziehen von offenen Angaben; 2. Einbeziehen von Angaben zu ausländischen Abschlüssen (fs).	Bei Ersterfassung: *PB0200; PB0220; PB0231; PB0300; PB0401 (PENDDAT)* Bei wiederholter Erfassung: *schul2* aus Vorwelle; *PB0200; PB0220; PB0231; PB0300; PB0401 (PENDDAT)*
schulabj	*Jahr des höchsten Schulabschlusses*: Jahr, in dem die Befragungsperson ihren höchsten Schulabschluss erworben hat (fs). *Anmerkung: Wiederholt Befragte, für die bereits Informationen zum Schulabschluss aus einer Vorwelle vorlagen, 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 19 vor Mai 2025 geführt, wurde davon ausgegangen, dass der Abschluss in 2024 erworben wurde, bei später geführten Interviews wurde 2025 als Abschlussjahr angenommen.*	Bei Ersterfassung: *PB0220; PB0230; PB0410; pintjahr; pintmon (PENDDAT)* Bei wiederholter Erfassung: *schulabj* aus Vorwelle; *PB0220; PB0230; PB0410; pintjahr; pintmon (PENDDAT)*
statakt	*Aktueller Hauptstatus, generiert (ab Welle 2)*: Zeigt an, welchen Hauptstatus die ZP zum Zeitpunkt des Personeninterviews der jeweiligen Welle hatte (neu).	*zensiert; spintegr; BIO0101; azges2 (bio_spells)*

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge) (Fortsetzung)

stib	*Stellung im Beruf, Kennziffer, aktuelle Erwerbstätigkeit, generiert*: Generierung der detaillierten Kennziffer der Stellung im Beruf aus den Einzelvariablen. Generierung der Variable mit Informationen aus dem Modul â€žErwerbstätigkeitâ€œ (*ET060*-ET120**). Bei mehreren aktuell andauernden Erwerbstätigkeitsepisoden wurde die Episode mit dem höheren Stundenumfang gewählt. Bei mehreren aktuell andauernden Episoden mit exakt gleichem Stundenumfang wurde die Episode gewählt, die eher begonnen hat (neu).	*ET0617; ET0717; ET0817; ET0917; ET1017; ET1117; ET1217 (bio_spells)*
stibeewt	*Stellung im Beruf, erste Erwerbstätigkeit, Kennziffer, generiert*: Detaillierte Kennziffer der Stellung im Beruf, in dem die Befragten erstmals hauptberuflich tätig waren. Für die Generierung wurden Informationen zur ersten Erwerbstätigkeit mit Informationen aus den Erwerbstätigkeitsepisoden kombiniert, wenn die Befragten angaben, die erste Erwerbstätigkeit bereits bei der Erhebung der Erwerbstätigkeitsepisoden seit Januar 2023 genannt zu haben (uv).	Bei Ersterfassung: *PET3300; PET3400; PET3500; PET3600; PET3700; PET3800; PET3900 (PENDDAT)* *ET0617; ET0717; ET0817; ET0917; ET1017; ET1117; ET1217 (bio_spells)* Nach Ersterfassung: *stibeewt* aus Vorwelle *(PENDDAT)*
stiblewt	*Stellung im Beruf, letzte Erwerbstätigkeit, Kennziffer, generiert*: Detaillierte Kennziffer der Stellung im Beruf, in dem die Befragten zuletzt erwerbstätig waren. Für die Generierung wurden Informationen aus den Erwerbstätigkeitsepisoden mit den Angaben zur letzten Erwerbstätigkeit kombiniert, wenn die Befragten seit Januar 2023 nicht erwerbstätig waren (fs).	Bei Ersterfassung: *PET1210; PET1220; PET1230; PET1240; PET1250; PET1260; PET1270 (PENDDAT)* *ET0617; ET0717; ET0817; ET0917; ET1017; ET1117; ET1217 (bio_spells)* Bei wiederholter Erfassung: *stiblewt* aus Vorwelle *(PENDDAT) ET0617; ET0717; ET0817; ET0917; ET1017; ET1117; ET1217 (bio_spells)*
vberuf1	*Höchster ber. Abschluss des Vaters, inkl. Vater im HH, ohne off. Ang., gen.*: Generierung für den höchsten Berufsabschluss des Vaters analog zu *mberuf1* (uv).	Bei Ersterfassung: *PSH0600a-i (PENDDAT)* Nach Ersterfassung: *mberuf1* aus Vorwelle *(PENDDAT)*

Tabelle 14: Einfache Generierungen für Welle 19 im Personendatensatz (*PENDDAT*) (alphabetische Reihenfolge) (Fortsetzung)

vberuf2	*Höchster ber. Abschluss des Vaters, inkl. Vater im HH, mit off. Ang., gen.*: Generierung für den höchsten Berufsabschluss des Vaters (unter Berücksichtigung der offenen Angaben) analog zu *mberuf1* (uv).	Bei Ersterfassung: *PSH0601a-i (PENDDAT)* Nach Ersterfassung: *mberuf1* aus Vorwelle *(PENDDAT)*
vhh	*Steuervariable: Vater lebt im HH*: Indikator, der anzeigt, dass der unspezifische, leibliche, Stief-/Adoptivvater der Befragungsperson im Haushalt lebt (neu).	Informationen zu Beziehungen im Haushalt *(Haushaltsgrid)*
vschul1	*Höchster allg. Schulabschluss d. Vaters, inkl. Vater im HH, o. off. Ang., gen.*: Generierung für den höchsten Schulabschluss des Vaters analog zu *mschul1* (uv).	Bei Ersterfassung: *PSH0500 (PENDDAT)* Nach Ersterfassung: *vschul1* aus Vorwelle *(PENDDAT)*
vschul2	*Höchster allg. Schulabschluss d. Vaters, inkl. Vater im HH, inkl. off. Ang., gen.*: Generierung für den höchsten Schulabschluss des Vaters (unter Berücksichtigung der offenen Angaben) analog zu *mschul2* (uv).	Bei Ersterfassung: *PSH0501 (PENDDAT)* Nach Ersterfassung: *vschul2* aus Vorwelle *(PENDDAT)*
vstib	*Stellung im Beruf des Vaters, Kennziffer, generiert*: Detaillierte Stellung des Vaters im Beruf, generiert aus den Einzelvariablen (uv).	Bei Ersterfassung: *PSH0620; PSH0630; PSH0640; PSH0660; PSH0670; PSH0680 (PENDDAT)* Nach Ersterfassung: *vstib* aus Vorwelle *(PENDDAT)*

Tabelle 15: Einfache Generierungen für Welle 19 im Arbeitslosengeld-II-Datensatz *(alg2_spells)* (Reihenfolge analog zur Reihenfolge im Datensatz)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 19
bmonat	*ALG2-Episode: Anfangsmonat, generiert*: Beginnmonat der ALG-II- Episode / Bürgergeld-Episode, in der für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden *Anmerkung: Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben der 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*	*AL20100 (alg2_spells)*
bjahr	*ALG2-Episode: Anfangsjahr, generiert*: Endjahr der Arbeitslosengeld-II- Episode / Bürgergeld-Episode Anmerkung: siehe *bmonat*	*AL20200 (alg2_spells)*
emonat	*ALG2-Episode: Endmonat, generiert*: Endmonat der Arbeitslosengeld-II- Episode / Bürgergeld-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. *Anmerkung: siehe bmonat*	*AL20300 (alg2_spells)* *hintmon (HHENDDAT)*
ejahr	*ALG2-Episode: Endjahr, generiert*: Endjahr der Arbeitslosengeld-II- Episode / Bürgergeld-Episode, in der bei rechtszensierten (d.h. zum Befragungszeitpunkt des Haushalts aktuell andauernden) Spells das Interviewjahr eingesetzt wurde. *Anmerkung: siehe bmonat*	*AL20400 (alg2_spells)* *hintjahr (HHENDDAT)*

Tabelle 15: Einfache Generierungen für Welle 19 im Arbeitslosengeld-II-Datensatz *(alg2_spells)*

(Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

alg2kbma- alg2kbmi	*ALG2: 1. Kürzung: Anfangsmonat, generiert* bis *ALG2: 9. Kürzung: Anfangsmonat, generiert*: Beginnmonat der Arbeitslosengeld II- Kürzung, bei der für die Generierung die Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden. *Anmerkung: Die Arbeitslosengeld-II- Kürzungen sind in die Arbeitslosengeld-II- Bezugsspells eingebettet. Über den Indikator am Ende der jeweiligen Variablen (a-i) können die Informationen zu den einzelnen Kürzungsepisoden unterschieden werden. Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben des Befragten sind ab der Welle 2 in den Quellvariablen enthalten.*	1. Kürzung: *AL21000a (alg2_spells)* bis 9. Kürzung: *AL21000i (alg2_spells)* (nur bis Welle 11 erhoben)
alg2kbja- alg2kbji	*ALG2: 1. Kürzung: Anfangsjahr, generiert* bis *ALG2: 9. Kürzung: Anfangsjahr, generiert*: Beginnjahr der Arbeitslosengeld II- Kürzung *Anmerkung: siehe alg2kma - alg2kbmi*	1. Kürzung: *AL21100a (alg2_spells)* bis 9. Kürzung: *AL21100i (alg2_spells)* (nur bis Welle 11 erhoben)
alg2kema- alg2kemi	*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 von den Befragten eine Dauer der Kürzung berichtet, wurde aus dieser auf Basis des generierten Beginndatums der Kürzung das jeweilige Enddatum berechnet. *Anmerkung: siehe alg2kma - alg2kbmi*	1. Kürzung: *alg2kbma; alg2kbja; AL21200a; AL21201a; AL21202a (alg2_spells)* bis 9. Kürzung: *alg2kbmi; alg2kbji; AL21200i; AL21201i; AL21202i (alg2_spells)* (nur bis Welle 11 erhoben)
alg2keja- alg2keji	*ALG2: 1. Kürzung: Endjahr, generiert* bis *ALG2: 9. Kürzung: Endjahr, generiert*: Endjahr der Arbeitslosengeld-II- Kürzung. Wurde von den Befragten eine Dauer der Kürzung berichtet, wurde aus dieser auf Basis des generierten Beginndatums der Kürzung das jeweilige Enddatum berechnet. *Anmerkung: siehe alg2kma - alg2kbmi*	1. Kürzung: *alg2kbma; alg2kbja; AL21200a; AL21201a; AL21202a (alg2_spells)* bis 9. Kürzung: *alg2kbmi; alg2kbji; AL21200i; AL21201i; AL21202i (alg2_spells)* (nur bis Welle 11 erhoben)

Tabelle 15: Einfache Generierungen für Welle 19 im Arbeitslosengeld-II-Datensatz *(alg2_spells)*

(Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

AL22150a-AL22150i	*ALG2: 1. Kürzung: Welchem HH-Mitglied wurde gekürzt, generiert* bis *ALG2: 9. Kürzung: Welchem HH-Mitglied wurde gekürzt, generiert*: 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 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, enthält einen der folgenden Codes, die den Kürzungsstatus der jeweiligen Person angeben. Codes: 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 3 erhoben)
zensiert	*ALG2-Episode: Spell dauert z. Zeitpkt. d. letzten HH-Int. an (Rechtszens.), gen.*: Der Zensierungsindikator gibt an, ob ein Spell zum Zeitpunkt des letzten Haushaltsinterviews noch andauerte. *Anmerkung: 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 19 noch andauert. Oder ein Enddatum wird genannt, das identisch mit dem Interviewdatum der Welle 19 ist, und in der Nachfrage wird bestätigt, dass der Bezug aktuell noch andauert.*	*AL20300; AL20400; AL20500 (alg2_spells)*

Tabelle 16: Einfache Generierungen für Welle 19 im BIO-Spelldatensatz *(bio_spells)* (Reihenfolge analog zur Reihenfolge im Datensatz)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 19
bmonat	*Alle Aktivitäten: Anfangsmonat, generiert*: Beginnmonat der Tätigkeitsepisode, bei dem für die Generierung Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden. *Anmerkung: Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben der Befragten sind in den Quellvariablen erhalten.Die Jahreszeitangaben wurden nach folgenden Regeln in Monatswerte recodiert:* *21 Jahresanfang/Winter: Januar* *24 Frühjahr/Ostern: April* *27 Jahresmitte/Sommer: Juli* *30 Herbst: Oktober* *32 Jahresende: Dezember*	*BIO0200 (bio_spells)*
bjahr	*Alle Aktivitäten: Anfangsjahr, generiert*: Beginnjahr der Tätigkeitsepisode *Anmerkung: siehe bmonat*	*BIO0300 (bio_spells)*
emonat	*Alle Aktivitäten: Endmonat, generiert*: Endmonat der Tätigkeitsepisode, wobei für deren Generierung die Jahreszeitangaben in konkrete Monatsangaben umgerechnet wurden und bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells der Interviewmonat eingesetzt wurde. *Anmerkung: siehe bmonat*	*BIO0400, BIO0600 (bio_spells); pintmon (PENDDAT)*
ejahr	*Alle Aktivitäten: Endjahr, generiert*: Endjahr der Tätigkeitsepisode, wobei bei rechtszensierten (d.h. zum Befragungszeitpunkt des Haushalts aktuell andauernden) Spells das Interviewjahr eingesetzt wurde. *Anmerkung: siehe bmonat*	*BIO0500, BIO0600 (bio_spells); pintjahr (PENDDAT)*

Tabelle 16: Einfache Generierungen für Welle 19 im BIO-Spelldatensatz *(bio_spells)* (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

zensiert	*Alle Aktivitäten: gegenwärtig noch andauernder Spell (Rechtszensierung)*: Der Zensierungsindikator gibt an, ob ein Spell zum Zeitpunkt des Personeninterviews in der letzten Befragungswelle noch andauerte, d.h. ob es sich um einen rechtszensierten Spell handelt. *Anmerkung: Ein Spell gilt als zensiert, wenn eine der beiden folgenden Bedingungen erfüllt ist:* *(a) Die Person gibt bezüglich des Enddatums des BIO-Spells an, zum Zeitpunkt des Interviews die Tätigkeit noch auszuüben.* *(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.*	*BIO0400; BIO0500; BIO0600 (bio_spells)*
stib	*Stellung im Beruf, Kennziffer, generiert*: Generierung der detaillierten Kennziffer der Stellung im Beruf aus den Einzelvariablen.	Erfassung von Spell-informationen in Welle 19: *ET0617; ET0717; ET0817; ET0917; ET1017; ET1117; ET1217 (bio_spells)*. Ansonsten bleibt der Wert der Vorwelle bestehen.
az1	*Vertragl. Arbeitszeit, gen.*	Erfassung von Spell-informationen in Welle 19: ET2017 *(bio_spells)*. Ansonsten bleibt der Wert der Vorwelle bestehen. Ausnahme: az1 wird mit Code -3 belegt, wenn eine Tätigkeit bislang eine abhängige Beschäftigung war und die Stellung im Beruf nun in Selbständigkeit/ Mithelfende Familienangehörige, Keine Angabe oder Weiß nicht abgeändert wird.

Tabelle 16: Einfache Generierungen für Welle 19 im BIO-Spelldatensatz *(bio_spells)* (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

az2	*Tatsächliche. Arbeitszeit inkl. Angaben bei unregelmäßiger Arbeitszeit, gen.*: 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 Spell-informationen in Welle 19: *ET2117; ET2217 (bio_spells)*. Ansonsten bleibt der Wert der Vorwelle bestehen.
alg1bm	*ALG1-Bezug: Anfangsmonat, generiert*: Beginnmonat des Arbeitslosengeld-I-Bezugs. Für die Generierung wurden die Jahreszeitangaben in konkrete Monatsangaben umgerechnet. *Anmerkung: Die Arbeitslosengeld-I-Bezüge sind in die Arbeitslosmeldungsspells eingebettet. Pro Arbeitslosmeldung ist maximal ein Arbeitslosengeld-I-Bezugszeitraum vorgesehen. Die generierten Datumsvariablen wurden plausibilitätsgeprüft und ggf. bereinigt. Die originalen Datumsangaben der Befragten sind in den Quellvariablen enthalten. Umrechnung der Monatsangaben: siehe bmonat*.	*AL0800 (bio_spells)*
alg1bj	*ALG1-Bezug: Anfangsjahr, generiert*: Beginnjahr des Arbeitslosengeld-I-Bezugs *Anmerkung: siehe alg1bm*	*AL0900 (bio_spells)*
alg1em	*ALG1-Bezug: Endmonat, generiert*: Endmonat des Arbeitslosengeld-I-Bezugs. Für die Generierung wurden die Jahreszeitangaben in konkrete Monatsangaben umgerechnet und bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells das Interviewdatum eingesetzt. *Anmerkung: siehe alg2kbma - alg2kbmi*	*AL1000; AL1200 (bio_spells); pintmon (PENDDAT)*
alg1ej	*ALG1-Bezug: Endjahr, generiert*: Endjahr des Arbeitslosengeld-I-Bezugs. Bei rechtszensierten (d.h. zum Befragungszeitpunkt der Person aktuell andauernden) Spells wurde das Interviewdatum eingesetzt. *Anmerkung: siehe alg2kbma - alg2kbmi*	*AL1100; AL1200 (bio_spells) pintjahr (PENDDAT)*

Tabelle 16: Einfache Generierungen für Welle 19 im BIO-Spelldatensatz *(bio_spells)* (Reihenfolge analog zur Reihenfolge im Datensatz) (Fortsetzung)

alg1akt	*ALG1-Bezug: gegenwärtig noch andauernder Spell (Rechtszensierung)*: Der Zensierungsindikator gibt an, ob der ALG-I-Bezug zum Zeitpunkt des Personeninterviews in der letzten Welle noch andauerte, d.h. ob es sich um einen rechtszensierten Spell handelt. *Anmerkung: Ein Spell gilt als zensiert, wenn eine der beiden folgenden Bedingungen erfüllt ist:* *(a) Die Person gibt bezüglich des Enddatums des Arbeitslosengeld-I-Bezug-Spells an, zum Zeitpunkt des Interviews noch Arbeitslosengeld I zu erhalten. (b) 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.*	*emonat; ejahr; AL1000; AL1100; AL1200 (bio_spells)*
br	*Bruttoeinkommen (inkl. kat. Ang.), generiert*: Für andauernde Spells aus Welle 19 erfolgt die Berechnung aus den Befragungdaten der aktuellen Welle. Für nicht andauernde oder nicht fortgeschriebene Spells wird in Welle 19 auf die Informationen aus Welle 18 zurückgegriffen.	*ET28*; ET29*; ET30*; ET31*; ET32*; ET33* (bio_spells)*
net	*Nettoeinkommen (inkl. kat. Ang.), generiert*: Für andauernde Spells aus Welle 19 erfolgt die Berechnung aus den Befragungsdaten der aktuellen Welle. Für nicht andauernde oder nicht fortgeschriebene Spells wird in Welle 19 auf die Informationen aus Welle 18 zurückgegriffen.	*ET34*; ET35*; ET36*; ET37*; ET38*; ET39* (bio_spells)*

Tabelle 17: Einfache Generierungen für Welle 19 im Personenregisterdatensatz *(p_register* (alphabetische Reihenfolge)

Variable	Label und Beschreibung	Quellvar. für Gen. in Welle 19
alter19	*Alter der Person in Welle 19 (2025)*: Die Variable enthält die „beste“ verfügbare Altersinformation einer Person. Dabei handelt es sich entweder (a) um das Alter, das anhand des in Welle 19 berichteten Geburtsdatums berechnet wurde, oder (b) um die Altersangabe aus dem Haushaltsinterview, wenn kein Geburtsdatum aus Welle 19 vorliegt. Die Information aus *alter19* wurde auch in den Haushaltsdatensatz übernommen und entspricht den Angaben in *HD0200a* bis *HD0200o*. 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 19 im Haushaltsdatensatz enthaltene „beste“ Altersinformation wurde in den Plausibilitätsprüfungen und bei der Generierung der Bedarfsgemeinschaften und Haushaltstypen berücksichtigt.	*PD0100; pintjahr; pintmon; pinttag (PENDDAT); HD0200a bis HD0200o (HHENDDAT)*
erwprox19	*Erwerbsstatus laut HH-Interview in Welle 19 (2025)*: Die Variable ist eine unveränderte Übernahme der *HD1101** der aktuellen Welle aus *HHENDDAT*.	*HD1101**
kinddat19	*Person im Kinderdatensatz enthalten für Welle 19 (2025)*: Zeigt an, ob die Person in der aktuellen Welle im KINDER-Datensatz enthalten ist. Alle Kinder im Alter von unter 15 Jahren sowie in den Wellen 6 bis 10 zusätzlich 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 *KINDER*-Datensatz enthalten.	*pnr (KINDER)*

Tabelle 17: Einfache Generierungen für Welle 19 im Personenregisterdatensatz *(p_register)*

(alphabetische Reihenfolge) (Fortsetzung)

korrsex	*Angabe zum Geschlecht wurde zwischen Befragungswellen korrigiert*: Die Variable gibt für Personen, die in mehr als einer Welle Mitglied eines Befragungshaushalts waren, an, ob im Rahmen des Haushaltsinterviews das Geschlecht angepasst wurde.	*HD0100a* bis *HD0100o* aller Wellen *(HENDDAT)*
lastint	*Befragungswelle des letzten Interviews auf Personenebene*: Die Variable gibt an, in welcher Welle das letzte Interview auf Personenebene (Personeninterview oder Senioreninterview) mit der Person geführt wurde.	Personeninterviews aller Wellen *PENDDAT*
neuj19	*Jahr, seit dem eine Person neu im aktuellen HH ist, berichtet in Welle 19 (2025)*: Die Variable zeigt an, seit welchem Jahr die Person neu zum Haushalt gehört, dessen Haushaltsmitglied sie in Welle 19 ist. *Anmerkung: Information zum Datum stammt aus dem Welle 19-Interview mit dem wiederbefragten Haushalt, in den die Person seit der Vorwelle zugezogen ist bzw. in dem sie geboren wurde*	Informationen zum Datum, seit dem eine Person neu im Haushalt ist. Erhoben im Haushaltsgrid.
neum19	*Monat, seit dem eine Person neu im aktuellen HH ist, berichtet in Welle 19 (2025)*: Die Variable zeigt an, seit welchem Monat die Person neu zum Haushalt gehört, dessen Haushaltsmitglied sie in Welle 19 ist. *Anmerkung: siehe neuj19*	Informationen zum Datum, seit dem eine Person neu im Haushalt ist. Erhoben im Haushaltsgrid.
wegj19	*Jahr, seit dem eine Person nicht mehr im vorherigen HH ist, berichtet in Welle 19 (2025)*: Die Variable zeigt an, seit welchem Jahr die Person nicht mehr zum Haushalt der Vorwelle gehört. *Anmerkung: Information zum Datum stammt aus dem Welle 19-Interview mit dem Haushalt, in dem die Person in der Vorwelle gelebt hat.*	Informationen zum Datum, seit dem eine Person neu im Haushalt ist. Erhoben im Haushaltsgrid.
wegm19	*Monat, seit dem eine Person nicht mehr im vorherigen HH ist, berichtet in Welle 19 (2025)*: Die Variable zeigt an, seit welchem Jahr die Person nicht mehr zum Haushalt der Vorwelle gehört. *Anmerkung: siehe wegj19*	Informationen zum Datum, seit dem eine Person neu im Haushalt ist. Erhoben im Haushaltsgrid.

Tabelle 17: Einfache Generierungen für Welle 19 im Personenregisterdatensatz *(p_register)*

(alphabetische Reihenfolge) (Fortsetzung)

zdub19	*Zeiger: Personennr. der durch die ZP gedoppelten Person in Welle 19 (2025)*: Zeigt an, dass eine aus einem Ur-HH bekannte Person aktuell in einem Split-HH lebt, ohne dass der Ur-HH den Auszug dieser Person gemeldet hat. Für Matchvorgänge mit dem *p_register* über die Personennummer muss zunächst eine Matchvariable generiert werden, die *zdub** entspricht, sofern diese größer 0 ist und ansonsten der pnr entspricht. *Anmerkung: Eine detaillierte Erläuterung zu den Gründen der Einführung dieser Variable findet sich in Kapitel 5.4.1.2 des Datenreports zu Welle 5 (vgl. Berg et al., FDZ-Datenreport 06/2012).*	Informationen zu allen Haushaltsmitgliedern eines Ur-HH und aller seiner Split-HH im Haushaltsgrid der aktuellen und der Vorwelle.
zmhh19	*Zeiger: Personennr. der Mutter der Zielperson im HH in Welle 19 (2025)*: Identifiziert die Personennummer der Mutter, sofern sie im Haushalt lebt. Als Mütter gelten: leibliche Mütter, Stiefmütter, Adoptivmütter oder Pflegemütter sowie Mütter mit nicht weiter spezifiziertem Status.	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
zparthh19	*Zeiger: Personennr. des Partners der Zielperson im HH in Welle 19 (2025)*: Identifiziert die Personennummer einer im Haushalt lebenden Partnerin oder eines im Haushalt lebenden Partners. Als Partnerin oder Partner gelten: Ehepartnerin/Ehepartner, eingetragene Lebenspartnerin/eingetragener Lebenpartner, unverheiratete Partnerin/unverheirateter Partner und Partnerin/Pertner mit nicht weiter spezifiziertem Status.	Informationen zu Beziehungen im Haushalt (Haushaltsgrid)
zupanel	*Befragungswelle des Zugangs der Person zum Panel*: Die Variable gibt an, in welcher Welle die Person erstmals Mitglied eines Befragungshaushaltes war.	Informationen zu den im Haushalt lebenden Personen aller Wellen (Haushaltsgrid)
zvhh19	*Zeiger: Personennr. des Vaters der Zielperson im HH in Welle 19 (2025)*: Identifiziert die Personennummer des Väters, 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 einfachen und komplexen generierten Variablen auf Personenebene.

Abbildung 3: Übersicht über generierte Variablen auf Personenebene in Welle 19

	PENDDAT						BIO-Spells
	Aktueller Status	Erwerbshistorie		Soziale Herkunft		Minijob	Erwerbs- und Arbeitslosig-keitsbiografie
		letzte ET	erste ET	Mutter	Vater		
Bildung	berabj						
	beruf1			mberuf1	vberuf1		
	beruf2			mberuf2	vberuf2		
	schulabj						
	schul1			mschul1	vschul1		
	schul2			mschul2	vschul2		
Bildungs-klassifikationen	casmin			mcasmin	vcasmin		
	isced97			misced97	visced97		
	bilzeit			mbilzeit	vbilzeit		
Informationen zum aktuellen Status	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
Stellung im Beruf	stib	stiblewt	stibeewt	mstib	vstib		stib
	stibkz						
Datierung der Erwerbstätigkeit			begmeewt			begmminj	bmonat
			begjeewt			begjminj	bjahr
		emonlewt					emonat
		ejhrlewt					ejahr
Datierung der Arbeitslosigkeit							alg1bm
							alg1bj
							alg1em
							alg1ej
Informationen zur Erwerbs-tätigkeit	befrist						
	azhpt1						az1
	azhpt2						az2
	azges1						
	azges2						
Berufliche Tätigkeit	isco88	isco88lewt	isco88eewt	misco88	visco88	isco88minj	isco88
	isco08	isco08lewt	isco08eewt	misco08	visco08	isco08minj	isco08
	kldb1992	kldb1992lewt	kldb1992eewt	mkldb1992	vkldb1992	kldb1992minj	kldb1992
	kldb2010	kldb2010lewt	kldb2010eewt	mkldb2010	vkldb2010	kldb2010minj	kldb2010
Tätig in Branche	branche1					brancheminj1	branche1
	branche2					brancheminj2	branche2

		PENDDAT					BIO-Spells
	Aktueller Status	Erwerbshistorie		Soziale Herkunft		Minijob	Erwerbs- und Arbeitslosig-keitsbiografie
		letzte ET	erste ET	Mutter	Vater		
Einkommen	netges						
	brges						
	netto						
	nettokat						
	brutto						
	bruttokat						
Leistungsbezug	alg1abez						alg1akt
	aktgefbesch						
Haushalts-kontext und Familienstand	hhgr						
	famstand						
	vhh						
	mhh						
	apartner						
	epartner						
	ekind						
	ekin614						
	ekinu15						
	ekinu18						
	ekin1517						
	kindzges						
	kindzihh						
Migrations-hintergrund	ogebland						
	ostaatan						
	ozulanda						
	ozulandb						
	ozulandc						
	ozulandd						
	ozulande						
	ozulandf						
	migration						
Informationen zur Person	gebhalbj						
	palter						
	zpalthh						
	zpsex						
Allgemein	altbefr						
	fb_vers						
	panel						
	pintdat						
	RegP0100						
	RegP0200						
	sample						

4.5 Konstruktvariablen

Unter Konstruktvariablen fallen diejenigen Variablen, für deren Generierung umfangreichere Umcodierungen und/oder Vercodungsarbeiten notwendig sind. In den meisten Fällen handelt es sich um bereits an anderer Stelle erprobte, theoretisch fundierte 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 18: Bildung in Jahren

Variablenname:	*bilzeit*
Variablenlabel:	Schul- und Berufsbildung in Jahren, generiert
Ausgangsvariablen:	*schul2; beruf2*
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 die tatsächliche 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 damit verbundenen – trotz vergleichbarer Dauer – teilweise recht stark divergierenden Einkommenserträge, ist eine einfache Umrechnung der Ausbildungsdauer nicht möglich.

Tabelle 18: Bildung in Jahren (Fortsetzung)

	Dieses Problem kann dadurch umgangen werden, indem 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, Deutschland: +1,5 Jahre Anderer Abschluss, Ausland: +1,5 Jahre
Literatur:	Helberger (1988)

Tabelle 19: Bildung in Jahren, Mutter

Variablenname:	*mbilzeit*
Variablenlabel:	Schul- und Berufsbildung der Mutter in Jahren, generiert
Ausgangsvariablen:	*mschul2; mberuf2*
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 addierten Jahre von denen der 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

Tabelle 19: Bildung in Jahren, Mutter (Fortsetzung)

	Meister: +3 Jahre Berufsakademie: +3 Jahre Fachhochschule: +3 Jahre Universität: +5 Jahre Anderer Abschluss, Deutschland: +1,5 Jahre Anderer Abschluss, Ausland: +1,5 Jahre
Literatur:	Helberger (1988)

Tabelle 20: Bildung in Jahren, Vater

Variablenname:	*vbilzeit*
Variablenlabel:	Schul- und Berufsbildung des Vaters in Jahren, generiert
Ausgangsvariablen:	*vschul2; vberuf2*
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 addierten Jahre von denen der 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, Deutschland: +1,5 Jahre Anderer Abschluss, Ausland: +1,5 Jahre
Literatur:	Helberger (1988)

Tabelle 21: CASMIN

Variablenname:	*casmin*
Variablenlabel:	Bildung klassifiziert nach CASMIN, aktual. Version, generiert
Ausgangsvariablen:	*schul2; beruf2*
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
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 (selbstver ständlich 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 mit positiven Werten, solche mit definierten missing values mit negativen Werten bezeichnet.
Literatur:	Brauns et al. (1999); Granato (2000); König et al. (1987); Lechert et al. (2006)

Schule / Beruf	nicht erhob.	Schüler	n. gest.	TNZ	KA	WN	ohne Abschl.	Sonder-schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. aus. Abschl.
nicht erhob.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-
unplaus. Wert	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Schüler	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
nicht gest.	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
TNZ	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
KA	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
WN	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
ohne Abschl.	-	-	-	-3	-2	-1	1a	1a	1b	2b	2c_gen	2c_gen	1b	1b
Anlernausbild.	-	-	-	-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
Dissert.	-	-	-	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b	3b
And. dt. Abschl.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c
And. aus. Abschl.	-	-	-	-3	-2	-1	1c	1c	1c	2a	2c_voc	2c_voc	1c	1c

Tabelle 22: MCASMIN

Variablenname:	*mcasmin*
Variablenlabel:	Bildung der Mutter, klassifiziert nach CASMIN, aktual. Version, generiert
Ausgangsvariablen:	*mschul2; mberuf2*
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	Generelle Darstellung: vgl. CASMIN Aufgrund der unterschiedlichen Erhebung der Bildungsvariablen für Befragte und deren Eltern weicht das Vercodungsschema von *mcasmin* und *vcasmin* leicht von dem bei *casmin* verwendeten ab. Zu Details vgl. die folgende Abbildung.
Literatur:	Brauns et al. (1999); Granato (2000); König et al. (1987); Lechert et al. (2006)

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

Tabelle 23: VCASMIN

Variablenname:	*vcasmin*
Variablenlabel:	Bildung des Vaters, klassifiziert nach CASMIN, aktual. Version, generiert
Ausgangsvariablen:	*vschul2; vberuf2*
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	Generelle Darstellung: vgl. CASMIN Aufgrund der unterschiedlichen Erhebung der Bildungsvariablen für Befragte und deren Eltern weicht das Vercodungsschema von *mcasmin* und *vcasmin* leicht von dem bei *casmin* verwendeten ab. Zu Details vgl. die folgende Abbildung.
Literatur:	Brauns et al. (1999); Granato (2000); König et al. (1987); Lechert et al. (2006)

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

Tabelle 24: ISCED-97

Variablenname:	*isced97*
Variablenlabel:	Bildung klassifiziert nach isced97, aktual. Version, generiert
Ausgangsvariablen:	*schul2; beruf2*
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)). 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 mit positiven Werten, solche mit definierten missing values mit negativen Werten bezeichnet.
Literatur:	BMBF (2003); OECD (1999)

Schule / Beruf	nicht erhob.	Schüler	nicht gest.	TNZ	KA	WN	ohne Abschl.	Sonder-schule	HS	RS	FHR	Abi	And. dt. Abschl.	And. aus. Abschl.
nicht erhob.	-10	-	-	-	-	-	-	-	-	-	-	-	-	-
unplaus. Wert	-	-	-	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8
Schüler	-	-5	-	-	-	-	-	-	-	-	-	-	-	-
nicht gest.	-	-	-4	-	-	-	-	-	-	-	-	-	-	-
TNZ	-	-	-	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
KA	-	-	-	-3	-2	-2	-2	-2	-2	-2	-2	-2	-2	-2
WN	-	-	-	-3	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1
ohne Abschl.	-	-	-	-3	-2	-1	1	1	2	2	3a	3a	2	2
Anlernausbild.	-	-	-	-3	-2	-1	2	2	2	2	3a	3a	2	2
Lehre	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Berufsfachsch.	-	-	-	-3	-2	-1	3b	3b	3b	3b	4a	4a	3b	3b
Schul. d. 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

Tabelle 25: MISCED-97

Variablenname:	*misced97*
Variablenlabel:	Bildung der Mutter klassifiziert nach *isced97*, aktual. Version, generiert
Ausgangsvariablen:	*mschul2; mberuf2*
Typ / Datensatz:	Bildung / Personendaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	Zum theoretischen Hintergrund und zur Generierung vgl. ISCED-97. Im Gegensatz zur Umsetzung von ISCED-97 bei den Befragten selbst lässt sich ISCED Stufe 6 für die Eltern nicht generieren, da die entsprechenden Abschlüsse (Promotion oder Habilitation) nicht separat erfragt wurden. Aus diesem Grund umfasst die hier vorgenommene Codierung lediglich die ISCED-Stufen 2-5. Zu Details der Vercodung vgl. die folgende Abbildung.
Literatur:	BMBF (2003); OECD (1999)

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

Tabelle 26: VISCED-97

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

Tabelle 26: 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.

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

Literatur:	BMBF (2003); OECD (1999)

Tabelle 27: International Standard Classification of Occupations 1988 (ISCO88)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	Aktuelle *(PENDDAT) - isco88 - ET2500, PET2500*
	Spelldaten *(bio_spells) - isco88 - ET2500*
	erste *(PENDDAT) - isco88eewt - ET2500, PET1280, PET3950*
	letzte *(PENDDAT) - isco88lewt - ET2500, PET1280*
	des Vaters *(PENDDAT) - visco88 - PSH0800*
	der Mutter *(PENDDAT) - misco88 - PSH0700*
	Minijob - *isco88minj - PMJ0900*
Variablenlabel:	aktuelle Ewt.: Intern. Standard Classification of Occupations 88, aktuelle Erwerbstätigkeit, generiert
	Spelldaten *(bio_spells)*: 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

Tabelle 27: Internat. Standard Class. of Occupations 1988 (ISCO88) (Fortsetzung)

	Mutter: ISCO 88 der Mutter, generiert Minijob: ISCO 88, aktueller Minijob, generiert
Typ / Datensatz:	Beruf / Personendaten
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 zur Entscheidung zwischen verschiedenen infrage kommenden Berufscodes genutzt wird, angenommen, dass bei der Tätigkeit keine Leitungsfunktion vorliegt. In diesen Fällen wurde immer der Beruf mit dem niedrigeren Prestige vercodet.
Literatur:	ILO (1990)

Tabelle 28: International Standard Classification of Occupations 2008 (ISCO08)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	aktuelle (PENDDAT) - *isco08 - ET2500, PET2500* Spelldaten *(bio_spells) - kldb - ET2500* erste (PENDDAT) - *isco08ewt - ET2500, PET1280, PET3950* letzte (PENDDAT) - *isco08ewt - ET2500, PET1280* des Vaters (PENDDAT) - *visco08 - PSH0800* der Mutter (PENDDAT) - *misco08 - PSH0700* Minijob - *isco08mini - PMJ0900* Ausbildungsaspiration, Wunschberuf - *isco08berufswunsch - PAA0100* Ausbildungsaspiration, angestrebter Beruf - *isco08angberuf - PAA1000*
Variablenlabel:	aktuelle Ewt.: Intern. Standard Classification of Occupations 08, aktuelle Erwerbstätigkeit, gen.

Tabelle 28: Internat. Standard Class. of Occupations 2008 (ISCO08) (Fortsetzung)

	Spelldaten *(bio_spells)*: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 (ab Welle 17 inkl. Senioren) 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 zur Entscheidung zwischen verschiedenen infrage kommenden Berufscodes genutzt wird, angenommen, dass bei der Tätigkeit keine Leitungsfunktion vorliegt. In diesen Fällen wurde immer der Beruf mit dem niedrigeren Prestige vercodet. Ab Welle 17 werden auch Angaben zum Minijob aus den Seniorenінterviews codiert.

Tabelle 28: Internat. Standard Class. of Occupations 2008 (ISCO08) (Fortsetzung)

	Auch bei der Codierung der Angaben zur Ausbildungsaspiration, die für die Wellen 11 bis 14 vorliegen, 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 29: Klassifizierung der Berufe 1992 (KldB1992)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	Aktuelle - kldb1992 - ET2500, PET2500
	Spelldaten - *(bio_spells) - kldb1992 - ET2500*
	Erste - *kldb1992eewt - ET2500, PET1280, PET3950*
	Letzte - *kldb1992lewt - ET2500, PET1280*
	des Vaters - *vkldb1992 - PSH0800*
	der Mutter - *mkldb1991 - PSH0700*
	Minijob - *kdlb1992mini- PMJ0900*
Variablenlabel	aktuelle Ewt.: Klassifizierung der Berufe 1992, aktuelle Erwerbstätigkeit, gen.
	Spelldaten - *(bio_spells)*: 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

Tabelle 29: Klassifizierung der Berufe 1992 (KldB1992) (Fortsetzung)

Ansprechpartner:	Bernhard Christoph
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 30: Klassifizierung der Berufe 2010 (KldB2010)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	Aktuelle - *kldb2010 - ET2500, PET2500*
	Spelldaten *(bio_spells) - kdlb2010 - ET2500*
	Erste - *kldb2010eewt - ET2500, PET1280, PET3950*
	Letzte - *kldb2010lewt - ET2500, PET1280*
	des Vaters - *vkldb2010 - PSH0800*
	der Mutter - *mkldb2010 - PSH0700*
	Minijob - *kldb2010mini - PMJ0900*
	Ausbildungsaspiration, Wunschberuf - *kldb2010berufswunsch - PAA0100*
	Ausbildungsaspiration, angestrebter Beruf - *kldb2010angberuf - PAA1000*
Variablenlabel	aktuelle Ewt.: Klassifizierung der Berufe *2010*, aktuelle Erwerbstätigkeit, gen.
	Spelldaten *bio_spells)*: 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. (ab Welle 17 inkl. Senioren)
	Ausbildungsaspiration, Wunschberuf: Klassifizierung der Berufe *2010*, Berufsvorstellungen: Berufswunsch, gen.
	Ausbildungsaspiration, angestrebter Beruf: Klassifizierung der Berufe *2010*, Berufsvorstellungen: Angestrebter Beruf, gen.

Tabelle 30: Klassifizierung der Berufe 2010 (KldB2010) (Fortsetzung)

Typ/ Datensatz	Beruf/ Personendaten
Ansprechpartner	Christian Dickmann
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 KldB2010 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. Ab Welle 17 werden auch Angaben zum Minijob aus den Seniorenінterviews codiert. Bei der Codierung der Angaben zur Ausbildungsaspiration, die für die Wellen 11 bis 14 vorliegen, 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 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	Bundesagentur für Arbeit (2011)

Tabelle 31: Klassenschema nach Erikson, Goldthorpe und Portocarrero (EGP)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	Aktuelle - *egp - isco88, stib*
	Spelldaten - *(bio_spells) - egp - isco88, stib*
	Erste - *egpeewt - isco88eewt, stibeewt*
	Letzte - *egplewt - isco88lewt, stiblewt*
	des Vaters - *vegp - visco88, vstib*
	der Mutter - *megp - misco88, mstib*
Variablenlabel:	aktuelle Ewt.: Klassenschema nach Erikson, Goldthorpe & Portocarrero (EGP), akt. Beruf, gen.
	Spelldaten *(bio_spells)*: 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 Portocarero (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.

Tabelle 31: Klassenschema (EGP) (Fortsetzung)

	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 32: European Socio-economic Classification (ESeC)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	Aktuelle - *esec - isco88, stib, PET2000, PET2700* Spelldaten *(bio_spells) - esec -* Erste - *eseceewt - isco88eewt, stibeewt, PET1261* Letzte - *eseclewt - isco88lewt, stiblewt, PET3801* des Vaters - *vesec - visco88, vstib, PSH0670* der Mutter - *mesec - misco88, mstib, PSH0370*
Variablenlabel:	aktuelle Ewt.: European Socio-economic Classification (ESeC), akt. Beruf, gen. Spelldaten *(bio_spells)*: 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

Tabelle 32: European Socio-economic Classification (ESeC) (Fortsetzung)

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 überlicherweise erfragten Informationen zur Stellung im Beruf im Minijobmodul nicht erhoben worden sind.
Literatur:	Fischer & Wirth (2007); Harrison & Rose (2006); Müller et al. (2006, 2007); Rose & Harrison (2007)

Tabelle 33: Magnitude-Prestigeskala (MPS)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	Aktuelle - *mps* - *isco88* Spelldaten *(bio_spells)* - *mps* - *isco88* Erste - *mpseewt* - *isco88eewt* Letzte - *mpslewt* - *isco88lewt* des Vaters - *vmps* - *visco88* der Mutter - *mmps* - *misco88*
Variablenlabel:	aktuelle Ewt.: Magnitude-Prestigeskala, aktuelle Ewt. gen. Spelldaten *(bio_spells)*: 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

Tabelle 33: Magnitude-Prestigeskala (MPS) (Fortsetzung)

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 Zuspielung 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 34: Standard International Occupational Prestige Scale (SIOPS/Treiman-Skala)- Basis ISCO-88

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariable(n)
	Aktuelle - *siops1 - isco88* Spelldaten *(bio_spells) - siops1 - isco88* Erste - *siopseewt1 - isco88eewt* Letzte - *siopslewt1 - isco88eewt* des Vaters - *vsiops1 - visco88* der Mutter - *msiops1 - misco88*
Variablenlabel:	aktuelle Ewt.: Standard International Occupational Prestige Scale (Basis ISCO-88), aktuelle Ewt., gen. Spelldaten *(bio_spells)*: 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

Tabelle 34: Standard Internat. Occ. Prestige Scale (SIOPS/Treiman-Skala) (Fortsetzung)

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 Zuspielung 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 35: Standard International Occupational Prestige Scale (SIOPS/Treiman-Skala) – Basis ISCO-08

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	Aktuelle - *siops2 - isco08*
	Spelldaten *(bio_spells)* - *siops2 - isco08*
	Erste - *siopseewt2 - isco08eewt*
	Letzte - *siopslewt2 - isco08lewt*
	des Vaters - *vsiops2 - visco08*
	der Mutter - *msiops2 - misco08*
Variablenlabel:	aktuelle Ewt.: Standard Intern. Occupational Prestige Scale (Basis ISCO08), aktuelle Ewt., gen.
	Spelldaten *(bio_spells)*: Standard Intern. Occupational Prestige Scale (Basis ISCO08), gen.
	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.

Tabelle 35: Standard Internat. Occ. Prestige Scale (SIOPS/Treiman-Skala) (Fortsetzung)

	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 (erhoben in den Wellen 11 bis 14) 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 36: International Socio-Economic Index (ISEI) – Basis ISCO-88

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	Aktuelle - *isei1 - isco88* Spelldaten *(bio_spells) - isei1 - isco88* Erste - *iseieewt1 - isco88eewt* Letzte - *iseilewt1 - isco88lewt* des Vaters - *visei1 - visco88* der Mutter - *misei1 - misco88*
Variablenlabel:	aktuelle Ewt.: International Socio-Economic Index (Basis ISCO88), aktuelle Ewt., gen. Spelldaten *(bio_spells)*: International Socio-Economic Index (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
Ansprechpartner:	Bernhard Christoph
Erläuterung:	Der International Socio-Economic Index ist 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. Dabei wird der Beruf, bzw. dessen sozioökonomischer Status, als „intervening variable“ zwischen Bildung und Einkommen verstanden.

Tabelle 36: International-Socio Economic Index (ISEI) - Basis ISCO-88

(Fortsetzung)

	Der ISEI wurde zunächst für ISCO-68 entwickelt (Ganzeboom et al. (1992)), später jedoch auch für ISCO-88 umgesetzt (Ganzeboom & Treiman (1996, 2003)). Die Zuspielung 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 37: International Socio-Economic Index (ISEI) – Basis ISCO-08

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	Aktuelle - *isei2 - isco08* Spelldaten *(bio_spells) - isei2 - isco08* Erste - *iseieewt2 - isco08eewt* Letzte - *iseilewt2 - isco08lewt* des Vaters - *visei2 - visco08* der Mutter - *misei2 - misco08*
Variablenlabel:	aktuelle Ewt.: International Socio-Economic Index (Basis ISCO08), aktuelle Ewt., gen. Spelldaten *(bio_spells)*: International Socio-Economic Index (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:	Die Basis des ISEI-08-Indizes sind die Datensätze des International Social Survey Programme (ISSP) der Jahre 2002 bis 2007. Die Zuspielung erfolgte im Rahmen der Berufsvercodung durch infas.

Tabelle 37: 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 (erhoben in den Wellen 11 bis 14) wurde ISEI nicht generiert, da die üblicherweise erfragten Informationen zur Stellung im Beruf in diesen Modulen nicht erhoben worden sind.
Literatur:	Ganzeboom (2010)

Tabelle 38: Klassifikation der Wirtschaftszweige 2003 (WZ2003)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	Aktuelle - *branche1 - ET2600* Spelldaten *(bio_spells) - branche1 - ET2600* Minijob - *brancheminj1 - PMJ1300*
Variablenlabel:	aktuelle Ewt.: Jetzige Tätigkeit: Wirtschaftszweig/Branche, aktuelle Ewt. (WZ2003) Spelldaten *(bio_spells)*: 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 39: Klassifikation der Wirtschaftszweige 2008 (WZ2008)

Gebildet:	Erwerbstätigkeit - Variablenname - Ausgangsvariablen
	Aktuelle - *branche2 - ET2600*
	Spelldaten *(bio_spells) - branche2 - ET2600*
	Minijob - *brancheminj2 - PMJ1300*
Variablenlabel:	aktuelle Ewt.: Jetzige Tätigkeit: Wirtschaftszweig/Branche, aktuelle Ewt. (WZ 2008)
	Spelldaten *(bio_spells)*: 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 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 40: Körperliche Skala des SF12v2 (SOEP-Version, NBS)

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

Tabelle 40: Körperliche Skala des SF12v2 (SOEP-Version, NBS) (Fortsetzung)

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 *pcs* in PASS basiert auf der bei Nübling et al. (2006) wiedergegebenen SPSS- Syntax. Die SF-12-Indikatoren wurden in PASS bislang in den Wellen 3, 6, 9, 12 und 16 erhoben.
Literatur:	Nübling et al. (2006); Andersen et al. (2007)

Tabelle 41: Psychische Skala des SF12v2 (SOEP-Version, NBS)

Variablenname:	*mcs*
Variablenlabel:	psychische Skala des SF12v2 (SOEP-Version, NBS), generiert
Ausgangsvariablen:	*PG1200; PG1205; PG1210; PG1215**
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 *mcs* in PASS basiert auf der bei Nübling et al. (2006) wiedergegebenen SPSS- Syntax. Die SF-12-Indikatoren wurden in PASS bislang in den Wellen 3, 6, 9, 12 und 16 erhoben.
Literatur:	Nübling et al. (2006); Andersen et al. (2007)

4.5.2 Haushalts- oder Bedarfsgemeinschaftsebene

Tabelle 42: Bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung

Variablenname:	*incoecda*
Variablenlabel:	Bedarfsgewichtetes Haushaltseinkommen, alte OECD-Gewichtung (gerundet)
Ausgangsvariablen:	*HD0200a-HD0200o; HA0100; hhincome*
Typ / Datensatz:	sozioökonomische Position / Haushaltsdaten
Bearbeitung:	Bernhard Christoph / Christian Dickmann
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 14 erhalten ein Bedarfsgewicht von 0,7; Personen bis einschl. 13 gehen mit einem Gewicht von 0,5 in die bedarfsgewichtete Haushaltsgröße ein. Die Variable incoecda ersetzt ab Welle 17 (auch rückwirkend) die frühere Variable oecdinca, bei der andere Altersgrenzen für die Äquivalenzgewichtung verwendet wurden (Berg et al., 2023).
Literatur:	OECD (1982)

Tabelle 43: Bedarfsgewichtetes Haushaltseinkommen, neue OECD-Gewichtung

Variablenname:	*incoecdn*
Variablenlabel:	Bedarfsgewichtetes Haushaltseinkommen, neue OECD-Gewichtung (gerundet)
Ausgangsvariablen:	*HD0200a-HD0200o; HA0100; hhincome*
Typ / Datensatz:	sozioökonomische Position / Haushaltsdaten
Bearbeitung:	Bernhard Christoph / Christian Dickmann
Erläuterung:	Generelle Darstellung: 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 14 Jahre erhalten ein Bedarfsgewicht von 0,5; Personen bis einschl. 13 Jahre gehen mit einem Gewicht von 0,3 in die bedarfsgewichtete Haushaltsgröße ein. Die Variable incoecdn ersetzt ab Welle 17 (auch rückwirkend) die frühere Variable oecdincn, die andere Altersgrenzen für die Äquivalenzgewichtung verwendet hat (Berg et al., 2023).
Literatur:	Hagenaars et al. (1994)

Tabelle 44: Deprivationsindex, ungewichtet

Variablenname:	*depindug3*
Variablenlabel:	Alle Wellen: Deprivationsindex, ungewichtet (Itemsumme: 22)
Ausgangsvariablen:	*HLS0100a-HLS0400a; HLS0100b-HLS0400b; HLS0600a-HLS0900a; HLS0600b-HLS0900b; HLS1100a-HLS1200a;* *HLS1100b-HLS1200b; HLS1400a-HLS2500a; HLS1400b-HLS2500b;*
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.

Tabelle 44: Deprivationsindex, ungewichtet (Fortsetzung)

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 22 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.

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 22 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 *depindug* eine Version des ungewichteten Deprivationsindex vor, der auf 26 Items basiert, nämlich zusätzlich zu den oben genannten 22 Items auch auf *HLS0500*, HLS1300** und *HLS2600** (diese drei Items werden seit Welle 5 nicht mehr erhoben) sowie auf *HLS1000**, das seit Welle 13 nicht mehr erhoben wird. Daher wurde *depindug2* in Welle 5 neu in den Datensatz eingefügt und rückwirkend seit Welle 1 generiert.

Tabelle 44: Deprivationsindex, ungewichtet (Fortsetzung)

	Für die Wellen 1 bis 12 liegt zudem mit der Variable *depindug2* eine weitere Version des ungewichteten Deprivationsindex vor, der auf 23 statt 22 Items basiert, nämlich zusätzlich zu den oben genannten Items auch auf *HLS1000**. Dieses Item wird seit Welle 13 nicht mehr erhoben. Daher wurde *depindug3* in Welle 13 neu in den Datensatz eingefügt und rückwirkend seit Welle 1 generiert.
Literatur:	Andreß & Lipsmeier (2001); Halleröd (1995); Nolan & Whelan (1996); Ringen (1988)

Tabelle 45: Deprivationsindex, gewichtet

Variablenname:	*depindg3*
Variablenlabel:	Alle Wellen: Deprivationsindex, gew. (Itemsumme bis W7: 11.03, ab W8: 10.53)
Ausgangsvariablen:	*HLS0100a-HLS0400a; HLS0100b-HLS0400b;* *HLS0600a-HLS0900a; HLS0600b-HLS0900b; HLS1100a-HLS1200a; HLS1100b-HLS1200b;* *HLS1400a-HLS2500a; HLS1400b-HLS2500b;* *PLS0100-PLS0400; PLS0600-PLS0900; PLS1100-PLS1200;* *PLS1400-PLS2500;*
Typ / Datensatz:	materielle Situation / Haushaltsdaten
Bearbeitung:	Bernhard Christoph
Erläuterung:	Für eine generelle Darstellung: vgl. Deprivationsindex, ungewichtet Ein Problem, welches häufig an ungewichteten Indizes, wie dem oben dargestellten k ritisiert 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 Fernseher, 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)).

Tabelle 45: Deprivationsindex, gewichtet (Fortsetzung)

	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.). Für die Wellen 1 bis 4 liegt mit der Variable depindg eine Version des gewichteten Deprivationsindex vor, der auf 26 Items basiert, nämlich zusätzlich zu den oben genannten 22 Items auch auf *HLS0500*, HLS1300** und *HLS2600*. Diese drei Items werden seit Welle 5 nicht mehr erhoben. Und zudem auf Item *HLS1000**, das seit Welle 13 nicht mehr erhoben wird. Daher wurde *depindg2* in Welle 5 neu in den Datensatz eingefügt und rückwirkend seit Welle 1 generiert. Für die Wellen 1 bis 12 liegt zudem mit der Variable *depindg2* eine weitere Version des gewichteten Deprivationsindex vor, der auf 23 statt 22 Items basiert, nämlich zusätzlich zu den oben genannten Items auch auf *HLS1000**. Dieses Item wird seit Welle 13 nicht mehr erhoben. Daher wurde depindg3 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 46: Haushaltstypologie

Variablenname:	*hhtyp*
Variablenlabel:	Haushaltstyp, generiert
Ausgangsvariablen:	Haushaltsinformationen zu Alter und Verwandtschaftsbeziehungen
Typ / Datensatz:	Haushaltsstruktur / Haushaltsdaten
Bearbeitung:	Daniel Gebhardt
Erläuterung:	Für die Typologisierung von Haushalten existiert eine Vielzahl von Varianten und Vorschlägen (vgl. z.B. Lengerer et al. (2005) für die Haushaltstypologie des Mikrozensus, Porst (1984) sowie Beckmann & Trometer (1991) für die Typologie des ALLBUS und Frick et al. (o.J.) für das SOEP). Die Haushaltstypologie des PASS orientiert sich an der letztgenannten Typologie. Die entscheidenden Differenzierungskriterien bilden dabei Partnerschaften sowie Anzahl und Alter der Kinder sowie das Vorhandensein von Generationenbeziehungen. Während im SOEP jedoch lediglich die Stellung aller Haushaltsmitglieder zum Haushaltsvorstand die Basis für die Generierung des Haushaltstyps bildet, werden im PASS die Angaben zu den Beziehungen zwischen allen Haushaltsmitgliedern untereinander für die Generierung verwendet. Daneben werden das Alter der Haushaltsmitglieder laut Haushaltsinterview und die Haushaltsgröße bei der Generierung mit einbezogen. Definition der Beziehungen bei d. Generierung d. Haushaltstyps: • Paare: Ehepaare; eingetragene Lebenspartnerschaften; unverheiratete Partnerschaften sowie Partnerschaften mit nicht weiter spezifiziertem Status (fehlender Wert in Nachfrage nach Art der Partnerschaft). • Kind einer Person: leibliches Kind; Stief-, Adoptiv- oder Pflegekind; Kind mit nicht weiter spezifiziertem Status (fehlender Wert in Nachfrage nach Art der Beziehung zum Kind). • Elternteil einer Person: leibliches Elternteil; Stief-, Adoptiv- und Pflegeelternteil; Eltern mit nicht weiter spezifiziertem Status (fehlender Wert in Nachfrage nach Art der Elternschaft).

Tabelle 46: Haushaltstypologie (Fortsetzung)

	Definition der Haushaltstypen: • 1-Personen Haushalt: Haushalt mit nur einer Person im Haushalt. • Paar ohne Kinder: Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen. • 1-Eltern-Haushalt: Haushalt besteht ausschließlich aus einem Elternteil und dessen Kindern. Hinsichtlich des Alters der Kinder werden keine Einschränkungen vorgenommen. • Paar mit Kindern $<$ 16 Jahre: Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen und deren jeweiligen bzw. gemeinsamen Kindern. Alle Kinder sind unter 16 Jahre alt. • Paar mit Kindern $>=$ 16 Jahre: Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen und deren jeweiligen bzw. gemeinsamen Kindern. Alle Kinder sind 16 Jahre oder älter. • Paar mit Kindern $<$ 16 Jahre und $>=$ 16 Jahre: Haushalt besteht ausschließlich aus zwei in Partnerschaft lebenden Personen und deren jeweiligen bzw. gemeinsamen Kindern. Es leben sowohl Kinder unter 16 Jahre als auch im Alter von 16 Jahren oder älter im Haushalt. • Mehrgenerationenhaushalt: Haushalt besteht aus Angehörigen mindestens dreier linear aufeinanderfolgender Generationen. Der Haushalt besitzt einen Mehrgenerationenkern, d.h. mindestens eine Person im Haushalt ist sowohl Kind als auch Elternteil einer anderen Person im Haushalt. Ausgehend von den zu diesem Mehrgenerationenkern gehörenden Personen leben im Haushalt ausschließlich Eltern, Kinder, Geschwister, Partnerinnen und Partner sowie Geschwister der Partnerinnen und Partner. • Anderer Haushaltstyp: Haushalt, der keinem der anderen definierten Haushaltstypen zugeordnet werden konnte. • Generierung nicht möglich (fehlende Werte): Grundsätzlich alle Haushalte mit mindestens einem fehlenden (-1, -2, -4) oder unplausiblen Wert (-8) in der Oberkategorie in einer Beziehungsvariable oder bei der Altersangabe (Ausnahme: In Haushalten mit drei und weniger Mitgliedern wurde in eindeutigen Konstellationen auch bei fehlenden Altersangaben der Haushaltstyp gebildet).
Literatur	Beckmann & Trometer (1991); Frick et al. (o.J.); Lengerer et al. (2005); Porst (1984)

Tabelle 47: Bedarfsgemeinschaftsnummer, Welle 19

Variablenname:	*bgnr19*
Variablenlabel:	Bedarfsgemeinschaftsnummer in Welle 19 (2025)
Ausgangsvariablen:	Haushaltsinformationen zu Alter und Verwandtschaftsbeziehungen
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Gerrit Müller
Erläuterung:	Die Variable *bgnr19* 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 *bgnr19* setzt sich aus der bekannten Haushaltsnummer sowie einem zweistelligen Bedarfsgemeinschafts-Indikator zusammen. 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.

Tabelle 47: Bedarfsgemeinschaftsnummer, Welle 19 (Fortsetzung)

	Inhaltlich erfolgt die Zuordnung von Personen zu Bedarfsgemeinschaften auf Basis der aktuell gültigen Fassung des §7 Abs. 3 SGB II. Danach bilden grundsätzlich alle Personen, die das 25. Lebensjahr vollendet die Altersgrenze nach §7a noch nicht erreicht haben, jeweils eine eigene Bedarfsgemeinschaft; es sei denn, diese Personen leben in einer Partnerschaft und/oder haben (ein) Kind(-er), das/die wiederum selbst das 25. Lebensjahr noch nicht vollendet hat/haben und ohne Partnerin oder Partner bzw. eigene Kinder sind. In letzterem Falle bilden Personen jeweils gemeinsam mit ihrer Partnerin oder ihrem Partner und Kind(ern) eine Bedarfsgemeinschaft. Sollten zwei Personen mit einem gemeinsamen Kind im gleichen Haushalt leben, aber im HH-Grid nicht angeben, Partnerin und Partner voneinander zu sein, wird dies im Sinne des §7 Abs. (3a) trotzdem unterstellt und die entsprechenden Personen gemeinsam mit dem/den 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 einer Partnerin oder einem Partner (bzw. einem eigenen Kind) gemeinsam im Haushalt. Sollten Personen zwischen 15 und 25 Jahre ohne Eltern (oder Partner/Partnerin bzw. Kind) leben, so bilden sie jeweils eine eigene Bedarfsgemeinschaft. Personen, die die Altersgrenze nach §7a erreicht haben, fallen nicht in den Geltungsbereich des SGB II und gehören daher keiner Bedarfsgemeinschaft an (Code 0); es sei denn, sie leben gemeinsam mit einer Partnerin oder einem Partner, die oder der die Altersgrenze nach §7a (bzw. einem Kind, welches das 25. Lebensjahr) noch nicht erreicht 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 48: Bedarfsgemeinschaftstypologie, Welle 19

Variablenname:	*bgtyp19*
Variablenlabel:	Typ der Bedarfsgemeinschaft in Welle 19 (2025)
Ausgangsvariablen:	Haushaltsinformationen zu Alter und Verwandtschaftsbeziehungen

Tabelle 48: Bedarfsgemeinschaftstypologie, Welle 19 (Fortsetzung)

Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Gerrit Müller
Erläuterung:	Der Bedarfsgemeinschaftstypologie liegt das gleiche Konzept der synthetischen Bedarfsgemeinschaft wie bei der Variable *bgnr19* zugrunde. Kinder gehören bis zur Vollendung des 25. Lebensjahres der Bedarfsgemeinschaft ihrer Eltern an, sofern sie nicht selbst Partnerin oder 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 *bgnr19*). Code -5, Generierung nicht möglich (fehlende Werte), wurde für Haushalte mit fehlenden Beziehungs- bzw. Altersangaben zu einzelnen Personen (vgl. Code 99 bei *bgnr19*) vergeben.
Literatur:	-

Tabelle 49: Bürgergeld-Bezug der Bedarfsgemeinschaft zum Zeitpunkt der Stichprobenziehung, Welle 19

Variablenname:	*bgbezs19*
Variablenlabel:	Bürgergeld-Bezug der Bedarfsgemeinschaft bei Stichprobenziehung in Welle 19 (2025)
Ausgangsvariablen:	*HA0250*, HA0300, AL20100, AL20200, AL20300, AL20400,* *AL20618, AL20718*, HA0400, sample, hnr, bgnr19, hhgr*
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Mark Trappmann
Erläuterung:	Diese Variable gibt für jede Bedarfsgemeinschaft, die entsprechend dem bei Variable *bgnr19* beschriebenen Vorgehen gebildet wurde, an, ob diese Bedarfsgemeinschaft zum Zeitpunkt der Stichprobenziehung im Juli des Vorjahres für Welle 19 Bürgergeld bezogen hat oder nicht.
Literatur:	-

Tabelle 50: Bürgergeld-Bezug der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 19

Variablenname:	*bgbezb19*
Variablenlabel:	Bürgergeld-Bezug d. Bedarfsgemeinschaft zum Befragungszeitpkt. in Welle 19 (2025)
Ausgangsvariablen:	*AL20618, AL20718*, zensiert (alg2_spells), sample, hhgr, bgnr19*
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Daniel Gebhardt
Erläuterung:	Diese Variable gibt für jede Bedarfsgemeinschaft, die entsprechend dem bei Variable *bgnr19* beschriebenen Vorgehen gebildet wurde, an, ob diese Bedarfsgemeinschaft zum Befragungszeitpunkt in Welle 19 (*pinttag, pintmonat, pintjahr*) Bürgergeld bezogen hat oder nicht.
Literatur:	-

Tabelle 51: Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 10

Variablenname:	*bgbezbkorr10*
Variablenlabel:	Korrektur des ALG2-Bezug d. BG zum Befragungszeitpkt. in W10 (2016)
Ausgangsvariablen:	*hintmon, hintjahr, kennungfbversW11 (HHENDDAT), AL20100-AL20400, AL20610, AL20710*, zensiert (alg2_spells), sample, hhgr, bgnr11, bgnr12, bgnr13, bgbezb10*
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Christian Dickmann
Erläuterung:	Siehe Abschnitt unten
Literatur:	-

Tabelle 52: Flag zur Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 10

Variablenname:	*bgbezbkorrflag10*
Variablenlabel:	Flag z. Korrektur des ALG2-Bezug d. BG zum Befragungszeitpkt. in W10 (2016)
Ausgangsvariablen:	*kennungfbversW11 (HHENDDAT), sample, bgbezb10, bgbezbkorr10*
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister

Tabelle 52: Flag zur Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 10 (Fortsetzung)

Bearbeitung:	Christian Dickmann
Erläuterung:	Siehe Abschnitt unten
Literatur:	-

Tabelle 53: Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 11

Variablenname:	*bgbezbkorr11*
Variablenlabel:	Korrektur des ALG2-Bezugs d. BG zum Befragungszeitpkt. in W11 (2017)
Ausgangsvariablen:	*hintmon, hintjahr, kennungfbversW11 (HHENDDAT), AL20100-AL20400, AL20610, AL20710*, zensiert (alg2_spells), sample, hhgr, bgnr12, bgnr13, bgbezb11*
Typ / Datensatz:	Bedarfsgemeinschaft / Personenregister
Bearbeitung:	Christian Dickmann
Erläuterung:	Siehe Abschnitt unten
Literatur:	-

Tabelle 54: Flag zur Korrektur des Arbeitslosengeld-II-Bezugs der Bedarfsgemeinschaft zum Befragungszeitpunkt, Welle 11

Variablenname:	*bgbezbkorrflag11*
Variablenlabel:	Flag zur Korrektur des ALG2-Bezugs d. BG zum Befragungszeitpkt. in W11 (2017)
Ausgangsvariablen:	*kennungfbversW11 (HHENDDAT), sample, bgbezb11, bgbezbkorr11*
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 Bevölkerungsstichproben, *sample*=2, *sample*=6 und *sample*=15), 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); Bruckmeier et al. (2018); 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 SGB-II-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 SGB-II-Zugangsstichproben) Änderungen im ALG-II-Modul vorgenommen. Die Änderungen betrafen ausschließlich die Zugangs-Teilstichprobe der syrischen und irakischen Haushalte der Wellen 10 und 11 (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. Version 1 beinhaltet das Frageprogramm vor der Änderung (während der laufenden Feldzeit), Version 2 das Frageprogramm danach. Anhand der Variable kennungfbversW11 im Haushaltsenddatensatz (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-Spelldatensatz (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 die Nutzerinnen und 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 55: Anzahl der Bedarfsgemeinschaften im Haushalt

Variablenname:	*anzbg*
Variablenlabel:	Anzahl der synthetischen Bedarfsgemeinschaften im HH, generiert
Ausgangsvariablen:	*bgnr19, hnr*
Typ / Datensatz:	Bedarfsgemeinschaft / Haushaltsdatensatz
Bearbeitung:	Daniel Gebhardt
Erläuterung:	Diese Variable gibt die Anzahl der Bedarfsgemeinschaften an, die innerhalb des Haushalts existieren. Die Identifizierung der Bedarfsgemeinschaften erfolgte dabei entsprechend der bei den Erläuterungen zur Variable *bgnr19* beschriebenen Vorgehensweise.
Literatur:	-

Tabelle 56: Anzahl der zum Stichprobenzeitpunkt Leistungen beziehenden Bedarfsgemeinschaften im Haushalt

Variablenname:	*nbgbezug*
Variablenlabel:	Anzahl zum Stichprobenzeitpunkt beziehender BGs im HH
Ausgangsvariablen:	*bgbezs19, bgnr19, hnr*
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 *bgbezs19* 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 19 erhobenen Informationen liegen bei infas zunächst als ASCII-Daten vor. In einer ersten Stufe erstellt infas aus diesen Rohdaten die folgenden Datensätze[38]:

- 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/Bürgergeld“)
- Datensatz über die Fortschreibung der Haushaltszusammensetzung (Matrix)
- Datensatz über die Fortschreibung der Verwandtschaftsbeziehungen im Haushalt (Beziehungsmatrix)
- Personen-/Seniorendatensatz für im Querschnitt erhobene Fragen
- Personendatensatz für im Längsschnitt erhobene Daten (Modul „Erwerbsbiografie“)
- 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.

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 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

[38] Im Rahmen der Datenaufbereitung kamen die Softwarepakete Stata Version 16 sowie SPSS Version 26 zum Einsatz.

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 57: Übersicht über die Schritte der Datenaufbereitung in der Welle 19 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 der Personen- und Seniorendatensätze
5	Korrektur der Haushaltsstruktur bei wiederbefragten Haushalten (siehe Kapitel 5.1)
6	Filterbereinigung auf 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 auf 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 und 5.7 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 alle 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 19 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 Panelhaushalt 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.

- Abgleich der Vornamen (Alter und Geschlecht): 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.

- Abgleich der Geburtsdaten: Weiterhin wurde geprüft, ob mehrere Personen mit demselben Geburtsdatum im Haushalt leben. Im Haushaltskontext der aktuellen Welle und der Vorwelle wurde hier jeweils entschieden, ob es sich um plausible oder unplausible Fälle handelte. Anschließend wurden die übrigen Fälle einer weiteren Prüfung unterzogen. Dabei wurden Haushalte identifiziert, bei denen ein Geburtsdatum in der aktuellen Welle und in der Vorwelle von Personen auf unterschiedlichen Positionen der Haushaltsstruktur berichtet wurde. Hier lag die Vermutung nahe, dass das entsprechende Personeninterview in der aktuellen Welle von einer anderen Person als in der Vorwelle geführt wurde. Im Kontext der Haushalts- und Personeninformationen der beiden Wellen wurden jeweils Einzelfallentscheidungen in Bezug auf die jeweiligen Haushalts- und Personeninterviews getroffen.
- Prüfung der Altersangabe im Personeninterview (PD0100): Generell gilt, dass das Geburtsdatum aus dem Personen-/Seniorenінterview 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.
- Zusammenspielen von Haushalts- und Personeninterviews: 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 vorlag[39].
- Plausible Ein- und Auszüge: Breiten Raum nehmen auch die Prüfungen von Ein- und

[39] 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.

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 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 eine Partnerin oder 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 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 lebt zumindest eine Person dennoch im neuen Splithaushalt, obwohl sie laut Panelhaushalt als nnicht ausgezogen"bzw. als "verzogen"(und zwar in einen anderen Splithaushalt als den betrachteten) markiert 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.

- Splithaushalte und Panelteilnehmende: In Splithaushalten wird bei Personen, die laut Feld neu über den Splithaushalt 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. Werden jedoch tatsächlich Personen in einem gemeinsamen Splithaushaushalt angetroffen, die laut Panelhaushalt in verschiedenen Splithaushalten leben, haben jene Personen, die durch den Panelhaushalt einem anderen Splithaushalt zugerechnet wurden, in diesem Splithaushalt keinen Preload. Sie werden dann nämlich 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-/Seniorenінterviews als neubefragt gesteuert werden, obwohl sie u.U. altbefragte Personen sind. Diese Abweichung wird grundsätzlich nicht korrigiert, siehe hierzu auch Kapitel 4.4.

- Wiedereinzug von Panelteilnehmenden im Zeitverlauf: 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 sowohl ein nachträgliches Ändern der Personennummer der zu verschiebenden Person als auch ein Verschieben der personenbezogenen Angaben im Haushaltsinterview z. B. zur Kinderbetreuung 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.

- Personenbewegungen zwischen Panel- und Splithaushalten (Netzwerke): 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 aus einem Splithaushalt wieder in einen Panelhaushalt zurückziehen oder Personen, die über einen Splithaushalt neu zum PASS hinzugekommen sind, in einen anderen Panelhaushalt umziehen. Es ist auch 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 im Längsschnitt: 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 vier Personen laut Befragung 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.
- Dopplung von Haushalten: In sehr seltenen Fällen werden durch die Rückmeldung der Feldinterviewer Haushalte entdeckt, die als identische Haushalte doppelt in der

Panelstichprobe enthalten sind. Dies betraf in Welle 4 zwei Haushalte, wobei der später befragte Haushalt (mit höherer hnr Nummer) für die Folgewellen aus der Stichprobe gelöscht wurde. In Welle 13 wurden zwei Dubletten und in Welle 17 vier Dubletten entdeckt. Auch hier wurden jeweils die später befragten Haushalte ab der Folgewelle aus der Panelstichprobe gelöscht. Eine rückwirkende Entfernung der Dubletten in den Vorwellen findet nicht statt, da dies beispielsweise Auswirkungen auf die Gewichtung hätte. Im hh_register sind Dublettenhaushalte in der Welle, in der die Dopplung entdeckt wird, in der Variable hnettod* 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 Welle, in der die Dopplung entdeckt wird, in der Variable pnettod* entsprechend mit Code 56 „HH nicht realisiert, Dublette“ gekennzeichnet.

- Fusion von Haushalten: Erstmals sind in Welle 15 zwei bislang voneinander unabhängige Haushalte zusammengezogen. Die beiden Haushalte standen bislang in keiner Beziehung zueinander. Nun zog die einzige Person des einen Haushalts in den anderen Haushalt ein. Sie übernimmt ihre Personennummer pnr und ihre Ursprungs-Haushaltsnummer uhnr in ihren neuen Haushalt.

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 19 unklar war, wurden alle Interviews der Welle 19 entfernt. Sie werden in der Welle 20 wie Haushalte behandelt, die in der Welle 19 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 19 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 beispielsweise 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. Im Falle des Alters wurde sie auf den Vorwellenwert plus die Zahl der Jahre seit der letzten gültigen Realisierung in diesem Haushalt korrigiert.

Darüber hinaus wurden alle Interviews von Personen entfernt, für deren Haushalt kein

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

Einen Hinweis auf entfernte Interviews geben die Nettovariablen *(hnettok19, hnettod19, pnettok19, pnettod19)* 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 Auffrischungsstichprobe der Welle 19, 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 und vergebene Missing Codes

In der Filterbereinigung wurde die in den jeweiligen Instrumenten vorgesehene Filterführung mithilfe eines Statistikprogramms nachvollzogen. Wurden bestimmte Fragen gestellt, obwohl der Wert der relevanten Filtervariable etwas anderes verlangt hätte (wurden also z. B. detaillierte Angaben zur Berufsausbildung erhoben, obwohl die/der Befragte angegeben hatte, über keinen Berufsabschluss zu verfügen), wurden die Variablen auf den Missing Code „-3“ (trifft nicht zu) gesetzt, den sie auch bei einer korrekten Ausführung der Filterführung erhalten hätten[40]. Des Weiteren kam es vor, dass in Einzelfällen Items nicht erhoben wurden, obwohl dies laut der entsprechenden Filtervariable notwendig gewesen wäre (z. B. wenn keine weiteren Angaben zur Berufsausbildung erhoben wurden, obwohl die/der Befragte angegeben hatte, eine solche Ausbildung gemacht zu haben). In diesen Fällen wurde der spezifische Missing Code „4“ (Frage irrtümlich nicht gestellt) vergeben. Die Vergabe des Codes „-4“ kann auch Folge der in Kapitel 5.1 beschriebenen Haushaltsstrukturprüfungen sein. Wenn der Auszug einer Person nachträglich als unplausibel verworfen und die Person nachträglich als noch zu ihrem bisherigen Haushalt gehörend gewertet wird, dann bedeutet dies auch, dass personenbezogene Angaben zu dieser Person im Haushaltsinterview nachträglich als irrtümlich nicht erhoben codiert werden müssen. Der Code „-4“ weist also nicht in allen

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

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 „trifft nicht zu“-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.

Tabelle 58: 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 [41]

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

[41] 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.

Welle 2 vergeben. 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 einer Welle in keiner Fragebogenversion mehr erfragt wurden, für die Beobachtungen ab dieser Welle den Wert „-9“. Ebenso verhält es sich mit erstmals ab einer bestimmten Welle erhobenen Variablen. Sie werden für Beobachtungen der vorherigen 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 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 19 erhobenen Daten für die Prüfung herangezogen. Eine Prüfung im Längsschnitt, also zwischen den in der aktuellen und der Vorwelle gemachten Angaben, wurde nicht durchgeführt.

Im Einzelnen wurden die folgenden Schritte ausgeführt:

- Widerspruchsprüfung: Generell wurden Widersprüche nur dann korrigiert, wenn 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 (z. B. „-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.
- Korrektur unplausibler Angaben bei Fehleingabe: 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.

- Vergabe des Missing Codes „-8“: 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 Informationen 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.
- Abgleich Beziehungsinformation und demografische Angaben: 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 zwei 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.
- Abgleich verwandtschaftlicher Beziehungen innerhalb des Haushalts: 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 Ehepartnerin oder Ehepartner von Person B. Person A ist leibliches Elternteil von Person C. Person C hat mit Person B eine geschwisterliche Beziehung. 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.

- Plausibilitätsprüfung der Spelldatensätze: Auch die Spelldatensätze wurden einer Reihe von Plausibilitätsprüfungen unterzogen, welche in Kapitel 5.6 und 5.7 vertieft dargestellt werden.

5.4 Rückwirkende Veränderungen der Wellen 1 bis 18

Im Rahmen der Datenaufbereitung für den Scientific Use File (SUF) der Welle 19 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 18 festgestellt wurden. Die korrigierten Daten können stets in den SUF-Datensätzen der aktuellen Welle, also derzeit Welle 19, genutzt werden. Die folgenden Tabellen geben einen Überblick über die rückwirkenden Veränderungen an bereits ausgelieferten Wellen des PASS[42].

Tabelle 59: Übersicht über rückwirkende Veränderungen im Haushaltsdatensatz *(HHENDDAT, KINDER)*

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

Tabelle 60: Übersicht über rückwirkende Veränderungen im Personendatensatz *(PENDDAT)*

geänderte Variable	betroffener Datensatz	geänderte Welle	Art der Änderung	Beschreibung der Änderung
netges	*PENDDAT*	11-17	Korrektur	Im Rahmen einer in Welle 18 vorgenommenen Korrektur wur-den 28 Fälle aus den Wellen 11 bis 17 auf Code -4 umgesetzt statt auf den korrekten Code -5.
PMI3000, PMI3001	*PENDDAT*	17, 18	Korrektur	Wegen eines Übersetzungsfehlers in der ukrainischen Fassung des Erhebungsinstruments wurden in den Variablen PMI3000 und PMI3001 die Angaben von 162 Fällen aus den Wellen 17 und 18 auf Code -4 umgesetzt.

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

Tabelle 61: Übersicht über rückwirkende Veränderungen in Spelldatensätzen *(bio_spells, alg2_spells, ee_spells)*

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

Tabelle 62: Übersicht über rückwirkende Veränderungen in Registerdatensätzen *(hh_register; p_register)*

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

Tabelle 63: Übersicht über rückwirkende 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 ist es notwendig, einige Variablen nur in vergröberter Form in den Scientific Use File aufzunehmen. Diese Variablen sind im Variablenlabel generell mit dem Zusatz „anonymisiert“ gekennzeichnet. Aus dem gleichen Grund wird auf verfügbare Regionalinformationen mit Ausnahme der Bundesländer und einer davon abgeleiteten Ost-/West-Information verzichtet. Auch die Angaben zu den Verwandtschaftsbeziehungen im Haushalt sowie die Vornamen der Haushaltsmitglieder sind aus Datenschutzgründen nicht Teil des Scientific Use Files. Hinweise auf die Haushaltsstruktur geben allerdings generierte Variablen, bspw. zum Haushalts- und Bedarfsgemeinschaftstyp (*hhtyp*[43], *bgtyp*[44]), Indikatorvariablen zu Partnerinnen und Partnern im Haushalt (*apartner; epartner*[45]), Zeigervariablen auf Eltern und Partnerinnen und Partner im Haushalt (*zmhh; zvhh; zparthh*[46]) sowie verschiedene Indikatorvariablen, die anzeigen ob Eltern (*mhh; vhh*[47]) oder Kinder der Zielperson (z. B. *ekind*[48]) im Haushalt leben. Nachfolgende Tabelle gibt einen Überblick über die betroffenen Variablen und die vorgenommenen Anonymisierungsschritte[49] im Personendatensatz. In den beiden darauf folgenden Tabellen sind die anonymisierten Variablen des Erwerbstätigkeitsspelldatensatzes und des KINDER-Datensatzes dargestellt.

[43] Enthalten im Haushaltsdatensatz *(HHENDDAT)*, siehe Kapitel 4.5.2.
[44] Wellenspezifische Variablen enthalten im Personenregister *(p_register)*, siehe Kapitel 4.4.
[45] Enthalten im Personendatensatz *(PENDDAT)*, siehe Kapitel 4.4.
[46] Wellenspezifische Variablen enthalten im Personenregister *(p_register)*, siehe Kapitel 4.4.
[47] Enthalten im Personendatensatz *(PENDDAT)*, siehe Kapitel 4.4.
[48] Enthalten im Personendatensatz *(PENDDAT)*, siehe Kapitel 4.4.
[49] Falls für Ihr Forschungsvorhaben die Verwendung der nicht anonymisierten Versionen einer oder 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 64: Übersicht über die in Welle 19 anonymisierten Variablen im Personendatensatz *(PENDDAT)*

Varname	Variablenlabel	Vorgehen
PD0100	Geburtsdatum (anon.)	Das genaue Geburtsdatum wurde zum Geburtsjahr verkürzt.
gebhalbj	Geburtshalbjahr, generiert	Das genaue Geburtsdatum wurde zu einem Indikator für das 1. bzw. 2. Halbjahr verkürzt.
PET1210	Letzte Stellung im Beruf, einfache Einteilung (ohne Spellinfo) (anon.)	In der Befragung wurden aus erhebungstechnischen Gründen Berufs und Zeitsoldatinnen und -soldaten 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 Beamtinnen und Beamten und Richterinnen und Richtern zu einer gemeinsamen Gruppe zusammengeführt.
PET1250	Letzte StiB Beamter: Det. Erfassung, inkl. Soldaten (ohne Spellinfo) (anon.)	Diese Variable enthält zusätzliche Fälle. Die Berufs- und Zeitsoldatinnen und -soldaten aus *PET1240* wurden der entsprechenden Beamtenkategorie zugeschlagen. Die Variable für Berufs- und Zeitsoldatinnen und -soldaten *PET1240* wird nicht ausgeliefert.
PET1211	Letzte Stellung im Beruf, einfache Einteilung (inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu *PET1210*.
PET1251	Letzte StiB Beamter: Det. Erf., inkl. Soldaten (inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu *PET1250*. Die Variable für Berufs- und Zeitsoldatinnen und -soldaten *PET1240* wird nicht ausgeliefert.
stiblewt	Stellung im Beruf, letzte Erwerbstätigkeit, Kennziffer, generiert	Bei der Generierung der Stellung im Beruf werden die Berufs- und Zeitsoldatinnen und -soldaten der entsprechenden Beamtenkategorie zugeordnet.
PET1510	Jetzige Stellung im Beruf, einf. Eint., Erh. ab W2 (anon.)	Vorgehen analog zu *PET1210*.
PET1900	Jetzige StiB Beamter: Det. Erfassung, inkl. Soldaten (anon.)	Vorgehen analog zu *PET1250*. Die in den Seniorenінterviews erhobene Variable für Berufs- und Zeitsoldatinnen und -soldaten *PET1800* wird nicht ausgeliefert. Für die Personeninterviews wird aus den Erwerbstätigkeitsspells *ET09** keine generierte Variable für Berufs- und Zeitsoldaten in den Personendatensatz übernommen.

Tabelle 64: Übersicht über die in Welle 19 anonymisierten Variablen im Personendatensatz *(PENDDAT)*

(Fortsetzung)

stibkz	Jetzige Stellung im Beruf, einf. Eint., harmonisiert (anon.)	Bei der Generierung der Stellung im Beruf werden die Berufs- und Zeitsoldatinnen und -soldaten der entsprechenden Beamtenkategorie zugeordnet.
stib	Stellung im Beruf, Kennziffer, aktuelle Erwerbstätigkeit, generiert	Vorgehen analog zu *stiblewt*.
PET3300	Erste Stellung im Beruf, einfache Einteilung (anon.)	Vorgehen analog zu *PET1210*.
PET3700	Erste StiB Beamter: Det. Erfassung, inkl. Soldaten	Vorgehen analog zu *PET1250*. Die Variable für Berufs- und Zeitsoldatinnen und -soldaten *PET3600* wird nicht ausgeliefert.
PET3301	Erste Stellung im Beruf, einf. Eint.(zus.gef., inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu *PET1210*.
PET3701	Erste StiB Beamter: Det. Erf., inkl. Soldaten, (zus.gef., inkl. Spellinfo) (anon.), gen.	Vorgehen analog zu *PET1250.* Die Variable für Berufs- und Zeitsoldatinnen und -soldaten *PET3600* wird nicht ausgeliefert.
stibeewt	Stellung im Beruf, erste Erwerbstätigkeit, Kennziffer, generiert	Vorgehen analog zu *stiblewt*.
PSH0320	Damalige Stellung im Beruf der Mutter, einfache Einteilung (anon.)	Vorgehen analog zu *PET1210*.
PSH0360	Damalige StiB d. Mutter Beamte, inkl. Soldaten: Det. Erfassung (anon.)	Vorgehen analog zu *PET1250*. Die Variable für Berufs- und Zeitsoldatinnen und -soldaten *PSH0350* wird nicht ausgeliefert.
mstib	Stellung im Beruf der Mutter, Kennziffer, generiert	Vorgehen analog zu *stiblewt*.
PSH0620	Damalige Stellung im Beruf des Vaters, einfache Einteilung (anon.)	Vorgehen analog zu *PET1210*.
PSH0660	Damalige StiB d. Vaters Beamter inkl., Soldaten: Det. Erfassung (anon.)	Vorgehen analog zu *PET1250*. Die Variable für Berufs- und Zeitsoldatinnen und -soldaten *PSH0650* wird nicht ausgeliefert.
vstib	Stellung im Beruf des Vaters, Kennziffer, generiert	Vorgehen analog zu *stiblewt*.
PMI0200	Nicht in D geboren: Geburtsland, Kategorien (anon.)	Länder mit zu geringen Fallzahlen wurden zu größeren Kategorien zusammengefasst.
ogebland	Geburtsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu *PMI0200*.

Tabelle 64: Übersicht über die in Welle 19 anonymisierten Variablen im Personendatensatz *(PENDDAT)*

(Fortsetzung)

PMI0500	Keine d. Staatsangehörigkeit: Welche Staatsangehörigkeit? (anon.)	Staatsangehörigkeiten von Ländern mit zu geringen Fallzahlen wurden zu größeren Kategorien zusammegefasst.
ostaatan	Staatsangehörigkeit, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu *PMI0500*.
ostaatansyr	Staatsangehörigkeit syr./irak. HH, inkl. off. Ang., Kategorien (anon.)	Für die Teilstichproben der syrischen und irakischen Haushalte wird die syrische Nationalität separat ausgewiesen.
ostaatanukr	Staatsangehörigkeit ukrain. HH, inkl. off. Ang., Kategorien (anon.)	Für die Teilstichproben der ukrainischen Haushalte wird die ukrainische Nationalität separat ausgewiesen.
PMI1000a	Vater: Zuzugsland (anon.)	Zuzugsländer mit zu geringen Fallzahlen wurden zu größeren Kategorien zusammengefasst.
PMI1000b	Mutter: Zuzugsland (anon.)	Vorgehen analog zu *PMI1000a*.
PMI1000c	Vater des Vaters: Zuzugsland (anon.)	Vorgehen analog zu *PMI1000a*.
PMI1000d	Mutter des Vaters: Zuzugsland (anon.)	Vorgehen analog zu *PMI1000a*.
PMI1000e	Vater der Mutter: Zuzugsland (anon.)	Vorgehen analog zu *PMI1000a*.
PMI1000f	Mutter der Mutter: Zuzugsland (anon.)	Vorgehen analog zu *PMI1000a*.
ozulanda	Vater: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu *PMI1000a*.
ozulandb	Mutter: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu *PMI1000a*.
ozulandc	Vater des Vaters: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu *PMI1000a*.
ozulandd	Mutter des Vaters: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu *PMI1000a*.
ozulande	Vater der Mutter: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu *PMI1000a*.
ozulandf	Mutter der Mutter: Zuzugsland, inkl. offene Angaben, Kategorien (anon.)	Vorgehen analog zu *PMI1000a*.

Tabelle 65: Übersicht über die in Welle 19 anonymisierten Variablen im BIO-Spelldatensatz *(bio_spells)*

Varname	Variablenlabel	Vorgehen
ET0617	Welle 19, Stellung im Beruf: Einfache Einteilung (anon.)	Vorgehen analog zu *PET1210*.
ET1017	Welle 19, Stellung im Beruf: Beamter/Richter/Soldat (anon.), detaillierte Erfassung	Vorgehen analog zu *PET1250*. Die Variable für Berufs- und Zeitsoldatinnen und -soldaten wird nicht ausgeliefert.
stib	Stellung im Beruf, Kennziffer, generiert	Vorgehen analog zu *stiblewt*.

Tabelle 66: Übersicht über die in Welle 19 anonymisierten Variablen im Kinder-Datensatz *(KINDER)*

Varname	Variablenlabel	Vorgehen
alter12u14m	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/ Bürgergeld-Bezüge

Bereits in den Wellen 1 bis 18 wurden die Arbeitslosengeld-II-Bezüge auf Haushaltsebene in Episodenform erfasst. Dieses Konzept wurde in Welle 19 mit einem leicht überarbeiteten Fragenprogramm fortgeführt (inkl. Umbenennung „Arbeitslosengeld II“ in „Bürgergeld“)[50].

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

Für die Fortschreibung der in der Vorwelle andauernden und damit im Spelldatensatz rechtszensierten SGB-II-Bezüge (Arbeitslosengeld II oder Bürgergeld) 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 19 erstmals befragten Haushalte der Auffrischungsstichprobe wurden nach

[50] Auch wenn durch die SGB-II-Reform seit 1.1.2023 das Bürgergeld das Arbeitslosengeld II abgelöst hat, wird im Rahmen von Kapitel 5.6 i.d.R. weiterhin der Begriff Arbeitslosengeld-II verwendet (oder allgemein von SGB-II-Bezügen gesprochen).

dem Bezug von SGB-II-Leistungen in der Zeit seit der letzten Veränderung des Haushalts gefragt. Lag diese Veränderung vor Januar 2023 oder wurde dazu keine Angabe gemacht, wurden die Bezüge des Haushalts seit Januar 2023 erhoben. Das Referenzdatum wird in jeder Welle um + 1 Jahr angepasst und liegt stets im Januar zwei Jahre vor dem aktuellen Erhebungsjahr.

5.6.2 Struktur des Arbeitslosengeld-II Spelldatensatzes

Durch die Integration der in Welle 19 berichteten SGB-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 19 nicht mehr erhobenen Variablen.

1. Auch in Welle 19 wurden mit *AL20618, AL20718a* bis *AL20718o* und *AL20818* neue, wellenspezifische Querschnittsvariablen in den Arbeitslosengeld-II Spelldatensatz eingefügt. Diese Variablen sind auf den Interviewzeitpunkt der Welle 19 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 Querschnittsinformationen.

Tabelle 67: Querschnittsvariablen im ALG II-Spelldatensatz *(alg2_spells)*

	Welle 1	Welle 2	...	Welle 11	Welle 12	...	Welle 19
Erhält der HH ALG II für alle HH-Mitglieder?	*AL20600*	*AL20601*	...	*AL20610*	*AL20611*	...	*AL20618*
Erhält der HH ALG II für die Personen 1 bis 15	*AL20700a* bis *AL20700o*	*AL20701a* bis *AL20701o*	...	*AL20710a* bis *AL20710o*	*AL20711a* bis *AL20711o*	...	*AL20718a* bis *AL20718o*
Höhe des monatlichen ALG II-Bezugs?	*AL20800*	*AL20801*	...	*AL20810*	*AL20811*	...	*AL20818*
Hat eine Kürzung des ALG II begonnen? (nur bis Welle 11 erhoben)	*AL20900*	*AL20901*	...	*AL20910*	—	...	—

2. Liegen in Welle 19 nicht vor.
3. Liegen in Welle 19 nicht vor.

5.6.3 Plausibilitätsprüfung und -bereinigung des Arbeitslosengeld-II Spelldatensatzes

Wie in den Wellen 1 bis 18 wurden die Informationen zu SGB-II-Bezügen auch in der Welle 19 einer Reihe von Plausibilitätsprüfungen unterzogen. Unzulässige Überschneidungen und Datierungen der Spells wurden dabei ggf. bereinigt. Veränderungen wurden dabei grundsätzlich nur in den generierten Datumsvariablen (*bmonat; bjahr; emonat; ejahr*) des Spells sowie im Zensierungsindikator des Arbeitslosengeld-II-Spells (*zensiert*) vorgenommen. Konnten fehlende Plausibilitäten nicht durch Korrektur der Datumsvariablen bereinigt werden, wurden in geringem Umfang Spells zusammengeführt bzw. gänzlich gelöscht.

5.6.4 Fortschreibung des Arbeitslosengeld-II-Spelldatensatzes

Im Anschluss an die Umstrukturierung der in Welle 19 berichteten Arbeitslosengeld-II-Episoden/Bürgergeld-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 SGB-II-Bezüge mit den in Welle 19 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 19-Informationen im engeren Sinne dar. Dabei werden neben dem Zensierungsindikator auch die generierten und erhobenen Enddaten sowie die wellenspezifischen Querschnittsinformationen (siehe oben) 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 19 berichtete Spells zugespielt (3). Diese drei Varianten sollen im Folgenden kurz dargestellt werden:

1. *Fälle, in denen der Haushalt in Welle 19 einem andauernden Arbeitslosengeld-II-Bezug zum Befragungszeitpunkt der Vorwelle widerspricht.*
   Hat der Haushalt der Information, dass zum Zeitpunkt der Vorwelle ein aktuell andauernder SGB-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 unverändert[51]. 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 19 Angaben zum Enddatum eines in der Vorwelle zensierten SGB-II-Bezugs vor, so wurde der in der Vorwelle zensierte Spell mit den aktuellen Angaben fortgeschrieben. Zunächst wurden das erhobene Enddatum (*AL20300; AL20400*), das generierte Enddatum (*emonat; ejahr*), die Nachfrage, ob der SGB-II-Bezug aktuell noch andauert (*AL20500*) sowie der Zensierungsindikator (*zensiert*) mit den in der Vorwelle erhobenen Informationen überschrieben. Weiterhin wurden die auf die Welle 18 bezogenen Querschnittsinformationen (*AL20618; AL20718a* bis *AL20718o, AL20818*) übernommen.
3. *In Welle 19 neu berichtete Arbeitslosengeld-II-Spells, die keine in der Vorwelle zensierte Episode fortschreiben.*
   In Welle 19 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

Seit Welle 4 wird das Konzept einer integrierten Erhebung der Erwerbsbiografie umgesetzt (bereits in den Wellen 2 und 3 wurden die Erwerbstätigkeits-, Arbeitslosigkeits- und Lückenphasen auf Personenebene in Episodenform erfasst). Für Personen, die in Welle 19 erstmals nach ihrer Erwerbsbiografie gefragt wurden, wurde erneut das Referenzdatum für den Beginn des Retrospektivintervalls angepasst. Das Referenzdatum wird in jeder Welle um + 1 Jahr angepasst und liegt stets im Januar zwei Jahre vor dem aktuellen Erhebungsjahr. In Welle 19 sollten demnach alle Erwerbstätigkeits- und Arbeitslosigkeitsepisoden seit Januar 2023 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.

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

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 Stunde 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 Abfragelogik 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 68: Generierungslogik *erwerb, erwerb2, nichterw, nichtew2*

Variable	Generierungslogik Welle 1	Generierungslogik Wellen 2 und 3
erwerb	*(1) Ausdifferenzierung Status Haupt-ET* - Keine Haupt-ET - Haupt-ET: nicht Lehre/ ABM/ Minijob - Haupt-ET: Teil einer Lehre - Haupt-ET: ABM etc. - Haupt-ET: Minijob *(2) Ausdifferenzierung Status Haupt-ET ist Basis für weitere Generierung* - Haupt-ET: nicht Lehre/ ABM/ Minijob → ET als Erwerbsstatus (Ausnahmen: Auszubildende (aus PB0100) mit arbzeit <21 → Auszubildende; Schülerinnen und Schüler (aus PB0100) mit arbzeit >0 & arbzeit <24 → Schülerinnen und Schüler; Studierende (aus PB0100) mit arbzeit >0 & arbzeit <21 → Studierende; Erwerbstätige mit arbzeit >0 & arbzeit <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.) *(3) Widersprüchliche Angaben entscheiden* - erwerb: ABM etc. + PB0100: Schülerinnen und Schüler/ Studierende/ Azubi → -8	nicht generiert (-9)

Tabelle 68: Generierungslogik *erwerb, erwerb2, nichterw, nichtew2* (Fortsetzung)

	- erwerb: Schülerinnen und Schüler + PB0100: Studierende → -8 - erwerb: Rentnerinnen und Rentner + PB0100: Azubi → -8 - erwerb: Schülerinnen und Schüler + PB0100: Azubi → -8 - erwerb: Sonstiges + PB0100: Schülerinnen und Schüler/ Studierende/ Azubi → Erwerbsstatus aus PB0100	
erwerb2	*(1) Recode von erwerb* - Zusammenlegung von Kategorien: - Arbeitslos + ABM/ 1-Euro-Job/ u.ä. → Arbeitslos - Lehre/ Ausbildung/ Weiterbildung/ Umschulung + Studierende → Lehre/ Ausbildung/ Studium	*(1) Recode von nichtew2* *(2) ETs integrieren* - Ersetzen der Werte, wenn aktuelle ET (>400 Euro aus ET-Spells) vorliegt *(3) Anpassungen vornehmen* - erwerb2: ET + PB0100: Studierende + Arbeitszeit <=20h → Studierende - erwerb2: AL + PB0100: Studierende → Studierende - erwerb2: Schülerinnen und Schüler + PB0100: Studierende → Unklarer Status
nichterw	*(1) Recode von PET0800*	*(1) Recode von LU0100 (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 PET0151/ PET0911 + Indikator für fehlerhaft nicht ins Lückenmodul gefilterte Fälle
nichtew2	*(1) Recode von PET0801* - Zusammenlegung von Kategorien: - Arbeitslos + ABM/ 1-Euro-Job/ u.ä → Arbeitslos	*(1) Recode von LU0101 (Lückenstatus mit offenen Angaben)* - Zusammenfassung von Kategorien - Arbeitslos gemeldet + nicht gemeldet →

Tabelle 68: Generierungslogik *erwerb, erwerb2, nichterw, nichtew2* (Fortsetzung)

	- Lehre/ Ausbildung/ Weiterbildung Umschulung + Studierende → Lehre/ Ausbildung/ Studium	Arbeitslos - Etwas anderes/Hauptstatus unklar → Sonstiges/ Hauptstatus unklar *(2) Schülerinnen und Schüler/Studierende/Azubis aus PB0100 berücksichtigen* - Wenn bisher kein gültiger Status vorhanden → Übernahme der Information aus PB0100

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 >= 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)
- Keine aktuelle Maßnahmeteilnahme (§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 69: Überarbeitung *erwerb, erwerb2, nichterw, nichtew2*

Variable	Entscheidung	Begründung
erwerb	beibehalten (Welle1: inhaltlich generiert) (Welle 2ff: -9)	Die Variable bildet das Erhebungskonzept der Welle 1 optimal ab. Der Schwerpunkt liegt auf ET (diese schlagen verein- facht gesprochen Arbeitslosigkeit, die wiederum alles andere schlägt.). Gegen eine Fortführung sprechen in- 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. inhalt-lich sinnvollere Entscheidungen zu treffen) Fazit: Eine harmonisierte Variable mit Schwerpunkt auf ET (wie bisher in erwerb2) ist die einzige Möglichkeit für eine 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.
nichterw	entfällt aus SUF	Die 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

Tabelle 69: Überarbeitung *erwerb, erwerb2, nichterw, nichtew2* (Fortsetzung)

		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.
nichtew2	entfällt aus SUF	(siehe nichterw)

Ab Welle 2 werden die folgenden Variablen generiert:

- *etakt: Aktuell erwerbstätig (oberhalb Minijobverdienstgrenze), 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 (oberhalb Minijobverdienstgrenze), 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 oberhalb der Minijobverdienstgrenze) 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. oberhalb der Minijobverdienstgrenze). Eine aktuelle ET liegt vor, wenn im Spelldatensatz der jeweiligen Welle ein zensierter ET-Spell vorhanden ist.

Ausprägungen der zu generierenden Variable:

- -10: Item für Fragebogenversion nicht relevant
- -5: Generierung nicht mgl. (fehlende Werte)
- -3: Trifft nicht zu (Filter)
- 1: ZP akt. erwerbstätig (über Minijobverdienstgrenze)
- 2: ZP akt. nicht erwerbstätig (über Minijobverdienstgrenze)

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 für 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 Schülerinnen und Schüler/Studierende/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 Schülerinnen und Schüler bzw. als Studierende/Azubi aus PB0100 wird berücksichtigt, als ob in der jeweiligen Welle ein aktuell andauernder Spell vorliegen würde.

Ausprägungen der zu generierenden Variable:

- -10: Item für Fragebogenversion nicht relevant
- -5: Generierung nicht mgl. (fehlende Werte)
- -3: Trifft nicht zu (Filter)
- 1: Erwerbstätig m. Eink. über Minijobverdienstgrenze
- 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 70: Grobzuweisung - Spell höherer Priorität schlägt Spell niedrigerer Priorität

Priorität eines aktuellen Spells (bzw. analoger Zustand aus PB0100)	Code in statakt (analog zu Variable spell-typ)	Bedeutung
1	2	Arbeitslos gemeldet/ Massnahmenteiln.
2	1	Erwerbstätig m. Eink. über Minijobverdienstgrenze
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 71: Feinzuweisung für Sonderfälle

Grobzuweisung	Zusatzinformation	Entscheidung
Arbeitslos gemeldet	Erwerstätig m. Eink. oberhalb der Minijobverdienstgrenze + Arbeitszeit (az2ges; tatsächliche Arbeitszeit, Summe über zensierte ETs)>=15h	Erwerbstätig m. Eink. oberhalb der Minijobverdienstgrenze
Erwerbstätig m. Eink. oberhalb der Minijob-verdienstgrenze	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 Haupterwerbstätigkeit[52]. Ausgenommen sind hier die Angaben zum Brutto-/Nettoeinkommen in den Wellen 2 bis 4 – diese beziehen sich auf alle aktuell andauernden Erwerbstätigkeiten oberhalb der Minijobverdienstgrenze (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-/Nettoeinkommen 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:

[52] Dabei handelt es sich bei der „Haupterwerbstätigkeit“ 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 72: Überarbeitung Einkommensvariablen

Variable - Inhalt - Datensatz	Generiert für	Basis
	W1 - W2 - W3 - W4 - W5ff.	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

* 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 Haupterwerbst. (ohne Minijobs, kategorisiert), gen.)*
- **brutto** *(Akt. Bruttoeinkommen Haupterwerbst. (ohne Minijobs, inkl. kat. Ang.), gen.)*
- **nettokat** *(Akt. Nettoeinkommen Haupterwerbst. (ohne Minijobs, kategorisiert), gen.)*
- **netto** *(Akt. Nettoeinkommen Haupterwerbst. (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 (oberhalb der Minijobverdienstgrenze) 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.)*
Die Variable brges enthält die über alle ETs (oberhalb der Minijobverdienstgrenze) 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 Über-arbeitung). 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 (oberhalb der Minijobverdienstgrenze). 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.)*
Die Variable netges enthält die über alle ETs (oberhalb der Minijobverdienstgrenze) 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 (oberhalb der Minijobverdienstgrenze). 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 73: Überarbeitung Arbeitszeitvariablen

Variable - Inhalt - Datensatz	Generiert für	Basis	Anmerkung
	W1 - W2 - W3	OFFA - KatA	
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 arbzeit (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 arbzeit (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):

arbzeit *(wöchentliche Arbeitszeit inkl. Angaben bei unregelmäßiger Arbeitszeit, gen.)*
Die Variable arbzeit entfällt aus *PENDDAT* und *BIO-Spells*. Sie wird inhaltlich durch *azhpt2* (*PENDDAT*) und *az2* (*BIO-Spells*) ersetzt.

Neue Variablen ab Welle 4:

az1 *(vertragl. Arbeitszeit, gen.)*
Die Variable az1 wird für alle Episoden in den *BIO-Spells* generiert. Sie enthält die jeweils für den Spell aktuellste Angabe zur vertraglichen Arbeitszeit (ET oberhalb der Minijobverdienstgrenze). 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
- Spell angelegt in Welle 9, fortgeschrieben in den Wellen 10, 11 und 12: Querschnittsvariablen Welle 12

azhpt1 *(vertragl. akt. Arb.zeit Haupterwerbst. (ohne Minijobs), gen.)*
Die Variable azhpt1 wird für den *PENDDAT* generiert. Sie enthält die vertragliche Arbeitszeit der in der jeweiligen Welle aktuell andauernden Haupt-ET aus den Spelldaten (ET oberhalb der Minijobverdienstgrenze). 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, in welchen die Variable *arbzeit* zur Bestimmung verwendet wurde).

azges1 *(vertragl. akt. Arb.zeit insgesamt (ohne Minijobs), gen.)*
Die Variable azges1 wird für den *PENDDAT* generiert. Sie enthält die kumulierte vertragliche Arbeitszeit aller in der jeweiligen Welle aktuell andauernden Erwerbstätigkeiten aus den Spelldaten (ET oberhalb der Minijobverdienstgrenze). 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 *(tatsächliche Arbeitszeit inkl. Angaben bei unregelmäßiger Arbeitszeit, gen.)*
Die Variable az2 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
- Spell angelegt in Welle 9, fortgeschrieben in den Wellen 10, 11 und 12: Querschnittsvariablen Welle 12

Die Variable az2 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 kom-menden Erwerbstätigkeiten die niedrigste Spell-Nummer.

azhpt2 *(tatsächl. akt. Arb.zeit Haupterwerbst. (ohne Minijobs, inkl. kat. Ang.), gen.)*
Die Variable azhpt2 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 Arbeits-zeit 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 *(tatsächl. akt. Arb.zeit insgesamt (ohne Minijobs, inkl. kat. Ang.), gen.)*
Die Variable azges2 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 19 wurden andauernde ET-, AL- und Lückenspells fortgeschrieben. Für die Fortschreibung der in der Vorwelle andauernden und damit im Spelldatensatz rechtszensierten 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 der 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 19 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 19 nicht mehr erhobenen Variablen.

1. Neue Variablen, die auf eine bestimmte Welle bezogen sind: Im BIO-Spelldatensatz werden die ET-spezifischen Variablen *ET0600* bis *ET2200* als wellenspezifische, auf Welle 2 bezogene Querschnittsinformationen 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, *ET0559* bis *ET2209* auf Welle 11, *ET0560* bis *ET2210* auf Welle 12, *ET0561* bis *ET2211* auf Welle 13, *ET0562* bis *ET2212* auf Welle 14, *ET0563* bis *ET2213* auf Welle 15, *ET0564* bis *ET2214* auf Welle 16, *ET0565* bis *ET2215* auf Welle 17, *ET0566* bis *ET2216* auf Welle 18 und *ET0567* bis *ET2217* auf Welle 19 bezogene Querschnittsinformationen. Seit Welle 5 werden zudem mit *ET28** bis *ET39** Variablen zum Einkommen für gegenwärtig andauernde Spells erhoben. Seit Welle 9 werden mit *ET41** und *ET42** Fragen zu Überstunden gestellt. Folgende Tabelle gibt einen Überblick über die im BIO-Spelldatensatz enthaltenen ET-spezifischen Querschnittsinformationen.

Tabelle 74: ET-spezifische Querschnittsvariablen im BIO-Spelldatensatz (*bio_spells*)

	Welle 2	Welle 3	Welle 4	Welle 5	...	Welle 9	...	Welle 19
Stellung im Beruf			*ET0552*	*ET0553*	...	*ET0557*	...	*ET0567*
(einfache und	*ET0600*	*ET0601*	*ET0602*	*ET0603*	...	*ET0607*	...	*ET0617*
detaillierte	*ET0700*	*ET0701*	*ET0702*	*ET0703*	...	*ET0707*	...	*ET0717*
Einteilung)	*ET0800*	*ET0801*	*ET0802*	*ET0803*	...	*ET0807*	...	*ET0817*
	ET1000	*ET1001*	*ET1002*	*ET1003*	...	*ET1007*	...	*ET1017*
	ET1100	*ET1101*	*ET1102*	*ET1103*	...	*ET1107*	...	*ET1117*
	ET1200	*ET1201*	*ET1202*	*ET1203*	....	*ET1207*	...	*ET1217*
Aufsichtsfunktion;	*ET1300*	*ET1301*	*ET1302*	*ET1303*	...	*ET1307*	...	*ET1317*

Tabelle 74: ET-spezifische Querschnittsvariablen im BIO-Spelldatensatz (*bio_spells*) (Fortsetzung)

Zahl der beaufsichtigten Arbeitskräfte	*ET1400*	*ET1401*	*ET1402*	*ET1403*	...	*ET1407*	...	*ET1417*
Entfristung	*ET1700*	*ET1701*	*ET1702*	*ET1703*	...	*ET1707*	...	*ET1717*
einer anfangs				*ET1753a*	...	*ET1757a*	...	*ET1767a*
unbefristeten Tätigkeit				*ET1753b*	...	*ET1757b*	...	*ET1767b*
Arbeitszeit			*ET1952*	*ET1953*	...	*ET1957*	...	*ET1967*
(vertraglich;	*ET2000*	*ET2001*	*ET2002*	*ET2003*	...	*ET2007*	...	*ET2017*
tatsächlich;	*ET2100*	*ET2101*	*ET2102*	*ET2103*		*ET2107*	...	*ET2117*
durchschnittlich bei unregelmäßiger)	*ET2200*	*ET2201*	*ET2202*	*ET2203*		*ET2207*	...	*ET2217*
Einkommen für gegenwärtig andauernde Spells				*ET2800-ET3907*	...	*ET2804-ET3904*	...	*ET2814-ET3914*
Überstunden					...	*ET4100*	...	*ET4110*
					...	*ET4200*	...	*ET4210*

Im BIO-Spelldatensatz gibt es darüber hinaus eine AL-spezifische Variable, die als wellenspezifische Querschnittsinformation 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, *AL1309* für Welle 11, *AL1310* für Welle 12, *AL1311* für Welle 13, *AL1312* für Welle 14, *AL1313* für Welle 15, *AL1314* für Welle 16, *AL1315* für Welle 17, *AL1316* für Welle 18 und *AL1317* für Welle 19). Nachfolgende Tabelle gibt einen Überblick über die im Spelldatensatz enthaltenen Querschnittsinformationen.

Tabelle 75: AL-spezifische Querschnittsvariablen im BIO-Spelldatensatz (*bio_spells*)

	Welle 2	Welle 3	...	Welle 19
Höhe des monatlichen ALG I-Bezugs?	*AL1300*	*AL1301*	..	*AL1317*

2. Neue Variablen, die nicht auf eine bestimmte Welle bezogen sind: Liegen in Welle 19 im Vergleich zu Welle 18 nicht vor.
3. Nicht mehr erhobene Variablen: Liegen in Welle 19 im Vergleich zu Welle 18 nicht vor.

5.7.5 Plausibilitätsprüfung und -bereinigung der Spelldatensätze

Auf Personenebene orientierten sich die Plausibilitätsprüfungen und Bereinigungsregeln an den Wellen ab Welle 2. Die Prüfungen fanden analog zu den Vorwellen nur innerhalb eines Spelltyps statt. Spelltypübergreifende Prüfungen wurden nicht durchgeführt. Bereinigungen und Umcodierungen wurden, wie bei den Arbeitslosengeld-II-Spelldaten, 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.

In einigen Fällen, wurden im Zuge der Datenaufbereiting ganze Episoden gelöschen. 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 19 berichteten Episoden ins Spellformat werden die zum Befragungszeitpunkt der Vorwelle andauernden Spells mit den in Welle 19 erhobenen Informationen fortgeschrieben. Dabei werden unzulässige Überschneidungen und Spells mit unplausiblen Datierungen bereinigt.

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 19-Informationen im engeren Sinne dar. Dabei werden neben dem Zensierungsindikator auch die generierten und erhobenen Enddaten und die wellenspezifischen Querschnittsinformationen (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 19 berichtete Spells zugespielt (3). Diese drei Varianten sollen im Folgenden kurz dargestellt werden:

1. *Veränderung des Zensierungsindikators: Fälle, in denen die Person in Welle 19 einem andauernden Spell zum Befragungszeitpunkt der Vorwelle widerspricht.*
   Hat die Person der Information widersprochen, dass zum Zeitpunkt der Vorwelle ein

aktuell andauernder Spell vorlag, und dies entweder explizit in der Fortschreibungsfrage oder implizit durch die Angabe eines vor dem Interviewdatum der Vorwelle liegenden Enddatums getan, 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 unverändert[53]. Das generierte Enddatum des Spells (*emonat; ejahr*) wurde bereits in der Vorwelle auf das Interviewdatum der Vorwelle gesetzt.

2. *Fortschreibung des in der Vorwelle zensierten Spells mit Welle 19-Informationen im engeren Sinne: Fälle, in denen die Person Angaben zum Enddatum eines in der Vorwelle aktuell andauernden Spells macht.*
   Liegen in Welle 19 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 19 erhobenen Informationen überschrieben. Weiterhin wurden die auf die Welle 19 bezogenen Querschnittsinformationen (*ET0567* bis *ET4210*) 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 19 erhobenen Informationen überschrieben. Weiterhin wurden die auf die Welle 19 bezogenen Querschnittsinformationen (*AL1317*) übernommen. Bei den AL-Spelldaten 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 angedauert 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

[53] 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.

(*alg1em, alg1ej*) sowie der Zensierungsindikator des Bezugs (*alg1akt*) mit den in Welle 19 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 19 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. *Zuspielung neu berichteter Spells: Fälle, in denen Personen in Welle 19 neue Spells berichten, also Spells, die keine in der Vorwelle zensierte Episode fortschreiben.*
In Welle 19 neu berichtete Spells wurden dem BIO-Spelldatensatz angehängt. Danach wurde der Spellzähler neu generiert, um eine lückenlose Variable *spellnr* zu schaffen. Das Fortschreiben der Spelldatensätze hat keine Auswirkung auf die Spellnummern des Vorwellen-SUF. Bereits im SUF zu Welle 18 enthaltene Spells (*spellnret, spellnral, spellnrlu, spellnr*) behalten ihre Spellnummer. Die neuen Spells aus Welle 19 werden an den entsprechenden Datensatz angehängt und die Spellnummern fortgeschrieben.

6 Gewichtung Welle 19

Die Konstruktion der Gewichte für die 19. Welle erfolgte im Wesentlichen analog zur Welle 18 (vgl. Anker et al., 2025). Ausgangspunkt für die Gewichtung der 19. Welle und des Längsschnitts von Welle 18 auf Welle 19 waren nunmehr die Querschnittsgewichte aus Welle 18 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 19 dar. Ein Überblick über das Gewichtungskonzept von PASS findet sich in Kapitel 3 des User Guides von PASS (Chauhan et al., 2026).

6.1 Designgewichte für die Bestandshaushalte der Welle 19

Unter Berücksichtigung von Zuzügen aus dem Inland wurden für die 19. Welle nunmehr aus den Haushaltsquerschnittsgewichten der 18. Welle neue Haushaltsdesigngewichte für die Welle 19 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 18 ins Sample zu gelangen. Das neue Designgewicht für Teilstichprobe i $dwihh_{19}$ errechnet sich also aus dem alten Querschnittsgewicht $wqihh_{18}$:

$$1/dw_ihh_{19} = 1/wq_ihh_{18} + (n_{sample_i}/n_{population_i})$$

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

6.2 Designgewichte für die Zugangsstichprobe und die Bestandsauffrischung der Welle 19

In der Welle 19 wurde das Panel erneut durch eine Nachziehung von Neuzugängen zum Leistungsbezug aufgefrischt. Eine Ziehungschance hatten alle Haushalte, die im Juli 2024 Leistungen bezogen haben, im selben Monat der Vorjahre 2006 bis 2023 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 19 befindet sich für Fälle in *sample* = 38 (normale Zugangsstichprobe), *sample* = 39 (Zugangsstichprobe – Syrer/Iraker) und *sample* = 40 (Zugangsstichprobe – Ukrainer) in der Variable *dw_ba*.

Des Weiteren wurde in Welle 19 das Panel durch eine Auffrischung des Bestands ergänzt. Eine Ziehungschance hatten alle Personen, die zum Ziehungszeitpunkt Leistungen bezogen und nicht zu den Neuzugängen zählten. Nachgezogen wurde gänzlich in den 100 neuen Points der EWO-Auffrischung aus Welle 17. Das Designgewicht der Bestandsauffrischung für Welle 19 wurde entsprechend getrennt von der Zugangsstichprobe der Neuzugänge berechnet. Für Fälle in *sample* = 41 (normale Bestandsauffrischung), *sample* = 42 (Bestandsauffrischung – Syrer/Iraker) und *sample* = 43 (Bestandsauffrischung – Ukrainer) befindet sich dieses Designgewicht jedoch – ebenso wie das Designgewicht der SGB-II-Neuzugänge – in der Variable *dw_ba*.

6.3 Wiederteilnahme-Propensity Haushalte

In diesem Schritt wurde wiederum analog zum Vorgehen für Welle 18 für jeden Haushalt, der in Welle 18 teilgenommen hatte, die Wahrscheinlichkeit zur Wiederteilnahme an Welle 19 über Logit-Modelle für Panelbereitschaft, Erreichbarkeit und Teilnahme geschätzt. Auch Haushalte mit letzter Teilnahme in Welle 17, ohne Teilnahme in Welle 18 (temporäre Ausfälle), wurden in der Modellierung für Welle 19 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 *hpbleib* zu finden. Das Längsschnittgewicht eines Haushalts aus einer der Stichproben der Welle 1 für den gesamten möglichen Zeitraum [t1 bis t19] zwischen allen 19 Wellen lässt sich dann als Produkt aus Querschnittsgewicht zu t1, *hpbleib* (Welle 1 zu Welle 2) und *hpbleib* (Welle 2 zu Welle 3 usw.) berechnen (vgl. hierzu auch Kapitel 3 des User Guides (Chauhan et al., 2026)).

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.

6.4 Teilnahme-Propensity erstbefragte Split-Haushalte

In diesem Schritt wurden die Teilnahme-Propensities für neue Split-Haushalte, d.h. Haushalte, die aufgrund des Auszugs einer Person der Panelstichprobe in einen neuen Haushalt in das Panel aufgenommen werden, berechnet. Hierbei wurden nur solche Split-Haushalte berücksichtigt, die in den vorausgegangenen Wellen noch nicht befragt wurden. Modelliert wurden also die Teilnahmepropensities für die erstbefragten Split-Haushalte, getrennt nach dem Kriterium in Welle 18 entstandener Split-Haushalt (Split-W18-Haushalte) und in Welle 19 entstandener Split-Haushalt (Split-W19-Haushalte). Die Wahrscheinlichkeit zur Wiederteilnahme wurde über Logit-Modelle geschätzt. Bei den Split-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. Der Kehrwert der vorhergesagten Teilnahmewahrscheinlichkeit ist für die Split-Haushalte ebenfalls in der Variable *hpbleib* zu finden.

Die Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der erstbefragten Split-Haushalte (Welle 18 und Welle 19) sind → **Tabelle A11** zu entnehmen.

Die Logit-Modelle zur Erstteilnahme von Split-W18-Haushalten sind → **Tabelle A12** zu entnehmen.

Die Logit-Modelle zur Erstteilnahme von Split-W19-Haushalten sind → **Tabelle A13** zu entnehmen.

6.5 Ausfallgewichtung für Haushalte aus der Auffrischungsstichprobe der SGB-II-Neuzugänge Welle 19

Für die Haushalte aus der Auffrischungsstichprobe der SGB-II-Neuzugänge zum ALG-II-Bezug, sowie der Bestandsauffrischung in Welle 19 (*sample* = 38, normale Zugangsstichprobe, *sample* = 39, Zugangsstichprobe – Syrer/Iraker und *sample* = 40, Zugangsstichprobe – Ukrainer, *sample* = 41, normale Bestandsauffrischung, *sample* = 42,

Bestandsauffrischung – Syrer/Iraker und sample = 43, Bestandsauffrischung – Ukrainer) wurde erneut eine Ausfallmodellierung (Teilnahme) analog zur Auffrischungsstichprobe der Welle 18 vorgenommen, jeweils für Erreichbarkeit und Teilnahme. Es wurde ein integriertes Modell für die drei Teilstichproben geschätzt, die Variable *(samaufftyp*)* unterscheidet die Zugehörigkeit zu den Teilstichproben. Die aus dem Modell abgeleitete Teilnahmewahrscheinlichkeit befindet sich in der Variable *prop_t0*.

Die Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der SGB-II-Auffrischungsstichprobe Welle 19 sind → **Tabelle A14** zu entnehmen.

Die Logit-Modelle zur Erstteilnahme der SGB-II-Auffrischungsstichprobe Welle 19 sind → **Tabelle A15** zu entnehmen.

6.6 Wiederteilnahme-Propensity Personen

Das entscheidende Längsschnittgewicht ist nicht dasjenige auf Haushalts-, sondern das auf Personenebene, da es sich hierbei um die über die Zeit stabilen Einheiten handelt. Die Teilnahmepropensities für Personen in Welle 19 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 t19] über alle 19 Wellen lässt sich dann als Produkt aus Querschnittsgewicht 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 A16** zu entnehmen.

Die Logit-Modelle zur Wiederteilnahme Personen für Panelbereitschaft, Erreichbarkeit und Teilnahme sind → **Tabelle A17** zu entnehmen.

6.7 Integration der Gewichte zum Gesamtgewicht vor Kalibrierung

In diesem Schritt ging es erneut darum, die durch die Ausfallmodellierung modifizierten Haushaltsgewichte der neuen Auffrischungsstichprobe der SGB-II-Neuzugänge und der SGB-II-Bestandsauffrischung Welle 19 sowie der Bestandsstichprobe (einschließlich der Auffrischer aus Welle 2 bis 18) zu kombinieren. Die SGB-II-Bestandsaufstockung (sample = 41, 42, 43) musste zusätzlich noch mit ALG II beziehenden Fällen aus allen SGB-II-Stichproben der Wellen 1-18 mittels Konvexkombination integriert werden. Die mehrfache Auswahlwahrscheinlichkeit, die eine nun gezogene Leistungsempfängerin oder ein nun gezogener Leistungsempfänger, die/der in den vorausgegangenen Jahren in einem gemeinsamen Haushalt mit Leistungsempfängenden 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 jeweils im Juli der Jahre 2006 bis 2024 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 2024. Dazu wurde die Inklusionswahrscheinlichkeit in der jeweils anderen Stichprobe für Fälle aus der Microm-Stichprobe (Welle 1), der EWO-Auffrischungen (Welle 5, Welle 11, Welle 17 und Welle 18) und der neuen Auffrischungsstichprobe der SGB-II-Neuzugänge sowie Bestandsauffrischung (Welle 19) geschätzt. Für Fälle aus den neuen Zugangsstichproben wurde die mittlere Welle 1-Auswahlwahrscheinlichkeit in der Microm-Stichprobe, die wellenspezifische mittlere Auswahlwahrscheinlichkeit der EWO-Auffrischungen in der jeweiligen Postleitzahl und jeweils die durchschnittliche Teilnahmewahrscheinlichkeit (für Welle 1 bis Welle 19) in dieser Stichprobe unterstellt. Für Fälle aus der Microm- oder den EWO-Auffrischungsstichproben wurde, falls es sich (laut Befragungsdaten) um Neuzugänge zum Arbeitslosengeld-II-Leistungsbezug handelt, die durchschnittliche Auswahlwahrscheinlichkeit eines Auffrischungshaushalts (SGB-II-Neuzugänge Welle 19) in der jeweiligen PLZ und die durchschnittliche Teilnahmewahrscheinlichkeit in dieser Stichprobe unterstellt. Danach wurden die daraus resultierenden Gewichte zu einem neuen Gesamtgewicht integriert.

6.8 Integration temporärer Ausfälle (Haushalte)

In Welle 19 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 Teilnehmenden 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 17 (die Welle 18 ist der temporäre Ausfall).

Für die temporären Ausfälle wurde dann die Wahrscheinlichkeit der Nichtteilnahme in Welle 18 gegeben die Teilnahme in Welle 17 (Nichtteilnahme-Propensities Welle 18) und die Wahrscheinlichkeit der Teilnahme in Welle 19 gegeben die Nichtteilnahme in Welle 18 (Teilnahme-Propensities Welle 19) bestimmt. Die Wahrscheinlichkeit der Nichtteilnahme in Welle 18 ergibt sich über 1-Teilnahmewahrscheinlichkeit in Welle 18.

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 17 mit dem Kehrwert dieses Produkts.

Die Variablenübersicht, Abkürzungen und Referenzkategorien für die Logit-Modelle der temporären Ausfälle sind → **Tabelle A18** zu entnehmen.

Die Logit-Modelle temporärer Ausfälle sind → **Tabelle A19** zu entnehmen.

Die Konvexkombination der Gewichte der Teilnehmenden über alle Wellen (Bestandsstichprobe) und der temporären Ausfälle erfolgte für die Gewichte aller drei Teilstichproben i (Microm, SGB II 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.

6.9 Kalibrierung zum Haushaltsgewicht Welle 19 Querschnitt

Anschließend fand wieder eine Kalibrierung der modifizierten Designgewichte inklusive der Ausfallgewichtung auf Haushaltsebene mittels GREG-Verfahren an die Eckwerte des Statistischen Bundesamts für 2024 und für Haushalte mit Leistungsbezug an die Eckwerte der Statistik der Bundesagentur für Arbeit für Juli 2024 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 (392.263) verwendet. Dabei werden diejenigen Fälle der bisherigen Stichproben aus den Wellen 1 bis 18 auf die Eckwerte zum Arbeitslosengeld-II-Bezug der Bundesagentur für Arbeit hochgerechnet, die laut Welle 19 des Surveys im Juli 2024 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 resultieren, bei zu eng gesetzten Grenzen kann allerdings die Kalibrierung aller Eckwerte unmöglich werden.

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

6.10 Kalibrierung der SGB-II-Stichprobe

Die Grundgesamtheit der kumulierten SGB-II-Stichprobe aus allen 19 Wellen bilden alle Haushalte in Deutschland mit mindestens einer Bedarfsgemeinschaft mit Leistungen nach dem SGB II zu einem der bisher 19 Ziehungszeitpunkte (jeweils im Juli der Jahre 2006 bis 2024). Eine Kalibrierung findet in Welle 19 dagegen nur an die Eckwerte der BA-Statistik vom Juli 2024 statt. Damit beeinflusst die Kalibrierung nur die Gewichte der Haushalte aus der SGB-II-Stichprobe, in denen auch im Juli 2024 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 10%-Perzentil und am 90%-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 218,954 bis 2237,689. 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 21,895 bis maximal 4369,696.

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 (392.263).

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 (SGB-II-Stichprobe, Haushalte) sind → **Tabelle A20** zu entnehmen.

Die Kenngrößen der Verteilung der Gewichte (SGB-II-Stichprobe, Haushalte) sind → **Tabelle A21** zu entnehmen.

6.11 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 10%-Perzentil und am 90%-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 4443,823 bis 37996,7. 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 3085,667 bis maximal 66578,11.

Für folgende Merkmale wurde kalibriert:

1. Haushalte: Grundlage Mikrozensus 2024:
    - 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 A22** zu entnehmen.

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

6.12 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 10%-Perzentil und am 90%-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 322,047 bis 18658,45. 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 32,2047 bis maximal 25893,99.

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 2024:

   - 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 (392.263) 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 A24** zu entnehmen.

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

6.13 Kalibrierung zum Personengewicht Welle 19 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 (510.819) 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 19 des Surveys im Juli 2024 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.

6.14 SGB-II-Stichprobe

Die Grundgesamtheit der kumulierten SGB-II-Stichprobe aller 19 Wellen bilden alle Personen im Alter ab 15 Jahre, die in einem Haushalt leben, in dem es zu einem der bisher 19 Ziehungszeitpunkte (jeweils im Juli 2006 bis 2024) 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 2024 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 SGB-II-Stichprobe. Diese wurden am 10%-Perzentil und am 90%-Perzentil ihrer Verteilung getrimmt und anschließend so reskaliert, dass ihre Summe wieder die Summe der ungetrimmten kalibrierten Haushaltsgewichte ergab. Die getrimmten Hochrechnungsfaktoren reichen von 135,253 bis 2203,617. 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 43,202 bis maximal 7658,056.

Für folgende Merkmale wurde kalibriert:

Leistungsempfangende 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 (510.819) 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 (SGB-II-Stichprobe, Personen) sind → **Tabelle A26** zu entnehmen.

Die Kenngrößen der Verteilung der Gewichte (SGB-II-Stichprobe, Personen) sind → **Tabelle A27** zu entnehmen.

6.15 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 10%-Perzentil und am 90%-Perzentil ihrer Verteilung getrimmt und anschließend so reskaliert, dass ihre Summe wieder die Summe der ungetrimmten kalibrierten Haushaltsgewichte ergab. Die getrimmten Hochrechnungsfaktoren reichen von 5616,657 bis 44730,62. 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 561,6657 bis maximal 223653,0922.

Für folgende Merkmale wurde kalibriert:

1. Bevölkerung Grundlage Mikrozensus 2024:
    - 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
2. 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 A28** zu entnehmen.

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

6.16 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 10%-Perzentil und am 90%-Perzentil ihrer Verteilung getrimmt und anschließend so reskaliert, dass ihre Summe wieder die Summe der ungetrimmten kalibrierten Haushaltsgewichte ergab. Die getrimmten Hochrechnungsfaktoren reichen von 282,5299 bis 24676,35. 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 28,2530 bis maximal 74029,0388.

Für folgende Merkmale wurde kalibriert:

1. Leistungsempfängende 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 2024:
    - 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 (510.819) 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 A30** zu entnehmen.

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

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

Schließlich verblieben auch in der Welle 19 noch Haushalte und Personen, denen durch Kalibrierung kein BA-Haushaltsquerschnittsgewicht oder BA-Personenquerschnittsgewicht

zugewiesen werden konnte. Die Anzahl dieser Haushalte ist nun in Welle 19 erneut größer als in den Wellen zuvor, da ein größerer Teil der SGB-II-Stichprobe der Wellen 1 bis 18 mittlerweile aus dem Leistungsbezug ausgeschieden ist. Es handelt sich um die folgenden drei Gruppen ohne Leistungsbezug im Juli 2024, die aber zur Grundgesamtheit der SGB-II-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, Juli 2016, Juli 2017, Juli 2018, Juli 2019, Juli 2020, Juli 2021, Juli 2022, Juli 2023 oder Juli 2024) gehören:

- Aus der Auffrischungsstichprobe: Nicht-BG-Mitglieder im Haushalt: Das Personengewicht wurde hier aus dem BA-Haushaltsgewicht in Welle 19 nach Kalibrierung (*wqbahh*) durch Division durch den Anteil dieser Personen, die – vorausgesetzt ihr Haushalt hat teilgenommen – ein Personen- oder Seniorenінterview gegeben haben, gewonnen.
- Bestandshaushalte, in denen im Juli 2024 niemand mehr Arbeitslosengeld-II-Leistungen 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 18 erhalten als neues BA-Personengewicht das BA-Haushaltsgewicht ihres Haushalts für Welle 19 dividiert durch den Anteil solcher Personen, die teilnehmen, vorausgesetzt ihr Haushalt nimmt teil.
- Nicht-BG-Mitglieder in Bestandshaushalten, die im Juli 2024 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 SGB-II-Teilstichproben abzüglich der aktuell im Bezug befindlichen Personen bzw. Bedarfsgemeinschaften näherungsweise bestimmen. Dieser Eckwert wird zusätzlich um die geschätzte kumulierte Sterberate dieser Personengruppe für den Zeitraum 2005 bis 2024 um 4,9% reduziert. Die entsprechende Zahl der aus dem Grundsicherungsbezug abgegangenen Personen bemisst sich damit für Welle 19 auf 8.949.764, die Zahl der aus dem Grundsicherungsbezug abgegangenen Bedarfsgemeinschaften auf 6.727.191.

7 Anhang: Kurzbeschreibung des Datensatzes

Die Kurzbeschreibung des Datensatzes ist in dem → **Tabellenanhang** dargestellt.

Literatur

AAPOR (2023). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 10. 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 Sozialstaatsfor-schung: Zur Konzeption der IAB-Panelerhebung „Arbeitsmarkt und Soziale Sicherung“, *IAB-Forschungsbericht 12/2007* (S. 11-32), Nürnberg.

Andersen, H. H., Mühlbacher, A., Nübling, M., Schupp, J., & Wagner, G. G. (2007). Computation of standard values for physical and mental health scale scores using the SOEP version of SF-12v2. *Schmollers Jahrbuch*, 127(1), 171-182

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.

Anker, F., Berg, M., Cramer, R., Dickmann, C., Gerber, V., Gilberg, R., Jesske, B., Kleudgen, M., Beste, J., Dummert, S., Frodermann, C., Malich, S., Prospero, V., Wenzig, C., Trappmann, M., Bähr, S., Collischon, M., Gundert, S., Mackeben, J., Müller, M., Stegmaier, J. & Wunder, A.: Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS) Band I: Datenreport Welle 18, *FDZ Datenreport, 10/2025*. Nürnberg

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., Kleudgen, M., Bethmann, A., Fuchs, B., Gebhardt, D.(2011). Codebuch und Dokumentation des „Panel Arbeitsmarkt und Soziale Sicherung“ (PASS). Bd. 1: Datenreport Welle 4. *FDZ Datenreport 08/2011*. Nürnberg.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, M., Bethmann, A.,

Fuchs, B., Trappmann, M. & Wurdack, A. (2012). Codebuch und Dokumentation des „Panel Arbeitsmarkt und Soziale Sicherung“ (PASS). Bd. 1: Datenreport Welle 5. *FDZ-Datenreport 06/2012*. Nürnberg.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, Beste, J., Dummert, 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.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, Beste, J., Dummert, S., Frodermann, C., Schwarz, S., Trappmann, M., Bähr, S., Coban, M., Friedrich, M., Gundert, S., Müller, B. Teichler, N., Unger, S. & Wenzig, C. (2019). Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS) Bd. 1: Datenreport Welle 12. *FDZ-Datenreport 09/2019*. Nürnberg.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, Beste, J., Dummert, S., Frodermann, C., Schwarz, S., Trappmann, M., Bähr, S., Coban, M., Friedrich, M., Gundert, S., Müller, B., Teichler, N., Unger, S. & Wenzig, C. (2020). Codebuch und Dokumentation des „Panel Arbeitsmarkt und soziale Sicherung“ (PASS) Bd. 1: Datenreport Welle 13. *FDZ-Datenreport 12/2020*. Nürnberg.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, M., Beste, J., Dummert, S., Frodermann, C., Wenzig, C., Trappmann, M., Bähr, S., Collischon, M., Coban, M., Friedrich, M., Gleiser, P., Gundert, S., Malich, S., Müller, B., Schwarz, S., Stegmaier, J., Teichler, N. & Unger, S. (2021): Codebuch und Dokumentation des Panel „Arbeitsmarkt und soziale Sicherung“ (PASS) Bd. 1 Datenreport Welle 14. *FDZ-Datenreport 14/2021*. Nürnberg.

Berg, M., Cramer, R., Dickmann, C., Gilberg, R., Jesske, B., Kleudgen, M., Beste, J., Dummert, S., Frodermann, C., Schwarz, S., Wenzig, C., Trappmann, M., Altschul, S., Bähr, S., Collischon, M., Coban, M., Gleiser, P., Gundert, S., Küfner, B., Mackeben, J., Malich, S., Müller, B., Stegmaier, J., Teichler, N. & Unger, S. (2022): Codebuch und Dokumentation des Panel „Arbeitsmarkt und soziale Sicherung“ (PASS) Bd. 1 Datenreport Welle 15. *FDZ-Datenreport 10/2022*. Nürnberg.

Berg, M., Cramer, R., Dickmann, C., Gerber, V., Gilberg, R., Jesske, B., Kleudgen, M., Beste, J., Dummert, S., Frodermann, C., Malich, S., Schwarz, S., Wenzig, C., Trappmann, M., Bähr, S., Bömmel, N., Coban, M., Collischon, M., Gundert, S., Küfner, B., Mackeben, J., Müller, B., Stegmaier, J., Teichler, N. & Wunder, A. (2023): Codebuch und Dokumentation des Panel „Arbeitsmarkt und soziale Sicherung“ (PASS) Bd. 1 Datenreport Welle 16. *FDZ-Datenreport 10/2023*. Nürnberg.

Berg, M., Cramer, R., Dickmann, Gerber, V., Gilberg, R., Jesske, B., Kleudgen, M., Beste, J., Dummert, S., Frodermann, C., Malich, S., Schwarz, S., Wenzig, C., Trappmann, M., Bähr, S., Bömmel, N., Coban, M., Collischon, M., Gundert, S., Küfner, B., Mackeben, J., Müller, B., Steg-maier, J., Teichler, N. & Wunder, A. (2024): Codebuch und Dokumentation des Panel „Arbeits-markt und soziale Sicherung“ (PASS) Bd. 1 Datenreport Welle 17. *FDZ-Datenreport 07/2024*. Nürnberg.

Beste, J., Olbrich, L. & Schwanhäuser, S. (2021). Interviewer:innenkontrolle im Panel Arbeitsmarkt und soziale Sicherung (PASS). *FDZ-Methodenreport 04/2021*. Nürnberg.

Bonin, H. & Rinne, U. (2017). Omnibusbefragung zur Verbesserung der Datenlage neuer Beschäftigungsformen. Kurzexpertise im Auftrag des Bundesministeriums für Arbeit und Soziales. IZA Research Report. Bonn. 80.

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.

Bruckmeier, K., Hohmeyer, K., Schwarz, S. (2018): Welfare receipt misreporting in survey data and its consequences for state dependence estimates: new insights from linked administrative and survey data. *Journal for Labour Market Research, Vol. 52, No. 1*, 21 S.

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-Prestige-skala auf die Berufsklassifikation ISCO88. *ZUMA-Nachrichten, 57*. S. 79-127.

Chauhan, M., Bähr, S., Beste, J., Collischon, M., Dummert, S., Fehn, A. , Frodermann, C., Malich, S., Müller, M., Prospero, V., Stegmaier, J., Trappmann, M. & Wenzig, C. (2026). Panel Study Labour Market and Social Security (PASS) User Guide, Version 3.0. *FDZ-Datenreport 06/2026*. Nürnberg.

Current Population Survey staff (2018). Electronically mediated work: New questions in the Contingent Worker Supplement. Monthly Labor Review, 1-32. Retrieved from https://www.jstor.org/stable/2650263

Dickmann, C., Gerber, V., Jesske, B., Beste, J., Dummert, S., Frodermann, C., Malich, S., Mühlbacher, V., Schwarz, S., Wenzig, C., Trappmann, M., Zins, S., Bähr, S., Collischon, M., Gundert, S., Küfner, B., Mackeben, J., Müller, B., Stegmaier, J., Teichler, N. & Wunder, A. (2023): Codebuch und Dokumentation des Websurveys im Panel „Arbeitsmarkt und soziale Sicherung“ (PASS). Datenreport Websurvey 2022. *FDZ-Datenreport 13/2023*. Nürnberg.

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)].

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 (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., Kiesl, 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.

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 13/2018*. Nürnberg.

Jesske, B., Schulz, S., Quandt, S. (2019). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 12. Erhebungswelle – 2018 (Haupterhebung), *FDZ Methodenreport 05/2019*. Nürnberg

Jesske, B., Schulz, S., Quandt, S. (2020). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 13. Erhebungswelle – 2019 (Haupterhebung), *FDZ Methodenreport 06/2020*. Nürnberg.

Jesske, B. & Schulz, S. (2021). Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 14. Erhebungswelle – 2020 (Haupterhebung), *FDZ-Methodenreport 06/2021*. Nürnberg.

Jesske, B., Schulz, S. & Gerber, V. (2022): Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 15. Erhebungswelle – 2021 (Haupterhebung), *FDZ-Methodenreport 06/2022*. Nürnberg.

Jesske, B., Schulz, S. & Gerber, V. (2023): Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 16. Erhebungswelle – 2022 (Haupterhebung), *FDZ-Methodenreport 06/2023*. Nürnberg.

Jesske, B., Schulz, S. & Gerber, V. (2024): Methodenbericht Panel Arbeitsmarkt und Soziale

Sicherung PASS. 17. Erhebungswelle – 2023 (Haupterhebung). *FDZ-Methodenreport 05/2024*. Nürnberg.

Jesske, B., Schulz, S. & Torregroza, S. (2025): Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 18. Erhebungswelle – 2024 (Haupterhebung). *FDZ-Methodenreport 08/2025*, Nürnberg.

Jesske, B., Schulz, S. & Bangel, P. (2026): Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS. 19. Erhebungswelle – 2025 (Haupterhebung). *FDZ-Methodenreport 03/2026*, Nürnberg.

König, W., Lüttinger, P. & Müller, W. (1987). Eine vergleichende Analyse der Entwicklung und Struktur von Bildungssystemen. Methodologische Grundlagen und Konstruktion einer vergleichbaren 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. *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 Mikrozensen 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.

Müller, B., Bähr, S., Gundert, S., Teichler, N., Unger, S. & Wenzig, C. (2020). PASS Scales and Instruments Manual, *FDZ-Methodenreport 07/2020 (en)*. Nürnberg.

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

Nübling, M., Andersen, H. H. & Mühlbacher, A. (2006). *Entwicklung eines Verfahrens zur Berechnung der körperlichen und psychischen Summenskalen auf Basis der SOEP-Version des SF 12 (Algorithmus)* (No.16). DIW Data Documentation

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.

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-Forschungs-bericht 12/2007* (S. 60-101), Nürnberg.

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

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.

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 10|2026 DE

Veröffentlichungsdatum
18. November 2026

Herausgeber
Forschungsdatenzentrum
der Bundesagentur für Arbeit
am Institut für Arbeitsmarkt- und Berufsforschung
Regensburger Straße 104
90478 Nürnberg

Nutzungsrechte
Diese Publikation ist unter folgender Creative-Commons-Lizenz veröffentlicht:
Namensnennung – Weitergabe unter gleichen Bedingungen 4.0 International (CC BY-SA 4.0)
https://creativecommons.org/licenses/by-sa/4.0/deed.de

Bezugsmöglichkeit dieses Dokuments
https://doku.iab.de/fdz/reporte/2026/DR_10-26.pdf

Dokumentation Version
PASS-SUF0625_DE_v1_dok1, DOI: 10.5164/IAB.FDZD.2610.de.v1

Dataset Version
PASS-SUF0625, v1, DOI: 10.5164/IAB.PASS-SUF0625.de.en.v1

Bezugsmöglichkeit aller Veröffentlichungen der Reihe „FDZ-Datenreport“
https://fdz.iab.de/forschung/publikationen/fdz-datenreporte/

Website
https://fdz.iab.de

Rückfragen zum Inhalt
Jonas Beste
Telefon +49 (0) 911/179-2279
E-Mail Jonas.Beste@iab.de