

Christian Gerhards | Anett Friedrich

BIBB-Qualifizierungspanel 2014

Version 1.1



Heft 3/2016

Christian Gerhards | Anett Friedrich

BIBB-Qualifizierungspanel 2014

Version 1.1

► Danksagung

Besonderer Dank für die Unterstützung bei der Durchführung der Studie geht zum einen an Dr. Sebastian Bechmann, Kerstin Sleik, Matthias Passmann und Josef Hartmann von TNS Infratest Sozialforschung. Zum anderen bedanken wir uns beim Institut für Arbeitsmarkt- und Berufsforschung (IAB), insbesondere bei Phillip Grunau, für die Bereitstellung der Adressbasis bei der Stichprobenziehung.

► Datenverfügbarkeit

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

► Hinweis zur Zählweise bei Versionsnummern

Änderungen gegenüber der Vorversion 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.

Bibliografische Information der Deutschen Nationalbibliothek

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

© 2016 by Bundesinstitut für Berufsbildung, Bonn

BIBB-FDZ Daten- und Methodenbericht 3/2016, Version 1.0

Herausgeber:

Bundesinstitut für Berufsbildung, Bonn
Internet: www.bibb.de

Publikationsmanagement:

AB 1.4. Ursula Knüpper-Heger/Sigrid Koppenhöfer

Gesamtherstellung:

W. Bertelsmann Verlag GmbH & Co. KG

Vertriebsadresse:

Bundesinstitut für Berufsbildung
Arbeitsbereich 1.4 – Publikationsmanagement/Bibliothek
– Veröffentlichungen –
Robert-Schuman-Platz 3
53175 Bonn
E-Mail: zentrale@bibb.de

Kontakt FDZ:

E-Mail: fdz@bibb.de
Tel.: +49 228 107-2041
Fax: +49 228 107-2020

Printed in Germany

ISSN: 2190-300X



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

Weitere Informationen finden Sie im Internet auf unserer Creative-Commons-Infoseite www.bibb.de/cc-lizenz.

Diese Netzpublikation wurde bei der Deutschen Nationalbibliothek angemeldet und archiviert: [urn:nbn:de:](http://nbn-resolving.org/urn:nbn:de:dnb:1-1072041)

Internet: www.bibb.de/veroeffentlichungen

Downloads unter: www.bibb-fdz.de

Inhaltsverzeichnis

Tabellen	5
1 Einleitung	6
1.1 Gegenstand des Handbuchs	6
1.2 Aufbau des Handbuchs	6
1.3 Kurzbeschreibung des Datensatzes der Welle 2014	6
1.4 Zentrale Eigenschaften im Überblick	7
2 Neuerungen in der Erhebungswelle 2014	8
2.1 Aufstockung der Fallzahl	8
2.2 Themenschwerpunkte	8
2.3 Vignettenmodul	8
3 Ergebnisse der Durchführung der Welle 2014	9
3.1 Bruttostichprobe und Rücklauf	9
3.2 Ergebnis zum Einsatz unterschiedlicher Erhebungsinstrumente	9
3.3 Panelmortalität	9
4 Item-Non-Response und Imputation	11
4.1 Item-Non-Response	11
4.2 Imputation	11
5 Gewichtung und Hochrechnung	13
6 Organisation der Mikrodaten	14
6.1 Variablenbeschreibung und Hinweise zu einzelnen Variablen	14
6.2 Fehlende Werte	15
6.3 Im Fragebogen nicht erkennbare Filter	15
6.4 Wiederholte Abfrage konstanter Betriebsmerkmale	16
6.5 Abgrenzung von Querschnitt- und Paneldatensatz	17
6.6 Vergleichbarkeit des Querschnitts mit den Vorjahren	18
7 Praktischer Umgang mit den Datensätzen	20
7.1 Praktischer Umgang mit dem Querschnittsdatsatz	20
7.2 Praktischer Umgang mit dem Panel-Datensatz	21
8 Datenanonymisierung und Datenzugang	24
8.1 Bestimmungen des Datenschutzes	24
8.2 Formale Datenanonymisierung	24
8.3 Datenzugang	24
9 Literaturverzeichnis	26

Anhang	27
Anhang A1 Variablen mit fehlenden Angaben über zehn Prozent	27
Anhang A2 Übersicht über Ursprungsvariablen und imputierte Variablen	28

Tabellen

Tab. 1: Selektivität über die Erhebungswellen 2011 bis 2014	10
Tab. 2: Fallzahlen für die unterschiedlichen Längsschnitte	13
Tab. 3: Beschäftigtenzahl laut D26xxx13 für das Jahr 2013 in Welle 2014	14
Tab. 4: Im Fragebogen nicht erkennbare Filter	15
Tab. 5: Unterschiede zwischen den Abfragen konstanter Betriebsmerkmale	16
Tab. 6: Querschnittsfälle und Wellenkennzeichen	17
Tab. 7: Panelfälle 2011–2014 und Wellenkennzeichen	18
Tab. 8: Häufigkeit der Durchführung von Weiterbildung aufgrund einer gesetzlichen Vorschrift	19
Tab. 9: Anzahl der Teilnahme der Betriebe an den Wellen des BIBB- Qualifizierungspanels	21
Tab. 10: Nicht verfügbare Variablen für das BIBB-Qualifizierungspanel 2014	24

1 Einleitung

1.1 Gegenstand des Handbuchs

Der vorliegende Daten- und Methodenbericht beschreibt das BIBB-Betriebspanel zu Qualifizierung und Kompetenzentwicklung 2014 (kurz: BIBB-Qualifizierungspanel 2014). Es handelt sich um die vierte Welle der Panelstudie. Der Bericht soll externen Wissenschaftlerinnen und Wissenschaftlern beim Umgang mit den Daten der Erhebungswelle 2014 helfen. Allgemeine Informationen zu Konzept und Methoden des BIBB-Qualifizierungspanels sind in den Daten- und Methodenberichten der Jahre 2011, 2012 und 2013¹ ausführlich beschrieben. Wichtige Kerngrößen, wie der Rücklauf der einzelnen Erhebungen, die Zahl der Panelfälle oder die Zuteilung der Branchen zu den Sichtungscodes, wurden für alle Wellen in der Excel-Datei *Querschnitt- und Panelcharakteristika des BIBB-Qualifizierungspanels* von HOHN u. a. 2016 gebündelt und stehen auf der BIBB-FDZ-Metadatenportalseite des BIBB-Qualifizierungspanels zur Verfügung.

1.2 Aufbau des Handbuchs

Im Abschnitt 1.3 wird zunächst ein Überblick über die Welle 2014 des BIBB-Qualifizierungspanels gegeben. In Kapitel 2 werden anschließend die Neuerungen der Erhebung in der Welle 2014 gegenüber den Vorjahren beschrieben. Kapitel 3 enthält die Ergebnisse der Durchführung der Welle 2014, Kapitel 4 gibt Auskunft über Item-Non-Response und die Imputation fehlender Werte. Das 5. Kapitel skizziert das Vorgehen bei der Gewichtung und Hochrechnung der Daten. Die Organisation der Mikrodaten wird in Kapitel 6 erläutert, und in Kapitel 7 werden anschließend praktische Tipps im Umgang mit den Daten gegeben. Das 8. und letzte Kapitel beschreibt die vom BIBB-FDZ vorgenommenen Datenanonymisierungsschritte sowie den Datenzugang für die BIBB-externe Forschung.

1.3 Kurzbeschreibung des Datensatzes der Welle 2014

Das BIBB-Qualifizierungspanel 2014 ist eine repräsentative Betriebsbefragung von 3.450 Betrieben in Deutschland zur betrieblichen Qualifizierung und Kompetenzentwicklung. Zur Grundgesamtheit gehören alle Betriebe mit mindestens einer sozialversicherungspflichtigen Beschäftigung (SVB). Die Gesamtstudie des BIBB-Qualifizierungspanels ist im Längsschnittdesign als Panel angelegt und wird seit 2011 durchgeführt.

Durchgeführt wurde das BIBB-Qualifizierungspanel 2014, wie in den Jahren zuvor, von TNS Infratest Sozialforschung als CAPI-Studie (Computer Assisted Personal Interview). Die Feldzeit erstreckte sich von September 2014 bis April 2015. Aus einer Bruttostichprobe von 15.419 Betrieben resultierte nach Abzug qualitätsneutraler Ausfälle (13.605 verbleibende Adressen) ein Rücklauf von 3.450 Betrieben im Querschnitt. Mittels eines Hochrechnungsgewichts sind repräsentative Aussagen für Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten in Deutschland für den Querschnitt und für Längsschnitte möglich.

¹ Die entsprechenden Berichte können auf den Internetseiten des BIBB-FDZ oder im Metadatenportal des BIBB-FDZ heruntergeladen werden.

Das BIBB-Qualifizierungspanel als Gesamtstudie liefert Informationen insbesondere zu Themen der betrieblichen Aus- und Weiterbildung, aber auch zur Personalstruktur und der Arbeitsorganisation in den Betrieben. Die Erhebung 2014 hat darüber hinaus den Forschungsschwerpunkt „latente Auswahlkriterien und praktizierte Auswahlverfahren in der betrieblichen Ausbildung“ und die bildungspolitischen Schwerpunkte „rückläufige Ausbildungsbeteiligung“ und „abnehmende Neuvertragszahlen“.

1.4 Zentrale Eigenschaften im Überblick

Befragungstitel	BIBB-Betriebspanel zu Qualifizierung und Kompetenzentwicklung 2014
DOI	10.7803/371.14.1.2.10
Kurzbeschreibung	Betriebspanel mit dem Ziel, detaillierte Informationen über die Strukturen, Entwicklungen und Zusammenhänge betrieblicher Qualifizierungsmaßnahmen und betrieblich-qualifikatorischer Arbeitskräftenachfrage zu liefern
Erhebungsjahr	2014
Erhebungseinheit	Betriebe
Themenschwerpunkt	Ausbildung, Weiterbildung
Datenzugangsmöglichkeiten	Gastwissenschaftleraufenthalt/Datenfernverarbeitung
Variablenanzahl	331 (GWA)
Grundgesamtheit	alle Betriebe in Deutschland mit mindestens einem sozialversicherungspflichtig Beschäftigten
Gewichtung/Hochrechnung	Hochrechnungsfaktoren (auf Basis der Betriebsdatei der BA), Längsschnitt-Hochrechnungsfaktoren (2011 bis 2014, 2012 bis 2014 und 2013 bis 2014)
repräsentative Region	Ost- und Westdeutschland
Fallzahl	3.450 (davon 557 Panelbetriebe zwischen 2011 und 2014)
Erhebungsverfahren	CAPI, PAPI (auf Wunsch des Betriebes)
Auswahlverfahren	disproportional geschichtete Zufallsauswahl
Erhebungsdesign	Längsschnittdesign (Panelstudie)
Bemerkung	Wiederbefragungsbereitschaft: 75,9 %
Links	BIBB-FDZ Metadatenportal: http://metadaten.bibb.de/metadaten/76 Projektseite: http://www.bibb.de/qp
Stichworte	Qualifizierung, Ausbildung, Personalbewegungen, Arbeitsorganisation, Personalstruktur, (vorzeitige) Vertragslösungen, Bildungswege, Organisation und Durchführung von Weiterbildung, Entwicklung der Ausbildungsbeteiligung, innerbetriebliche Karrierewege

2 Neuerungen in der Erhebungswelle 2014

2.1 Aufstockung der Fallzahl

Bei den drei bisherigen Erhebungswellen von 2011 bis 2013 handelte es sich um Befragungen von ca. 2.000 Betrieben in Deutschland. Für die Welle 2014 fand eine Aufstockung der Netto-stichprobe auf 3.500 Betriebe statt. Damit sind nun für noch detailliertere Teilgruppen, z. B. nach einzelnen Branchen, differenzierte Analysen möglich. Das Stichprobeverfahren wurde durch die Aufstockung nicht verändert, es handelt sich nach wie vor um eine disproportional geschichtete Zufallsauswahl.

2.2 Themenschwerpunkte

Zusätzlich zu den jährlich wiederkehrenden Inhalten des BIBB-Qualifizierungspanels enthält die Welle 2014 folgende Themenschwerpunkte:

- ▶ die Entwicklung der Ausbildungsbeteiligung; mit Fragen zu den mit dem Fachkräftebedarf, den Bewerbern, den Kosten bzw. der Organisation und betrieblichen Entwicklungen zusammenhängenden Gründen für den Rückgang von Neuvertragszahlen,
- ▶ innerbetriebliche Karrierewege; mit Fragen zu Aufstiegsfortbildungen, Beförderungen und Aufgaben- bzw. Tätigkeitswechseln im Betrieb.

2.3 Vignettenmodul

Eine Besonderheit der Erhebungswelle 2014 ist das zusätzlich zum Fragebogen implementierte Vignettenmodul. Die Planung und Umsetzung dieses Vignettenexperiments erfolgte in enger Zusammenarbeit mit dem Wissenschaftszentrum Berlin (WZB), welches sowohl für die inhaltliche Ausrichtung als auch für methodische Konzeptionierung der Vignetten verantwortlich war. Untersucht wurden latente Auswahlkriterien und praktizierte Auswahlverfahren in der betrieblichen Ausbildung.

Das Ziel des Vignettenmoduls ist es, Aussagen darüber zu erlangen, welche Personenmerkmale der Bewerber für die Besetzung von Ausbildungsplätzen für Betriebe entscheidend sind. Dazu wurden zwei verschiedene Modelle mit jeweils unterschiedlichen Schwerpunkten entwickelt.

Für die Erfassung der Rekrutierung von Auszubildenden wurde das Vignettenmodul sowohl ausbildenden Betrieben als auch nicht ausbildenden Betrieben, welche versucht haben, Ausbildungsstellen zu besetzen, vorgelegt. Das Vignettenmodul wurde durch den Interviewten eigenständig am Computer (Computer Assisted Self-Interview – CASI) durchgeführt. Aus diesem Grund kamen nur die 2.247 computergestützt interviewten Betriebe für die Teilnahme am Vignettenexperiment infrage, von diesen wiederum erfüllten 1.468 Betriebe die weiteren Auswahlbedingungen.

Eine detaillierte Beschreibung des Vignettenmoduls wird in einem gesonderten Daten- und Methodenbericht erfolgen, da die entsprechenden Daten nicht Teil des Datensatzes des Qualifizierungspanels 2014 sind und gesondert veröffentlicht werden.

3 Ergebnisse der Durchführung der Welle 2014

Mit der Feldarbeit zur vierten Welle des BIBB-Qualifizierungspanels wurde am 6. September 2014 begonnen. Es wurden zwei Feldeinsatzstichproben eingesetzt. Mit der ersten wurde zunächst versucht, 13.015 Betriebe zu kontaktieren. Die wesentlich kleinere zweite Feldeinsatzstichprobe mit 2.404 Betriebsadressen wurde ab dem 21. März 2015 im Feld eingesetzt. Das Feldende war schließlich der 27. April 2015.

3.1 Bruttostichprobe und Rücklauf

Insgesamt wurden 15.419 Betriebsadressen eingesetzt. Nach Abzug qualitätsneutraler Ausfälle blieben 13.605 Betriebsadressen. Insgesamt konnten in 3.461² Betrieben erfolgreich Interviews durchgeführt werden. Davon können 3.450 Fälle im Querschnitt ausgewertet werden.

Bei 13.605 eingesetzten Adressen ohne qualitätsneutrale Ausfälle und 3.552 durchgeführten Interviews³ ergab sich eine Teilnahmequote von 26,1 Prozent. Die Teilnahmequote wird definiert als der Quotient aus der Anzahl auswertbarer Interviews und der Anzahl nicht auswertbarer Interviews geteilt durch die Anzahl der eingesetzten Adressen ohne qualitätsneutrale Ausfälle:

$$\text{Teilnahmequote} = \frac{\text{auswertbare} + \text{nicht auswertbare Interviews}}{\text{eingesetzte Adressen} - \text{qualitätsneutrale Ausfälle}}$$

3.2 Ergebnis zum Einsatz unterschiedlicher Erhebungsinstrumente

Das Haupterhebungsverfahren der Studie waren computerunterstützte persönliche Interviews (CAPI-Interviews). Daneben war es den Betrieben möglich, das Interview ganz oder teilweise als Papierversion selbst auszufüllen. Dadurch ergaben sich drei mögliche Interview szenarien:

- ▶ als reine CAPI-Interviews wurden 2.167 Interviews (62,5 Prozent) durchgeführt;
- ▶ als teilweise CAPI, teilweise Selbstausfüller wurden 80 Interviews (2,3 Prozent) durchgeführt;
- ▶ ausschließlich als Selbstausfüller wurden 1.219 Interviews (35,2 Prozent) durchgeführt.

3.3 Panelmortalität

Die Indikatoren zur Panelmortalität der Erhebungswelle 2014 sind nur bedingt mit denen der vorherigen Wellen zu vergleichen, da in diesem Jahr die Nettostichprobe von rund 2.000 auf rund 3.500 Betriebe erhöht wurde (vgl. Abschnitt 2.1). Daher ist die Auffrischerstichprobe mit neuen Betrieben wesentlich erweitert worden. Gleichzeitig zeigt sich dennoch, dass die Wiederbefragungsbereitschaft für die Teilnahme an der Welle 2015 geringfügig nachgelassen hat.

Tabelle 1 zeigt die erfolgreich durchgeführten Interviews des Querschnitts (3.450), gegliedert nach den Teilgruppen der alten Betriebe, die bereits in vorangegangenen Wellen teilgenommen haben, und neuen Betrieben. Aufgrund der Aufstockung der Stichprobe stammen nur 38,7 Pro-

² Hinweis: Im Infratest-Bericht sind 3.466 Betriebe als zum Längsschnitt zugehörig ausgewiesen. Gegenüber den 3.461 Fällen, die hier ausgewiesen sind, handelt es sich um Betriebe, die nachträglich nicht für Auswertungen in der Welle 2014 genutzt werden können.

³ Zu den durchgeführten Interviews zählen alle auswertbaren und nicht auswertbaren Interviews.

zent dieser Betriebe aus der bisherigen Stichprobe. Von den wiederbefragungsbereiten Betrieben der letzten Erhebungswelle (1.775) nahmen 1.245 wieder teil (70,1 Prozent).

Tabelle 1

Selektivität über die Erhebungswellen 2011 bis 2014

	<i>Welle 2011</i>	<i>Welle 2012</i>	<i>Welle 2013</i>	<i>Welle 2014</i>
Teilnehmer (Querschnittsfälle)	2.004	2.006	2.026	3.450
davon alte Betriebe	–	1.168	1.321	1.336
% alte Betriebe	–	58,2	65,2	38,7
davon neue Betriebe	2.004	838	705	2.114
% neue Betriebe	100	41,8	34,8	61,3
Wiederbefragungsbereit	1.618	1.726	1.775	2.617
% an Teilnehmern	80,7	86,0	87,6	75,9
davon aus alten Betrieben	–	1.078	1.217	1.163
% an alten Betrieben	–	92,3	92,1	87,1
davon aus neuen Betrieben	1.618	648	558	1.454
% an neuen Betrieben	80,7	77,3	79,1	68,8
Teilnahme bei Wiederbefragungsbereitschaft in der letzten Welle	–	1.168	1.256	1.245
% an den Wiederbefragungsbereiten der letzten Welle		72,1	72,8	70,1
% an Teilnehmern letzte Welle		58,3	62,6	61,5

4 Item-Non-Response und Imputation

4.1 Item-Non-Response

Im Verlauf der Interviews kam es in Einzelfällen bei bestimmten Fragen zu Antwortausfällen (Item-Non-Response). Wie auch in anderen Studien und in den vergangenen Erhebungswellen zeigen sich bei Abfragen von Geschäftszahlen (Anteil von Vorleistungen und Fremdkosten, Höhe des Geschäftsvolumens, Summe aller Investitionen und der durchschnittliche Bruttolohn nach Beschäftigtengruppen) die größten Anteile an Item-Non-Response. Dies ist mit der Vertraulichkeit und der Verfügbarkeit der Informationen verbunden. Ebenfalls typisch sind hohe Missing-Anteile bei Abfragen zu Personalbestand und -bewegung (insbesondere Anzahl der Bewerber um Ausbildungsstellen und die Anzahl Beschäftigter zum 31. Dezember 2013 nach berufsbildenden Abschlüssen).

Hohe Raten von Item-Non-Response zeigten sich auch bei der Abfrage spezieller, nicht kursförmiger Weiterbildungsmaßnahmen. Eine Tabelle mit allen Variablen, welche fehlende Anteile über 10 Prozent aufweisen, ist im Anhang A1 zu finden.

4.2 Imputation

Für eine Reihe kontinuierlicher Merkmale mit fehlenden Angaben wurden regressionsanalytische Imputationen durchgeführt, wodurch das Problem des Item-Non-Response verringert wird. Ziel war, für weiter gehende Analysen eine möglichst vollständige Wertematrix nutzbar zu machen. Zur Ersetzung von fehlenden Werten durch Antwortausfälle wurden zu den betreffenden Variablen daher imputierte Versionen mit ergänzten Werten generiert. Für die jeweilige Imputation wurde eine Regressionsanalyse durchgeführt. Dazu wurde wie folgt vorgegangen: Imputiert wurden ausschließlich kontinuierliche Merkmale, welche für weitere Auswertungsmerkmale eine wichtige Basis darstellen, wie z. B. die Gesamtzahl von Weiterbildungsteilnehmern in einem Betrieb als Basis für die Berechnung des Anteils an Weiterbildungsteilnehmern nach Beschäftigtengruppen.

In den meisten Fällen wurden für die Imputation zu den zu imputierenden Variablen zunächst Anteilswerte berechnet, die sich auf eine übergeordnete Basis beziehen (z. B. zur Gesamtzahl der Beschäftigten im Betrieb). Diese Variablen in Form von Anteilswerten dienten als abhängige Variablen der Imputationen. Nach der Imputation wurden diese Variablen wieder in Absolutwerte umgerechnet. Dieses Vorgehen hat den Vorteil, dass die Imputationen auf den Wertebereich von 0 bis 100 Prozent beschränkt werden und durch vorangegangene Filterführungen wesentlich plausiblere Werte berechnet werden, denn häufig interessieren gerade diese Anteilswerte statt der Absolutwerte.

Für die Imputationen wurde grundsätzlich ein gemeinsames Modell gewählt, das wichtige Strukturvariablen berücksichtigt. Bei den Modellen stand weniger im Vordergrund, dass sie inhaltlich informativ sind, sondern vielmehr statistische Interdependenzen (Korrelationen) einzelner Variablen berücksichtigt sind (zum Verfahren vgl. ALDA/ROHRBACH-SCHMIDT 2011). Die unabhängigen Variablen für das (maximale) Modell waren Ausbildungsbetrieb (ja/nein), Größenklasse, West/Ost, Branche, Anteil Auszubildender, Anzahl Neueinstellungen, Anteile der Tätigkeitsgruppen und Weiterbildungsbeteiligung.

Um möglichst viele der ausgewählten unabhängigen Variablen nutzen, aber gleichzeitig auch möglichst viele Fälle imputieren zu können, wurde ein Verfahren in drei Schritten gewählt. Dazu wird zunächst ein maximales Modell berechnet, das alle der ausgewählten unabhängigen Variablen zur Imputation nutzt. Für Fälle, bei denen dieses Modell aufgrund fehlender Werte bei den unabhängigen Variablen keinen Imputationswert liefert, wurde ein vereinfachtes und ggf. anschließend ein minimales Modell berechnet. Das minimale Modell beinhaltet die unabhängigen Variablen Ausbildungsbetrieb (ja/nein), Größenklasse, West/Ost, Branche.

Bei einigen Modellen wurden teils spezifische Variablen als unabhängige Variablen berücksichtigt bzw. weggelassen, wenn diese abhängige Variablen der Imputation waren oder sonst zu Problemen der Heteroskedastizität geführt hätten.

Zur Imputation wird der lineare Vorhersagewert des Regressionsmodells unter Addition eines Störterms bei je 1.000 Wiederholungen verwendet (vgl. ALDA/ROHRBACH-SCHMIDT 2011, Kapitel 2.2.2). Gegenüber den nicht mehr dem State of the Art entsprechenden Verfahren der einfachen Imputation hat dieses Verfahren mit Störterm den Vorteil, dass dadurch die Varianz nicht unterschätzt wird. Zudem hat dieses Vorgehen gegenüber nutzerspezifischen multiplen Imputationen für das Angebot eines für die wissenschaftliche Öffentlichkeit verfügbaren Datensatzes den Vorteil, dass reproduzierbare (d. h. für alle Nutzer gleiche) imputierte Werte zur Verfügung gestellt werden.

Insgesamt wurden für 81 Variablen Imputationen durchgeführt (vgl. im Anhang Tabelle A2). Bei nur wenigen, aber gleichzeitig zentralen Variablen wurden mehr als 10 Prozent der Werte der resultierenden Variablen imputiert.

Nach der Imputation wurden, sofern Anteilswerte an Referenzvariablen für die Imputation gebildet wurden, diese wieder in Absolutzahlen zurückgerechnet (mit entsprechender ganzzahliger Rundung). Dabei wurden erneut Prüfungen durchgeführt, ob die Werte zwischen 0 und 100 Prozent lagen (und ggf. auf 0 bzw. 100 Prozent angepasst) und ob bei Summenvariablen die Summanden zusammenpassten (und ggf. proportional angepasst). Im finalen Datensatz, der im BIBB-FDZ zur wissenschaftlichen Nutzung zur Verfügung gestellt wird, sind neben den Ursprungsvariablen diese imputierten Variablen enthalten (jeweils mit dem Kürzel „_imp“ versehen).

5 Gewichtung und Hochrechnung

Durchgeführt wurde die Gewichtung und Hochrechnung von TNS Infratest Sozialforschung in Absprache mit dem BIBB. Ziel ist es, zum einen den disproportionalen Stichprobenansatz auszugleichen und zum anderen möglichen Verzerrungen durch disproportionale Antwortausfälle entgegenzuwirken. Grundlage für die Gewichtung ist die Betriebsdatei der Bundesagentur für Arbeit zum 31. Dezember 2013. Alle Gewichtungen und Hochrechnungen wurden jeweils betriebs- und beschäftigtenproportional vorgenommen. Die für die Gewichtungen maßgeblichen Sollvorgaben können HOHN u. a. (2016) entnommen werden.

Im Querschnitt wurden unterschiedliche Gewichte berechnet. Zunächst wurden Design- und Ausfallgewichte ermittelt. Anschließend wurde darauf aufbauend eine Strukturanpassungsgewichtung berechnet und eine Hochrechnung durchgeführt, wobei nur der finale Hochrechnungsfaktor (*hr2014q*) im Datensatz enthalten ist.

Wie genau die einzelnen Gewichte berechnet wurden und wie der anschließende Hochrechnungsfaktor gebildet wurde, kann TNS INFRATEST SOZIALFORSCHUNG (2015, S. 25 ff.) entnommen werden.

Ergänzend zum Hochrechnungsfaktor des Querschnitts sind im Datensatz ebenfalls drei Panelhochrechnungsfaktoren enthalten (*hr11_14p*, *hr12_14p*, *hr13_14p*), diese unterscheiden sich danach, welcher Zeitraum maßgeblich für die Definition von Panelfällen ist:

- ▶ *hr11_14p* für Fälle, die in den Wellen 2011 bis 2014 teilgenommen haben,
- ▶ *hr12_14p* für Fälle, die von 2012 bis 2014 teilgenommen haben, und
- ▶ *hr13_14p* für Fälle, die von 2013 bis 2014 teilgenommen haben.

Dabei werden Unterbrechungen zugelassen.

Die Gewichte passen die unterschiedlichen Fallgruppen jeweils so an, dass für die Angaben zu den einzelnen überdeckten Jahren die Ränder möglichst gut getroffen werden. Für die Fälle der Welle 2014 ergibt sich, dass 1.076 Fälle für den Längsschnitt 2013 bis 2014 betrachtet werden können, 791 Fälle für den Längsschnitt 2012 bis 2014 und 557 Fälle für den Längsschnitt 2011 bis 2014 (vgl. Tabelle 2).

Tabelle 2

Fallzahlen für die unterschiedlichen Längsschnitte

Variable	N	Mittelwert	Standardabw.	Minimum	Maximum
<i>hr11_14p</i>	557	3.672	7.868	8,3	4.0226
<i>hr12_14p</i>	791	2.600	4.863	40,1	19.662
<i>hr13_14p</i>	1.076	1.818	3.715	40,0	19.979

6 Organisation der Mikrodaten

6.1 Variablenbeschreibung und Hinweise zu einzelnen Variablen

Die Benennung der inhaltlichen Variablen folgt der Fragennummerierung aus dem Fragebogen. Die Variablennamen der inhaltlichen Variablen beginnen mit „d“ für Welle 4 („a“ stand für Welle 1 und „b“ für Welle 2 etc.), gefolgt von der zweistelligen Fragennummer im Fragebogen. Besteht eine Frage aus Unterfragen, so sind diese durch einen Zusatz gesondert aufgeführt (z. B. „d01a“, „d01b“ und „d01c“). Zum Teil wurden „sprechende“ Kürzel für diese Zusätze verwendet. Wiederkehrende Kürzel können HOHN u. a. (2016) entnommen werden. Die Besonderheiten von einzelnen Variablen in der Welle 2014 werden im Folgenden aufgeführt.

Bei der Variable d19 („Ausbildungsverträge davor neu abgeschlossen“) wurden 13 Werte nachträglich als „-3. unplausibler Wert“ codiert. Es handelt sich dabei um Fälle, die weder laut Frage D18 noch D19 in den letzten drei Jahren oder davor Auszubildende eingestellt haben. Zudem haben die Betriebe in den Fragen nach den Vertragslösungen (D14) oder Abschlussprüfungen (D15) gültige Angaben gemacht. Es handelt sich also um Betriebe, die scheinbar nie Ausbildungsverträge abgeschlossen haben, aber gleichzeitig Vertragslösungen oder Abschlussprüfungen hatten. Dies wird als unplausibel angesehen, und die Fälle wurden entsprechend umcodiert. Für die folgenden Fragen (D20 bis D25) weisen die Betriebe jeweils einen Filter auf.

Die Angaben aus dem Textfeld der Frage D25 zu sonstigen Gründen für Ausbildungsrückgänge sind nicht im Datensatz enthalten. Sofern möglich, wurden die genannten Gründe den Items der Fragen D21 bis D24 zugeordnet und die entsprechenden Variablen umcodiert. Weiterhin wurden häufig „ungeeignete Bewerber“ als Grund genannt. Diese Information wurde in der zusätzlich erstellten Variabel *d22c_z* gebündelt.

Die Gesamtzahlen der Beschäftigten zum 31. Dezember 2013 und zum 31. Dezember 2012 sind in Frage D26 erfasst. Die einzelnen Variablen beinhalten die in Tabelle 3 beschriebenen Informationen.

Tabelle 3

Beschäftigtenzahl laut D26xxx13 für das Jahr 2013⁴ in Welle 2014

Variable	Bedeutung
d26svb13	Anzahl 31.12.13: Sozialversicherungspflichtig Beschäftigte
d26ger13	Anzahl 31.12.13: Geringfügig Beschäftigte
d26nsv13	Anzahl 31.12.13: Nicht sozialversicherungspflichtig Beschäftigte
d26zwi13	Summe 31.12.13: Alle Beschäftigten ohne Auszubildende (d26svb13 + d26ger13 + d26nsv13)
d26aus13	Anzahl 31.12.13: Auszubildende nach BBiG/HwO lt. D02.a
d26soz13	Anzahl 31.12.13: Auszubildende nach Ausbildungsregelung des Gesundheitswesens oder sozialpflegerischer und -pädagogischer Berufe

⁴ Die Benennung der Variablen für das Jahr 2012, bis auf den Suffix „12“, sind identisch.

Bei Frage D52 zum durchschnittlichen Bruttolohn bzw. -gehalt eines vollzeitbeschäftigten Mitarbeiters nach Beschäftigtengruppen wurden Betriebe, die die Antwort verweigert haben, bei CAPI-Interviews gebeten, diesen kategorial anzugeben. Der Datensatz beinhaltet deswegen zusätzlich die Variablen *d52einfk*, *d52qualk* und *d52hochk* („k“ für kategorial). Diese enthalten zum einen die Angaben aus den CAPI-Interviews und zum anderen die originalen Angaben, welche den entsprechenden Kategorien zugeordnet wurden.

6.2 Fehlende Werte

Fehlende und ungültige Werte wurden nach der Standarddeklarierung fehlender Werte des BIBB-FDZ, unter Beachtung von Besonderheiten der Codierung von TNS Infratest Sozialforschung, für alle Variablen des BIBB-Qualifizierungspanels umcodiert, um so den Umgang mit den fehlenden Werten zu erleichtern. HOHN u. a. (2016) enthält die Werte und Wertelabels für die fehlenden Werte bei allen Variablen im BIBB-Qualifizierungspanel.

In Datensätzen des BIBB-FDZ werden grundsätzlich keine Missing Values definiert – die Vergabe wird den Anwendenden überlassen. Eine Ausnahme bilden die Längsschnittgewichtungen (*hr11_14p*, *hr12_14p*, *hr13_14p*), diese enthalten für die entsprechenden Nichtpanelfälle Missing Values. Darüber hinaus enthalten alle Fälle ein Missing Value, welche nicht zum Querschnittsdatsatz, aber zum Längsschnittsdatsatz gehören (vgl. Abschnitt 6.5).

6.3 Im Fragebogen nicht erkennbare Filter

Zu einigen wenigen Variablen wurden nachträglich im Fragebogen nicht ersichtliche Filter codiert. Der Grund ist, dass spezielle Fragen, die bestimmte Beschäftigtengruppen betreffen, den Betrieben nur dann gestellt wurden, wenn diese Beschäftigtengruppe im Betrieb auch vorhanden war, ohne dies explizit im Fragebogen zu vermerken. Um diese Fälle von den anderen fehlenden Werten leichter abgrenzen zu können, wurde für diese Fälle der Wert „9999. Beschäftigtengruppe nicht vorhanden“ vergeben. Der Tabelle 4 kann entnommen werden, welche Variablen betroffen sind und nach welchen Variablen gefiltert wurde.

Tabelle 4

Im Fragebogen nicht erkennbare Filter

Variable ⁵	Variableninhalt	Filtervariable	Inhalt der Filtervariable	Regel der Umcodierung	Anzahl umcodierter Fälle
d31bac	Anzahl 31.12.13: Beschäftigte mit Bachelor als höchstem beruflichem Abschluss	d30uni	Anzahl 31.12.13: Beschäftigte mit Fachhochschul- oder Universitätsabschluss	Umcodiert wurden alle Fälle, die bei der Filtervariable den Wert null aufweisen	1.039
d52einf	Ø Bruttolohn/-gehalt: Beschäftigte mit einfachen Tätigkeiten	d29einf	Anzahl Beschäftigte 31.12.13: einfache Tätigkeiten	Umcodiert wurden alle Fälle, die bei der Filtervariable keinen gültigen Wert aufweisen (d. h. kleiner als null sind)	1.402

⁵ Für die imputierten Variablen gilt die Umcodierung ebenfalls.

(Fortsetzung Tab. 4)

Variable ⁵	Variableninhalt	Filtervariable	Inhalt der Filtervariable	Regel der Umcodierung	Anzahl umcodierter Fälle
d52qual	Ø Bruttolohn/-gehalt: Beschäftigte mit qualifizierten Tätigkeiten	d29qual	Anzahl Beschäftigte 31.12.13: qualifizierte Tätigkeiten	Umcodiert wurden alle Fälle, die bei der Filtervariable keinen gültigen Wert aufweisen (d. h. kleiner als null sind)	286
d52hoch	Ø Bruttolohn/-gehalt: Beschäftigte mit hoch qualifizierten Tätigkeiten	d29hoch	Anzahl Beschäftigte 31.12.13: hoch qualifizierte Tätigkeiten	Umcodiert wurden alle Fälle, die bei der Filtervariable keinen gültigen Wert aufweisen (d. h. kleiner als null sind)	774

6.4 Wiederholte Abfrage konstanter Betriebsmerkmale

Für wiederholt teilnehmende Betriebe wurden eine Reihe von Angaben nicht erneut abgefragt, da diese, wie bspw. das Jahr der Betriebsgründung, über die Jahre konstant sind. Damit die Informationen für Querschnittanalysen dennoch zur Verfügung stehen, werden Angaben aus den Vorjahren in den Datensatz integriert. Um welche Variablen es sich handelt, kann HOHN u. a. (2016) entnommen werden.

Mit der Erhebung 2014 wurden ein Teil der in der Regel konstanten Betriebsmerkmale erneut abgefragt, und zwar bei Merkmalen, die sich doch langfristig verändern können. Es handelt sich dabei um die Frage zu den Tarifverträgen, der Personalvertretung und der Kammerzugehörigkeit. Der Datensatz beinhaltet für alle wiederholt befragten Betriebe sowohl die Angaben der ersten als auch der zweiten Abfrage, sodass Veränderungen beobachtbar sind. Die Informationen der ersten Abfrage sind in den angespielten Variablen *p_bhtv*, *p_br*, *p_hk*, *p_ihk* sowie *p_sk* und die aus der aktuellen sind in den Variablen *d53* bis *d55c* enthalten. Insgesamt sind die Angaben der Betriebe stringent, nur zwischen drei und zwölf Prozent der Betriebe weisen unterschiedliche Angaben auf (vgl. Tabelle 5).

Tabelle 5

Unterschiede zwischen den Abfragen konstanter Betriebsmerkmale⁶

Variablenname	Variableninhalt	Unterschiede absolut	gleich absolut	Unterschiede in %
<i>p_bhtv</i> und <i>d53</i>	Branchen- oder Haus- bzw. Firmentarifvertrag	130	990	11,6
<i>p_br</i> und <i>d54</i>	Betriebsrat/Personalrat nach BetrVG/PersVG	33	1.092	2,9
<i>p_hk</i> und <i>d55a</i>	Betrieb Mitglied in der Handwerkskammer	42	1.038	3,8
<i>p_ihk</i> und <i>d55b</i>	Betrieb Mitglied in der IHK	61	1.041	5,5
<i>p_sk</i> und <i>d55c</i>	Betrieb Mitglied in einer anderen Kammer	36	1.042	3,3

⁶ Fehlende Werte wurden in die Auswertung nicht einbezogen.

6.5 Abgrenzung von Querschnitt- und Paneldatensatz

Der Datensatz des BIBB-Qualifizierungspanels 2014 enthält insgesamt 3.518 Betriebe, von denen 68 nicht zum Querschnittsdatensatz gehören. Der Querschnitt umfasst entsprechend 3.450 Betriebe. Es gibt drei verschiedene Gründe, warum ein Betrieb im Datensatz enthalten ist, aber nicht zum Querschnitt gehört (vgl. Tabelle 6).

Tabelle 6

Querschnittsfälle und Wellenkennzeichen

Wellenkennzeichen 2014 (well2014)	Querschnittsfall 2014 (quer2014)		
	1. Ja	2. Nein	Gesamt
1. Erstinterview	2.332	0	2.332
10. Wiederholer, gleiche Erhebungseinheit, mind. 1 SVB	1.054	0	1.054
11. Wiederholer, gleiche Erhebungseinheit, keine SVB	0	11	11
20. Vorjahresausfall, mind. 1 SVB	64	0	64
50. erloschen in der laufenden Welle	0	28	28
51. erloschen in einer früheren Welle	0	29	29
Gesamt	3.450	68	3.518

Erstens, der Betrieb gehört für die Erhebung 2014 nicht zur Grundgesamtheit, er hatte also zum 31. Dezember 2013 keine sozialversicherungspflichtig Beschäftigten, ist aber dennoch Panelbetrieb. In diesem Fall wurde das Interview dennoch geführt, und es liegen gültige Angaben vor, welche für Längsschnittanalysen genutzt werden können, aber eben nicht zum Querschnitt gehören. Dies trifft für elf Betriebe zu.

Zweitens, der Betrieb ist in der aktuellen Welle erloschen, d.h., der entsprechende Betrieb existiert im Jahr 2014 nicht mehr. Diese Betriebe gehören nicht zum Querschnitt, es konnte kein Interview realisiert werden, und entsprechend liegen keine gültigen Angaben für die inhaltlichen Variablen vor. Da die Betriebe zum Panel gehören (vgl. TNS INFRATEST 2015, S. 28), gibt es entsprechende Längsschnittgewichte. Tabelle 7 kann entnommen werden, welche der insgesamt 28 Betriebe zu welchen Panelzeiträumen gehören und daher für diesen Zeitraum Längsschnittgewichte haben.

Drittens, der Betrieb ist in einer der vorhergehenden Wellen erloschen, womit folglich ebenfalls kein Interview geführt wurde und keine gültigen Angaben für die inhaltlichen Variablen vorliegen. Für diese 29 Betriebe gibt es Längsschnitt-Hochrechnungsfaktoren bis zu dem Jahr, in welchem der Betrieb erloschen ist. Tabelle 7 kann entnommen werden, für welchen Zeitraum die Betriebe als Panelbetriebe zählen.⁷

⁷ Zu beachten ist, dass ein einzelner Betrieb zu keinem Zeitpunkt als Panelbetrieb ausgewiesen wird und somit kein Längsschnittgewicht besitzt, auch wenn dies theoretisch nicht der Fall sein sollte.

Tabelle 7

Panelfälle 2011–2014 und Wellenkennzeichen

	Panelfall						Gesamt
	2013–2014		2012–2014		2011–2014		
	(pan13_14)	(pan12_14)	(pan12_14)	(pan12_14)	(pan11_14)	(pan11_14)	
Wellenkennzeichen 2014	1. Ja	2. Nein	1. Ja	2. Nein	1. Ja	2. Nein	
1. Erstinterview	0	2.332	0	2.332	0	2.332	2.332
10. Wiederholer, gleiche Erhebungseinheit, mind. 1 SVB	1.045	9	744	310	520	534	1.054
11. Wiederholer, gleiche Erhebungseinheit, keine SVB	7	4	7	4	6	5	11
20. Vorjahresausfall, mind. 1 SVB	0	64	0	64	0	64	64
50. erloschen in der laufenden Welle	24	4 ⁸	21	7	11	17	28
51. erloschen in einer früheren Welle	0	29	19	10	20	9	29
Gesamt	1.076	2.442	791	2.727	557	2.961	3.518

6.6 Vergleichbarkeit des Querschnitts mit den Vorjahren

Eine Aufschlüsselung, welche Fragen der Welle 2014 mit den Fragen der Vorjahre vergleichbar sind, ist in HOHN u. a. (2016) dokumentiert. Dort werden von allen in mehr als einer Welle auftauchenden Fragen die korrespondierenden Fragennummern aufgeführt, wobei Abweichung in den konkreten Frageformulierungen und Antwortvorgaben ggf. in den entsprechenden Fußnoten erläutert werden.

Neben den Formulierungen wurden in zwei Fällen auch die Filterführungen im Fragebogen geändert. Dies muss beachtet werden, wenn man in diesen Fällen Zeitvergleiche anstellt. Betroffen sind die Fragen D15/C22 (Abschlussprüfungen für Auszubildende) und D41/C48 (Weiterbildungsmaßnahmen aufgrund gesetzlicher Vorschriften).

2014 wurde in der neu auftretenden Frage D14b (Vertragslösungen von Ausbildungsverträgen im Kalenderjahr 2013) gefiltert und somit zum nächsten Frageblock geleitet, wenn die Betriebe angegeben haben, dass sie 2013 gar keine Auszubildenden nach BBiG oder HwO hatten. Dadurch beantworten diese Betriebe Frage D15 nicht (ob Abschlussprüfungen für Auszubildende nach BBiG oder HwO stattfanden) und weisen bei der entsprechenden Variable den Wert -1 „Filter“ auf. 2013 dagegen wurde die Frage nach den Abschlussprüfungen ohne vorherige Filter gestellt. Aus diesem Grund ist die Kategorie 3 „Keine Auszubildenden“ 2013 wesentlich stärker besetzt als 2014.⁹ Besonders bei der Analyse der Anteile muss darauf geachtet werden, entweder 2013 auch den Wert 3 als ungültig zu codieren oder 2014 den Wert 3 aus der vorherigen Variable zu übertragen. Die entsprechenden Stata- bzw. SPSS-Befehle lauten:

⁸ Diese vier Fälle müssten theoretisch auch Panelfälle sein. Da dies ist jedoch nicht der Fall ist, haben sie kein Längsschnittgewicht für den Zeitraum von 2013–2014 (drei Fälle) bzw. von 2011–2014, 2012–2014 und 2013–2014 (ein Fall).

⁹ Theoretisch dürfte die Kategorie sogar 2014 gar nicht besetzt sein, doch in einem Fall wurde die Frage D14b nicht beantwortet, weist also einen ungültigen Wert auf. Bei den anderen Fällen handelt es sich um PAPI-Fälle, sodass die Filterführung nicht automatisch geprüft wurde.

```
- 2014: replace d15=d14b if d14b==3
      if d14b eq 3 c15=3.
      exe.
```

```
- 2013: replace c22=-1 if c22==3
      if c22 eq 3 c22=-1.
      exe.
```

Die Frage, ob Weiterbildungsmaßnahmen aufgrund gesetzlicher Vorschriften angeboten wurden (D41), wurde 2014 allen Betrieben gestellt. Im Jahr davor dagegen wurden durch die Frage C46 alle Betriebe gefiltert, die keine Weiterbildung angeboten haben (also weder laut Frage C43 noch laut Frage C44). Demzufolge würde man die „Nein“-Antworten 2014 gegenüber 2013 überschätzen, weil 362 Betriebe die Frage D41 beantwortet haben, die im Jahr zuvor gar nicht bis zu dieser Frage gelangt wären. In Tabelle 8 sind die Häufigkeiten der Variable *c48* sowie *d41* vor und nach der Anpassung der Filterführung dargestellt. Es ist deutlich zu erkennen, dass man ohne die Anpassung den Eindruck gewinnt, der Anteil an Betrieben, die Weiterbildung aufgrund einer gesetzlichen Vorschrift anbieten, würde abnehmen. Der Stata- bzw. SPSS-Befehl zum Umcodieren der Variable lautet:

```
- replace d41=-1 if d37a==2 & d39==2
- if (d37a eq 2 and d39 eq 2) d41=-1.
      exe.
```

Tabelle 8

Häufigkeit der Durchführung von Weiterbildung aufgrund einer gesetzlichen Vorschrift¹⁰

<i>Durchführung der Weiterbildung aufgrund gesetzlicher Vorschrift</i>	2013		2014 vor Anpassung		2014 nach Anpassung	
	Absolut	Prozent	Absolut	Prozent	Absolut	Prozent
ja	1.254	70,81	2.200	64,52	2.200	72,18
nein	517	29,19	1.210	35,48	848	27,82
Gesamt	1.771	100,00	3.410	100,00	3.048	100,00

¹⁰ Fehlende Werte wurden in die Auswertung nicht einbezogen.

7 Praktischer Umgang mit den Datensätzen

Die Datensätze zum Qualifizierungspanel können, wie in Abschnitt 8.3 beschrieben, vor Ort als Gastwissenschaftler (der Datensatz trägt daher das Suffix „gwa“) oder via Datenfernverarbeitung genutzt werden. Um eine Vorstellung für den praktischen Umgang mit dem Datensatz zu bekommen, wird im Folgenden beispielhaft und auszugsweise die Stata-Syntax zur Erstellung grundlegender Auswertungen mit dem BIBB-Qualifizierungspanel beschrieben.

\$dtapfad ist das gewählte Arbeitsverzeichnis. „//“ leitet Kommentare ein.

7.1 Praktischer Umgang mit dem Querschnittsdatsatz

Zunächst muss der Datensatz geladen werden. Im Beispiel wird dazu zunächst der Arbeitsspeicher geleert (Befehl `clear all`) und dann der Datensatz der Welle 2014 geladen (Befehl `use`):

```
clear all // Löschen ggf. offener Datensätze aus dem Arbeitsspeicher
use "$dtapfad\Qualifizierungspanel_2014_gwa.dta" // Laden des Datensatzes
```

Möchte man den Querschnittsdatsatz für das BIBB-Qualifizierungspanel 2014 auswerten, benötigt man die Variable `quer2014`. Nur wenn diese den Wert 1 annimmt, gehört der Betrieb zum Querschnittsdatsatz. Der entsprechende Stata-Befehl lautet:

```
keep if quer2014==1
```

Werden deskriptive Analysen mit dem Querschnitt gerechnet, sollte darauf geachtet werden, dass das Hochrechnungsgewicht `hr2014q` verwendet wird (Befehl `weight`). Andernfalls beziehen sich die erstellten Ergebnisse auf die disproportional gezogenen Fälle und sind nicht repräsentativ.

Hier soll beispielsweise die hochgerechnete Fallzahl nach Betriebsgrößenklassen tabellarisch (Befehl `tabulate`) dargestellt werden. Für die hochgerechnete Fallzahl nach Betriebsgrößenklassen (Variable `grkl2014`) wird das Querschnitt-Hochrechnungsgewicht `hr2014q` verwendet. Es ergibt sich:

```
tab grkl2014 [iweight = hr2014q]
```

Es ergibt sich folgender Output:

Größenklasse (SVB Gesamt 31.12.2013)	Freq.	Percent	Cum.
1. 1 bis 19 SVB	1.835.260	88,67	88,67
2. 20 bis 99 SVB	188.648,524	9,11	97,78
3. 100 bis 199 SVB	26.195,41	1,27	99,05
4. 200 und mehr SVB	19.761,868	0,95	100,00
Total	2.069.865	100,00	

Würde man die Auszählung dagegen ohne Hochrechnung durchführen, käme man zu folgendem Ergebnis:

Größenklasse (SVB Gesamt 31.12.2013)	Freq.	Percent	Cum.
1. 1 bis 19 SVB	1.032	29,91	29,91
2. 20 bis 99 SVB	1.003	29,07	58,99
3. 100 bis 199 SVB	416	12,06	71,04
4. 200 und mehr SVB	999	28,96	100,00
Total	3.450	100,00	

Es ist deutlich zu erkennen, dass man aufgrund des Stichprobendesigns die Anzahl der Kleinbetriebe unter- und die der Großbetriebe überschätzen würde. Da sich gerade Groß- und Kleinbetriebe in vielen inhaltlichen Punkten, wie etwa beim Weiterbildungsverhalten, stark unterscheiden, würde man ohne Hochrechnung zu verzerrten Aussagen bezüglich der Betriebe in Deutschland kommen.

7.2 Praktischer Umgang mit dem Panel-Datensatz

Beispielhaft ist im Folgenden die Stata-Syntax vorgestellt, welche zeigt, wie Analysen mit dem BIBB-Qualifizierungspanel im Längsschnitt durchgeführt werden können.

Möchte man Panelauswertungen durchführen, benötigt man neben den Querschnittsfällen auch die Nichtquerschnittsbetriebe. Die genaue Zusammenstellung der Fälle richtet sich jeweils nach dem zu untersuchenden Zeitraum. Der Tabelle 9 kann entnommen werden, wie viele Betriebe an wie vielen Wellen teilgenommen haben. Insgesamt haben 529 Betriebe an allen bisher durchgeführten Betrieben des BIBB-Qualifizierungspanels teilgenommen.

Tabelle 9

Anzahl der Teilnahme der Betriebe an den Wellen des BIBB-Qualifizierungspanels

Anzahl Wellen	Anzahl absolut	Anzahl in Prozent
1	1.617 ¹¹	45,6
2	844	23,8
3	557	15,7
4	529	14,9
Gesamt	3.547	100,0

Zunächst müssen für eine Analyse im Längsschnitt die einzelnen Querschnittsdatsätze der Erhebungswellen des BIBB-Qualifizierungspanels verknüpft werden. Der Vollständigkeit halber wird die Verknüpfung aller bisherigen Erhebungswellen vorgestellt. Mit dem „merge“-Befehl werden dann die einzelnen zurückliegenden Wellen angespielt. „id2“ wird dabei als eindeutiger Indikator bei der Verknüpfung benutzt. „1:1“ gibt an, dass die Verknüpfung eineindeutig ist. Mit der Option „generate“ wird festgelegt, dass eine Variable erzeugt werden soll, in der das Ver-

¹¹ Die Betriebe, die 2014 erstmalig am BIBB-Qualifizierungspanel teilgenommen haben, wurden nicht beachtet.

knüpfungsergebnis auf Fallebene dokumentiert wird. Ohne diese Option würde standardmäßig eine Variable „_merge“ erzeugt, und es würde nach dem zweiten Verknüpfungsbefehl zu einer Fehlermeldung kommen, da die Variable „_merge“ schon vorhanden wäre. Außerdem kann mit diesen Variablen komfortabel der Verknüpfungserfolg der einzelnen Schritte verfolgt werden.

```
clear all
use "$pfad\Qualifizierungspanel_2011_gwa_v2.dta"
merge 1:1 id2 using "$pfad\Qualifizierungspanel_2012_gwa_v2.dta",
    gen(_merge11_12)
merge 1:1 id2 using "$pfad\Qualifizierungspanel_2013_gwa_v2.dta",
    gen(_merge11_13)
merge 1:1 id2 using "$pfad\Qualifizierungspanel_2014_gwa.dta",
    gen(_merge11_14)
```

Will man eine Längsschnittauswertung für den Längsschnitt 2013 bis 2014 gewichtet durchführen, ist der Längsschnitt-Hochrechnungsfaktor `hr13_14p` zu verwenden. Im Folgenden wird allerdings nicht gewichtet. Für eine Analyse des Längsschnitts 2013 bis 2014 wäre natürlich eine Verknüpfung der Datensätze nur dieser beiden Wellen ausreichend.

Eine denkbare Längsschnittanalyse wäre, die Entwicklung der Ausbildungsbeteiligung (nach BBiG und HwO) 2013 zu 2014 zu betrachten. Für eine Panelanalyse müssen dazu zunächst die Variablenamen harmonisiert werden. Nur so kann Stata im nächsten Schritt erkennen, dass es sich um dieselbe Variable zu unterschiedlichen Zeitpunkten handelt. Dazu wird die Variable `c05` (c=Welle 3, Frage 05) in `ausb2013` umbenannt, `d01` in `ausb2014`. Der Befehl „clonevar“ legt dabei ein Duplikat der Ursprungsvariablen an und nimmt alle Eigenschaften mit.

```
clonevar ausb2013 = c05
clonevar ausb2014 = d01
```

Gegebenenfalls notwendige Recodierungen zur Harmonisierung müssen vorgenommen werden. Erforderlich ist, dass in allen verwendeten Variablen die gleichen Wertebereiche gelten und die Werte der einzelnen Wellen jeweils die gleiche Bedeutung haben. Im Beispiel betrachtet werden soll die Ausbildungsaktivität in Abhängigkeit von der Betriebsgrößenklasse („ausb“ und „grkl“). Mit dem folgenden Recode-Befehl werden im Beispiel der Variable `ausb2013` und `aus2014` alle Ja-Antworten auf 1 gesetzt und alle übrigen Antworten auf „0“. In allen `grkl`-Variablen (* ist ein Platzhalter für 2011 etc.) wird der Wert 0 auf „0“ gesetzt.

```
recode ausb2013 ausb2014 (1=1) (else = 0)
recode grkl*(0=.)
```

Im Folgenden soll beispielhaft eine regressionsanalytische Panelanalyse durchgeführt werden.

Der Datensatz kann zunächst ins Longformat mit dem Befehl „reshape“ umgewandelt werden. Das hat den Vorteil, dass dabei automatisch ein Wellenindikator erzeugt wird, der im nächsten Schritt benötigt wird.

```
reshape long ausb grkl, i(id2) j(welle)
```

Dabei ist „i“ der Fall-Indikator und „j“ der zu generierende Wellen-Indikator. „ausb“ und „grkl“ sind die vorher vereinheitlichten Stämme der Variablennamen ohne den Wellenindex als Zahl (also „ausb“ statt „ausb2014“).

Mit dem „xtset“-Befehl wird dann definiert, dass es sich um einen Längsschnittdatensatz handelt und welche Variable als Indikator zur Identifizierung der Welle dient (hier die vorher im „reshape“-Befehl generierte Variable „welle“).

```
xtset welle
```

Es könnte beispielsweise die Entwicklung der Ausbildungsbeteiligung (ausb) in Abhängigkeit von der Entwicklung der Betriebsgrößenklasse (grkl) als Panel-Logit-Modell (xtlogit) untersucht werden. Via „i.“ werden die einzelnen Ausprägungen der Variable „grkl“ als Dummy-Variablen berücksichtigt. Für den Einfluss der Größenklasse auf die Ausbildungsbeteiligung ergibt sich (hier als Random-Effects-Modell, Option „re“):

```
xtlogit ausb i.grkl, re
```

8 Datenanonymisierung und Datenzugang

8.1 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, DS-GVO) 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 Forscher/-innen auf die 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.

8.2 Formale Datenanonymisierung

Vor der Erhebung wurde den freiwillig teilnehmenden Betrieben Datenschutz und Anonymität zugesichert (siehe Datenschutzblatt und Fragebogen). Die wenigen Strukturinformationen im Datensatz erlauben keine Identifikation der betreffenden Betriebe.

Im Datensatz werden keine eindeutigen Identifikatoren für eine Auswertung angeboten. Zur Unterscheidung der Fälle steht lediglich ein systemfreies Betriebskennzeichen (id2) zur Verfügung. Zudem sind die in Tabelle 10 aufgeführten Variablen, die sich aus dem Fragebogen ergeben, nicht im Datensatz enthalten.

Tabelle 10

Nicht verfügbare Variablen für das BIBB-Qualifizierungspanel 2014

Fragebogennummer	Variableninhalt	Anonymisierung
D00	Bereitschaft zur Teilnahme	gelöscht
D61	Bereitschaft Datenverknüpfung	gelöscht

8.3 Datenzugang

Die Daten des BIBB-Qualifizierungspanels 2014 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 Nutzer(n)/-innen erstellter Syntaxprogramme im BIBB-FDZ. Die Datennutzer/-innen 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 Nutzer/-innen zurückgeschickt.

Gastwissenschaftleraufenthalte erlauben die Analyse schwach anonymisierter Forschungsdaten an den abgeschotteten PC-Arbeitsstationen in einem eigenen Gästeraum des BIBB-FDZ. Über einen Gastaufenthalt lassen sich beispielsweise auch vergrößerte Merkmale in ihrem originalen Wertebereich sowie Volltexte analysieren.

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

9 Literaturverzeichnis

- ALDA, Holger; ROHRBACH-SCHMIDT, Daniela (2011): Imputation fehlender Werte für die Einkommensvariable in der BIBB/BAuA-Erwerbstätigenbefragung 2006. In: BIBB-FDZ Daten- und Methodenberichte, Nr.2/2011.
- HOHN, Christine; FRIEDRICH, Anett; GERHARDS, Christian (2016): Querschnitt- und Panelcharakteristika des BIBB-Qualifizierungspanels. Onlinedokument: <https://metadaten.bibb.de/metadatengruppe/16> (Stand: 31.10.2016).
- TNS INFRATEST SOZIALFORSCHUNG (2015): BIBB-Qualifizierungspanel. Qualifizierung und Kompetenzentwicklung. Methodenbericht zur vierten Welle.

Anhang

Anhang A1 Variablen mit fehlenden Angaben über zehn Prozent

Variable	Variableninhalt	N gesamt	N gültig	N fehlend	% fehlend
d50	Anteil Vorleistungen und Fremdkosten am Umsatz 2013 in Prozent	2.593	1.531	1.062	41,0
d48a	Höhe Geschäftsvolumen	2.853	1.834	1.019	35,7
d51	Summe aller Investitionen im letzten Geschäftsjahr in Euro	3.461	2.348	1.113	32,2
d52hoch	Durchschnittlicher/-s Bruttolohn/-gehalt: Beschäftigte mit hoch qualifizierten Tätigkeiten	3.461	2.689	772	22,3
d52qual	Durchschnittlicher/-s Bruttolohn/-gehalt: Beschäftigte mit qualifizierten Tätigkeiten	3.461	2.800	661	19,1
d11ohne	Anzahl Bewerber Ausbildungsstellen: Ohne Schulabschluss oder Abschluss unterhalb der Hauptschule	1.782	1.470	312	17,5
d40einf	Es gab im Jahr 2013 Beschäftigte mit einfachen Tätigkeiten	2.494	2.066	428	17,2
d52einf	Durchschnittlicher/-s Bruttolohn/-gehalt: Beschäftigte mit einfachen Tätigkeiten	3.461	2.885	576	16,6
d11haupt	Anzahl Bewerber Ausbildungsstellen: Hauptschulabschluss oder vergleichbar	1.782	1.488	294	16,5
d11abi	Anzahl Bewerber Ausbildungsstellen: Abitur, Fachhochschulreife oder vergleichbar	1.782	1.489	293	16,4
d11real	Anzahl Bewerber Ausbildungsstellen: Real- schulabschluss, mittlere Reife oder vergleichbar	1.782	1.506	276	15,5
d30schul	Anzahl 31.12.13: Beschäftigte mit schulischer Berufsausbildung	3.461	2.942	519	15,0
d30ber	Anzahl 31.12.13: Beschäftigte mit betrieblicher Berufsausbildung	3.461	2.943	518	15,0
d30tec	Anzahl 31.12.13: Beschäftigte mit Meister- oder Technikerabschluss o. Ä.	3.461	2.968	493	14,2
d52hochk	Nur CAPI: Bruttolohn/-gehalt: Beschäftigte mit hoch qualifizierten Tätigkeiten	3.461	2.975	486	14,0
d30ohne	Anzahl 31.12.13: Beschäftigte ohne Berufsausbildung, Fachhochschul- oder Universitätsabschluss	3.461	2.981	480	13,9
d31bac	Anzahl 31.12.13: Beschäftigte mit Bachelor als höchstem beruflichem Abschluss	3.461	2.989	472	13,6
d30uni	Anzahl 31.12.13: Beschäftigte mit Fachhochschul- oder Universitätsabschluss	3.461	2.993	468	13,5

(Fortsetzung Anhang 1)

Variable	Variableninhalt	N gesamt	N gültig	N fehlend	% fehlend
d49	Anteil Auslandsumsatz im letzten Geschäftsjahr in Prozent	2.593	2.295	298	11,5
d52qualk	Nur CAPI: Bruttolohn/-gehalt: Beschäftigte mit qualifizierten Tätigkeiten	3.461	3.066	395	11,4
d38hoch	Anzahl Weiterbildungsteilnehmer unter Beschäftigten für hoch qualifizierte Tätigkeiten	2.851	2.532	319	11,2
d38qual	Anzahl Weiterbildungsteilnehmer unter Beschäftigten für qualifizierte Tätigkeiten	2.851	2.534	317	11,1
d52einfk	Nur CAPI: Bruttolohn/-gehalt: Beschäftigte mit einfachen Tätigkeiten	3.461	3.085	376	10,9
d38einf	Anzahl Weiterbildungsteilnehmer unter Beschäftigten für einfache Tätigkeiten	2.851	2.547	304	10,7
d44wech	Anzahl Beschäftigter, die 2013 ihren Aufgaben-/Tätigkeitsbereich wechselten	849	759	90	10,6
d40hoch	Es gab im Jahr 2013 Beschäftigte mit hoch qualifizierten Tätigkeiten	2.494	2.238	256	10,3
d36qual	Anzahl ausgeschiedener Beschäftigter für qualifizierte Tätigkeiten	2.457	2.206	251	10,2
d36hoch	Anzahl ausgeschiedener Beschäftigter für hoch qualifizierte Tätigkeiten	2.457	2.210	247	10,1
d59	Jahr der Betriebsgründung	2.114	1.903	211	10,0

Anhang A2 Übersicht über Ursprungsvariablen und imputierte Variablen

Imputierte Variable		N inkl. imp.	N gültig	N Imp.	Imp. %
d02ausf	Anzahl der weiblichen Auszubildenden am 31.12.13	2.058	2.054	4	0,2
d04b1	Anzahl Azubis in Ausbildungsberuf 1 zum 31.12.13	2.085	2.078	7	0,3
d04b2	Anzahl Azubis in Ausbildungsberuf 2 zum 31.12.13	1.087	1.080	7	0,6
d04b3	Anzahl Azubis in Ausbildungsberuf 3 zum 31.12.13	672	665	7	1,0
d04b4	Anzahl Azubis in Ausbildungsberuf 4 zum 31.12.13	418	409	9	2,2
d04b5	Anzahl Azubis in Ausbildungsberuf 5 zum 31.12.13	276	268	8	2,9
d05gt	Anzahl gew.-techn. Azubis duales Studium am 31.12.13	382	360	22	5,8

(Fortsetzung Anhang A2)

Imputierte Variable		N inkl. imp.	N gültig	N Imp.	Imp. %
d05kv	Anzahl kauf.-verw. Azubis duales Studium am 31.12.13	382	370	12	3,1
d06beh	Anzahl Azubis mit amtlich anerkannter Behinderung am 31.12.13	202	200	2	1,0
d09neu	Anzahl neu eingestellter Auszubildender	1.537	1.523	14	0,9
d10ohne	Anzahl Azubis: Ohne Schulabschluss oder Abschluss unterhalb Hauptschule	1.537	1.499	38	2,5
d10haupt	Anzahl Azubis: Hauptschulabschluss oder vergleichbarer Abschluss	1.537	1.459	78	5,1
d10real	Anzahl Azubis: Realschulabschluss, mittlere Reife oder vergleichbarer Abschluss	1.537	1.493	44	2,9
d10abi	Anzahl Azubis: Abitur, Fachhochschulreife oder vergleichbarer Abschluss	1.537	1.494	43	2,8
d11ohne	Anzahl Bewerber Ausbildungsstellen: Ohne Schulabschluss oder Abschluss unterhalb der Hauptschule	1.782	1.470	312	17,5
d11haupt	Anzahl Bewerber Ausbildungsstellen: Hauptschulabschluss oder vergleichbarer Abschluss	1.782	1.488	294	16,5
d11real	Anzahl Bewerber Ausbildungsstellen: Realschulabschluss, mittlere Reife oder vergleichbarer Abschluss	1.782	1.506	276	15,5
d11abi	Anzahl Bewerber Ausbildungsstellen: Abitur, Fachhochschulreife oder vergleichbarer Abschluss	1.782	1.489	293	16,4
d13unb	Anzahl unbesetzter Ausbildungsstellen nach BBiG/HwO	519	487	32	6,2
d14unb	Anzahl unbesetzter Ausbildungsstellen nach BBiG/HwO wegen gelöster Ausbildungsverträge	519	500	19	3,7
d14vor	Anzahl vorzeitig gelöster Ausbildungsverträge Kalenderjahr 2013	427	419	8	1,9
d16gt	Anzahl gew.-techn. Azubis zur Abschlussprüfung 2013	1.463	1.404	59	4,0
d16kv	Anzahl kauf.-verw. Azubis zur Abschlussprüfung 2013	1.463	1.431	32	2,2
d17gta	Anzahl gew.-techn. Auszubildender, die unbefristet übernommen wurden	1.463	1.387	76	5,2
d17gtb	Anzahl gew.-techn. Auszubildender, die befristet übernommen wurden	1.463	1.387	76	5,2
d17gtc	Anzahl kauf.-verw. Auszubildender, die den Betrieb verlassen haben	1.463	1.389	74	5,1
d17gtd	Anzahl gew.-techn. Auszubildender, die betriebsbedingt oder wg. Prüfungsnote nicht übernommen	1.463	1.391	72	4,9
d17gte	Anzahl gew.-techn. Auszubildender, die die Abschlussprüfung nicht bestanden haben	1.463	1.394	69	4,7

(Fortsetzung Anhang A2)

Imputierte Variable		N inkl. imp.	N gültig	N Imp.	Imp. %
d17kva	Anzahl kauf.-verw. Auszubildender, die unbefristet übernommen wurden	1.463	1.417	46	3,1
d17kvb	Anzahl kauf.-verw. Auszubildender, die befristet übernommen wurden	1.463	1.417	46	3,1
d17kvc	Anzahl kauf.-verw. Auszubildender, die den Betrieb verlassen haben	1.463	1.417	46	3,1
d17kvd	Anzahl kauf.-verw. Auszubildender, die betriebsbedingt oder wg. Prüfungsnote nicht übernommen wurden	1.463	1.417	46	3,1
d17kve	Anzahl kauf.-verw. Auszubildender, die die Abschlussprüfung nicht bestanden haben	1.463	1.419	44	3,0
d26ger13	Anzahl 31.12.13: Geringfügig Beschäftigte	3.450	3.419	31	0,9
d26nsv13	Anzahl 31.12.13: Nicht sozialversicherungspflichtig Beschäftigte	3.450	3.437	13	0,4
d26zwi13	Summe 31.12.13: Alle Beschäftigten ohne Auszubildende	3.450	3.413	37	1,1
d26svb12	Anzahl 31.12.12: Sozialversicherungspflichtig Beschäftigte	3.450	3.426	24	0,7
d26ger12	Anzahl 31.12.12: Geringfügig Beschäftigte	3.450	3.381	69	2,0
d26nsv12	Anzahl 31.12.12: Nicht sozialversicherungspflichtig Beschäftigte	3.450	3.407	43	1,2
d26zwi12	Summe 31.12.12: Alle Beschäftigten ohne Auszubildende	3.450	3.381	69	2,0
d26aus12	Anzahl 31.12.12: Auszubildende nach BBiG/HwO	3.450	3.428	22	0,6
d26soz12	Anzahl 31.12.12: Auszubildende nach Ausbildungsregelung des Gesundheitswesen oder sozialpflegerischer und -pädagogischer Berufe	3.450	3.428	22	0,6
d27leih	Anzahl 31.12.13: Leiharbeitskräfte	676	642	34	5,0
d28fr	Anzahl 31.12.13: Frauen	3.450	3.338	112	3,2
d28tz	Anzahl 31.12.13: Teilzeitbeschäftigte	3.450	3.317	133	3,9
d28bef	Anzahl 31.12.13: Befristet Beschäftigte	3.450	3.329	121	3,5
d29einf	Anzahl Beschäftigte 31.12.13: Einfache Tätigkeiten	3.450	3.283	167	4,8
d29qual	Anzahl Beschäftigte 31.12.13: Qualifizierte Tätigkeiten	3.450	3.278	172	5,0
d29hoch	Anzahl Beschäftigte 31.12.13: Hoch qualifizierte Tätigkeiten	3.450	3.283	167	4,8
d30ohne	Anzahl 31.12.13: Beschäftigte ohne Berufsausbildung, Fachhochschul- oder Universitätsabschluss	3.450	2.970	480	13,9
d30ber	Anzahl 31.12.13: Beschäftigte mit betrieblicher Berufsausbildung	3.450	2.932	518	15,0
d30schul	Anzahl 31.12.13: Beschäftigte mit schulischer Berufsausbildung	3.450	2.931	519	15,0

(Fortsetzung Anhang A2)

Imputierte Variable		N inkl. imp.	N gültig	N Imp.	Imp. %
d30tec	Anzahl 31.12.13: Beschäftigte mit Meister- oder Technikerabschluss o. Ä.	3.450	2.957	493	14,3
d30uni	Anzahl 31.12.13: Beschäftigte mit Fachhochschul- oder Universitätsabschluss	3.450	2.982	468	13,6
d31bac	Anzahl 31.12.2013: Beschäftigte mit Bachelor als höchstem beruflichem Abschluss	3.450	2.978	472	13,7
d33neu	Anzahl neu eingestellter Beschäftigter 2013	2.554	2.384	170	6,7
d34einf	Anzahl der neu eingestellten Beschäftigten mit einfachen Tätigkeiten	2.554	2.346	208	8,1
d34qual	Anzahl der neu eingestellten Beschäftigten mit qualifizierten Tätigkeiten	2.554	2.340	214	8,4
d34hoch	Anzahl der neu eingestellten Beschäftigten mit hoch qualifizierten Tätigkeiten	2.554	2.342	212	8,3
d35unb	Anzahl unbesetzter Stellen 2013	619	600	19	3,1
d35einf	Anzahl unbesetzter Stellen für einfache Tätigkeiten	619	598	21	3,4
d35qual	Anzahl unbesetzter Stellen für qualifizierte Tätigkeiten	619	596	23	3,7
d35hoch	Anzahl unbesetzter Stellen für hoch qualifizierte Tätigkeiten	619	596	23	3,7
d36aus	Anzahl ausgeschiedener Beschäftigter 2013	2.453	2.257	196	8,0
d36einf	Anzahl ausgeschiedener Beschäftigter für einfache Tätigkeiten	2.453	2.211	242	9,9
d36qual	Anzahl ausgeschiedener Beschäftigter für qualifizierte Tätigkeiten	2.453	2.204	249	10,2
d36hoch	Anzahl ausgeschiedener Beschäftigter für hoch qualifizierte Tätigkeiten	2.453	2.208	245	10,0
d37wei	Anzahl Teilnehmer an Weiterbildungsmaßnahmen 2013	2.848	2.677	171	6,0
d38einf	Anzahl Weiterbildungsteilnehmer unter Beschäftigten für einfache Tätigkeiten	2.848	2.544	304	10,7
d38qual	Anzahl Weiterbildungsteilnehmer unter Beschäftigten für qualifizierte Tätigkeiten	2.848	2.531	317	11,1
d38hoch	Anzahl Weiterbildungsteilnehmer unter Beschäftigten für hoch qualifizierte Tätigkeiten	2.848	2.529	319	11,2
d42fort	Anzahl Beschäftigter mit Teilnahme an Aufstiegsfortbildungen 2013	1.311	1.242	69	5,3
d43befo	Anzahl beförderter Beschäftigter 2013	1.278	1.170	108	8,5
d44wech	Anzahl Beschäftigter, die 2013 ihren Aufgaben-/ Tätigkeitsbereich wechselten	849	759	90	10,6
d48vol	Höhe Geschäftsvolumen	2.843	1.825	1.018	35,8
d49prz	Anteil Auslandsumsatz im letzten Geschäftsjahr in Prozent	2.583	2.285	298	11,5

(Fortsetzung Anhang A2)

Imputierte Variable		N inkl. imp.	N gültig	N Imp.	Imp. %
d50prz	Anteil Vorleistungen und Fremdkosten am Umsatz 2013 in Prozent	2.583	1.524	1.059	41,0
d51inv	Summe aller Investitionen im letzten Geschäftsjahr in Euro	3.450	2.338	1.112	32,2
d52einf	Durchschnittlicher/-s Bruttolohn/-gehalt: Beschäftigte mit einfachen Tätigkeiten	2.057	1.483	574	27,9
d52qual	Durchschnittlicher/-s Bruttolohn/-gehalt: Beschäftigte mit qualifizierten Tätigkeiten	3.170	2.512	658	20,8
d52hoch	Durchschnittlicher/-s Bruttolohn/-gehalt: Beschäftigte mit hoch qualifizierten Tätigkeiten	2.681	1.914	767	28,6



Bundesinstitut für Berufsbildung
Robert-Schuman-Platz 3
53175 Bonn

Telefon: (0228) 107-0

Internet: www.bibb.de
E-Mail: zentrale@bibb.de

Bundesinstitut
für Berufsbildung **BiBB** ▶

- ▶ Forschen
- ▶ Beraten
- ▶ Zukunft gestalten