

Anett Friedrich | Kevin Ord

# BIBB-Qualifizierungspanel 2011 bis 2017 Längsschnittdatensatz

Version 1.0



Heft 4/2020

Anett Friedrich | Kevin Ord

# **BIBB-Qualifizierungspanel 2011 bis 2017 Längsschnittdatensatz**

Version 1.0

## Danksagung

Besonderer Dank für die Unterstützung bei der Erstellung des Datensatzes geht an Christian Gerhards für die Hilfe bei der Harmonisierung der Variablenamen und an Christine Hohn für die Erstellung der SPSS-Programmcodes.

## Datenverfügbarkeit

Der Zugang zu den in diesem Handbuch beschriebenen Forschungsdaten erfolgt – wie bei allen Betriebsdaten des BIBB-FDZ – über die Datenfernverarbeitung (DFV) und das Gastwissenschaftlermodell (GWA).

## Hinweis zur Zählweise bei Versionsnummern

Änderungen gegenüber einer möglicherweise vorhandenen Vorversion dieses Berichts ohne größere inhaltliche Relevanz werden durch fortlaufende Nummern nach dem Punkt dokumentiert (zweite Ebene). Inhaltlich relevante Änderungen führen demgegenüber zu einer fortlaufenden Nummerierung auf der ersten Ebene.

## Impressum

### Zitiervorschlag:

Friedrich, Anett; Ord, Kevin: BIBB-Qualifizierungspanel 2011 bis 2017. Längsschnittdatensatz. BIBB-FDZ Daten- und Methodenbericht 4/2020. Bonn 2020

1. Auflage 2020

### Herausgeber:

Bundesinstitut für Berufsbildung  
Robert-Schuman-Platz 3  
53175 Bonn  
Internet: [www.bibb.de](http://www.bibb.de)

### Publikationsmanagement:

Stabsstelle „Publikationen und wissenschaftliche Informationsdienste“  
E-Mail: [publikationsmanagement@bibb.de](mailto:publikationsmanagement@bibb.de)  
[www.bibb.de/veroeffentlichungen](http://www.bibb.de/veroeffentlichungen)

### Gesamtherstellung:

Verlag Barbara Budrich  
Stauffenbergstraße 7  
51379 Leverkusen  
Internet: [www.budrich.de](http://www.budrich.de)  
E-Mail: [info@budrich.de](mailto:info@budrich.de)

### Kontakt FDZ:

E-Mail: [fdz@bibb.de](mailto:fdz@bibb.de)  
Tel.: +49 228 107-2041  
Fax: +49 228 107-2020

### Lizenzierung:

Der Inhalt dieses Werkes steht unter einer Creative-Commons-Lizenz (Lizenztyp: Namensnennung – Keine kommerzielle Nutzung – Keine Bearbeitung – 4.0 International).



Weitere Informationen zu Creative Commons und Open Access finden Sie unter [www.bibb.de/oa](http://www.bibb.de/oa).

ISSN: 2190-300X

urn:nbn:de: 0035-0880-2

### Bibliografische Information der Deutschen Nationalbibliothek

Die deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b> .....	<b>5</b>
1.1.	Gegenstand des Handbuchs .....	5
1.2	Aufbau des Handbuchs .....	5
1.3	Kurzbeschreibung des Datensatzes .....	5
1.4	Zentrale Eigenschaften im Überblick .....	6
<b>2</b>	<b>Überblick über das BIBB-Qualifizierungspanel</b> .....	<b>7</b>
<b>3</b>	<b>Organisation der Datensätze</b> .....	<b>9</b>
3.1	Datensätze .....	9
3.2	Benennung der inhaltlichen Variablen .....	13
3.3	Harmonisierung der Variablen für die Längsschnittdatensätze .....	14
<b>4</b>	<b>Organisation der Mikrodaten</b> .....	<b>15</b>
4.1	Imputierte Variablen .....	15
4.2	Im Fragebogen nicht erkennbare Filter .....	15
4.3	Fehlende Werte .....	16
4.4	Anzahl der Beschäftigten im Betrieb .....	19
4.5	Besonderheiten einzelner Jahre .....	19
4.6	Englische Übersetzung .....	24
<b>5</b>	<b>Gewichtung und Hochrechnung</b> .....	<b>25</b>
5.1	Berechnung der Gewichtungs- und Hochrechnungsfaktoren .....	25
5.2	Gewichtungs- und Hochrechnungsfaktoren im Datensatz .....	26
<b>6</b>	<b>Weiterführung in den Folgejahren</b> .....	<b>30</b>
<b>7</b>	<b>Anonymisierung und Datenzugang</b> .....	<b>31</b>
7.1	Bestimmungen des Datenschutzes .....	31
7.2	Formale Datenanonymisierung .....	31
7.3	Datenzugang .....	31
<b>8</b>	<b>Literaturverzeichnis</b> .....	<b>33</b>
	<b>Anhang</b> .....	<b>34</b>



# 1 Einleitung

## 1.1. Gegenstand des Handbuchs

Der vorliegende Daten- und Methodenbericht beschreibt den Längsschnittdatensatz des BIBB-Betriebspanels zu Qualifizierung und Kompetenzentwicklung 2011 bis 2017 (kurz: *BIBB-Qualifizierungspanel 2011 bis 2017 long*). Der Bericht soll externen Wissenschaftlerinnen und Wissenschaftlern beim Umgang mit den Daten helfen. Neben der Dokumentation der Datensatzstruktur und Harmonisierung enthält der Bericht Stata- und SPSS-Befehle, die die Arbeit mit den Daten erleichtern sowie zusätzliche Harmonisierungsmöglichkeiten aufzeigen sollen.

Der Datensatz ist speziell für Längsschnittauswertungen konzipiert. Für Auswertungen einzelner Wellen des BIBB-Qualifizierungspanels ist es sinnvoll, auf den entsprechenden Querschnittdatensatz zurückzugreifen. Die jährlichen Querschnittdatensätze sind ebenfalls im BIBB-FDZ verfügbar.

Wichtige Kerngrößen wie der Rücklauf der einzelnen Erhebungen, die Zahl der Panelfälle oder Forschungs- und bildungspolitische Themenschwerpunkte wurden für alle Wellen in der Exceldatei „Querschnitt- und Panelcharakteristika des BIBB-Qualifizierungspanels“ von HOHN u. a. 2020 gebündelt und stehen auf der BIBB-FDZ Metadatenportalseite des BIBB-Qualifizierungspanels zur Verfügung.

## 1.2 Aufbau des Handbuchs

In Abschnitt 1.3 erfolgt eine steckbriefartige Beschreibung des Datensatzes und Abschnitt 1.4 gibt einen Überblick über die zentralen Eigenschaften des Datensatzes. In Kapitel 2 wird das BIBB-Qualifizierungspanel beschrieben, u. a. das Stichprobenkonzept erläutert und der Ablauf der Befragung dargelegt. Kapitel 3 umfasst den Aufbau der Datensätze, erläutert, welche Teildatensätze es gibt und wodurch sie gekennzeichnet sind, und skizziert die Harmonisierung der Variablen über die Zeit. Darauf folgen in Kapitel 4 die Ausführungen zur Organisation der Mikrodaten, im Einzelnen zu fehlenden Werten, zur Imputation, zu im Fragebogen nicht erkennbaren Filtern und zu Besonderheiten einzelner Jahre sowie zu der Möglichkeit, den Datensatz ins Englische zu übersetzen. Kapitel 5 thematisiert die Gewichtung- und Hochrechnungsfaktoren. Abschließend informiert Kapitel 6 über die Fortschreibung des long-Datensatzes und Kapitel 7 über die Anonymisierung und den Datenzugang.

## 1.3 Kurzbeschreibung des Datensatzes

Der Längsschnittdatensatz des BIBB-Betriebspanels zu Qualifizierung und Kompetenzentwicklung 2011 bis 2017 umfasst einen Hauptdatensatz und zehn ihn ergänzende Teildatensätze (Details siehe Abschnitt 3). Der Hauptdatensatz enthält die Variablen, die zwischen 2011 und 2017 regelmäßig abgefragt wurden. Ein erster Teildatensatz enthält die sporadisch mehrfach abgefragten Variablen, ein zweiter Teildatensatz enthält die Gewichtung- und Hochrechnungsfaktoren und die übrigen acht Teildatensätze<sup>1</sup> enthalten die Variablen, die nur im entsprechenden Erhebungsjahr abgefragt wurden. Der Gesamtdatensatz umfasst 24.075 Beobachtungen von 13.473 Betrieben und insgesamt 1.172 Variablen.

---

1 Für das Jahr 2016 gibt es zwei Teildatensätze, da in diesem Jahr eine zusätzliche CATI-Befragung erfolgte.

## 1.4 Zentrale Eigenschaften im Überblick

Befragungstitel	BIBB-Betriebspanel zu Qualifizierung und Kompetenzentwicklung 2011 bis 2017 long
DOI	10.7803/371.1117.1.2.10
Kurzbeschreibung	Betriebspanel mit dem Ziel, detaillierte Informationen über die Strukturen, Entwicklungen und Zusammenhänge betrieblicher Qualifizierungsmaßnahmen und betrieblich-qualifikatorischer Arbeitskräftenachfrage zu liefern
Erhebungsjahr	2011 bis 2017
Erhebungseinheit	Betriebe
Themenschwerpunkt	Ausbildung, Weiterbildung
Datenzugangsmöglichkeiten	Gastwissenschaftleraufenthalt/Datenfernverarbeitung
Variablenanzahl	1.172
Grundgesamtheit	Alle Betriebe in Deutschland mit mindestens einem sozialversicherungspflichtig Beschäftigten
Gewichtung/Hochrechnung	Hochrechnungsfaktoren (auf Basis der Betriebsdatei der BA) Längsschnittgewichte
Repräsentative Region	Ost- und Westdeutschland
Fallzahl	24.075 Beobachtungen von 13.473 Betrieben
Erhebungsverfahren	CAPI (PAPI oder CAWI; auf Wunsch des Betriebes)
Auswahlverfahren	Disproportional geschichtete Zufallsauswahl
Erhebungsdesign	Längsschnittdesign (Panelstudie)
Bemerkung	Der Datensatz ist in elf Teildatensätze zerlegt. Der Hauptdatensatz ist im long-Format gespeichert.
Links	BIBB-FDZ Metadatenportal: <a href="https://metadaten.bibb.de/metadatengruppe/16">https://metadaten.bibb.de/metadatengruppe/16</a> Projektseite: <a href="http://www.bibb.de/qp">http://www.bibb.de/qp</a>
Stichworte	Qualifizierung, Ausbildung, Personalbewegungen, Arbeitsorganisation, Personalstruktur, Ausbildung von Geflüchteten, Digitalisierung, Tätigkeitsprofile, Vertragslösungen, Organisation und Durchführung von Weiterbildung

## 2 Überblick über das BIBB-Qualifizierungspanel

Das BIBB-Qualifizierungspanel ist eine repräsentative, seit 2011 jährlich durchgeführte Panelbefragung von etwa 2.000, ab 2014 etwa 3.500 Betrieben (vgl. HOHN u. a. 2020) in Deutschland. Der Fokus der Studie liegt auf betrieblichen Rahmenbedingungen und Entwicklungen in Bezug auf betriebliche Aus- und Weiterbildung, der Personalstruktur und der Arbeitsorganisation innerhalb eines Betriebs. Allerdings kommen in allen Befragungswellen weitere Themenschwerpunkte hinzu, wie etwa die betriebliche Digitalisierung oder die betriebliche Ausbildung und Beschäftigung von Menschen mit Fluchthintergrund.<sup>2</sup>

In der sechsten Erhebungswelle (2016) wurde die Erhebung zum Schwerpunkt Digitalisierung in besonderer Weise durchgeführt: Der zusätzliche Themenschwerpunkt wurde parallel zur Hauptstudie in computergestützten Telefoninterviews (Computer Assisted Telephone Interview, „CATI“) bei 3.500 zusätzlichen Betrieben, also insgesamt rund 7.000 Betrieben, abgefragt.

Die Grundgesamtheit, aus der die Stichproben des BIBB-Qualifizierungspanels gezogen werden, umfasst alle Betriebe in Deutschland mit mindestens einem sozialversicherungspflichtig Beschäftigten zum 31.12. des Vorjahres der Erhebung. Ein Betrieb wird als wirtschaftlich und regional abgegrenzte Einheit mit mindestens einem sozialversicherungspflichtig Beschäftigten definiert (vgl. FISCHER u. a. 2008, S. 7; GERHARDS u. a. 2012). Die Stichprobe, genauer: zunächst die Bruttostichprobe wird aus der Betriebsdatei der Bundesagentur für Arbeit (BA) gezogen. Die Nettostichprobe unterteilt sich in wiederbefragte Betriebe und sogenannte Auffrischer (Betriebe), die benötigt werden, um die Panelmortalität auszugleichen. Im Jahr 2017 wurden für die Auffrischungstichprobe ausschließlich Adressen von wiederbefragungsbereiten Teilnehmern und Teilnehmerinnen der genannten gesonderten CATI-Befragung aus der Welle 2016 verwendet.

Die Bruttostichprobe, bestehend aus allen Betrieben, die kontaktiert werden, wird nach vier Kriterien disproportional geschichtet gezogen:

- ▶ die Regionalverteilung der Betriebe nach Ost- bzw. Westdeutschland,
- ▶ die Betriebsgröße, unterteilt in vier Klassen (1–19 Beschäftigte, 20–99 Beschäftigte, 100–199 Beschäftigte sowie 200 und mehr Beschäftigte),
- ▶ dem Wirtschaftszweig, unterteilt in acht Klassen (Landwirtschaft/Bergbau/Energie; verarbeitendes Gewerbe; Bau; Handel und Reparatur; unternehmensnahe Dienstleistungen; personennahe Dienstleistungen; medizinische Dienstleistungen; öffentlicher Dienst und Erziehung),
- ▶ ob es sich um einen Ausbildungsbetrieb handelt oder nicht.

Die Stichprobenziehung des BIBB-Qualifizierungspanels erfolgt disproportional geschichtet, da ansonsten zu kleine Fallzahlen auftreten können, um sinnvolle statistische Analysen durchführen zu können. Konkret wird der Anteil von Ausbildungsbetrieben, Betrieben in Ostdeutschland, Betrieben aus bestimmten Wirtschaftszweigen, etwa dem verarbeitenden Gewerbe, und von Betrieben mit mindestens 200 sozialversicherungspflichtig Beschäftigten in der Stichprobe höher als der Anteil dieser Gruppen in der untersuchten Grundgesamtheit angesetzt.

Die Befragungen werden in Form von Computer Assisted Personal Interviews (CAPI) durchgeführt, wobei Betrieben, die kein persönliches Interview durchführen wollen, seit 2016 die Möglichkeit zur Teilnahme via Computer Assisted Web Interview (CAWI) mit identischen

---

2 Für eine Auflistung aller thematischen Schwerpunkte vgl. HOHN u. a. 2020.

Fragen angeboten wird. Bis 2015 wurde statt des CAWI ein schriftlicher Selbstausfüllerbogen (Paper and Pencil Interview, PAPI) als Alternative zum CAPI angeboten.

Vor der Feldphase wird der gesamte Fragebogen in einem Pretest mit ca. 30 Betrieben auf Verständnisschwierigkeiten, Länge, Komplexität und Programmierfehler getestet. Die daraufhin erstellte finale Fassung des CAPI- und CAWI-Fragebogens wird anschließend im BIBB getestet, bevor er im Feld eingesetzt wird. Die Feldphase beginnt in der Regel Anfang Februar und läuft bis Anfang September.

Die Interviews des Qualifizierungspanels wurden bis 2014 vom Forschungsinstitut „TNS Infratest Sozialforschung“<sup>3</sup>, ab 2015 vom „infas Institut für angewandte Sozialwissenschaft“ durchgeführt. Das Erhebungsinstitut schickt zuvor geschulte Interviewer/-innen zu den teilnehmenden Betrieben, um die Interviews dort abzuhalten. Als Interviewpartner/-innen vonseiten der befragten Betriebe werden Personen mit ausreichendem Überblick über die innerbetrieblichen Aktivitäten benötigt, um den Fragebogen ausfüllen zu können. Typischerweise handelt es sich dabei um Eigentümer/-innen, Anteilseigner/-innen oder die Leitung der jeweiligen Personalabteilung. Die Auswahl des Interviewpartners/der Interviewpartnerin innerhalb des Betriebs wird dem Betrieb selbst überlassen.

Die Feldphase beginnt mit der schriftlichen Kontaktierung der Betriebe aus der Bruttostichprobe, in der sie über den Start der nächsten Erhebungswelle und den aktuellen thematischen Schwerpunkt informiert und um Teilnahme gebeten werden. Außerdem wird die bevorstehende Kontaktaufnahme durch eine/-n Interviewer/-in mit dem Betrieb angekündigt. Zusätzlich enthält das Schreiben eine Datenschutzerklärung vom BIBB und dem durchführenden Erhebungsinstitut. Seit 2016 werden die Betriebe der Auffrischer-Stichprobe und Betriebe, für die kein/-e Ansprechpartner/-in bekannt war, telefonisch vorkontaktiert.

Im Rahmen der Kontaktaufnahme durch die Interviewer/-innen lassen diese den Betrieben auf Wunsch das Schreiben erneut zukommen. Bei Bedarf wird auch eine Übersicht der Erhebungsfragen übergeben, damit sich der/die jeweilige Interviewpartner/-in ausreichend auf das Interview vorbereiten kann. Sollte eine Teilnahme am CAPI vonseiten des Betriebes abgelehnt werden, bieten die Interviewer/-innen eine Teilnahme an der CAWI-Version der Befragung an (in früheren Wellen PAPI).

Durchschnittlich hatte das BIBB-Qualifizierungspanel zwischen 2011 und 2017 einen Rücklauf<sup>4</sup> von etwa 36,8 Prozent und eine Panelbereitschaft der Betriebe von etwa 85 Prozent (HOHN u. a. 2020).

---

3 TNS Infratest Social Research hat sich mittlerweile zu „Kantar TNS“ umfirmiert.

4 Rücklauf wird im BIBB-Qualifizierungspanel definiert als die Anzahl der auswertbaren Interviews geteilt durch die Anzahl der eingesetzten Adressen.

## 3 Organisation der Datensätze

### 3.1 Datensätze

Insgesamt umfasst das *BIBB-Qualifizierungspanel 2011 bis 2017 long* elf Teildatensätze (vgl. im Folgenden Tabelle 1), zwei Längsschnittdatensätze, einen Datensatz mit den Gewichtung- sowie Hochrechnungsfaktoren und je einen Querschnittdatensatz pro Erhebungsjahr. Aufgrund der zusätzlichen CATI-Befragung bildet das Jahr 2016 eine Ausnahme, da es für dieses Erhebungsjahr zwei Querschnittdatensätze gibt.

Tabelle 1: Übersicht Teildatensätze

	Datensatz	Inhalt	Variablen (Spalten)	Beobachtungen (Zeilen)	Betriebe
1	Qualifizierungspanel_11bis17_long_gwa	<b>Hauptdatensatz</b> , Längsschnittdatensatz, alle Variablen, die mindestens zwischen 2015 und 2017 enthalten sind oder seit 2016 dauerhaft abgefragt werden	260	24.075	13.473
2	Qualifizierungspanel_11bis17_long2_gwa	Längsschnittdatensatz, alle anderen Variablen, die diesen Kriterien nicht entsprechen, aber mindestens zweimal erhoben wurden	122	24.075	13.473
3	Qualifizierungspanel_11bis17_weight_gwa	Gewichtungs- sowie Hochrechnungsfaktoren	27	24.075	13.473
4	Qualifizierungspanel_11bis17_quer_2011_gwa	Querschnittdatensatz 2011	137	2.004	2.004
5	Qualifizierungspanel_11bis17_quer_2012_gwa	Querschnittdatensatz 2012	118	2.023	2.023
6	Qualifizierungspanel_11bis17_quer_2013_gwa	Querschnittdatensatz 2013	102	2.063	2.063
7	Qualifizierungspanel_11bis17_quer_2014_gwa	Querschnittdatensatz 2014	40	3.523	3.523
8	Qualifizierungspanel_11bis17_quer_2015_gwa	Querschnittdatensatz 2015	67	3.589	3.589
9	Qualifizierungspanel_11bis17_quer_2016_gwa	Querschnittdatensatz 2016_gwa	41	3.616	3.616
10	Qualifizierungspanel_11bis17_11bis17_quer_2016_CATI_gwa	Querschnittdatensatz 2016 CATI-Zusatzerhebung	45	3.521	3.521
11	Qualifizierungspanel_11bis17_quer_2017_gwa	Querschnittdatensatz 2017_gwa	213	3.727	3.727

Alle Variablen, die mindestens zwischen 2015 und 2017 enthalten sind<sup>5</sup> oder seit 2016 dauerhaft abgefragt werden, sind im Datensatz „Qualifizierungspanel\_11bis17\_long\_gwa“ enthalten. Alle anderen Variablen, die diesen Kriterien nicht entsprechen, aber mindestens zweimal erhoben wurden, sind im Datensatz „Qualifizierungspanel\_11bis17\_long2\_gwa“ enthalten. Alle Variablen, die lediglich einmal erhoben wurden, sind in einem Jahresdatensatz „Qualifizierungspanel\_11bis17\_quer\_20\*\_gwa“ enthalten (\* = 11 bis 17).

Der Hauptdatensatz „Qualifizierungspanel\_11bis17\_long\_gwa“ bildet das Kernstück des *BIBB-Qualifizierungspanel 2011 bis 2017 long*. Der Datensatz ist im sogenannten langen (long) Format gespeichert. Das bedeutet, dass Betriebe im Datensatz so oft enthalten sind, wie sie an Befragungswellen teilgenommen haben. Jede Variable (z. B. well (Wellenkennzeichen)) ist dagegen nur einmal vorhanden, d. h., die Informationen sind zeilenweise abgelegt<sup>6</sup>.

Insgesamt umfasst der Hauptdatensatz 260 Variablen und 24.075 Beobachtungen von 13.473 verschiedenen Betrieben, denn Betriebe haben unterschiedlich oft an Erhebungswellen teilgenommen. Die Anzahl der Erhebungswellen, an denen diese Betriebe teilgenommen haben, kann Tabelle 2 entnommen werden. Im Durchschnitt haben die Betriebe an 2,5 Befragungen teilgenommen, wobei etwa die Hälfte nur einmal teilnahm, und ca. ein Prozent an allen sieben Wellen.

**Tabelle 2: Häufigkeit der Teilnahme der Betriebe am BIBB-Qualifizierungspanel zwischen 2011 und 2017**

Summe der Teilnahmen bis 2017	Absolute Anzahl	Prozent
1	7.471	55,45
2	3.390	25,16
3	1.496	11,10
4	670	4,97
5	170	1,26
6	126	0,94
7	150	1,11
<b>Gesamt (Anzahl Betriebe)</b>	<b>13.473</b>	<b>100,00</b>

Quelle: *BIBB-Qualifizierungspanel 2011 bis 2017 long*

Insgesamt gibt es 71 Muster für die Teilnahme der Betriebe; in Stata kann man sie sich mit dem Befehl `xtdes, pattern(71)` anzeigen lassen. Tabelle 3 enthält häufige Muster und die Anzahl, wie viele Betriebe in den entsprechenden Zeiträumen teilgenommen haben. So haben 2.097 Betriebe nur zwischen 2016 und 2017 teilgenommen. Bezieht man jedoch alle andern möglichen Muster mit ein, stehen für die Auswertung der beiden Jahre 3.607 Betriebe zur Verfügung.

5 Es gibt vier Ausnahmen, bei denen die Variablen im long2-Datensatz enthalten sind, obwohl sie genau zwischen 2015 und 2017 abgefragt wurden. Der Grund ist, dass die Variablen seit 2018 nicht mehr erhoben werden.

6 Im sogenannten breiten (wide) Format dagegen wäre jeder Betrieb nur einmal vorhanden und jede Variable für jedes Erhebungsjahr einmal (mit entsprechend unterschiedlichen Namen, z.B. well2011, well2012...well2017) genannt, d. h., die Informationen wären spaltenweise abgelegt.

**Tabelle 3: Muster der Teilnahme der Betriebe am BIBB-Qualifizierungspanel zwischen 2011 und 2017**

Zeitspanne	Anzahl teilnehmender Betriebe, nur in dieser Zeitspanne	Anzahl teilnehmender Betriebe, insgesamt in dieser Zeitspanne
2016-2017	2.097	3.607
2015-2017	906	1.496
2015-2016	302	1.969
2014-2017	294	582
2011-2014	287	561
2011-2013	285	858
2011-2017	150	150

Quelle: BIBB-Qualifizierungspanel 2011 bis 2017 long

Der Hauptdatensatz enthält, wie bereits ausgeführt, die Panelvariablen, welche regelmäßig abgefragt werden. Um welche es sich handelt, ist in einem eigens für die Dokumentation erstellten Fragebogen nachvollziehbar, der auf den Seiten des BIBB-FDZ Metadatenportals heruntergeladen werden kann. Doch auch die Variablen des long-Datensatzes sind nicht durchgängig für alle Erhebungsjahre abgefragt worden. So werden beispielsweise Fragen zu Geflüchteten oder zu Digitalisierung erst ab 2015 bzw. 2016 gestellt oder die Anzahl der Weiterbildungsteilnehmer 2011 und 2013 nicht getrennt nach Beschäftigungsgruppen erhoben. Welche Variablen und Items nicht für alle Jahre zu Verfügung stehen, kann im Anhang Tabelle A1 entnommen werden.

Der zweite long-Datensatz „Qualifizierungspanel\_11bis17\_long2\_gwa“ umfasst 122 Variablen und ebenfalls 24.075 Beobachtungen von 13.473 verschiedenen Betrieben. Dieser Datensatz enthält alle Variablen, die mehr als einmal, aber nicht dauerhaft abgefragt wurden, etwa die Fragen zu den Tätigkeiten der Beschäftigten oder zum prozentualen Leistungsniveau von Auszubildenden. Welche Variablen genau im Datensatz enthalten sind und in welchem Jahr sie abgefragt wurden, kann der Variablenliste im Anhang (vgl. Tabelle A1) entnommen werden. Der Datensatz ist ebenfalls im long-Format gespeichert.

Der Datensatz „Qualifizierungspanel\_11bis17\_weight\_gwa“ enthält 27 Variablen und dieselbe Anzahl an Beobachtungen und Betrieben wie die beiden long-Datensätze und ist ebenfalls im long-Format gespeichert. In Kapitel 5 „Gewichtung und Hochrechnung“ wird der Datensatz genauer beschrieben.

Die Querschnittdatensätze enthalten alle Variablen, die einmalig im entsprechenden Jahr erhoben bzw. in besonderer Art und Weise erfragt wurden. Die Anzahl an Variablen nimmt im Zeitverlauf tendenziell ab (vgl. Tabelle 1), da das Set an Panelfragen über die Jahre stabiler und umfangreicher geworden ist. Der jeweilige Querschnittdatensatz enthält neben den eigentlichen Querschnittsfällen auch jene Fälle, die nicht zum eigentlichen Querschnitt des entsprechenden Jahres gehören, weil sie nicht Teil der Grundgesamtheit sind oder der Betrieb erloschen ist. Man erkennt diese Betriebe daran, dass sie für alle inhaltlichen Variablen den Wert -11 ‚Filter (Betrieb erloschen)‘<sup>7</sup> enthalten. Einen Überblick über die Variablen der Querschnitte bieten die Variablenlisten, die im BIBB-FDZ Metadatenportal zu finden sind. Die dem

7 Im Datensatz ‚Qualifizierungspanel\_quer\_2012\_11bis17\_gwa‘ ist bei den Variablen b43\*p\_gen der Wert ‚-1111. Filter (Betrieb erloschen)‘, da -11 theoretisch ein gültiger Wert sein könnte.

Längsschnittdatensatz zugeordneten Querschnittdatensätze sind speziell für das Anspielen von Variablen an den Längsschnitt gedacht. Für die Auswertung des Querschnitts einzelner Jahre sollte auf die separat im BIBB-FDZ bisher veröffentlichten Querschnittdatensätze zurückgegriffen werden. Diese Datensätze bieten den Vorteil, dass sie alle in einem Jahr erhobenen Variablen für in diesem Jahr teilnehmende Betriebe enthalten.

In den long-Datensätzen definieren die Variablen *id2* und *jahr* eindeutig eine Datenzeile. Diese beiden Variablen müssen genutzt werden, um alle Datensätze aneinanderzuhängen. Die entsprechenden Befehle lauten:

#### Stata

```
use „Qualifizierungspanel_11bis17_long_gwa.dta“
sort id2 jahr
merge 1:1 id2 jahr using „Qualifizierungspanel_11bis17_long2_gwa.dta“,
keepusing(varlist)
cap drop _merge
merge 1:n id2 jahr using „Qualifizierungspanel_11bis17_quer_2011_gwa.
dta“, keepusing(varlist)
.
.
.
cap drop _merge
merge 1:n id2 jahr using „Qualifizierungspanel_11bis17_quer_2017_gwa.
dta“, keepusing(varlist)
cap drop _merge
merge 1:1 id2 jahr using „Qualifizierungspanel_11bis17_weight_gwa.dta“,
keepusing(varlist)
cap drop _merge
```

#### SPSS

```
GET FILE=“*\Qualifizierungspanel_11bis17_long_gwa.sav“.
SORT CASES BY id2 jahr.
MATCH FILES /FILE=*
  /FILE='*\Qualifizierungspanel_11bis17_long2_gwa.sav'
  /RENAME (pan_ges quer well = d0 d1 d2)
  /BY id2 jahr
  /DROP= d0 d1 d2.
EXECUTE.
MATCH FILES /FILE=*
  /TABLE='*\Qualifizierungspanel_11bis17_quer_2011_gwa.sav'
  .
  .
  .
  /TABLE='*\Qualifizierungspanel_11bis17_quer_2017_gwa.sav'
  /FILE='*\Qualifizierungspanel_11bis17_weight_gwa.sav'
  /BY id2 jahr.
EXECUTE.
```

## 3.2 Benennung der inhaltlichen Variablen

Die Benennung der inhaltlichen Variablen folgt bis 2015 der Fragennummerierung aus dem Fragebogen. Die Namen der inhaltlichen Variablen beginnen für Welle 1 mit „a“, für Welle 2 mit „b“ bis „e“ für Welle 5. Gefolgt wird dieses Wellenkennzeichen von der zweistelligen Fragennummer im Fragebogen. Besteht eine Frage aus Unterfragen, so sind diese durch einen Zusatz gesondert aufgeführt (z.B. „a01a“ und „a01b“). Ab 2016 wurde dieses Benennungssystem abgelöst, und es wurde ein Variablenbenennungssystem entwickelt, welches über die Wellen stabil bleibt. Jede Variable besteht aus zwei bis drei Teilen:

1. einem Präfix aus zwei Buchstaben, welche den Themenbereich kennzeichnen, dem die Variable zugeordnet werden kann (vgl. Tabelle 4);
2. einer laufenden Nummer;
3. ggf. einem Suffix, bei dem es sich um eines aus den bisherigen Wellen genutzten Kürzel handelt (vgl. HOHN u. a. 2020).

**Tabelle 4: System der Variablennamen**

Präfixe Variablen	Themenbereich
ab	Ausbildung
ps	Personalstruktur
pb	Personalbewegung
wb	Weiterbildung
mo	Wechselnde Module
di	Digitalisierung
be	Allgemeine Informationen zum Betrieb

In den Jahresdatensätzen vor 2016 wurde die originale Benennung beibehalten, wenn die Frage nur einmalig gestellt wurde. Dagegen wurde die neue Benennung genutzt, d.h., es wurden Variablen rückwirkend umbenannt, wenn die Frage der neuen Logik zuordenbar war. Das ist auch der Fall, wenn die gesamte Frage dem neuen Benennungssystem zuordenbar ist, aber ein einzelnes Item aus einer Batterie sich nicht anderen Items aus anderen Wellen zuordnen lässt. Um das Auffinden in den ursprünglichen Fragebögen zu erleichtern, wurde die Fragennummer im Variablenlabel ergänzt. Es folgen zwei Beispiele aus dem Querschnittdatensatz 2011 für das Benennungsschema und die Harmonisierung von Variablennamen (vgl. Tabelle 5):

**Tabelle 5: Beispiel für das System der Variablennamen im Querschnitt**

Variable	Variablenlabel	Erläuterung
a02	Bedarf an weiterer Qualifizierung der beschäftigten Mitarbeiter	Frage wurde nur 2011 gestellt, sie folgt dem Schema a= Welle2011 (b=2012)
ab058_4	@A03d: Ausbildung verursacht hohe Kosten	Frage ab058 wurde auch in anderen Wellen gestellt, sie folgt dem Schema „ab“= Frage zum Thema Ausbildung. Das Item _4 ist im Querschnitt enthalten, weil es nur einmalig in Welle 2011 gestellt wurde. Die Benennung mit dem Präfix ab058 erleichtert das Wiederauffinden zusammengehöriger Items über mehrere Wellen.

### 3.3 Harmonisierung der Variablen für die Längsschnittdatensätze

Damit alle Variablen in den long-Datensätzen einheitlich benannt sind, wurden die Variablenamen der Wellen vor 2016 an die 2016 eingeführte aktuelle Benennungssystematik angepasst. Um das Wiederauffinden der Variablen aus den Jahren vor 2016 in den Fragebögen zu erleichtern, steht für die Variablen des long- und long2-Datensatzes je eine Zuordnungstabelle auf den Seiten des BIBB-FDZ Metadatenportals zur Verfügung.

Weiterhin wurden in allen Variablen Fälle in 9997 bzw. 997 oder 97 codiert, wenn die entsprechende Frage in der jeweiligen Welle nicht abgefragt wurde. Weiterhin wurden alle Fälle in 9996 bzw. 96 (in einem Fall auch 95) codiert, wenn zwar die Frage, aber das entsprechende einzelne Item in der jeweiligen Welle nicht abgefragt wurde. Darüber hinaus wurden in allen imputierten Variablen Fälle in 998, 9998, 99998 oder in -98<sup>8</sup> codiert, wenn in der jeweiligen Welle keine Imputation dieser Variablen stattfand. Diese Kategorien sollen bei einer ersten Sichtung der Daten dabei helfen, sich darin zurechtzufinden. Für die Analyse der Daten kann es sinnvoll sein diese Werte als System Missings zu definieren (für eine entsprechende Programmierhilfe siehe Kapitel 4.3).

In den Jahren 2012 bis 2014 waren im Querschnitt Betriebe enthalten, die bereits erloschen waren, aber per Definition dennoch zum Panel gehörten und entsprechend in die Gewichtung einbezogen wurden (vgl. GERHARDS/FRIEDRICH 2016). Die betroffenen Betriebe wurden jeweils mit ‚-11. Filter (Betrieb erloschen)‘ gekennzeichnet, wenn die entsprechende Variable für das jeweilige Jahr gefüllt ist. Beginnend mit der Welle 2015 wurde die Definition, welche Betriebe Panelbetriebe sind, geändert. Erlöschene Betriebe gehören demnach nicht mehr zu den Panelbetrieben und werden nicht in die Berechnung der Gewichte einbezogen (vgl. FRIEDRICH/GERHARDS 2017).

Zum Teil gab es leichte Änderungen in Frageformulierungen von Panelfragen, entsprechende Hinweise stehen (bei Stata) in den Notes dokumentiert (Befehle zum Anzeigen der notes: `notes varname`) und können darüber hinaus HOHN u. a. 2020 entnommen werden.

Die über die bisher genannten Schritte hinausgehende Harmonisierung einzelner Variablen der beiden long-Datensätze auf Variablenebene ist im Anhang in Tabelle A3 dokumentiert. Die Variablen wurden so verändert, dass sie trotz jährlicher Unterschiede in der Erhebung (etwa unterschiedliche Anzahl an Antwortkategorien oder getrennte Erfassung von Angaben für gewerblich-technische und kaufmännisch-verwaltende Auszubildende) in den long-Datensätzen für alle Jahre einheitlich sind.

Die Variablen der Querschnitte sind nicht harmonisiert worden. Eine Ausnahme bilden die Variablen *wb011\** aus dem Jahr 2011. In diesen wurde die Kategorie ‚9999. Beschäftigtengruppe nicht vorhanden‘ nachträglich codiert.

Zusätzlich zu den aufgeführten Harmonisierungen erfolgten Anpassungen der internen Forschungsdaten des BIBB-Qualifizierungspanels gegenüber den bisher veröffentlichten Querschnittdatensätzen. Dabei handelt es sich um Korrekturen bzw. Aktualisierungen der Daten, die im Zuge der Projektarbeit von den Projektmitarbeiterinnen und -mitarbeitern vorgenommen wurden. Diese Anpassungen sind in den hier behandelten Daten enthalten, werden aber nicht im Einzelnen aufgeführt.

8 Für das imputierte Geschäftsvolumen wurde ein negativer Wert codiert, weil Wert, der hoch genug wäre, um höchster Wert zu sein, sich systembedingt nicht labeln lässt.

## 4 Organisation der Mikrodaten

### 4.1 Imputierte Variablen

Eine Imputation erfolgt im BIBB-Qualifizierungspanel ausschließlich bei kontinuierlichen Variablen mit hoher Item-Non-Response (i. d. R. über 10 %). Imputiert wurden alle Fälle, die entweder keine Angabe gemacht haben oder mit ‚weiß nicht‘ geantwortet haben; Missings oder Extremwerte wurden nicht imputiert. Das Ziel der Imputation ist es, zum einen Verzerrungen durch Item-Non-Response auszugleichen und zum anderen die Fallzahl für Analysen zu erhöhen.

Anhand eines gemeinsamen Regressionsmodells, welches u. a. wichtige Strukturvariablen berücksichtigt, wird ein linearer Vorhersagewert berechnet. Panelinformationen, also Werte aus den Vorjahren, werden für die Imputation teilweise genutzt. Als imputierter Wert wird der lineare Vorhersagewert des Regressionsmodells unter Addition eines Störterms bei je 1.000 Wiederholungen verwendet (für eine ausführliche Beschreibung des Imputationsverfahrens siehe GERHARDS/FRIEDRICH 2016, Kapitel 4).

Die imputierten Variablen sind jeweils mit dem Kürzel „\_imp“ versehen und zusätzlich zu den ursprünglichen Variablen im Datensatz enthalten. Es bleibt jedem/jeder Nutzenden selbst überlassen, ob er/sie die imputierte oder originale Variable benutzt. Da immer die metrischen Variablen mit einem hohen Item-Non-Response imputiert werden und kein festes Set an Variablen, ist es über die Jahre nicht einheitlich, welche Variablen imputiert wurden. Sobald eine Variable für ein Jahr imputiert wurde, ist sie im long- bzw. long2-Datensatz enthalten.

### 4.2 Im Fragebogen nicht erkennbare Filter

Zu einigen Variablen wurden bereits im jeweiligen Jahr bei der Aufbereitung der Querschnitte nachträglich im Fragebogen nicht ersichtliche Filter codiert. Dafür gibt es zwei Gründe:

Erstens sollen bei der Frage nach den Ausbildungsberufen und der Anzahl der in diesen beschäftigten Auszubildenden Betriebe ohne Auszubildende von solchen Betrieben abgegrenzt werden, die in weniger als fünf Berufen ausbilden. Zur leichteren Unterscheidung dieser Fälle von den anderen fehlenden Werten wurde für diese Fälle der Wert ‚99999. kein weiterer Ausbildungsberuf genannt‘ vergeben.

Zweitens sollten von den Befragten bei Fragen, die sich auf bestimmte Beschäftigtengruppen beziehen, solche Betriebe abgegrenzt werden, denen die Frage aufgrund des Fehlens der entsprechenden Beschäftigtengruppe nicht gestellt wurde, ohne dass es im Fragebogen einen expliziten Filter gibt. Gefiltert wurde nach der Anzahl der Beschäftigten in der jeweiligen Beschäftigtengruppe (ps20\*) und der Anzahl ausgeschiedener Beschäftigter in der jeweiligen Beschäftigtengruppe (pb017\*). Die Befehle zum Nachvollziehen der Filter lauten:

#### Stata

```
gen filter_gering=.
replace filter_gering=1 if ps020_einf>0 | pb017_einf>0 & pb017_
einf<9997
replace filter_gering=0 if ps020_einf==0 & pb017_einf==0

gen filter_qual=.
replace filter_qual=1 if ps020_qual>0 | pb017_qual>0 & pb017_einf<9997
```

```

replace filter_qual=0 if ps020_qual==0 & pb017_qual==0
gen filter_hoch=.
replace filter_hoch=1 if ps020_hoch>0 | pb017_hoch>0 & pb017_einf<9997
replace filter_hoch=0 if ps020_hoch==0 & pb017_hoch==0

```

### SPSS

```

compute filter_gering=$Sysmis.
if ps020_einf>0 | pb017_einf>0 & pb017_einf<9997 filter_gering=1.
if ps020_einf=0 & pb017_einf=0 filter_gering=0.

compute filter_qual=$Sysmis.
if ps020_qual>0 | pb017_qual>0 & pb017_einf<9997 filter_qual=1.
if ps020_qual=0 & pb017_qual=0 filter_qual=0.

compute filter_hoch=$Sysmis.
if ps020_hoch>0 | pb017_hoch>0 & pb017_einf<9997 filter_hoch=1.
if ps020_hoch=0 & pb017_hoch=0 filter_hoch=0.

EXECUTE.

```

Alle Fälle, die auf den entsprechenden Variablen nicht Wert 1 aufweisen, wurden jeweils mit dem Code 9999, 999 oder 99 ‚Beschäftigtengruppe nicht vorhanden‘ oder ‚versehen. Eine Ausnahme in der Codierung bildet das Jahr 2012, in welchem die Information zu den ausgeschiedenen Beschäftigten fehlt (vgl. Kapitel 4.5.2).

## 4.3 Fehlende Werte

Fehlende und ungültige Werte wurden nach der Standarddeklarierung fehlender Werte des BIBB-FDZ umcodiert, um so den Umgang mit den fehlenden Werten zu erleichtern. Dabei wurden Besonderheiten der Codierung aus vorherigen Wellen für alle Variablen des BIBB-Qualifizierungspanels beachtet. HOHN u. a. 2020 (Tabellenblatt „Fehlende Werte 13“) enthält die Werte und Wertelabels für die fehlenden Werte aller Variablen des BIBB-Qualifizierungspanels.

In Datensätzen des BIBB-FDZ werden grundsätzlich keine System Missings definiert – deren Vergabe wird den Anwendenden überlassen. System Missings erhalten eindeutige nicht fehlende Werte. Eine Ausnahme bilden die Gewichtung- und Hochrechnungsfaktoren, welche System Missings enthalten.

Nach der ersten Sichtung der Daten kann es für Analysen sinnvoll sein, sowohl fehlende Werte als auch die Kategorien für im entsprechenden Jahr nicht gestellte Fragen, nicht erfolgte Imputationen und im Fragebogen nicht erkennbare Filter als System Missings zu definieren. Die entsprechenden Befehle lauten:

### Stata

```

*****long
***Frage im entsprechendem Jahr nicht gestellt
mvdecode ab002 ab036 ab057 ab065 ab069 ps036 pb001_ang wb014 wb001
wb017_* di001_* be001_* be010_* be043 be026*k ,mv(97=.a)
mvdecode ab034 ab037 ab067 ab068 ps037 wb003*,mv(997=.a)

```

```
mvdecode ab023_* ps001aus ps001soz ps001ger_v ps001aus_v ps001soz_v
pb005* pb011* pb017* wb002* wb010*,mv(9997=.a)
mvdecode ps001zwi ps001bea ps001svb_v ps001nsv_v ps001zwi_v ps001bea_v
ps030_sch ps030_sch_imp ,mv(99997=.a)
***Im entsprechenden Jahr keine Imputation
mvdecode ab034_imp ab037_imp wb003*_imp ,mv(998=.b)
mvdecode ab017_imp ab018*_imp ab023*_imp ab033_imp ab039*_imp
ab043*_imp pb003_imp pb016_imp ps015_2_imp pb003_f_imp pb005*_imp
pb010_imp pb011*_imp pb017*_imp wb002_imp wb010*_imp,mv(9998=.b)
mvdecode ab004_b*_imp ps001ger_imp ps001nsv_imp ps001soz_imp ps001zwi_
imp ps001svb_v_imp ps001ger_v_imp ps001nsv_v_imp ps017*_imp ps020*_
imp ps030_uni_imp wb009_imp,mv(99998=.b)
mvdecode be026_* ,mv(999998=.b)
mvdecode be023_imp ,mv(-98=.b)
***Beschäftigtengruppe nicht vorhanden
mvdecode be026_*k,mv(99=.c)
mvdecode pb003_f pb005* wb010* ,mv(9999=.c)
mvdecode ab004*pr_dl ,mv(99=.d)
mvdecode be026_* ,mv(999999=.c)
***Kein weiterer Ausbildungsberuf
mvdecode ab004* kldb2010_5er ab004_b* ,mv(99999=.d)
mvdecode ab004*pr_dl ,mv(99=.d)
***Item nicht abgefragt
mvdecode ab004*pr_dl ab058_5 ab058_6 ab058_7 wb017_4 be001_6
be001_7 be001_8 be001_9 be010_7 be010_8 be035_3 be035_4 be035_5
be035_6,mv(96=.e)
mvdecode ab043_*,mv(9996=.e)
***WB Teilnehmerfälle
mvdecode wb010*,mv(9995=.f)
***Missing
mvdecode *,mv(-12/-1)

*****long2
***Frage im entsprechendem Jahr nicht gestellt
mvdecode ab009 ab027* ab035 ab056 ab064 ab083 ab085* ab089 wb015_1
wb016_1 wb017_* wb018 di018* mo001* be022 be042_*,mv(97=.a)
mvdecode ab055* ab083* ab086* ab087* wb015_2* wb016_*,mv(997=.a)
mvdecode ab003* ab082* ps039* ,mv(9997=.a)
***Im entsprechenden Jahr keine Imputation
mvdecode ab083*_imp ,mv(998=.b)
mvdecode ab003*imp ab082*imp ,mv(9998=.b)
***Beschäftigtengruppe nicht vorhanden
mvdecode wb017* mo*,mv(99=.c)
mvdecode ps039*,mv(9999=.c)
***Ausbildungsberuf nicht besetzt
mvdecode ab086* ab087*,mv(999=.d)
***Item nicht abgefragt
mvdecode ab027_2 ab027_5 ab027_6 ,mv(96=.e)
mvdecode ab027_7 ab027_8,mv(95=.e)
***Missing
```

```
mvdecode *,mv(-12/-1)
```

### SPSS

```
*****long
```

```
***Frage im entsprechendem Jahr nicht gestellt.
```

```
missing values ab002 ab036 ab057 ab065 ab069 ps036 pb001_ang wb014  
wb001 wb017_1 to wb017_5 di001_2 to di001_8 be001_1 to be001_9 be010_1  
to be010_8 be043 be026_einfk to be026_hochk (97)
```

```
ab034 ab037 ab067 ab068 ps037 wb003kv to wb003so_imp (997)
```

```
ab023_1 to ab023_4_imp ps001aus ps001soz ps001ger_v ps001aus_v  
ps001soz_v pb005* pb011_einf to pb011_hoch_imp pb017_einf to pb017_  
hoch_imp wb002 wb002_imp pb005* wb010einf to wb010hoch_imp (9997).
```

```
***Im entsprechenden Jahr keine Imputation.
```

```
missing values ab034_imp ab037_imp wb003gt_imp to wb003so_imp (998)  
be023_imp (-98)
```

```
ab017_imp ab018_1_imp to ab018_4_imp ab023_1_imp to ab023_4_imp ab033_  
imp ab039_gt_imp ab039_kv_imp ab043_gt1_imp to ab043_kv3_imp pb003_imp  
pb016_imp ps015_2_imp pb003_f_imp
```

```
pb005_einf_imp to pb005_hoch_imp pb010_imp pb011_einf_imp to pb011_  
hoch_imp pb017_einf_imp to pb017_hoch_imp wb002_imp wb010einf_imp to  
wb010hoch_imp (9998).
```

```
***Beschäftigtengruppe nicht vorhanden.
```

```
missing values be026_einfk to be026_hochk (99)
```

```
pb003_f pb005_einf to pb005_hoch_imp wb010einf to wb010hoch_imp ab004_  
b2 to ab004_b5_imp (9999).
```

```
***Kein weiterer Ausbildungsberuf.
```

```
MISSING VALUES ab004_a1_kldb2010_5er to ab004_b5_imp (99999).
```

```
***Item nicht abgefragt.
```

```
MISSING VALUES ab004*pr_d1 ab058_5 ab058_6 ab058_7 wb017_4 be001_6  
be001_7 be001_8 be001_9 be010_7 be010_8 be035_3 be035_4 be035_5  
be035_6 (96)
```

```
ab043_gt1 to ab043_kv3_imp (9996).
```

```
***WB Teilnehmerfälle
```

```
MISSING VALUES wb010* (9995).
```

```
***Missing.
```

```
MISSING VALUES all (-12 thru-1).
```

```
*****long2.
```

```
***Frage im entsprechendem Jahr nicht gestellt.
```

```
MISSING VALUES ab009 ab027_1 to ab027_8 ab035 ab056 ab064 ab083  
ab085gt ab085kv ab089 wb015_1 wb016_1 wb017_einf1 to wb017_hoch5 wb018  
di018_1 to di018_6 mo001einf_1 to mo001hoch_8 be022 be042_1 to  
be042_3 (97)
```

```
ab055 ab055_imp ab083 to ab083kv_imp ab086gt1 to ab086kv4 ab087gt1 to  
ab087kv4 wb015_2 wb015_2_imp wb016_1 to wb016_2_imp (997)
```

```
ab003f to ab003kv_imp ab082gt1 to ab082kv4_imp ps039 ps039_imp (9997).
```

```
***Im entsprechenden Jahr keine Imputation.
```

```
MISSING VALUES ab083gt_imp ab083kv_imp (998)
```

```
ab003f_imp ab003gt_imp ab003kv_imp ab082gt1_imp to ab082kv4_imp  
(9998).
```

```

***Beschäftigtengruppe nicht vorhanden.
MISSING VALUES  wb017_einf1 to wb017_hoch5 mo001einf_1 to mo001hoch_8
(99)
ps039 ps039_imp (9999).
***Ausbildungsberuf nicht besetzt.
MISSING VALUES  ab086gt1 to ab086kv4 ab087gt1 to ab087kv4 (999).
***Item nicht abgefragt.
MISSING VALUES ab027_2 ab027_5 ab027_6 (96)
ab027_7 ab027_8 (95).
***Missing.
MISSING VALUES all (-12 thru -1).

```

## 4.4 Anzahl der Beschäftigten im Betrieb

Im BIBB-Qualifizierungspanel werden verschiedene Angaben zu den Beschäftigten erhoben, und zwar:

- ▶ sozialversicherungspflichtig Beschäftigte,
- ▶ geringfügig Beschäftigte,
- ▶ tätige Inhaber/-innen und mithelfende Familienangehörige,
- ▶ Auszubildende nach BBiG/HwO,
- ▶ Auszubildende des Gesundheitswesens,
- ▶ Beamtenanwärter/-innen.

Zusätzlich gibt es die Variable *ps001zwi*, welche die Summe aus sozialversicherungspflichtig und geringfügig Beschäftigten sowie tätigen Inhabern/-innen und mithelfenden Familienangehörigen enthält. Diese Variable ist als Basis für die Berechnung von Anteilen etwa für die Variablen *ps017\** (Anzahl Frauen, Teilzeitbeschäftigte und befristete Beschäftigte), *ps020\** (Anzahl nach Beschäftigtengruppe), *ps030\** (Anzahl nach Schulabschlüssen), *ps037* (Anzahl Geflüchtete), *wb002* und *wb003* (Anzahl Aufstiegsfortbildungen) sowie *wb009* und *wb010\** (Anzahl Weiterbildung) geeignet. In Einzelfällen kann es jedoch zu Werten über 1 kommen.

## 4.5 Besonderheiten einzelner Jahre

In einzelnen Wellen gibt es Abweichungen in der Bedeutung bestimmter Variablen oder Unterschiede in deren Erhebung. Die folgenden Kapitel informieren über die entsprechenden Besonderheiten und die sich daraus ergebenden Konsequenzen.

### 4.5.1 2011

Bei der Frage nach dem Umfang der Weiterbildung wurde 2011 nach der Anzahl von Teilnehmern und Teilnehmerinnen (Personen) bzw. Teilnahmefällen (*wb011\**) gefragt, in den übrigen Jahren hingegen nach der Anzahl der Weiterbildungsteilnehmer/-innen (*wb010\**). Anhand der Variable *a44* ist es möglich zu unterscheiden, ob 2011 Teilnehmer/-innen oder Teilnahmefälle berichtet wurden. Die Angaben zu den Teilnehmenden wurden für 2011 in die Variable *wb010\** übernommen (vgl. Anhang Tabelle A3), die Angaben für die Teilnahmefälle dagegen wurden in *wb010\** nicht berücksichtigt.

Würde man alle Werte aus *wb011\** in *wb010\** übertragen, würde man die Teilnehmer/-innen 2011 überschätzen, da Personen, die mehrfach an Weiterbildungen teilgenommen haben,

anders als in den Folgejahren, auch mehrfach gezählt wurden. Dennoch kann es für bestimmte Fragestellungen sinnvoll sein, die Variablen *wb011\** in *wb010\** in Gänze zu kombinieren, etwa wenn es ‚nur‘ darum gehen soll, ob beispielsweise Beschäftigte mit einfachen Tätigkeiten an einer Weiterbildung teilgenommen haben. Eine solche Kombination wurde vom FDZ nicht programmiert, sondern bleibt jedem/jeder Nutzenden selbst überlassen. Die entsprechenden Befehle zur Kombination der Variablen lauten:

### Stata

```
***Mergen der Variable wb011* (Teilnehmer bzw. teilnahmefälle von
Beschäftigten an Weiterbildung) aus 2011er Querschnittsdatsatz
use „*\Qualifizierungspanel_11bis17_long_gwa.dta“
sort id2 jahr
merge 1:n id2 jahr using „*\Qualifizierungspanel_11bis17_quer_2011_
gwa.dta“, keepusing(wb011einf* wb011qual* wb011hoch*)
drop _merge
***Ersetzen der Variablen
foreach x in einf qual hoch{
replace wb010`x'=wb011`x' if jahr==2011
replace wb010`x'_imp=wb011`x'_imp if jahr==2011
}
drop wb011*
```

### SPSS

```
get file = „*\Qualifizierungspanel_11bis17_long_gwa.sav“.
DATASET NAME long.
SORT CASES by id2 jahr.

get file =„*\Qualifizierungspanel_11bis17_quer_2011_gwa.sav“.
save OUTFILE= „*\Qualifizierungspanel_11bis17_quer_2011_wb011.sav“
/keep =id2 jahr wb011einf wb011einf_imp wb011qual wb011qual_imp
wb011hoch wb011hoch_imp.
get file= „*\Qualifizierungspanel_11bis17_quer_2011_wb011.sav“.
DATASET NAME wb011.
SORT CASES by id2 jahr.
DATASET ACTIVATE long.

MATCH FILES /FILE=*
/TABLE='wb011'
/by id2 jahr.
EXECUTE.

***Ersetzen der Variablen.
if jahr=2011 wb010einf=wb011einf.
if jahr=2011 wb010qual=wb011qual.
if jahr=2011 wb010hoch=wb011hoch.
if jahr=2011 wb010einf_imp=wb011einf_imp.
if jahr=2011 wb010qual_imp=wb011qual_imp.
if jahr=2011 wb010hoch_imp=wb011hoch_imp.
EXECUTE.
```

```
DELETE VARIABLES wb011einf wb011einf_imp wb011qual wb011qual_imp
wb011hoch wb011hoch_imp.
EXECUTE.
```

#### 4.5.2 2012

Die Frage nach der Bezeichnung des Ausbildungsberufs, in welchem der Betrieb am 31.12. des entsprechenden Jahres ausbildet hat (*ab004\_a\*\_kldb2010\_5er*), und die Frage nach der Anzahl der Auszubildenden im entsprechenden Beruf (*ab004\_b1*) wurden 2012 nur an Nicht-panelbetriebe gestellt (*b44a\** und *b44b\**). Diese Angaben sind im long-Datensatz in der Variable *ab004\_a\*\_kldb2010\_5er* enthalten. Die 727 Panelbetriebe, welche die Frage B44 nicht gestellt bekamen, aber Ausbildungsbetriebe waren, sind mit dem Wert ,99996. 2012 nur neue Betriebe dazu befragt gekennzeichnet. Um die Variablen *ab004\_a\*\_kldb2010\_5er* für das Jahr 2012 zu ergänzen, ist es möglich, die Informationen aus dem vorherigen Befragungsjahr (2011) und den Angaben zu den 2012 neu besetzten Ausbildungsberufen (*b11a\**) zu nutzen. Dies wurde vom BIBB-FDZ im long-Datensatz jedoch nicht umgesetzt, sondern bleibt jedem/jeder Nutzenden selbst überlassen. Eine Möglichkeit zum Auffüllen der Angaben ist:

#### Stata

```
***Mergen der Variable b11a* (Ausbildungsberuf nach BBiG/HwO 1
Kldb2010) aus 2012er Querschnittdatensatz
use „*\Qualifizierungspanel_11bis17_long_gwa.dta“
sort id2 jahr
merge 1:n id2 jahr using „*\Qualifizierungspanel_11bis17_quer_2012_
gwa.dta“, keepusing(b11a*)
drop _merge

***Ausbildungsberuf aus dem Vorjahr übernehmen, wenn neu besetzte
Ausbildungsberufe den Ausbildungsberufen aus dem Vorjahr entsprechen
foreach var of varlist ab004_a*_kldb2010_5er{
  forvalues x=1/5{
    replace `var'=b11a`x' if `var'==99996 & `var'[_n-1]==b11a`x' &
id2==id2[_n] & b11a`x'>0
  }
}

*wenn nur 99999. kein weiterer Ausbildungsberuf genannt geändert
wurde, wieder in 99996. 2012 nur neue Betrieb dazu befragt codieren
recode ab004_a*_kldb2010_5er (99999=99996) if ab004_a1_
kldb2010_5er>99995 & ab004_a2_kldb2010_5er>99995 & ab004_a3_
kldb2010_5er>99995 & ab004_a4_kldb2010_5er>99995 & ab004_a5_
kldb2010_5er>99995

*wenn ein Ausbildungsberuf ersetzt wurde, die übrigen in 99999. kein
weiterer Ausbildungsberuf genannt statt 99996. 2012 nur neue Betrieb
dazu befragt
recode ab004_a*_kldb2010_5er (99996=99999) if ab004_a1_
kldb2010_5er<99996
```

```
*Angaben zu Ausbildungsberufen so tauschen, dass gültige Werte zuerst
und danach 99999. kein weiterer Ausbildungsberuf genannt
//swapval-ado muss heruntergeladen werden, der Befehl tauscht die
Werte der angegebenen Variablen
forvalues x=3/5{
if ab004_a2_kldb2010_5er==99999{
cap swapval ab004_a`x`_kldb2010_5er ab004_a2_kldb2010_5er if
ab004_a`x`_kldb2010_5er<99996
}
}
forvalues x=4/5{
if ab004_a3_kldb2010_5er==99999{
cap swapval ab004_a`x`_kldb2010_5er ab004_a3_kldb2010_5er if
ab004_a`x`_kldb2010_5er<99996
}
}
if ab004_a4_kldb2010_5er==99999{
cap swapval ab004_a5_kldb2010_5er ab004_a4_kldb2010_5er if
ab004_a5_kldb2010_5er<99996
}
}
drop b11a*
```

#### SPSS

```
get file ="*\Qualifizierungspanel_11bis17_long_gwa.sav".
DATASET NAME long.
SORT CASES by id2 jahr.

GET file = ,*\Qualifizierungspanel_11bis17_quer_2012_gwa.sav".
SORT CASES by id2 jahr.
save OUTFILE= ,*\Qualifizierungspanel_11bis17_quer_2012_b11a.sav"
/KEEP= id2 jahr b11a1 b11a2 b11a3 b11a4 b11a5.
get file= ,*\Qualifizierungspanel_11bis17_quer_2012_b11a.sav".
DATASET NAME b11a.
EXECUTE.
DATASET ACTIVATE long.

MATCH FILES /FILE=*
/TABLE='b11a'
/BY id2 jahr.
EXECUTE.

***Ausbildungsberuf aus dem Vorjahr übernehmen, wenn neu
besetzte Ausbildungsberufe den Ausbildungsberufen aus dem
Vorjahr entsprechen.
do repeat ab004_a = ab004_a1_kldb2010_5er ab004_a2_kldb2010_5er
ab004_a3_kldb2010_5er ab004_a4_kldb2010_5er ab004_a5_
kldb2010_5er.
if ab004_a=99996 & LAG(ab004_a)=b11a1 & id2=LAG(id2) & b11a1>0
ab004_a=b11a1.
if ab004_a=99996 & LAG(ab004_a)=b11a2 & id2=LAG(id2) & b11a2>0
```

```

ab004_a=b11a2.
  if ab004_a=99996 & LAG(ab004_a)=b11a3 & id2=LAG(id2) & b11a3>0
ab004_a=b11a3.
  if ab004_a=99996 & LAG(ab004_a)=b11a4 & id2=LAG(id2) & b11a4>0
ab004_a=b11a4.
  if ab004_a=99996 & LAG(ab004_a)=b11a5 & id2=LAG(id2) & b11a5>0
ab004_a=b11a5.
EXECUTE.

```

Die im Jahr 2012 zusätzlichen Angaben zu den neu besetzten (*b11\**) und unbesetzten (*b16\**) Ausbildungsberufen sind ursprünglich als KldB-1992-Codes erhoben worden. Um Vergleichbarkeit mit den andern Berufscodes des BIBB-Qualifizierungspanels herzustellen, wurden diese in die KldB 2010 umcodiert. Damit unterscheiden sich die Berufscodes des zum long-Datensatz gehörenden Querschnittdatensatz 2012 von den Codes im ursprünglichen Querschnittdatensatz.

Die Fragen zu pb017\* („Anzahl ausgeschiedener Beschäftigter nach Beschäftigtengruppe“) wurde 2012 nicht gestellt. Die Frage ist jedoch Grundlage für die Codierung der Kategorie ‚Beschäftigtengruppe nicht vorhanden‘. Aus diesem Grund muss die Codierung für 2012 ohne die Information zu Anzahl ausgeschiedener Beschäftigter erfolgen.

#### 4.5.3 2011 und 2015

In den Jahren 2011 und 2015 wurde abgefragt, welche Ziele mit den Weiterbildungsmaßnahmen des Vorjahrs verbunden waren. Drei der Items sind zwischen den Jahren vergleichbar (vgl. Tabelle 6). Da nur drei der bis zu sieben Items harmonisierbar wären, wurden diese nicht in den long2-Datensatz aufgenommen. Die Harmonisierung bleibt den jeweiligen Nutzerinnen und Nutzern selbst überlassen.

**Tabelle 6: Vergleichbare Items zu Zielen von Weiterbildungsmaßnahmen 2011 und 2015**

Variablenname 2011	Variablenname 2015	Label
a47b	e44_3	Ziele der Weiterbildung (WB): Anpassung fachlicher Kenntnisse an veränderte Arbeitsabläufe
a47c	e44_7	Ziele der WB: Beschäftigte darin fördern, sich selbstständig Wissen anzueignen
a47e	e44_2	Ziele der WB: Vorbereitung für innerbetriebliche Aufstiege/ Beförderungen

#### 4.5.4 2016 und 2017

Wie bereits in Kapitel 2 beschrieben, wurde 2016 ergänzend zur üblichen Erhebung des BIBB-Qualifizierungspanels eine CATI-Befragung mit Schwerpunkt Digitalisierung durchgeführt, aus welcher 2017 die Auffrischungstichprobe gezogen wurde. Ein Teil der Variablen wurde sowohl in der CAPI- als auch in der CATI-Befragung erhoben. Entsprechend ist ein Teil der Variablen der CATI- und CAPI-Befragungen identisch – um welche es sich handelt, kann FRIEDRICH/GERHARDS 2018 entnommen werden.

Die Variable *quer* enthält neben der Information, ob es sich um einen Querschnittsfall einer jeweiligen Welle handelt, für die Welle 2016 speziell die Information, welche 3.521 Betriebe

2016 an der CATI-Befragung teilgenommen haben. Es ist zu beachten, dass für diese Betriebe nicht alle Variablen im long- und long2-Datensatz gefüllt sind. Die entsprechenden Fälle weisen bei allen Variablen, die nicht zum Frageprogramm der CATI-Befragung gehören, den Wert ‚-12 CATI Fall‘ auf.

Etwa 38 Prozent (1.345) der CATI-Betriebe aus 2016 haben 2017 erneut am Panel teilgenommen. Für die meisten Panelanalysen ist es sinnvoll, alle reinen CATI-Betriebe auszuschließen. Die entsprechenden Befehle dazu lauten:

**Stata**

```
drop if quer==2 & pan_ges==1
```

**SPSS**

```
compute filter=1.
if quer=2 & pan_ges=1 filter=2.
EXECUTE.

select if filter=1.
EXECUTE.
DELETE VARIABLES filter.
EXECUTE.
```

Ob es sinnvoll ist, die CATI-Fälle, die 2017 erneut am BIBB-Qualifizierungspanel teilgenommen haben, auch auszuschließen, hängt vom genauen Auswertungsinteresse ab. Einige Variablen wie die Anzahl der Beschäftigten sind auch für die CATI-Betriebe gefüllt und können somit im Längsschnitt ausgewertet werden.

## 4.6 Englische Übersetzung

Die drei Längsschnittdatensätze des BIBB-Qualifizierungspanel 2011 bis 2017 wurden jeweils auch in Englisch gelabelt. In Stata sind die englischen Label direkt implementiert. Mit folgendem Befehl kann die Sprache der Label angepasst werden:

**Stata**

```
***Zeigt vorhandene Sprachen an
label language
***Ändert die Sprache in Englisch
label language en
***Ändert die Sprache in Deutsch
label language de
```

Für SPSS gibt es drei Syntaxen, welche die Variablen und Kategorien umlabeln. Diese stehen im Metadatenportal des BIBB-FDZ zur Verfügung.

# 5 Gewichtung und Hochrechnung

## 5.1 Berechnung der Gewichtungs- und Hochrechnungsfaktoren

Bei der Gewichtung ist zu beachten, dass 2015, wie bereits beschrieben, das Erhebungsinstitut gewechselt wurde, was wiederum eine Änderung in der Berechnungsweise der Gewichte nach sich zog. Die genaue Beschreibung der Gewichtung kann in den jeweiligen Feldberichten der Erhebungsinstitutionen nachgelesen werden (z. B. HÄRING u. a. 2017 und TNS INFRATEST 2015). Im Folgenden wird die Gewichtung ab 2015 in Anlehnung an HÄRING u. a. 2017 in Kürze beschrieben.

Die Querschnittgewichte sind eine Kombination aus (adjustierter) Design- und Strukturanpassungsgewichtung. „Allgemein spricht man von Designgewichtung, wenn man die sich aus dem Auswahlverfahren ergebenden Auswahlwahrscheinlichkeiten im Rahmen der Schätzung in Form von Gewichten als der Inversen der Auswahlwahrscheinlichkeit berücksichtigt“ (GABLER/GANNINGER 2010, S. 147). „Durch Nonresponse kann man in der Stichprobe zudem eine Verzerrung bei der Verteilung gewisser Merkmale erhalten. [...] Anpassungsgewichte sollen helfen, dieses Ungleichgewicht zu reparieren“ (GABLER/GANNINGER 2010, S. 153). Durch die (Struktur-)Anpassung werden die Randverteilungen der Stichprobe an die Randverteilungen der Grundgesamtheit angepasst (Kalibrierung).

Das Designgewicht wird getrennt nach Auffrischungs- und Panelstichprobe berechnet. Aus den Schichtungsmerkmalen Region, Betriebsgröße, Branche und Ausbildungsbetrieb (vgl. Kapitel 2) ergeben sich 128 Schichtungszellen. Mithilfe eines Horvitz-Thompson-Schätzers (vgl. GABLER/GANNINGER 2010) werden für alle Schichten und Stufen die inversen Auswahlwahrscheinlichkeiten berechnet, deren Reziprok das Designgewicht darstellt. Um extreme Gewichte zu vermeiden, wurden die Designgewichte fallzahlnormiert.

Für die Panelstichprobe wird das kalibrierte Gewicht aus der letzten bzw. für temporäre Ausfälle aus der vorletzten Erhebungswelle als Designgewicht definiert. Beide Designgewichte werden jeweils getrennt mittels Regression adjustiert. Die Nonresponse-Adjustierung der Designgewichte erfolgt erneut getrennt für Auffrischer- und Panelbetriebe. Dafür werden Ausfallmodelle berechnet, deren unabhängige Variablen die Strukturmerkmale darstellen. Die berechnete Teilnahmewahrscheinlichkeit wird jeweils mit dem Designgewicht multipliziert, um das adjustierte Designgewicht zu berechnen.

Das Ausfallmodell für Panelbetriebe schätzt ebenfalls die Bleibewahrscheinlichkeit ( $p_{\text{bleib}}$ ), welche für die Berechnung der Längsschnittgewichtung benötigt wird. Das Längsschnittgewicht ( $dw_p$ ) wird analog zum Designgewicht geschätzt, nur dass statt der Auswahlwahrscheinlichkeit die Bleibewahrscheinlichkeit für die Berechnung genutzt wird. Zusätzlich wurde aus dem Längsschnittgewicht der Hochrechnungsfaktor ( $hr_p$ ) gebildet.

Anschließend werden die adjustierten Designgewichte der Auffrischer- und Panelbetriebe kombiniert ( $adjdw_{\text{ges}}$ ). Dafür wird jeweils das adjustierte Designgewicht mit dem Quotienten aus der Fallzahl der jeweiligen Teilstichprobe (Auffrischer- oder Panelbetriebe) und dem Gesamtstichprobenumfang multipliziert.

Danach erfolgt die Kalibrierung des adjustierten Gesamtdesigngewichts mittels IPF (Iterative Proportional Fitting) der kombinierten Schichtungsmerkmale. Das Resultat ist das Strukturanpassungsgewicht ( $sw_q$ ) und der Hochrechnungsfaktor ( $hr_q$ ). Die Kalibrierung erfolgte zusätzlich für Branchen ( $hr_{\text{br}}$  und  $sw_{\text{br}}$ ) und Beschäftigte ( $hr_{\text{bes}}$  und  $sw_{\text{bs}}$ ).

## 5.2 Gewichtungs- und Hochrechnungsfaktoren im Datensatz

Die meisten Querschnitt- und Längsschnittgewichte, Hochrechnungsfaktoren sowie die Bleibewahrscheinlichkeit sind in einem gesonderten Datensatz „Qualifizierungspanel\_11bis17\_weight\_gwa“, die Variablen *hr\_q* und *hr\_p* im long-Basisdatensatz enthalten (vgl. Tabelle 7). Der Datensatz ist ebenfalls im long-Format abgespeichert, die Gewichtungs- und Hochrechnungsvariablen enthalten jeweils die Werte aus dem Erhebungsjahr.

Die zusätzliche CATI-Befragung im Jahr 2016 hat auch Auswirkungen auf die Gewichtung. Es wurden jeweils getrennte Gewichte für die CAPI- und CATI- Betriebe sowie ein integriertes Gewicht berechnet. Im Datensatz enthalten die Variablen *\*\_q*, *\*\_bes*, *\*\_br*, *\*\_p* und *adjdw\_ges* jeweils das integrierte Gewicht. Die einzelnen Gewichte sind ebenfalls im Datensatz „Qualifizierungspanel\_11bis17\_weight\_gwa“ enthalten.

Tabelle 7: Übersicht über im Datensatz enthaltene Gewichte bzw. Hochrechnungsfaktoren

Variable	Datensatz	Variableninhalt	Funktion	Jahre	Anwendung	Fallzahl	Mittelwert	Minimum	Maximum
sw_q	weight	Strukturangepassungsgewicht für den Querschnitt	Adjustierte Designgewichtung und Ausgleich der Verzerrung bei der Verteilung der Strukturmerkmale Branche, Größenklasse, Region und Ausbildungsbetrieb ja/nein für die Grundgesamtheit der Betriebe	Ab 2015	<u>Analysen auf Betriebsebene für einzelne Jahre</u> etwa wie viel Prozent der Betriebe in Deutschland hat Weiterbildung angeboten	14.416	1	0,001	14,16
sw_bes	weight	Beschäftigtenproportionales Strukturangepassungsgewicht	Adjustierte Designgewichtung und Ausgleich der Verzerrung bei der Verteilung der Beschäftigten, in der Grundgesamtheit	Ab 2015	<u>Analyse auf Beschäftigtenebene für einzelne Jahre</u> etwa wie viel Prozent der Beschäftigten an Weiterbildung teilnehmen	14.416	1	0,001	11,6
sw_br	weight	Strukturangepassungsgewicht 20er Branchen	Adjustierte Designgewichtung und Ausgleich der Verzerrung bei der Verteilung der Branche, in der Grundgesamtheit	Ab 2015	<u>Analysen auf Branchenebene für einzelne Jahre</u> etwa über den Prozentsatz an Weiterbildung in der Bauwirtschaft	14.416	1	0,001	24,7
hr_q	long	Hochrechnungsfaktor für Querschnitt des jeweiligen Jahres	Adjustierte Designgewichtung und Ausgleich der Verzerrung bei der Verteilung der Strukturmerkmale Branche, Größenklasse, Region und Ausbildungsbetrieb ja/nein sowie Hochrechnung auf Grundgesamtheit der Betriebe	2011-2017	<u>Analysen auf Betriebsebene für einzelne Jahre</u> etwa wie viele Betriebe in Deutschland haben Weiterbildung angeboten	23.902	607,6	0,39	11.984,1
hr_bes	weight	Beschäftigtenproportionaler Hochrechnungsfaktor	Adjustierte Designgewichtung und Ausgleich der Verzerrung bei der Verteilung der Beschäftigten, in der Grundgesamtheit sowie Hochrechnung auf die Beschäftigten in der Grundgesamtheit	Ab 2015	<u>Analyse auf Beschäftigtenebene für einzelne Jahre</u> etwa wie viele Beschäftigte nehmen an Weiterbildung teil	14.416	398,6	0,3	6.061,1
hr_br	weight	Hochrechnungsfaktor, 20er-Brancheneinteilung	Adjustierte Designgewichtung und Ausgleich der Verzerrung bei der Verteilung der Branche, in der Grundgesamtheit sowie Hochrechnung auf die Branchen in der Grundgesamtheit	Ab 2013	<u>Analysen auf Branchenebene für einzelne Jahre</u> etwa wie viel Weiterbildung wird in der Bauwirtschaft angeboten	19.892	523,8	0,6	3.2167,2

Variable	Datensatz	Variableninhalt	Funktion	Jahre	Anwendung	Fallzahl	Mittelwert	Minimum	Maximum
dw_p	weight	Längsschnittdesigngewicht der jeweiligen Jahre	Adjustierte Designgewichtung und Ausgleich von Panelmortalität als auch nicht erklärter Panelbereitschaft	Ab 2015	<u>Analysen über mehrere Jahre</u> etwa wie verändert sich der Prozentsatz der Betriebe, die Weiterbildung anbieten zwischen 2011 und 2017	6.935	1	0,001	23,5
hr_p	long	Hochrechnungsfaktor für Längsschnitt der jeweiligen Jahre	Adjustierte Designgewichtung und Ausgleich von Panelmortalität als auch nicht erklärter Panelbereitschaft sowie Hochrechnung auf Grundgesamtheit der Betriebe	2012-2017	<u>Analysen über mehrere Jahre</u> etwa verändert sich die Anzahl der Betriebe, die Weiterbildung anbieten zwischen 2011 und 2017	10.472	1.187,8	0,7	32.167,2
adjdw_ges	weight	Modifiziertes Designgewicht integrierte Gesamtstichprobe	Designgewichtung und Nonresponse-Adjustierung	Ab 2015 <sup>9</sup>	Analysen ohne Kalibrierung der Schichtungsmerkmale	10.735	1	0,001	32,6
pbleib	weight	Bleibwahrscheinlichkeit		Ab 2015		2.021	0,6	0,1	0,8
<b>Sondervariablen für einzelne Jahre</b>									
hr11_13p									
hr11_14p	weight	Hochrechnungsfaktor für Längsschnitt der jeweiligen Jahre		2013, 2014	Analysen für die Betrachtung der Jahre 2001 bis 2013 bzw. 2011 bis 2014 oder 2012 bis 2014	860	-	-	-
hr12_14p						553			
*_capi	weight	Hochrechnungsfaktoren, Strukturgewichtungen, Designgewichtungen (adjustiert und Längsschnitt)		2016	Analysen nur für CATI- bzw. CAPI-Betriebe	3.601	-	-	-
*_cati						3.521			

9 Für 2017 gibt es ein entsprechendes Gewicht nicht, da es keine Auffrischungsstichprobe gab.

Sobald deskriptive Veränderungen von einem auf das andere Jahr berechnet werden, sollte ein Längsschnittgewicht benutzt werden. Die in den Daten enthaltenen Längsschnittgewichte bzw. -hochrechnungsfaktoren enthalten jeweils die Faktoren zum Ausgleich der Bleibewahrscheinlichkeit vom Vorjahr (bzw. von zwei Jahren zuvor, falls der Betrieb im Vorjahr nicht an der Befragung teilgenommen hat) zum Befragungsjahr. Bei der Längsschnittgewichtung muss darauf geachtet werden, dass für das Jahr des Erstinterviews kein Längsschnittgewichts- bzw. Längsschnitthochrechnungsfaktor zur Verfügung steht. Das Jahr des Erstinterviews sollte mit dem jeweiligen Querschnittsfaktor gewichtet bzw. hochgerechnet werden. Die Befehle (nachdem der `weight` Datensatz an den `long`-Datensatz gemerget wurde) zur Kombination der Gewichte lauten:

**Stata**

```
replace hr_p= hr_q if well==1 | well==-3  
replace dw_p= sw_q if well==1 | well==-3
```

**SPSS**

```
if well==1 or well==-3 hr_p= hr_q  
if well==1 or well==-3 dw_p= sw_q  
EXECUTE.
```

## 6 Weiterführung in den Folgejahren

Die zukünftigen Wellen des BIBB-Qualifizierungspanels werden wie bisher als jährlicher Querschnitt Datensatz veröffentlicht. Zusätzlich werden die Forschungsdaten ab sofort in den bestehenden Längsschnitt integriert. Somit wird der Längsschnitt Datensatz des BIBB-Qualifizierungspanels jedes Jahr aktualisiert und ‚neu‘ veröffentlicht. Eine gesonderte Dokumentation bzw. Aktualisierung des Daten- und Methodenberichts zum Längsschnitt Datensatz wird nicht erfolgen.

# 7 Anonymisierung und Datenzugang

## 7.1 Bestimmungen des Datenschutzes

Der Zugang zu den Forschungsdaten des BIBB-FDZ erfolgt ausschließlich unter dem Gebot der Datensparsamkeit und der Einhaltung der geltenden Datenschutzregeln nach der Verordnung (EU) 2016/679 (Datenschutz-Grundverordnung, DSGVO) sowie ergänzender Vorschriften des Bundesdatenschutzgesetzes (BDSG 2018). Demnach dürfen für die Durchführung wissenschaftlicher Vorhaben Daten an die unabhängige wissenschaftliche Forschung weitergegeben werden, wenn die Herstellung eines Bezugs zu einer Erhebungseinheit nicht möglich ist („Anonymität“). Es wird empfohlen, dass zugangs- und nutzungsberechtigte Forscherinnen und Forscher zur Einhaltung von Datenschutzbestimmungen besonders verpflichtet sind (vgl. Nutzungsrichtlinien des BIBB-FDZ für die Datenfernverarbeitung und Gastaufenthalte).

Die Mitarbeiter/-innen des BIBB-FDZ nehmen nur zum Zwecke der Beratung, der Verbesserung des BIBB-FDZ-Services sowie zur Gewährleistung der Einhaltung des Datenschutzes Einblick in Forschungsfragen, Methoden und Analysen der Wissenschaftlerinnen und Wissenschaftler. Mitarbeiterinnen und Mitarbeiter des BIBB, die nicht dem BIBB-FDZ angehören, erhalten keinen Einblick in die Tätigkeiten der Wissenschaftlerinnen und Wissenschaftler.

## 7.2 Formale Datenanonymisierung

Vor der Erhebung wurde den freiwillig teilnehmenden Betrieben Datenschutz und Anonymität zugesichert (siehe Datenschutzblatt in HÄRING u. a. 2017). Die wenigen Strukturinformationen im Datensatz erlauben keine Re-Identifikation der befragten Betriebe. Zur Unterscheidung der Fälle steht ein wellenübergreifendes, systemfreies Betriebskennzeichen (id2) zur Verfügung.

## 7.3 Datenzugang

Die Daten des BIBB-Qualifizierungspanels 2011 bis 2017 können über die Datenfernverarbeitung (DFV) und den Gastwissenschaftleraufenthalt (GWA) im BIBB in Bonn analysiert werden.

Die kontrollierte Datenfernverarbeitung erlaubt die Analyse schwach anonymisierter Forschungsdaten über die Verarbeitung lauffähiger, von den Nutzerinnen und Nutzern erstellter Syntaxprogramme im BIBB-FDZ. Die Datennutzerinnen und -nutzer können hierfür im Metadatenportal des BIBB-FDZ Strukturdatensätze (Spieldaten) herunterladen. Die Spieldaten gleichen in Aufbau und Merkmalsausprägungen den Originaldaten. Sie ermöglichen somit das Erstellen von Auswertungsprogrammen (in den Analyseprogrammen Stata oder SPSS), mit denen das BIBB-FDZ anschließend die Originaldaten auswertet. Der auf Vertraulichkeit geprüfte Output wird schließlich an die Nutzerinnen und Nutzer zurückgeschickt. Gastwissenschaftleraufenthalte erlauben die Analyse schwach anonymisierter Forschungsdaten an den abgeschotteten PC-Arbeitsstationen in einem eigenen Gästeraum des BIBB-FDZ.

Im Jahr 2021 wird das BIBB-FDZ ein automatisiertes Datenfernverarbeitungssystem in Betrieb nehmen, mit dem Datennutzer/-innen ihre Syntaxprogramme eigenständig auf Originaldaten anwenden können. Auswertungsergebnisse können dann unmittelbar nach dem Durchlauf der Auswertungsprogramme (personalisiert) eingesehen werden. Zudem werden auf diese Weise die beiden Datenzugangswege Datenfernverarbeitung und Gastaufenthalte projektbezogen miteinander verbunden, d. h., auf beiden Datenzugangswegen ist ein Zugriff

auf das gleiche und jeweils aktuelle Projektverzeichnis möglich. Mit der Einführung des automatisierten Datenfernverarbeitungssystems wird das im vorherigen Absatz geschilderte manuelle Vorgehen der derzeitigen Datenfernverarbeitung im BIBB-FDZ vollständig abgelöst. Datennutzer/-innen bekommen dadurch eine größere Autonomie und mehr Freiheitsgrade bei der Auswertung von sensitiven Datenbeständen.

Die Nutzung der Daten erfordert für alle hier beschriebenen Datenzugangswege eine förmliche Beantragung. Die entsprechenden Nutzungsanträge können von den Internetseiten des BIBB-FDZ heruntergeladen werden.

## 8 Literaturverzeichnis

- FISCHER, Gabriele u.a.: Das IAB-Betriebspanel – von der Stichprobe über die Erhebung bis zur Hochrechnung. FDZ Methodenreport 01/2008, URL: [http://doku.iab.de/fdz/reporte/2008/MR\\_01-08.pdf](http://doku.iab.de/fdz/reporte/2008/MR_01-08.pdf) (Abruf: 10.08.2020)
- FRIEDRICH, Anett; GERHARDS, Christian: BIBB-Qualifizierungspanel 2015. Daten- und Methodenbericht Nr. 2/2017. Bonn 2017
- FRIEDRICH, Anett; GERHARDS, Christian: BIBB-Qualifizierungspanel 2016. Daten- und Methodenbericht Nr. 1/2018. Bonn 2018
- GABLER, Siegfried; GANNINGER, Matthias: Gewichtung. In: WOLF, Christof; BEST, Henning (Hrsg.): Handbuch der sozialwissenschaftlichen Datenanalyse. Wiesbaden 2010, S. 143–164
- GERHARDS, Christian; FRIEDRICH, Anett: BIBB-Qualifizierungspanel 2014. Daten- und Methodenbericht Nr. 3/2016. Bonn 2016
- GERHARDS, Christian u.a.: The BIBB Training Panel. An Establishment Panel on Training and Competence Development. Schmollers Jahrbuch, 132(4), S. 635–652. 2012
- GERHARDS, Christian u.a.: BIBB-Qualifizierungspanel 2013. Daten- und Methodenbericht Nr. 3/2015. Bonn 2015
- HÄRING, Armando; SCHIEL, Stefan; KLEUDGEN, Martin: Methodenbericht Erhebungswelle 2017 des BIBB-Qualifizierungspanels. Bonn 2017
- HOHN, Christine; FRIEDRICH, Anett; GERHARDS, Christian: Querschnitt- und Panelcharakteristika des BIBB-Qualifizierungspanels. URL: <https://metadaten.bibb.de/metadatengruppe/16>. Bonn 2020 (Abruf: 12.10.2020)
- TNS INFRATEST SOZIALFORSCHUNG: BIBB-Qualifizierungspanel. Qualifizierung und Kompetenzentwicklung. Methodenbericht zur vierten Welle. München 2015

# Anhang

**Tabelle A1: Variablen, die im Qualifizierungspanel\_11bis17\_long\_gwa nicht in allen Jahren abgefragt wurden**

Variable(n)	Label	Nicht abgefragte Jahre
ab023_*	Abschlussart Bewerber	2011, 2013
ab034*	Anzahl unbesetzter Ausbildungsstellen wegen vorzeitiger Vertragslösung	2011, 2012
ab036	Ausbildungsverträge nach BBiG/HwO im Kalenderjahr vorzeitig gelöst worden	2012, 2013
ab037	Anzahl vorzeitig gelöster Ausbildungsverträge im Kalenderjahr	2012, 2013
ab057	Hat Ihr Betrieb in den letzten 3 Jahren nach BBiG/HwO ausgebildet?	2011 bis 2014
ab065	Zum 31.12. des Vorjahres Flüchtlinge als Azubis im Betrieb?	2011 bis 2015
ab066	Wie viele Ihrer Azubis nach BBiG oder HwO am 31.12. des Vorjahres sind Flüchtlinge?	2011 bis 2015
ab067	Anzahl auszubildende Flüchtlinge in sozialpflegerischen Berufen	2011 bis 2015
ab068	Anzahl neu eingestellter Flüchtlinge in BBiG/HwO zum Ausbildungsjahr	2011 bis 2015
ab069	Zum 31.12. des Vorjahres mit Flüchtlingen geschlossene Ausbildungsverträge wieder gelöst	2011 bis 2015
ps001zwi*	Summe aus ps001svb ps001ger ps001nsv	2011
ps001aus*	Anzahl Azubis nach BBiG/HwO	2011
ps001soz*	Anzahl Azubis nach Ausbildungsregelungen des Gesundheitswesens o. Ä.	2011
ps001bea*	Anzahl Beamtenanwärter/-innen	2011 bis 2014
ps001svb_v	Anzahl sozialversicherungspflichtig Beschäftigte (31.12. des Vorjahres)	2011
ps001ger_v	Anzahl geringfügig Beschäftigte (31.12. des Vorjahres)	2011
ps001nsv_v	Anzahl tätige Inhaber/-innen und mithelfende Familienangehörige (31.12. des Vorjahres)	2011
ps030_sch	Anzahl Beschäftigte mit einer schulischen Berufsausbildung	2011, 2012
ps036	Flüchtlinge als Beschäftigte im Betrieb	2011 bis 2014
ps037	Anzahl der beschäftigten Flüchtlinge	2011 bis 2014
pb001_ang	Hat Ihr Betrieb im Vorjahr Stellen angeboten?	2011
pb005*	Anzahl neu eingestellte Beschäftigte nach Tätigkeitsgruppe	2012
pb011_*	Anzahl unbesetzter Stellen nach Tätigkeitsgruppe	2012
pb017	Anzahl ausgeschiedener Beschäftigter nach Tätigkeitsgruppe	2012
wb002	Anzahl Beschäftigte mit Teilnahme an Aufstiegsfortbildungen	2011, 2012

Variable(n)	Label	Nicht abgefragte Jahre
wb010*	Anzahl der WB-Teilnehmer/-innen nach Tätigkeitsgruppe	2011, 2013
di001_*	Nutzung digitaler Technologien	2011 bis 2015
be001_*	Welche der folgenden Regelungen für Beschäftigte gibt es in Ihrem Betrieb?	2011, 2013
be010*	Veränderungen im Betrieb	2011
<b>Einzelne Items, die nicht durchgängig abgefragt wurden</b>		
ab043_*1	Anzahl Azubis vom Betrieb unbefristet übernommen	2012
ab043_*2	Anzahl Azubis vom Betrieb befristet übernommen	2012
ab043_*4	Anzahl Azubis betriebsbedingt nicht übernommen	2012 bis 2014
ab043_*5	Anzahl Azubis wegen Prüfungsnoten nicht übernommen	2012 bis 2014
ab043_*6	Anzahl Azubis die Abschlussprüfung nicht bestanden haben	2017
ab058_5	Gründe selbst auszubilden: Tradition des Betriebes	2011
ab058_6	Gründe selbst auszubilden: Fachkräfte ausbilden, die langfristig bleiben	2011 bis 2015
ab058_7	Gründe selbst auszubilden: Ausbildung als Gemeinschaftsaufgabe der Wirtschaft	2011 bis 2015
be001_6	Regelung: Langzeitarbeitskonten	2012, 2014
be001_7	Regelung: wechselnde Zusammensetzung von Projektteams	2012, 2014
be001_8	Regelung: Maßnahmen zur Vereinbarkeit von Familie und Beruf	2012, 2014
be001_9	Regelung: Regelungen zur Altersvorsorge	2012, 2014
be010_7	Veränderungen im Vorjahr: Schließung/Ausgliederung/Ausgründung von Betriebsteilen	2012, 2013
be010_8	Veränderungen im Vorjahr: Eingliederungen von Betrieben/Unternehmensteilen	2012, 2013

**Tabelle A2: Variablenliste für den Datensatz Qualifizierungspanel\_11bis17\_long2\_gwa**

Variable	Variablenlabel	Abgefragte Jahre
id2	systemfreie Identifikationsnummer	
jahr	Jahr der Erhebung	
well	Wellenkennzeichen	
quer	Querschnittsfall	
pan_ges	Summe der Teilnahmen bis 2017	
ab003f	Anzahl weiblicher Azubis nach BBiG oder Hw0 zum 31.12. des Vorjahres	2011–2015
ab003f_imp	Anzahl der weiblichen Auszubildenden am 31.12. des Vorjahres	2011, 2012, 2014
ab003gt	Anzahl gewerblich–technische Azubis zum 31.12. des Vorjahres	2011–2015
ab003kv	Anzahl kaufmännisch–verwaltende Azubis zum 31.12. des Vorjahres	2011–2015
ab003gt_imp	Anzahl gewerblich–technische Azubis zum 31.12. des Vorjahres (imputiert)	2012
ab003kv_imp	Verteilung: Kaufmännisch–verwaltender Ausbildungsberuf (imputiert)	2012
ab009	Betrieb hat zum 31.12. des Vorjahres in mehr als 5 Ausbildungsberufen ausgebildet	2016, 2017
ab027_1	Angebote für Azubis: Übertarifliche Ausbildungsvergütung	2015, 2016
ab027_2	Angebote für Azubis: Möglichkeiten zum Erwerb zertifizierter Zusatzqualifikationen	2015
ab027_3	Angebote für Azubis: Übernahmegarantie bei erfolgreichem Abschluss	2015, 2016
ab027_4	Angebote für Azubis: Geldwerte Vorteile (z.B. Firmenwagen)	2015, 2016
ab027_5	Angebote für Azubis: Zusätzliche Maßnahmen zur Unterstützung leistungsschwacher Jugendlicher	2015
ab027_6	Angebote für Azubis: Zuschuss zur Nutzung des Nahverkehrs/ Verpflegung/Unterkunft	2015
ab027_7	Angebote für Azubis: Zuschuss zur Nutzung des öffentlichen Nahverkehrs	2016
ab027_8	Angebote für Azubis: Zuschuss für Verpflegung oder Unterkunft	2016
ab035	Auszubildende nach BBiG/Hw0 im Kalenderjahr beschäftigt	2015–2017
ab055	Anzahl befristet übernommener Azubis, die voraussichtlich unbefristeten Vertrag erhalten	2015, 2016
ab055_imp	Anzahl befristet übernommener Azubis, die voraussichtlich unbefristeten Vertrag erhalten (imputiert)	2015, 2016
ab056	Ausbildungsstellen nach BBiG/Hw0 geplant?	2015–2017
ab064	Erfüllt Ihr Betrieb die gesetzlichen Voraussetzungen zur Ausbildung nach BBiG/Hw0?	2015–2017
ab082gt1	Anzahl gewerblich–technischer Auszubildender im 1. Ausbildungsjahr	2011–2013
ab082gt2	Anzahl gewerblich–technischer Auszubildender im 2. Ausbildungsjahr	2011–2013

Variable	Variablenlabel	Abgefragte Jahre
ab082gt3	Anzahl gewerblich-technischer Auszubildender im 3. Ausbildungsjahr	2011-2013
ab082gt4	Anzahl gewerblich-technischer Auszubildender im 4. Ausbildungsjahr	2011-2013
ab082kv1	Anzahl kaufmännisch-verwaltender Auszubildender im 1. Ausbildungsjahr	2011-2013
ab082kv2	Anzahl kaufmännisch-verwaltender Auszubildender im 2. Ausbildungsjahr	2011-2013
ab082kv3	Anzahl kaufmännisch-verwaltender Auszubildender im 3. Ausbildungsjahr	2011-2013
ab082kv4	Anzahl kaufmännisch-verwaltender Auszubildender im 4. Ausbildungsjahr	2011-2013
ab082gt1_imp	Auszubildende im 1. Ausbildungsjahr gewerblich-technisch (imputiert)	2011
ab082gt2_imp	Auszubildende im 2. Ausbildungsjahr gewerblich-technisch (imputiert)	2011
ab082gt3_imp	Auszubildende im 3. Ausbildungsjahr gewerblich-technisch (imputiert)	2011
ab082gt4_imp	Auszubildende im 4. Ausbildungsjahr gewerblich-technisch (imputiert)	2011
ab082kv1_imp	Auszubildende im 1. Ausbildungsjahr kaufmännisch-verwaltend (imputiert)	2011
ab082kv2_imp	Auszubildende im 2. Ausbildungsjahr kaufmännisch-verwaltend (imputiert)	2011
ab082kv3_imp	Auszubildende im 3. Ausbildungsjahr kaufmännisch-verwaltend (imputiert)	2011
ab082kv4_imp	Auszubildende im 4. Ausbildungsjahr kaufmännisch-verwaltend (imputiert)	2011
ab083	Azubis duales Studium am 31.12. des Vorjahres	2011, 2014
ab083gt	Anzahl gewerblich-technischer Azubis duales Studium am 31.12. des Vorjahres	2011, 2014
ab083kv	Anzahl kaufmännisch-verwaltender Azubis duales Studium am 31.12. des Vorjahres	2011, 2014
ab083gt_imp	Anzahl gewerblich-technischer Azubis duales Studium am 31.12. des Vorjahres	2014
ab083kv_imp	Anzahl kaufmännisch-verwaltender Azubis duales Studium am 31.12. des Vorjahres	2014
ab085gt	Beschäftigung in Lehrgängen oder Kursen, gewerblich-technische Auszubildende	2011-2013
ab085kv	Beschäftigung in Lehrgängen oder Kursen, kaufmännisch-verwaltende Auszubildende	2011-2013
ab086gt1	Anteil produktiver Tätigkeiten während der Anwesenheit im Betrieb im 1. Ausbildungsjahr, gewerblich-technische Auszubildende	2011-2013
ab086gt2	Anteil produktiver Tätigkeiten während der Anwesenheit im Betrieb im 2. Ausbildungsjahr, gewerblich-technische Auszubildende	2011-2013
ab086gt3	Anteil produktiver Tätigkeiten während der Anwesenheit im Betrieb im 3. Ausbildungsjahr, gewerblich-technische Auszubildende	2011-2013
ab086gt4	Anteil produktiver Tätigkeiten während der Anwesenheit im Betrieb im 4. Ausbildungsjahr, gewerblich-technische Auszubildende	2011-2013

Variable	Variablenlabel	Abgefragte Jahre
ab086kv1	Anteil produktiver Tätigkeiten während der Anwesenheit im Betrieb im 1. Ausbildungsjahr, kaufmännisch-verwaltende Auszubildende	2011-2013
ab086kv2	Anteil produktiver Tätigkeiten während der Anwesenheit im Betrieb im 2. Ausbildungsjahr, kaufmännisch-verwaltende Auszubildende	2011-2013
ab086kv3	Anteil produktiver Tätigkeiten während der Anwesenheit im Betrieb im 3. Ausbildungsjahr, kaufmännisch-verwaltende Auszubildende	2011-2013
ab086kv4	Anteil produktiver Tätigkeiten während der Anwesenheit im Betrieb im 4. Ausbildungsjahr, kaufmännisch-verwaltende Auszubildende	2011-2013
ab087gt1	Prozentuales Leistungsniveau der gewerblich-technischen Auszubildenden im 1. Ausbildungsjahr	2011-2013
ab087gt2	Prozentuales Leistungsniveau der gewerblich-technischen Auszubildenden im 2. Ausbildungsjahr	2011-2013
ab087gt3	Prozentuales Leistungsniveau der gewerblich-technischen Auszubildenden im 3. Ausbildungsjahr	2011-2013
ab087gt4	Prozentuales Leistungsniveau der gewerblich-technischen Auszubildenden im 4. Ausbildungsjahr	2011-2013
ab087kv1	Prozentuales Leistungsniveau der kaufmännisch-verwaltenden Auszubildenden im 1. Ausbildungsjahr	2011-2013
ab087kv2	Prozentuales Leistungsniveau der kaufmännisch-verwaltenden Auszubildenden im 2. Ausbildungsjahr	2011-2013
ab087kv3	Prozentuales Leistungsniveau der kaufmännisch-verwaltenden Auszubildenden im 3. Ausbildungsjahr	2011-2013
ab087kv4	Prozentuales Leistungsniveau der kaufmännisch-verwaltenden Auszubildenden im 4. Ausbildungsjahr	2011-2013
ab089	Prognose: Anzahl der Ausbildungsplätze in den nächsten 3 Jahren	2011-2013
ps039	Anzahl 31.12. des Vorjahres: Beschäftigte mit Bachelor als höchstem beruflichem Abschluss	2013, 2014, 2017
ps039_imp	Anzahl 31.12. des Vorjahres: Beschäftigte mit Bachelor als höchstem beruflichem Abschluss (imputiert)	2013, 2014
wb015_1	Beförderungen	2013, 2014
wb015_2	Anzahl beförderter Beschäftigter	2013, 2014
wb015_2_imp	Anzahl beförderter Beschäftigter (imputiert)	2013, 2014
wb016_1	Wechsel des Aufgaben-/Tätigkeitsbereichs	2013, 2014
wb016_2	Anzahl Beschäftigter, die ihren Aufgaben-/Tätigkeitsbereich wechseln	2013, 2014
wb016_2_imp	Anzahl Beschäftigter, die ihren Aufgaben-/Tätigkeitsbereich wechseln (imputiert)	2013, 2014
wb017_einf1	Beschäftigte mit einfachen Tätigkeiten: Teilnahme an Fachvorträgen, Tagungen	2012, 2014, 2015, 2017
wb017_einf2	Beschäftigte mit einfachen Tätigkeiten: Teilnahme an Qualitätszirkeln, Lernstattgruppen	2012, 2014, 2015, 2017
wb017_einf3	Beschäftigte mit einfachen Tätigkeiten: Unterweisungen/Anlernen am Arbeitsplatz	2012, 2014, 2015, 2017
wb017_einf4	Beschäftigte mit einfachen Tätigkeiten: Mentorenprogramme, Einzelberatung, Coaching	2012, 2014, 2015, 2017
wb017_einf5	Beschäftigte mit einfachen Tätigkeiten: Selbstgesteuertes Lernen	2012, 2014, 2015, 2017

Variable	Variablenlabel	Abgefragte Jahre
wb017_qual1	Beschäftigte mit qualifizierten Tätigkeiten: Teilnahme an Fachvorträgen, Tagungen	2012, 2014, 2015, 2017
wb017_qual2	Beschäftigte mit qualifizierten Tätigkeiten: Teilnahme an Qualitätszirkeln, Lernstattgruppen	2012, 2014, 2015, 2017
wb017_qual3	Beschäftigte mit qualifizierten Tätigkeiten: Unterweisungen/Anlernen am Arbeitsplatz	2012, 2014, 2015, 2017
wb017_qual4	Beschäftigte mit qualifizierten Tätigkeiten: Mentorenprogramme, Einzelberatung, Coaching	2012, 2014, 2015, 2017
wb017_qual5	Beschäftigte mit qualifizierten Tätigkeiten: Selbstgesteuertes Lernen	2012, 2014, 2015, 2017
wb017_hoch1	Beschäftigte mit hochqualifizierten Tätigkeiten: Teilnahme an Fachvorträgen, Tagungen	2012, 2014, 2015, 2017
wb017_hoch2	Beschäftigte mit hochqualifizierten Tätigkeiten: Teilnahme an Qualitätszirkeln, Lernstattgruppen	2012, 2014, 2015, 2017
wb017_hoch3	Beschäftigte mit hochqualifizierten Tätigkeiten: Unterweisungen/Anlernen am Arbeitsplatz	2012, 2014, 2015, 2017
wb017_hoch4	Beschäftigte mit hochqualifizierten Tätigkeiten: Mentorenprogramme, Einzelberatung, Coaching	2012, 2014, 2015, 2017
wb017_hoch5	Beschäftigte mit hochqualifizierten Tätigkeiten: Selbstgesteuertes Lernen	2012, 2014, 2015, 2017
wb018	Durchführung von Weiterbildungsmaßnahmen aufgrund gesetzlicher Vorschriften	2013, 2014
di018_1	Auswirkungen digitaler Technologien: Entlastungen der Beschäftigten von Routinetätigkeiten	2016, 2017
di018_2	Auswirkungen digitaler Technologien: Geringqualifizierte bessere Chancen durch digitale Unterstützung	2016, 2017
di018_3	Auswirkungen digitaler Technologien: Wegfall bestimmter Tätigkeiten und Berufe	2016, 2017
di018_4	Auswirkungen digitaler Technologien: Hochqualifizierte werden von Routinetätigkeiten entlastet	2016, 2017
di018_5	Auswirkungen digitaler Technologien: Nachfrage nach betrieblich ausgebildeten Fachkräften wird sinken	2016, 2017
di018_6	Auswirkungen digitaler Technologien: Es werden Arbeitsplätze wegfallen	2016, 2017
mo001einf_1	Beschäftigte mit einfachen Tätigkeiten: Häufigkeit: Tätigkeiten, alle Einzelheiten vorgeschrieben	2012, 2015, 2016
mo001einf_2	Beschäftigte mit einfachen Tätigkeiten: Häufigkeit: Tätigkeiten, sich wiederholende Abläufe	2012, 2015, 2016
mo001einf_3	Beschäftigte mit einfachen Tätigkeiten: Häufigkeit: Werkzeuge/Maschinen benutzen	2012, 2015, 2016
mo001einf_4	Beschäftigte mit einfachen Tätigkeiten: Häufigkeit: Fingerfertigkeiten/handwerkliches Geschick	2012, 2015, 2016
mo001einf_5	Beschäftigte mit einfachen Tätigkeiten: Häufigkeit: Kunden- oder Patientenberatung	2012, 2015, 2016
mo001einf_6	Beschäftigte mit einfachen Tätigkeiten: Häufigkeit: andere überzeugen/Kompromisse aushandeln	2012, 2015, 2016
mo001einf_7	Beschäftigte mit einfachen Tätigkeiten: Häufigkeit: Abläufe organisieren oder forschen	2012, 2015, 2016
mo001einf_8	Beschäftigte mit einfachen Tätigkeiten: Häufigkeit: Verfahren und Prozesse verbessern/neu erproben	2012, 2015, 2016

Variable	Variablenlabel	Abgefragte Jahre
mo001qual_1	Beschäftigte mit qualifizierten Tätigkeiten: Häufigkeit: Tätigkeiten, alle Einzelheiten vorgeschrieben	2012, 2015, 2016
mo001qual_2	Beschäftigte mit qualifizierten Tätigkeiten: Häufigkeit: Tätigkeiten, sich wiederholende Abläufe	2012, 2015, 2016
mo001qual_3	Beschäftigte mit qualifizierten Tätigkeiten: Häufigkeit: Werkzeuge/Maschinen benutzen	2012, 2015, 2016
mo001qual_4	Beschäftigte mit qualifizierten Tätigkeiten: Häufigkeit: Fingerfertigkeiten/handwerkliches Geschick	2012, 2015, 2016
mo001qual_5	Beschäftigte mit qualifizierten Tätigkeiten: Häufigkeit: Kunden- oder Patientenberatung	2012, 2015, 2016
mo001qual_6	Beschäftigte mit qualifizierten Tätigkeiten: Häufigkeit: andere überzeugen/Kompromisse aushandeln	2012, 2015, 2016
mo001qual_7	Beschäftigte mit qualifizierten Tätigkeiten: Häufigkeit: Abläufe organisieren oder forschen	2012, 2015, 2016
mo001qual_8	Beschäftigte mit qualifizierten Tätigkeiten: Häufigkeit: Verfahren und Prozesse verbessern/neu erproben	2012, 2015, 2016
mo001hoch_1	Beschäftigte mit hochqualifizierten Tätigkeiten: Häufigkeit: Tätigkeiten, alle Einzelheiten vorgeschrieben	2012, 2015, 2016
mo001hoch_2	Beschäftigte mit hochqualifizierten Tätigkeiten: Häufigkeit: Tätigkeiten, sich wiederholende Abläufe	2012, 2015, 2016
mo001hoch_3	Beschäftigte mit hochqualifizierten Tätigkeiten: Häufigkeit: Werkzeuge/Maschinen benutzen	2012, 2015, 2016
mo001hoch_4	Beschäftigte mit hochqualifizierten Tätigkeiten: Häufigkeit: Fingerfertigkeiten/handwerkliches Geschick	2012, 2015, 2016
mo001hoch_5	Beschäftigte mit hochqualifizierten Tätigkeiten: Häufigkeit: Kunden- oder Patientenberatung	2012, 2015, 2016
mo001hoch_6	Beschäftigte mit hochqualifizierten Tätigkeiten: Häufigkeit: andere überzeugen/Kompromisse aushandeln	2012, 2015, 2016
mo001hoch_7	Beschäftigte mit hochqualifizierten Tätigkeiten: Häufigkeit: Abläufe organisieren oder forschen	2012, 2015, 2016
mo001hoch_8	Beschäftigte mit hochqualifizierten Tätigkeiten: Häufigkeit: Verfahren und Prozesse verbessern/neu erproben	2012, 2015, 2016
be022	Entwicklung des Geschäftsvolumens im Vergleich zum Vorjahr	2015–2017
be042_1	Wichtigkeit als Wettbewerbsfaktor: Qualitativ hochwertige Produkte/Dienstleistungen	2011, 2013
be042_2	Wichtigkeit als Wettbewerbsfaktor: Neuartige Produkt- und Dienstleistungsangebote	2011, 2013
be042_3	Wichtigkeit als Wettbewerbsfaktor: Niedriger Preis	2011, 2013

Tabelle A3: Vorgenommene Harmonisierungen an allen Variablen und Hinweise für den Zeitvergleich

Variable(n)	Label	Jahr(e)	Unterschiede	Vorgehen
<b>BIBB Qualifizierungspanel_11bis17_long_gwa</b>				
<b>Themenblock: Ausbildung</b>				
ab001	Zum 31.12. des Vorjahres Azubis im Betrieb beschäftigt?	2012 bis 2014	Die Frage wurde kombiniert mit ab002 gestellt.	Die Variable wurde auf das System ab 2015 codiert. Es wurden die Kategorien 1 bis 3 zu ‚1=ja‘ zusammengefasst.
ab002	Art der Azubis (nach BBig, Hw0 oder anderen Ausbildungsberufen)?	2011	Die Frage wurde nicht gestellt.	Die Fälle wurden entsprechend codiert.
		2012–2014	Die Frage wurde kombiniert mit ab001 gestellt und die Formulierung der Antwortkategorien weicht von den Folgejahren ab.	Die Informationen aus ab001 wurden übertragen.
ab004a*kldb2010_5er	Bezeichnung Ausbildungsberuf nach BBig oder Hw0 (1) Kldb20	2011 bis 2013	Die Variable enthält die Kldb 1992.	Alle Variablen wurden auf die Kldb 2010 umcodiert. Die Kldb 1992 ist im Datensatz nicht enthalten.
		2011	In diesem Jahr weicht die Frageformulierungen von den anderen Jahren ab: Es wurde nicht explizit nach Auszubildenden in BBig/Hw0-Berufen gefragt.	
		alle	Die Kategorie ‚keine weiteren Ausbildungsberufe‘ ist in den Querschnitten nicht einheitlich codiert.	In allen Wellen wurde der Wert ‚99999, kein weiterer Ausbildungsberuf genannt‘ codiert.
ab004*	PR DI	alle	Die Variablen wurden zum Teil aus der Kldb 1992 und zum Teil aus der Kldb 2010 codiert.	Alle Variablen wurden aus der im Datensatz enthaltenen Kldb 2010 codiert.
ab004_b*	Anzahl Azubis am 31.12. des Vorjahres	alle	Die Kategorie ‚keine weiteren Ausbildungsberufe‘ ist in den Querschnitten nicht einheitlich codiert.	In allen Wellen wurde der Wert ‚99999, kein weiterer Ausbildungsberuf genannt‘ codiert.

<i>Variable(n)</i>	<i>Label</i>	<i>Jahr(e)</i>	<i>Unterschiede</i>	<i>Vorgehen</i>
ab015	Neue Azubis nach BBiG/Hw0 zum Ausbildungsjahr ein- gestellt?	2011	Die Frage wurde in 2011 nicht so gestellt wie in den Folge- jahren.	Es wurde die Frage a29a genutzt, um die Information für die Variable zu erzeugen. Alle Fälle, in denen a29a==1, d.h. Ausbildungsstellen wurden angeboten und Auszubildende eingestellt, wurden zu ab015==1, d.h. Ja, neue Auszubildende eingestellt, codiert. Alle Fälle, in denen a29a==2, d.h. Ausbildungsstellen wur- den angeboten aber keine Auszubildenden eingestellt, wurden zu ab015==2, d.h. nein, keine neuen Aus- zubildenden eingestellt, codiert. Alle Fälle, in denen a29a==3, d.h. keine Ausbildungsstellen wurden ange- boten, zu ab015==2, d.h. nein, keine neuen Auszubil- denden eingestellt, codiert; unter der Annahme, dass, wenn keine Ausbildungsstellen angeboten werden, auch keine Auszubildenden eingestellt werden können.
		2011; 2012	Die Filterführung weicht von den Folgejahren ab. Die Frage wurde in diesen Jahren auch Betrieben gestellt, die an- gegeben haben, keine Auszubildenden nach BBiG/Hw0 zu beschäftigen (ab001==1 oder ab002==3). Diese Betriebe werden ab 2013 direkt auf die nächste Frage gefiltert.	
ab016	Zum Ausbildungs- Jahr Ausbildungs- stellen nach BBiG/ Hw0 angeboten?	2011	Statt einer der Antwortoptionen ja und nein, gibt es drei Möglichkeiten. 1: Ja, hat angeboten und eingestellt 2: Ja, hat angeboten und niemanden eingestellt 3: Nein	Die erste Antwortkategorie wurde als Filter codiert, weil die Frage in den Folgejahren nur Betrieben gestellt wurde, die keine neuen Auszubildenden eingestellt haben. Die zweite Antwortkategorie wurde als ja und die dritte Antwortkategorie als nein codiert.  Durch die Uneinheitlichkeit der Frageführung ist es mög- lich, dass 2011 tendenziell mehr Betriebe gefiltert wur- den als in den Folgejahren und die Antwortkategorie ,2 nein' unterschätzt wird. Die Originalvariable a29a verbleibt im entsprechenden Querschnittsdatensatz.
ab017	Anzahl neu einge- stellter Azubis nach BBiG/Hw0 zum Aus- bildungsjahr	2011	Die Werte wurden getrennt nach kaufmännisch-verwaltenden und gewerblich-technischen Berufen abgefragt.	Die Werte wurden addiert.

<i>Variable(n)</i>	<i>Label</i>	<i>Jahr(e)</i>	<i>Unterschiede</i>	<i>Vorgehen</i>
ab018*	Abschlussart neue Azubis	2011, 2012	Es wurde nach den Abschlüssen aller Auszubildenden gefragt. In den Folgejahren wurde nur nach Auszubildenden in BBiG/Hw0-Berufen gefragt. Demzufolge ist die Anzahl in den Jahren 2011 und 2012 tendenziell höher.	
ab027*	Angebote für Azubis	2015, 2016	Die abgefragten Items sind über die Jahre nicht einheitlich.	Es wurden die Kategorien ,9996. Item 2016 nicht erhoben' und 9995. Item 2015 nicht erhoben" codiert. Theoretisch ließen sich die Variablen ab027_7 und ab027_08 zur Variable ab027_6 kombinieren, dies wurde in den Daten aber nicht gemacht, sondern wird den Nutzenden überlassen.
ab032	Gab es Ausbildungsstellen nach BBiG/Hw0, die nicht besetzt wurden?	Ab 2013	Die Frage wurde um folgenden Hinweis ergänzt: „Berücksichtigen Sie bitte hier auch die Ausbildungsstellen für das Ausbildungsjahr, die wegen gelöster Ausbildungsverträge nicht besetzt sind.“	
ab033	Anzahl Ausbildungsstellen nach BBiG/Hw0, die nicht besetzt wurden?	2011	Die Werte wurden getrennt nach kaufmännisch-verwaltenden und gewerblich-technischen Berufen abgefragt.	Die Werte wurden addiert.
		Ab 2013	Die Frage wurde um folgenden Hinweis ergänzt: „Berücksichtigen Sie bitte hier auch die Ausbildungsstellen für das Ausbildungsjahr, die wegen gelöster Ausbildungsverträge nicht besetzt sind.“	
ab036	Ausbildungsverträge nach BBiG/Hw0 im Kalenderjahr vorzeitig gelöst	2011	Die Angaben beziehen sich nicht wie in den anderen Jahren auf das Kalenderjahr, sondern auf das Ausbildungsjahr.	
		2014	Die Kategorien weichen von den anderen Jahren ab. Es gibt die zusätzliche Kategorie ,3 Trifft nicht zu, hatten im Jahr 2013 keine Auszubildenden nach BBiG oder Hw0'.	Die Kategorie 3 wurde zu , -1 Filter' codiert.

<i>Variable(n)</i>	<i>Label</i>	<i>Jahr(e)</i>	<i>Unterschiede</i>	<i>Vorgehen</i>
ab037	Anzahl vorzeitig gelöster Ausbildungsverträge im Kalenderjahr	2011	Die Angaben beziehen sich nicht wie in den anderen Jahren auf das Kalenderjahr, sondern auf das Ausbildungsjahr.  Die Abfrage erfolgte getrennt für gewerblich-technische und kaufmännisch-verwaltende Berufe.	Beide Werte wurden addiert. In einem Fall entsteht aus der Kombination , -9 keine Angabe' bei gewerblich-technisch und null Vertragslösungen bei kaufmännisch-verwaltend der Wert 0.
ab038	Haben Azubis nach BBiG/Hw0 an der Abschlussprüfung teilgenommen?	2011  2013	Die Frage war anders formuliert.  Ab 2014 wurde in der Frage E14 (Vertragslösungen von Ausbildungsverträgen im Kalenderjahr 2015) gefiltert und somit zum nächsten Fragenblock geleitet, wenn die Betriebe angegeben haben, dass sie keine Auszubildenden nach BBiG oder Hw0 hatten. Dadurch beantworten diese Betriebe Frage E16 nicht (ob Abschlussprüfungen für Auszubildende nach BBiG oder Hw0 stattfanden) und weisen bei der entsprechenden Variable den Wert , -1 Filter' auf, wodurch die Kategorie ,3 Keine Auszubildenden' entfällt. Im Jahr 2013 dagegen wurde die Frage nach den Abschlussprüfungen ohne vorherige Filter gestellt.	Die Fälle dieser Kategorie 3 wurden zu , -1 Filter' codiert.
ab041*	Anzahl Azu- bis: Bestandene Abschlussprüfung	2012 bis 2014	Die Frage wurde in diesen Jahren nicht gestellt.	Die Werte wurden als Differenz der Auszubildenden, die zur Abschlussprüfung angetreten sind (ab039*), und den nicht bestandenen Prüfungen (ab043*) berechnet.  Die Filterführung wurde angepasst, sodass alle Fälle den Wert , -1 Filter' aufweisen, die auch bei ab039* gefiltert wurden.

Variable(n)	Label	Jahr(e)	Unterschiede	Vorgehen
ab043*	Übernahme der Auszubildenden	alle	Die Itemformulierung variiert zwischen den Jahren. Gleich sind die Formulierungen in den Jahren 2011, 2015 und 2016. Ebenfalls gleich sind die Formulierungen in den Jahren 2013 und 2014 (aber ein Item mehr). Das 3. Item ist für alle Jahre gleich. Die sonstigen Unterschiede sind in den folgenden Zeilen dokumentiert.	
		2015	Nur 2015 wurde die Kategorie ,9999, Keine gewerblich-technischen (bzw. kaufmännisch-verwaltenden) Auszubildenden' codiert.	Die Kodierung wurde angepasst, indem die Kategorie ,9999, Keine gewerblich-technischen (bzw. kaufmännisch-verwaltenden) Auszubildenden' in ,-1 Filter' codiert wurde.
ab043_1 und ab043_2	Vom Betrieb befristet übernommen und Vom Betrieb befristet übernommen	2012	Die Items wurden nicht erhoben. Dafür wurden folgende Items abgefragt: „Wurden als Fachkraft übernommen“ und „Wurden in andere Position übernommen“.	Als Variablen b19*a und b19*b im entsprechenden Querschnitt Datensatz enthalten.
ab043_2	Vom Betrieb befristet übernommen	2013, 2014	Die Items wurden abweichend formuliert.	
ab043_4	Betriebsbedingt nicht übernommen	2012 bis 2014	Das Item wurde nicht erhoben.	
		2012 bis 2013	Das Item wurde nicht erhoben.	
ab043_5	Wegen Prüfungsnoten nicht übernommen	2014	Das Item wurde zusammen mit „haben die Prüfung bestanden und wurden betriebsbedingt nicht übernommen“ erhoben.	Als Variable d17*d im entsprechenden Querschnitt Datensatz enthalten.
ab043_6	Abschlussprüfung nicht bestanden	2011, 2015, 2016	Das Item wurde nicht erhoben.	Der Wert wurde berechnet, indem von der Anzahl der zur Prüfung angetretenen Auszubildenden (ab039) die Anzahl der bestandenen Prüfungen (ab041*) subtrahiert wurde.

<i>Variable(n)</i>	<i>Label</i>	<i>Jahr(e)</i>	<i>Unterschiede</i>	<i>Vorgehen</i>
ab058_*	Gründe selbst auszubilden	2011	Die Frage war anders formuliert; es wurde eine andere Skala verwendet und die Frage wurde allen Betrieben gestellt, nicht nur solchen, die ausbilden, da die Frage sich nicht im Ausbildungsteil des Fragebogens befindet. Jeder Nutzende sollte kritisch prüfen, ob ein Zeitvergleich sinnvoll ist. Weiterhin wurden das Item _4 „Ausbildung verursacht hohe Kosten“ nur 2011 erhoben, das Item _5 „Tradition des Betriebes“ dagegen nicht.	
		Ab 2015	Es gibt die zusätzliche 5. Kategorie	Für einen Zeitvergleich können die Kategorien 4 und 5 zusammengefasst werden, dies wurde im Datensatz jedoch nicht getan, sondern bleibt dem/der Nutzenden selbst überlassen.
		2016	Die Items _6 „Fachkräfte ausbilden, die langfristig bleiben“ und _7 „Ausbildung als Gemeinschaftsaufgabe der Wirtschaft“ gibt es nur in diesem Jahr.	
ab087*	Prozentuales Leistungsniveau der Auszubildenden	2011, 2012, 2013	Im Querschnitt war die Kategorie „Ausbildungsjahr nicht besetzt“ nicht einheitlich codiert.	Die Kategorie „9999. Ausbildungsjahr nicht besetzt“ wurde einheitlich codiert.
<b>Themenblock: Personal</b>				
ps020*	Anzahl Beschäftigte nach Beschäftigten-Gruppe	2011	Bei der Anzahl der Beschäftigten wird zwischen vier (einfache, qualifizierte für gewöhnlich abgeschlossene Lehre, qualifizierte für gewöhnlichen Meister- oder Technikerabschluss, hochqualifizierte Tätigkeiten) statt drei (einfache, qualifizierte, hochqualifizierte Tätigkeiten) Beschäftigten-Gruppen unterschieden.	Die Kategorien qualifizierte Tätigkeiten für gewöhnlich abgeschlossene Lehre, qualifizierte Tätigkeiten für gewöhnlichen Meister- oder Technikerabschluss wurden zu qualifizierten Tätigkeiten zusammengefasst.
pb005*	Anzahl neu eingestellter Beschäftigter	2011	Bei den neu eingestellten Beschäftigten wird zwischen vier (einfache, qualifizierte für gewöhnlich abgeschlossene Lehre, qualifizierte für gewöhnlichen Meister- oder Technikerabschluss, hochqualifizierte Tätigkeiten) statt drei (einfache, qualifizierte, hochqualifizierte Tätigkeiten) Beschäftigten-Gruppen unterschieden.	Die Kategorien qualifizierte Tätigkeiten für gewöhnlich abgeschlossene Lehre, qualifizierte Tätigkeiten für gewöhnlichen Meister- oder Technikerabschluss wurden zu qualifizierten Tätigkeiten zusammengefasst.

<i>Variable(n)</i>	<i>Label</i>	<i>Jahr(e)</i>	<i>Unterschiede</i>	<i>Vorgehen</i>
pb011*	Anzahl unbesetzter Stellen	2011	Bei der Anzahl unbesetzter Stellen wird zwischen vier (einfache, qualifizierte für gewöhnlich abgeschlossene Lehre, qualifizierte für gewöhnlichen Meister- oder Technikerabschluss, hochqualifizierte Tätigkeiten) statt drei (einfache, qualifizierte, hochqualifizierte Tätigkeiten) Beschäftigtengruppen unterschieden.	Die Kategorien qualifizierte Tätigkeiten für gewöhnlich abgeschlossene Lehre, qualifizierte Tätigkeiten für gewöhnlichen Meister- oder Technikerabschluss wurden zu qualifizierten Tätigkeiten zusammengefasst.
pb017*		2017	Frage eindeutig gefiltert, deswegen Programmierung 9999 nicht nötig	Die Kategorie ,9999. keine Beschäftigten ausgeschließen' wurde durch -1 ersetzt.
<b>Themenblock: Weiterbildung</b>				
wb008, wb009	Haben Beschäftigte in an betrieblichen geförderten WB-Maßnahmen teilgenommen, Anzahl Beschäftigte mit Teilnahme an Weiterbildungsmäßnahmen	Ab 2015	Der Frage ist ab 2015 die Frage nach der Teilnahme von Beschäftigten an Aufstiegsfortbildungen (wb001) vorgestellt. Es ist davon auszugehen, dass Aufstiegsfortbildungen in wb008 nicht miterfasst werden. Vor 2015 war die Frage nach Aufstiegsfortbildungen nachgelagert oder wurde noch nicht gestellt. Es ist zu vermuten, dass ohne spezielle Nachfrage die Aufstiegsfortbildungen in wb008 enthalten sind. Zudem wird erst ab 2015 bei der Frage nach der Anzahl der Weiterbildungsteilnehmer/-innen (wb009) auf Folgendes hingewiesen: ,Teilnehmer an Aufstiegsfortbildungen oder einem berufsbegleitenden Studium bitte nicht berücksichtigen'. Für einen Zeitvergleich kann es also sinnvoll sein, die Variablen wb001 und wb008 bzw. wb002 und wb009 (und e40b ,berufsbegleitendes Studium' für 2015) zu verbinden. Diese Entscheidung bleibt aber jedem/jeder Nutzenden selbst überlassen. Es findet keine Harmonisierung vom FDZ statt.	
wb010*	Anzahl der WB-Teilnehmer	2011	Bei der Frage nach dem Umfang der Weiterbildung wurde 2011 nach der Anzahl von Teilnehmern und Teilnehmerinnen (Personen) bzw. Teilnahmefällen (wb011*) gefragt, in den übrigen Jahren hingegen nach der Anzahl der Weiterbildungsteilnehmer/-innen (wb010*).	Anhand der Information, ob Personen oder Teilnehmerfälle berichtet wurden (a44), werden die Angaben aus der Variable wb011* in die Variable wb010* übernommen. Für die anderen Fälle wird der Code ,9995. Teilnehmerfälle berichtet' vergeben.

<i>Variable(n)</i>	<i>Label</i>	<i>Jahr(e)</i>	<i>Unterschiede</i>	<i>Vorgehen</i>
wb017_1-wb017_5	Formen der betrieblichen Weiterbildung	2012, 2014, 2015, 2017 2011	Die Items wurden getrennt für einfache, qualifizierte und hochqualifizierte Beschäftigte abgefragt. Die Codierung der Antwortmöglichkeiten lautet ,0 nein' statt ,2 nein'.	Die Werte der drei Variablen wurden zusammengefasst. Die originalen Variablen sind im long2-Datensatz enthalten. Die Codierung wird angepasst ,0 nein' wird zu ,2 nein'.
<b>Themenblock: Allgemeine Angaben zum Betrieb</b>				
wz	Wirtschaftszweig Betrieb	2011	Im Jahr 2011 wurden noch 45 statt 43 Wirtschaftszweige abgefragt.	Die Wirtschaftscodes 2011 wurden folgendermaßen umcodiert (alt=neu): 39=41; 40=42; 41=43; 42=44; 43=45; 44=46; 38=47 (wobei 47 ein Sondercode ist, den es nur für das Jahr 2011 gibt). Die WZ-Codes 38, 39 und 40 sind für 2011 dagegen nicht vergeben.
bra	Branchenzugehörigkeit (Serien-Gliederung)	2011-2013	Die Zuordnung der Wirtschaftszweige zu den Branchen (welche auch zur Schichtung dienen) wurde geändert.	Die Jahre 2011-2013 wurden nach der aktuellen Einteilung neu codiert. Der entsprechende Code ist in FRIEDRICH/GERHARDS 2017.
tj_ki	technological intensity_knowledge-intensive services	Alle		Die Variable wurde nicht aus dem Querschnitt übernommen, sondern für den Längsschnitt neu gebildet.
be001*	Regelungen für Beschäftigte	2012, 2014	Die abgefragten Items sind über die Jahre nicht einheitlich.	Es wurde die Kategorien ,9996 Item 2012 und 2014 nicht erhoben' codiert.
be010_*	Veränderungen im Vorjahr	2012 und 2013	Die Angaben für 2012 und 2013 beziehen sich auf Veränderungen in den letzten zwei Jahren. Ab 2014 wurde nach Veränderungen im Vorjahr gefragt. Außerdem wurde die Frage 2013 nur neuen Betrieben gestellt, also keinen Panelbetrieben.	
be018_1	Art des Betriebs	2012 und 2013	Die abgefragten Items sind über die Jahre nicht einheitlich.	Es wurde die Kategorien ,9996 Item 2012 und 2013 nicht erhoben' codiert.
		2012	Die Fragen wurden 2012 nur an Panelbetriebe gestellt.	Für Panelbetriebe wurde die Information aus dem Vorjahr übernommen.
		Ab 2015	Die Reihenfolge der abgefragten Kategorien wurde geändert.	Die Jahre vor 2015 wurden an die geänderte Reihenfolge angepasst.

<i>Variable(n)</i>	<i>Label</i>	<i>Jahr(e)</i>	<i>Unterschiede</i>	<i>Vorgehen</i>
be018_2	Geschäftsvolumen	2012	Panelbetriebe wurden nicht direkt nach ihrem Geschäftsvolumen gefragt, sondern ob es zurückgegangen oder gestiegen ist und, wenn ja, um wie viel Prozent.	Mithilfe des Geschäftsvolumens 2011 und der Angaben zur Veränderung aus 2012 wurde das Geschäftsvolumen berechnet (vom Projekt). Die Angaben der Panelbetriebe und der erstbefragten Betriebe wurden kombiniert.
be020	Anteil des Auslandsumsatzes am Gesamtumsatz	2012	Panelbetriebe wurden nicht direkt nach dem Anteil des Auslandsumsatzes gefragt, sondern danach, ob er zurückgegangen oder gestiegen ist und, wenn ja, um wie viel Prozent.	Mithilfe des Anteils des Auslandsumsatzes 2011 und der Angaben zur Veränderung aus 2012 wurde der Anteil des Auslandsumsatzes am Gesamtumsatz berechnet (vom Projekt). Die Angaben der Panelbetriebe und der erstbefragten Betriebe wurden kombiniert.
be021	Anteil von Vorleistungen und Fremdkosten am Gesamtumsatz	2012	Panelbetriebe wurden nicht direkt nach dem Anteil der Vorleistungen gefragt, sondern danach, ob er zurückgegangen oder gestiegen ist und, wenn ja, um wie viel Prozent.	Mithilfe des Anteils der Vorleistungen 2011 und der Angaben zur Veränderung aus 2012 wurde der Anteil der Vorleistungen und Fremdkosten am Gesamtumsatz berechnet (vom Projekt). Die Angaben der Panelbetriebe und der erstbefragten Betriebe wurden kombiniert.
be023	Summe aller Investitionen im Betrieb im letzten Geschäftsjahr	2012	Panelbetriebe wurden nicht direkt nach der Summe der Investitionen gefragt, sondern danach, ob sie zurückgegangen oder gestiegen sind und, wenn ja, um wie viel Prozent.	Mithilfe der Summe der Investitionen 2011 und der Angaben zur Veränderung aus 2012 wurde die Summe aller Investitionen im Betrieb im letzten Geschäftsjahr berechnet (vom Projekt). Die Angaben der Panelbetriebe und der erstbefragten Betriebe wurden kombiniert.

<i>Variable(n)</i>	<i>Label</i>	<i>Jahr(e)</i>	<i>Unterschiede</i>	<i>Vorgehen</i>
		2012	Panelbetriebe wurden nicht direkt nach dem Bruttolohn gefragt, sondern danach, ob er zurückgegangen oder gestiegen ist und, wenn ja, um wie viel Prozent.	Mithilfe des Bruttolohns 2011 und der Angaben zur Veränderung aus 2012 wurde der durchschnittliche Bruttolohn berechnet (vom Projekt). Die Angaben der Panelbetriebe und der erstbefragten Betriebe wurden kombiniert.
be026*	Durchschnittlicher Bruttolohn	Alle	Die Kategorie ‚Beschäftigtengruppe nicht vorhanden‘ wurde in den einzelnen Jahren uneinheitlich codiert.	Die Kategorie ‚999999. Beschäftigtengruppe nicht vorhanden‘ wurde einheitlich codiert. Für die kategoriale Einteilung des durchschnittlichen Bruttolohns wurde die Kategorie ‚99. Beschäftigtengruppe nicht vorhanden‘ codiert. Die Codierung der Kategorie ‚Beschäftigtengruppe nicht vorhanden‘ erfolgte 2012 ohne die Information zu ausgeschiedenen Beschäftigten. Fälle, die einen ‚-1 Filter‘ enthalten – also nicht in ‚Beschäftigtengruppe nicht vorhanden‘ codiert wurden – wurden in ‚-6 weiß nicht/keine Angabe‘ codiert, da es laut Fragebogen bei dieser Frage keine Filter gibt.
be026_qual_imp	Durchschnittlicher Bruttolohn Beschäftigter mit qualifizierten Tätigkeiten (imputiert)	2012	Die Variable enthält nur imputierte Werte und keine originalen Werte.	In der Variable wurden alle ‚-1 Filter‘-Werte durch die Originalwerte ersetzt.
be026*k	Durchschnittlicher Bruttolohn , gruppiert	2013 2014	Die Kategorie ‚-1 Filter‘ ist nicht einheitlich codiert.	Alle Fälle, die gültige metrische Werte aufweisen, werden auf ‚-1 Filter‘ gesetzt.

Variable(n)	Label	Jahr(e)	Unterschiede	Vorgehen
be032	Branchen- oder Haus- bzw. Firmentarifvertrag	2012 bis 2014 Ab 2015	Die Frage nach dem Tarifvertrag wird nur neuen Betrieben gestellt. Mit der Welle 2015 ermöglicht die Frage nach dem Tarifvertrag Mehrfachantworten.	Für die Jahre 2012, 2013 und 2014 wurde die Variable mit den Informationen aus dem Vorjahr gefüllt. Die beiden Variablen zu b032_1 und b032_2 wurden in einer Variablen zusammengeführt. Dafür wurde die zusätzliche Kategorie, 4. Branchen- und Haus-/Firmentarifvertrag' codiert.
be034	Vorhandensein eines Betriebs- oder Personalrats	2012 bis 2014	Die Frage nach dem Betriebs-/Personalrat wird nur neuen Betrieben gestellt.	Die beiden Variablen a73 und a74 werden in be032 kombiniert.
be035*	Kammerzugehörigkeit	Alle	Die abgefragten Items sind über die Jahre nicht einheitlich.	Für die Jahre 2012, 2013 und 2014 wurde die Variable mit den Informationen aus dem Vorjahr gefüllt. Es wurde die Kategorie ,96. Item bis 2016 nicht abgefragt' codiert.
be035_1; be035_2; be035_7	Kammerzugehörigkeit	2012 bis 2014 2011	Die Frage nach der Kammerzugehörigkeit wird nur neuen Betrieben gestellt. Die Codierung der Antwortmöglichkeiten lautet ,0 nein' statt ,2 nein'.	Für die Jahre 2012, 2013 und 2014 wurde die Variable mit den Informationen aus dem Vorjahr gefüllt. Die Codierung wird angepasst ,0 nein' wird zu ,2 nein'.
be035_7	Betrieb Mitglied in einer anderen Kammer	Ab 2016	Mit der Welle 2016 wird die Kammerzugehörigkeit genauer erfasst, sprich mehr Kategorien werden abgefragt und das Item ,Sonstiges' entfällt.	Wenn eine der Variablen be035_3 bis be035_6 den Wert ,1 ja' enthält, wird be035_7 in ,1 ja' umcodiert. Wenn alle Variablen den Wert ,2 nein' oder einmal einen ungültigen Wert aufweisen, wird be035_7 in ,2 nein' codiert.
be041	Gründungsjahr des Betriebs	Ab 2011	Die Frage wird nur an neue Betriebe gestellt.	Für Panelbetriebe wird die Variable mit den Informationen aus dem Vorjahr gefüllt. Alle Fälle, die noch den Wert , -1 Filter' aufweisen, werden in , -6 weiß nicht/keine Angabe' codiert, da die Frage laut Fragebogen keinen Filter aufweisen soll.
be043	Betriebsart	2012 bis 2014	Die Frage nach der Betriebsart wird nur neuen Betrieben gestellt.	Für die Jahre 2012, 2013 und 2014 wurde die Variable mit den Informationen aus dem Vorjahr gefüllt. Alle Fälle, die noch den Wert , -1 Filter' aufweisen, werden in , -6 weiß nicht/keine Angabe' codiert, da die Frage laut Fragebogen keinen Filter aufweisen soll.

<i>Variable(n)</i>	<i>Label</i>	<i>Jahr(e)</i>	<i>Unterschiede</i>	<i>Vorgehen</i>
<b>BIBB Qualifizierungspanel_11bis17_long2_gwa</b>				
<b>Themenblock: Ausbildung</b>				
ab086			Anteil produktiver Tätigkeiten an der Anwesenheit im Betrieb 2011, 2012, 2013 Im Querschnitt war die Kategorie ‚Ausbildungsjahr nicht besetzt‘ nicht einheitlich codiert.	Die Kategorie ,999. Ausbildungsjahr nicht besetzt‘ wurde einheitlich codiert.
ab087*			Prozentuales Leistungsniveau der Auszubildenden 2011, 2012, 2013 Im Querschnitt war die Kategorie ‚Ausbildungsjahr nicht besetzt‘ nicht einheitlich codiert.	Die Kategorie ,999. Ausbildungsjahr nicht besetzt‘ wurde einheitlich codiert.
<b>Themenblock: Personal</b>				
ps039	Beschäftigte mit Bachelor als höchstem beruflichem Abschluss	2013	Die Kategorie ,9999. Beschäftigtengruppe nicht vorhanden‘ ist nicht codiert. Die entsprechenden Fälle enthalten eine 0.	Für alle Fälle, die bei ps030_uni (Anzahl Beschäftigte mit (fach-)Hochschulabschluss) eine 0 aufweisen, wird bei Variable ps039 die 0 in ,9999. Beschäftigtengruppe nicht vorhanden‘ codiert.
<b>Themenblock: Weiterbildung</b>				
wb017_einf* wb017_qual* wb017_hoc*	Beschäftigte mit * Tätigkeiten: Selbstgesteuertes Lernen	2012, 2014, 2015, 2017	Die Kategorie ‚Beschäftigtengruppe nicht vorhanden‘ ist nicht einheitlich codiert. 2012 gibt es die Kategorie in den Querschnittsdaten nicht.	Für alle Jahre wurde einheitlich der Wert ,99. Beschäftigtengruppe nicht vorhanden‘ codiert. Die Codierung der Kategorie ‚Beschäftigtengruppe nicht vorhanden‘ erfolgte 2012 ohne die Information zu ausgeschiedenen Beschäftigten.

Variable(n)	Label	Jahr(e)	Unterschiede	Vorgehen
<b>Themenblock: Sondermodule</b>				
mo001*	Häufigkeit Tätigkeiten	2012, 2014, 2015	Die Kategorie ‚Beschäftigtengruppe nicht vorhanden‘ ist nicht einheitlich codiert. 2012 gibt es die Kategorie gar nicht, die Fälle enthalten den Wert ‚-1. Filter‘.	Für alle Jahre wurde einheitlich der Wert ‚99. Beschäftigtengruppe nicht vorhanden‘ codiert.  Die Codierung der Kategorie ‚Beschäftigtengruppe nicht vorhanden‘ erfolgte 2012 ohne die Information zu unterschiedlichen Beschäftigten.  Fälle, die eine ‚-1 Filter‘ enthalten – also nicht in ‚Beschäftigtengruppe nicht vorhanden‘ codiert wurden – wurden in ‚-6 weiß nicht/keine Angabe‘ codiert, da es laut Fragebogen bei dieser Frage keine Filter gibt.
		2012	2012 ist die Antwortskala: ‚1. Häufig‘; ‚2. Manchmal‘ und ‚3. Nie‘; 2014 und 2015 ist die Antwortskala: ‚1. Nie‘; ‚2. Sehr selten‘; ‚3. Eher selten‘; ‚4. Eher häufig‘; und ‚5. Sehr häufig‘.	Die Antwortskala für 2011 wurde angepasst und sieht jetzt folgendermaßen aus: ‚1. Nie‘; ‚6. Manchmal‘; ‚7. Häufig‘.



Bundesinstitut für Berufsbildung  
Robert-Schuman-Platz 3  
53175 Bonn

Telefon (0228) 107-0

Internet: [www.bibb.de](http://www.bibb.de)  
E-Mail: [zentrale@bibb.de](mailto:zentrale@bibb.de)