

Paula Protsch | Anett Friedrich

WZB Vignettenmodul im BIBB- Qualifizierungspanel 2014

Version 1.0



Heft 3/2021

Paula Protsch | Anett Friedrich

WZB Vignettenmodul im BIBB-Qualifizierungspanel 2014

Version 1.0

Danksagung

Die beteiligten Wissenschaftlerinnen am Wissenschaftszentrum Berlin für Sozialforschung (WZB), Dorothea Kübler, Paula Protsch, Julia Schmid und Heike Solga, danken Christian Gerhards, Sabine Mohr und Klaus Troltsch vom Team des BIBB-Betriebspanels zu Qualifizierung und Kompetenzentwicklung und dem BIBB Forschungsdatenzentrum für die gute Kooperation. Weiterer Dank gilt Sebastian Bechmann und Matthias Passmann von TNS Infratest Sozialforschung sowie Katrin Auspurg für die D-effiziente Stichprobenziehung und methodische Beratung. Des Weiteren wird Sajoscha Engelhardt, Manuela Ludwig und Michaela Zwiebel für ihre Unterstützung am WZB gedankt.

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:

Protsch, Paula; Friedrich, Anett: WZB Vignettenmodul im BIBB-Qualifizierungspanel 2014. BIBB-FDZ Daten- und Methodenbericht 3/2021. Bonn 2021

1. Auflage 2021

Herausgeber:

Bundesinstitut für Berufsbildung
Robert-Schuman-Platz 3
53175 Bonn
Internet: www.bibb.de

Publikationsmanagement:

Stabsstelle „Publikationen und wissenschaftliche Informationsdienste“
E-Mail: publikationsmanagement@bibb.de
www.bibb.de/veroeffentlichungen

Gesamtherstellung:

Verlag Barbara Budrich
Stauffenbergstraße 7
51379 Leverkusen
Internet: www.budrich.de
E-Mail: info@budrich.de

Kontakt FDZ:

E-Mail: 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.

ISSN: 2190-300X

urn:nbn:de:0035-0945-5

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

1	Einleitung	5
1.1	Gegenstand des Daten- und Methodenberichts	5
1.2	Aufbau des Daten- und Methodenberichts	5
1.3	Kurzbeschreibung des Datensatzes	5
1.4	Zentrale Eigenschaften im Überblick	6
2	Grundlegende Aspekte	7
3	Spezifika des Erhebungsdesigns	8
4	Bewerbermerkmale und deren Ausprägungen in den Einleitungstexten und Vignetten	9
5	Vorbereitungen zur Analyse der Faktoriellen Surveys	12
6	Organisation der Mikrodaten	13
6.1	Fehlende Werte	13
6.2	Verknüpfung mit dem Hauptinterview	13
7	Anonymisierung und Datenzugang	14
7.1	Bestimmungen des Datenschutzes	14
7.2	Formale Datenanonymisierung	14
7.3	Datenzugang	15
	Literaturverzeichnis	16
	Anhang	17

1 Einleitung

1.1 Gegenstand des Daten- und Methodenberichts

Der vorliegende Daten- und Methodenbericht beschreibt das WZB Vignettenmodul im BIBB-Qualifizierungspanel 2014. Die Planung und Umsetzung des Vignettenmoduls erfolgte in enger Zusammenarbeit des Wissenschaftszentrum Berlin (WZB) mit dem BIBB-Qualifizierungspanel. Der Bericht soll externen Wissenschaftlerinnen und Wissenschaftlern den Umgang mit den Daten erleichtern. Neben der Dokumentation der Datensatzstruktur enthält der Bericht Stata-Befehle, die die Arbeit mit den Daten erleichtern.

1.2 Aufbau des Daten- und Methodenberichts

In Abschnitt 1.3 erfolgt eine steckbriefartige Beschreibung des Datensatzes und Abschnitt 1.4 gibt einen Überblick über dessen zentrale Eigenschaften. In Kapitel 2 werden grundlegende Aspekte der beiden Vignettenstudien des WZB Vignettenmoduls beschrieben. Kapitel 3 beschreibt Spezifika des Erhebungsdesigns. In Kapitel 4 werden die Merkmale, die in den Vignetten variiert bzw. konstant gehalten wurden, dargestellt. In Kapitel 5 folgt eine kurze Erläuterung zur Vorbereitung der Datenanalyse. Kapitel 6 umfasst Ausführungen zur Organisation der Daten, also zu fehlenden Werten und der Verknüpfungsmöglichkeit mit den Daten des Hauptinterviews des BIBB-Qualifizierungspanels 2014. Abschließend informiert Kapitel 7 über die Anonymisierung und den Datenzugang. In Anhang A1 wird ein Stata do-file zur Verfügung gestellt, mit dem die Setup-Daten der beiden Vignettenstudien an die Befragungsdaten herangespielt werden können.

1.3 Kurzbeschreibung des Datensatzes

Das WZB Vignettenmodul im BIBB-Qualifizierungspanel 2014 wurde zusätzlich zum regulären Fragebogen des BIBB-Qualifizierungspanel 2014 erhoben. Untersucht wurden Auswahlkriterien nach denen Betriebe ihre zukünftigen Auszubildenden rekrutieren. Die Erhebung deckt zum einen allgemeine Personenmerkmale von typischen Bewerberinnen und Bewerbern ab und fokussiert zum anderen auf Bewerberinnen und Bewerber mit Migrationshintergrund aus Spanien, die nach Deutschland kommen möchten, um eine Ausbildung zu beginnen oder in Deutschland geboren und aufgewachsen sind. Den Betrieben wurden jeweils fünf kurze Beschreibungen von fiktiven Bewerberinnen oder Bewerbern präsentiert, die sich in den Untersuchungsmerkmalen unterscheiden.

1.4 Zentrale Eigenschaften im Überblick

Befragungstitel	WZB Vignettenmodul im BIBB-Qualifizierungspanel 2014	
DOI	10.7803/371.14.2.2.10	
Kurzbeschreibung	Vignettenstudien zur betrieblichen Rekrutierung und Auswahl von Auszubildenden	
Erhebungsjahr	2014	
Erhebungseinheit	Betriebe	
Themenschwerpunkt	Ausbildung	
Datenzugangsmöglichkeiten	Gastwissenschaftleraufenthalt/Datenfernverarbeitung	
Variablenanzahl	GWA: 108	VT: 2
Grundgesamtheit	Alle Ausbildungsbetriebe und Betriebe in Deutschland mit mindestens einem/ einer sozialversicherungspflichtig Beschäftigten, die zum 31.12.2013 Auszubildende hatten oder versucht haben für das Ausbildungsjahr 2013/2014 Ausbildungsstellen zu besetzen.	
Gewichtung/Hochrechnung	Hochrechnungsfaktoren (auf Basis der Betriebsdatei der BA) Längsschnittgewichte	
repräsentative Region	Ost- und Westdeutschland	
Fallzahl	1.468	
Erhebungsverfahren	CASI (Computer Assisted Self-Interview)	
Auswahlverfahren	Disproportional geschichtete Zufallsauswahl	
Erhebungsdesign	Querschnittsdesign	
Bemerkung	Die Daten können mit dem BIBB-Qualifizierungspanel 2014 (10.7803/371.14.1.2.10) sowie mit dem jeweils aktuellen Längsschnittdatensatz verknüpft werden (vgl. BIBB-Qualifizierungspanel long).	
Links	BIBB-FDZ Metadatenportal: https://metadaten.bibb.de/metadatengruppe/16 Projektseite: http://www.bibb.de/qp	
Stichworte	Ausbildung, Auswahlprozesse, betriebliche Rekrutierungs- und Auswahlprozesse, Bewerberinnen und Bewerber, Faktorieller Survey, spanischer Migrationshintergrund, Qualifizierung, Vignettenstudie, Zugang zu Ausbildungsplätzen	

2 Grundlegende Aspekte

In die Erhebungswelle 2014 des BIBB-Betriebspanels zu Qualifizierung und Kompetenzentwicklung wurde zusätzlich zum regulären Fragebogen ein Vignettenmodul implementiert. Das Modul enthält zwei Faktorielle Surveys, auch Vignettenstudien genannt, sowie zusätzliche Fragen zur betrieblichen Rekrutierung und Auswahl von Auszubildenden. Teilnehmenden Betrieben wurde zufällig entweder der WZB Faktorielle Survey A oder der WZB Faktorielle Survey B zugewiesen, die jeweils andere Schwerpunkte setzen (siehe GERHARDS/FRIEDRICH 2016).

Die Planung und Umsetzung dieses Moduls erfolgte in enger Zusammenarbeit zwischen dem Team des BIBB-Betriebspanels zu Qualifizierung und Kompetenzentwicklung und dem Projekt „Rekrutierungsverhalten von Unternehmen auf Ausbildungs- und Arbeitsmärkten“ der Abteilungen *Ausbildung und Arbeitsmarkt* und *Verhalten auf Märkten* am Wissenschaftszentrum Berlin (WZB), welches sowohl für die inhaltliche Ausrichtung als auch die methodische Konzeptionierung der Vignetten verantwortlich war.

Zur Untersuchung von Auswahlkriterien für die betriebliche Berufsausbildung wurden den Betrieben jeweils fünf kurze Beschreibungen von fiktiven Bewerberinnen oder Bewerbern präsentiert, die sich in einigen Merkmalen unterscheiden. Der Faktorielle Survey A deckt allgemeine Personenmerkmale von typischen Bewerberinnen und Bewerbern ab, die aus Deutschland kommen. Der Faktorielle Survey B fokussiert auf Bewerberinnen und Bewerber mit Migrationshintergrund aus dem EU-Ausland. Unterschieden wird zwischen jungen Menschen aus Spanien, die nach Deutschland kommen möchten, um eine Ausbildung zu beginnen, und Bewerberinnen und Bewerbern, deren Familien aus Spanien stammen, die jedoch in Deutschland geboren und aufgewachsen sind. Bei beiden Faktoriellen Surveys wurden die Befragten gebeten, sich vorzustellen, dass es sich um Bewerbungen für Ausbildungsplätze in dem Ausbildungsberuf handelt, in dem ihr Betrieb die meisten Auszubildenden hat, oder – wenn sie gegenwärtig keine Auszubildenden haben – für den Beruf, für den der Betrieb am meisten Ausbildungsplätze angeboten hatte. Die Befragten konnten auf einer Skala von 1 bis 10 angeben, wie wahrscheinlich es ist, dass die Bewerberin oder der Bewerber eine Einladung zum nächsten Auswahlschritt für den entsprechenden Beruf erhalten würde. Mit der Methode des Faktoriellen Surveys können die Bewerbermerkmale in den kurzen Beschreibungen (Vignetten) experimentell variiert werden, sodass deren jeweiliger Einfluss auf die Einschätzung der Befragten für sich identifiziert werden kann, selbst wenn manche dieser Merkmale in der Realität stark miteinander korrelieren.

Das Vignettenmodul wurde sowohl ausbildenden Betrieben als auch nicht ausbildenden Betrieben vorgelegt, die versucht haben, Ausbildungsstellen zu besetzen. Die einzelnen Vignetten wurden durch die Interviewte oder den Interviewten eigenständig am Computer (Computer Assisted Self-Interview – CASI) bewertet. Daher kamen nur die 2.247 computergestützt interviewten Betriebe für die Teilnahme infrage; von diesen wiederum erfüllten 1.468 Betriebe die oben genannten Auswahlbedingungen. Für diese Teilstichprobe wurde ein gesonderter Hochrechnungsfaktor erstellt (siehe GERHARDS/FRIEDRICH 2016; TNS INFRATEST SOZIALFORSCHUNG 2015).

3 Spezifika des Erhebungsdesigns

Nach Abschluss der regulären Befragung wurde den Betrieben der soeben beschriebenen Teilstichprobe eine Reihe von Fragen zu den Themen Rekrutierung und Auswahl von Auszubildenden gestellt. Der entsprechende Fragebogen wurde im Methodenbericht von TNS Infratest Sozialforschung veröffentlicht (TNS INFRATEST SOZIALFORSCHUNG 2015) und ist beim BIBB-FDZ abrufbar. Im Anschluss wurden die Befragten gebeten, sich einen Einleitungstext durchzulesen und dann fünf Bewerberprofile – die Vignetten – zu bewerten. Per Zufallsmodus wurde den Betrieben einer der beiden Faktoriellen Surveys zugewiesen. Innerhalb der beiden Faktoriellen Surveys wurde eine Reihe von Bewerbermerkmalen (Dimensionen) berücksichtigt, die verschiedene Ausprägungen aufweisen (Levels). Die Dimensionen und ihre Levels werden in den Tabellen 1 und 2 im folgenden Abschnitt dargestellt. Aus den Vignettenuniversen, also der Gesamtheit der möglichen Kombinationen dieser Ausprägungen, wurden sogenannte D-effiziente Vignettenstichproben à 250 Vignetten – verteilt auf 50 Sets pro Faktoriellen Survey – gezogen. Welches Set die Betriebe jeweils erhielten, wurde ebenso zufällig ausgewählt. Durch dieses Vorgehen kann darauf verzichtet werden, die gesamten Vignettenuniversen bewerten zu lassen, und dennoch können die Effekte der Vignettendimensionen orthogonal, d. h., unkorreliert, geschätzt und eine ausgeglichene und somit effiziente Verteilung der Ausprägungen berücksichtigt werden. Für beide Faktoriellen Surveys wurden die Haupteffekte der Dimensionen und Zweifachinteraktionen (Resolution V Design) berücksichtigt und gute Werte für die D-Effizienz erzielt. Für WZB Faktorieller Survey A konnte eine D-Effizienz von 94,51, für WZB Faktorieller Survey B eine D-Effizienz von 82,23 erzielt werden. Der höchstmögliche Wert für die D-Effizienz ist 100. Die D-Effizienz für B fällt etwas geringer aus, da aus inhaltlichen Gründen einige unlogische Kombinationen von Merkmalsausprägungen ausgeschlossen werden mussten (siehe Tabelle 2). Außerdem wurde hier darauf geachtet, dass in jedem Vignettenset sowohl Personen vorkommen, die nach Deutschland kommen wollen, als auch Personen, die in Deutschland geboren und aufgewachsen sind. Um ein möglicherweise sozial erwünschtes Antwortverhalten im Zusammenhang mit dem Geschlecht der Bewerberinnen und Bewerber zu vermeiden, wurde die Vignettendimension Geschlecht in beiden Faktoriellen Surveys nur zwischen den Betrieben variiert. Das heißt, dass den Betrieben jeweils entweder ein Set von fünf Vignetten mit jungen Frauen oder ein Set von fünf Vignetten mit jungen Männern zur Bewertung präsentiert wurde. Aus diesem Grund sind auch alle Dreifachinteraktionen mit der Variable Geschlecht unkorreliert.

Durch die zufällige Zuweisung der Sets zu den Befragten wird erreicht, dass alle Vignettensets mehrfach bewertet werden. Darüber hinaus wurde die Reihenfolge, in der die Vignetten gezeigt wurden, ebenfalls zufällig gewählt. Die Befragten wurden außerdem darauf hingewiesen, dass sie zwischen den Vignetten hin- und herschalten und auch ihre Bewertungen bei Bedarf noch einmal ändern können. Zusammen mit der fein differenzierten Bewertungsskala von 1 (sehr unwahrscheinlich) bis 10 (sehr wahrscheinlich), dass der oder die Bewerber/-in zum nächsten Auswahlschritt eingeladen werden würde, konnte auf diese Weise erreicht werden, dass Befragte die gesamte Bewertungsskala nutzen, um ihre Antworten zu differenzieren (für eine ausführliche Beschreibung und Diskussion der Methodik siehe AUSPURG/HINZ 2015).

4 Bewerbermerkmale und deren Ausprägungen in den Einleitungstexten und Vignetten

Die in den Vignettenstudien präsentierten fiktiven Bewerber/-innen unterscheiden sich in einigen Merkmalen, haben aber auch einige Gemeinsamkeiten. Die konstant gehaltenen Merkmale werden in den einleitenden Texten zu den Vignetten genannt und teilweise in den Vignetten mit abgebildet. Im WZB Faktoriellen Survey A haben beispielsweise alle Personen einen mittleren Schulabschluss oder werden diesen erreichen, und im WZB Faktoriellen Survey B sind beispielsweise alle Personen ledig. Die Vignettendimensionen und ihre Ausprägungen sind in den Tabellen 1 und 2 dargestellt, der Einleitungstext folgt im nächsten Absatz.

Einleitende Texte zu den Vignetten

Wir präsentieren Ihnen nun kurze Beschreibungen von fünf fiktiven Personen. Stellen Sie sich vor, dass sich diese Personen bei Ihnen schriftlich oder online um einen Ausbildungsplatz für den Beruf [Einblendung genannter oder gesetzter Ausbildungsberuf aus Frage EV1 oder Einblendung Ausbildungsberuf aus Frage EV3] bewerben. Schätzen Sie anhand der Angaben aus dem jeweiligen Lebenslauf ein, wie wahrscheinlich es ist, dass die jeweiligen Bewerber nach Durchsicht der Bewerbungsunterlagen eine Einladung zum nächsten Auswahlschritt erhalten. Auf einer Skala von 1 bis 10 (sehr unwahrscheinlich bis sehr wahrscheinlich) können Sie Ihre Antworten abstimmen.

Spezifischer Teil für A:

Die Bewerber unterscheiden sich in einigen Merkmalen. Gemeinsam ist allen Bewerbern, dass sie aus Deutschland kommen und ein ansprechendes Bewerbungsschreiben formuliert haben.

Spezifischer Teil für B:

Die Bewerber unterscheiden sich in einigen Merkmalen. Alle Bewerbungsschreiben sind ansprechend gestaltet. Ebenso haben alle Bewerber Zeugnisse im Zweier- bis Dreierbereich. Ihr Arbeits- und Sozialverhalten wurde gut bewertet. Alle Bewerber sind volljährig und haben spanische Wurzeln.

Tabelle 1: WZB Faktorieller Survey A

Variable	Vignetten Dimension	Levels und Vignettentexte
x0	Geschlecht (nur zwischen den Betrieben variiert)	1 Geschlecht: männlich 2 Geschlecht: weiblich
x1	Geburtsjahr (Alter)/ Jahr Mittlerer Schulabschluss	1 Schulbildung: Mittlerer Schulabschluss, erwartet 2014 Geburtsdatum: 24.11.1997 2 Schulbildung: Mittlerer Schulabschluss 2012 Geburtsdatum: 03.11.1995 3 Schulbildung: Mittlerer Schulabschluss 2010 Geburtsdatum: 18.10.1993
x2	Tätigkeit seit Schulabschluss	1 Tätigkeit nach Schulabschluss: einjährige berufsvorbereitende Maßnahme, dann Ausbildung (abgebrochen); seit Dez. 2013 Aushilfsjob (nur wenn x1 ungleich 1) 2 Tätigkeit nach Schulabschluss: bis Dez. 2013 keine Angabe, dann Aushilfsjob (nur wenn x1 ungleich 1) 3 Tätigkeit nach Schulabschluss: trifft nicht zu, da Schüler/Schülerin (nur wenn x1 gleich 1)
x3	Notendurchschnitt	1 Aktueller Notendurchschnitt: 2,8 (Variante, wenn x1 gleich 1) 1 Notendurchschnitt Abschlusszeugnis: 2,8 (Variante, wenn x1 ungleich 1) 2 Aktueller Notendurchschnitt: 3,4 (Variante, wenn x1 gleich 1) 2 Notendurchschnitt Abschlusszeugnis: 3,4 (Variante, wenn x1 ungleich 1)
x4	Beruf des Vaters*	1 Beruf des Vaters: Lagerist 2 Beruf des Vaters: Versicherungskaufmann 3 Beruf des Vaters: Lehrer 4 Beruf des Vaters: Mitarbeiter im Betrieb
x5	Beruf der Mutter	1 Beruf der Mutter: Erzieherin 2 Beruf der Mutter: Altenpflegerin
x6	Sozialverhalten	1 Sozialverhalten laut Zeugnisanlage: sehr gut 2 Sozialverhalten laut Zeugnisanlage: überwiegend gut
x7	Unentschuldigte Fehltage	1 Unentschuldigte Fehltage laut Zeugnis: keine 2 Unentschuldigte Fehltage laut Zeugnis: drei

* Die Levels „Lagerist“, „Versicherungskaufmann“ und „Lehrer“ sind overgesampled, d. h., sie kommen im Vergleich zum Level „Mitarbeiter im Betrieb“ häufiger vor. Das Oversampling sollte möglichen Irritationen seitens der Befragten über zu viele Bewerbungen von Mitarbeiterkindern vorbeugen.

Tabelle 2: WZB Faktorieller Survey B

Variable	Vignetten Dimension	Levels und Vignettentexte
x0	Geschlecht (nur zwischen den Betrieben variiert)	1 Geschlecht: männlich 2 Geschlecht: weiblich
x1	Migrationsstatus	1 Herkunft: Spanien, Bewerber/Bewerberin möchte nach Deutschland kommen 2 Herkunft: Bewerber/Bewerberin in Deutschland geboren und aufgewachsen, Familie stammt aus Spanien
x2	Deutschkenntnisse	1 Deutschkenntnisse: fließend (verhandlungssicher) durch Besuch deutscher Schule in Spanien (Variante, wenn x1 gleich 1) 1 Muttersprachen: Deutsch und Spanisch (Variante, wenn x1 gleich 2) 2 Deutschkenntnisse: mittleres Niveau durch Schulunterricht und Intensivkurs (nur, wenn x1 gleich 1) 3 Deutschkenntnisse: Grundkenntnisse durch Schulunterricht (nur, wenn x1 gleich 1)
x3	Schul- und Berufsabschlüsse	1 Bildung: Mittlerer Schulabschluss 2 Bildung: Hochschulreife 3 Bildung/Ausbildung: Mittlerer Schulabschluss u. Abschluss schulische Ausbildung (technischer Beruf) 4 Bildung/Ausbildung: Mittlerer Schulabschluss u. Abschluss schulische Ausbildung (kaufmännischer Beruf) 5 Bildung/Ausbildung: Hochschulreife u. Bachelorstudium mit Abschluss (ingenieurwissenschaftliches Fach) 6 Bildung/Ausbildung: Hochschulreife u. Bachelorstudium mit Abschluss (wirtschaftswissenschaftliches Fach)
x4	Soziale Situation	1 Lebenssituation: ledig; Verwandte leben im Ort 2 Lebenssituation: ledig
x5	Finanzielle Situation	1 ...; wird Beihilfe zum Lebensunterhalt beantragen; 2 ...(leer)

5 Vorbereitungen zur Analyse der Faktoriellen Surveys

Um die Daten analysieren zu können, muss die Information darüber, welche Ausprägungen die spezifischen Vignetten haben, die die jeweiligen Befragten erhalten haben, aus den sogenannten „Setup-Datensätzen“ an die Befragungsdaten herangespielt werden. Die Befragungsdaten enthalten Identifikationsnummern, die angeben, welches Vignettenset aus welchem der beiden Faktoriellen Surveys jeder Betrieb erhalten hat, in welcher Reihenfolge die Vignetten innerhalb des Sets präsentiert wurden und wie bzw. ob die Vignetten jeweils bewertet wurden, d. h. die Einschätzung der Befragten auf der Skala von 1 bis 10. Über die Identifikationsnummern können die Vignettenvariablen (Dimensionen) und ihre Ausprägungen (Levels) aus den Setup-Daten hinzugespielt werden. Da die Identifikationsnummern in den beiden Datensätzen unterschiedliche Variablenamen haben, wird empfohlen den im Anhang A1 zur Verfügung gestellten Stata Do-file zu verwenden oder zu konsultieren.

6 Organisation der Mikrodaten

6.1 Fehlende Werte

In Datensätzen des BIBB-FDZ werden grundsätzlich keine System Missings definiert – deren Vergabe wird den Anwendenden überlassen. Fehlende und ungültige Werte wurden nach der Standarddeklarierung fehlender Werte des BIBB-FDZ umcodiert, um so den Umgang mit den fehlenden Werten zu erleichtern. Tabelle 3 gibt die Werte und Wertelabels für die fehlenden Werte bei allen Variablen im BIBB-Qualifizierungspanel wieder.

Tabelle 3: Deklaration fehlender Werte

Wert	Wertelabel
-9	Weiß nicht
-8	Keine Angabe
-7	Trifft nicht zu
-6	Weiß nicht/keine Angabe
-2	Ungültig (z. B. Ausschluss durch Datenbereinigung)
-1	Filter

6.2 Verknüpfung mit dem Hauptinterview

Die Daten des WZB Vignettenmoduls im BIBB-Qualifizierungspanel 2014 können über die *id2* mit den Daten des Hauptinterviews verknüpft werden. Die Daten des BIBB-Qualifizierungspanels 2014 müssen allerdings gesondert beantragt werden. Darüber hinaus ist es ebenfalls möglich, die Vignetten an den Längsschnittdatensatz des BIBB-Qualifizierungspanels zu hängen, sodass für Panelbetriebe Informationen aus mehreren Jahren vorhanden sind und Entwicklungen betrachtet werden können. Auch hier ist eine zusätzliche Beantragung des Längsschnittdatensatzes nötig.

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 TNS Infratest Sozialforschung 2015). Zur Unterscheidung der Fälle steht ein wellenübergreifendes, systemfreies Betriebskennzeichen (id2) zur Verfügung.

Der Datensatz zum WZB Vignettenmodul im BIBB-Qualifizierungspanel 2014 enthält standardmäßig keine Volltexte. Diese können nach einer gesonderten Begründung allerdings ebenfalls über die Datenfernverarbeitung oder einen Gastwissenschaftleraufenthalt ausgewertet werden. Die vorgenommenen Vergrößerungen und Löschungen der Volltextvariablen können Tabelle 4 entnommen werden.

Tabelle 4: Anonymisierungen der Daten

Variablenname	Variableninhalt	Anonymisierung	Neue Variable
ev1	Ausbildungsberuf mit den meisten Auszubildenden am 31.12.2013	Umschlüsselung in KldB 2010 Code ¹ , Original in VT Datensatz überführt	ev1_kldb2010
ev3	Ausbildungsberuf, in dem die meisten Ausbildungsstellen angeboten wurden	Umschlüsselung in KldB 2010 Code ¹ , Original in VT Datensatz überführt	ev3_kldb2010

1 Die Umschlüsselung wurde vom BIBB-FDZ vorgenommen. Bei unklaren Angaben wurden z. T. Sondercodes vergeben (negative Werte).

7.3 Datenzugang

Die Daten des WZB Vignettenmoduls im BIBB-Qualifizierungspanel 2014 können über die Datenfernverarbeitung (DFV) und den Gastwissenschaftleraufenthalt (GWA) im BIBB in Bonn analysiert werden.

Auswertung von Forschungsdaten, die das BIBB-FDZ aus Datenschutzgründen nicht über SUF (Scientific-Use-File) bereitstellen kann, erfolgt über das automatisierte Datenfernverarbeitungssystem JoSuA (Job Submission Application). JoSuA erlaubt Datennutzern und -nutzerinnen, schwach anonymisierte BIBB-Forschungsdaten entweder selbstständig vom eigenen Arbeitsplatz aus oder an den abgeschotteten PC-Arbeitsstationen im BIBB-FDZ auszuwerten. Dieses webbasierte Datenfernverarbeitungssystem ermöglicht mithilfe von lauffähigen, anhand von Strukturdatensätzen (Spieldaten) erstellten Syntaxprogrammen für die Softwarepakete Stata, SPSS und R die autarke Verarbeitung nicht vollständig anonymisierter Originaldaten. Die Datennutzer/-innen können im Metadatenportal Spieldaten herunterladen bzw. diese beim BIBB-FDZ anfragen. Die Spieldaten gleichen in Aufbau und Merkmalsausprägungen den Originaldaten. Der auf Datenschutzkonformität geprüfte Output wird den Nutzerinnen und Nutzern anschließend zum Download bereitgestellt.

Datennutzer/-innen erhalten mit JoSuA eine nahezu vollständige Autonomie bei der Datenauswertung und somit eine ortsunabhängige Möglichkeit zur Auswertung von BIBB-Forschungsdatensätzen. Zusätzlich steht den Forschenden an den Gastarbeitsplätzen des BIBB-FDZ ein Windows-Computer bereit auf dem sie Zugriff auf die Originaldaten über die Benutzeroberfläche mit Dateneditor der jeweiligen Statistiksoftware haben, um Syntaxen schreiben, testen und auswerten zu können. Datenfernverarbeitungen und Gastwissenschaftleraufenthalte sind über JoSuA eng miteinander verknüpft und erlauben Nutzenden, vor einem GWA bereits erarbeitete Syntaxen hochzuladen und anschließend Forschungsergebnisse, nach erfolgter Datenschutzprüfung, herunterzuladen. Die Nutzung von JoSuA ist nur über ein projektbezogenes und personalisiertes Log-in möglich. Für den Erhalt der Zugangsdaten ist das Durchlaufen eines formellen Antrags- und Vertragsverfahrens erforderlich. Die entsprechenden Formulare sowie Erläuterungen zum Antrags- und Vertragsverfahren können von den Internetseiten des BIBB-FDZ heruntergeladen werden.

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.

Literaturverzeichnis

AUSPURG, Katrin; HINZ, Thomas: Factorial Survey Experiments. Thousand Oaks: Sage 2015

GERHARDS, Christian; FRIEDRICH, Anett: BIBB-Qualifizierungspanel 2014. Daten- und Methodenbericht Nr. 3/2016. Bonn 2016 – URL: <https://www.bibb.de/dienst/veroeffentlichungen/de/publication/show/8216>

TNS INFRATEST SOZIALFORSCHUNG: BIBB-Qualifizierungspanel. Qualifizierung und Kompetenzentwicklung. Methodenbericht zur vierten Welle. München 2015

A1: Stata do-file zum Zusammenspielen der Vignettenvariablen (Dimensionen) und ihrer Ausprägungen (Levels) aus den Setup-Datensätzen an die Befragungsdaten

```
clear
version 15.1
capture log close
set more off

*****
/* Dieser Do-file bringt die Setup-Daten für die beiden Faktoriellen Surveys A und B mit
den Befragungsdaten zusammen. Die Faktoriellen Surveys A und B werden im Folgenden auch Modul A und Modul B
genannt

- Die Befragungsdaten enthalten Identifikationsnummern, die angeben welches Vignettenset (auch Deck genannt)
aus welchem der
beiden Module jeder Betrieb erhalten hat, in welcher Reihenfolge die Vignetten innerhalb des Sets präsentiert
wurden und
wie bzw. ob die Vignetten jeweils bewertet wurden (Einschätzung der Befragten auf Skala 1 bis 10).

- Über diese Identifikationsnummern werden die Vignettenvariablen (Dimensionen) und ihre Ausprägungen (Levels)
aus den Setup-Daten hinzugespielt.
*/
*****
* Dateipfade festlegen
```

```

global qualipanel ""
global input ""
global output ""

* Befragungsdatensatz. Datensatz ist im wide-Format. Jedem Betrieb wurden 5 Vignetten präsentiert.
use "$qualipanel\WZB_qp_2014_gwa.dta", clear
order id2

* Variablen des Datensatzes:
//id2 ist die ID der befragten Betriebe
//Variablen, deren Name mit vn* beginnt sowie ausw_ sind vignettenspezifische Variablen
//Variablen, deren Name mit ev* beginnt sind weitere Items, die zusammen mit den Vignetten erhoben wurden
//Variable hr2014vig ist ein Gewichtungsfaktor, um auf die Grundgesamtheit der Betriebe dieses Subsamples des
//QP2014 hochrechnen zu können

* Bedeutung der vignettenspezifischen Variablen:
//v#_1 : Modul A oder B
//v#_2 : ID des Vignettensets
//v#_3 : Setspezifische ID der Vignette
//v#_4 : Modulspezifische ID der Vignette über Sets hinweg
//v#_5 : ID generiert von Infratest, die die Info aus den var v#_1 bis v#4 enthält
//vn#_e : Einschätzung der Befragten, d. h. Rating auf Skala 1 bis 10
//vn#_woe: Hat Wert 1, wenn Befragte Vignette nicht bewertet haben
//ausw_ : Gibt an, welche Vignette innerhalb eines Sets die Befragten bei der randomisierten Reihenfolge als
//erstes, zweites, etc. erhalten haben

```

```
* Generierung Variable modul mit den Werten 1 und 2, Modul A==1, Modul B==2
gen modul =.
replace modul=1 if vn1_1==1 | vn2_1==1 | vn3_1==1 | vn4_1==1 | vn5_1==1
replace modul=2 if vn1_1==2 | vn2_1==2 | vn3_1==2 | vn4_1==2 | vn5_1==2

* Daten müssen vom wide- ins long-Format transformiert werden

* Umbenennung der Variablen für long-Format
rename vn1_e evn1
rename vn1_1 einsvn1
rename vn1_2 zweivn1
rename vn1_3 dreivn1
rename vn1_4 viervn1
rename vn1_5 fuenfvn1
rename vn1_woe woevn1

rename vn2_e evn2
rename vn2_1 einsvn2
rename vn2_2 zweivn2
rename vn2_3 dreivn2
rename vn2_4 viervn2
rename vn2_5 fuenfvn2
rename vn2_woe woevn2
```

```
rename vn3_e evn3  
rename vn3_1 einsvn3  
rename vn3_2 zweivn3  
rename vn3_3 dreivn3  
rename vn3_4 viervn3  
rename vn3_5 fuenfvn3  
rename vn3_woe woenvn3
```

```
rename vn4_e evn4  
rename vn4_1 einsvn4  
rename vn4_2 zweivn4  
rename vn4_3 dreivn4  
rename vn4_4 viervn4  
rename vn4_5 fuenfvn4  
rename vn4_woe woenvn4
```

```
rename vn5_e evn5  
rename vn5_1 einsvn5  
rename vn5_2 zweivn5  
rename vn5_3 dreivn5  
rename vn5_4 viervn5  
rename vn5_5 fuenfvn5  
rename vn5_woe woenvn5
```

* Transformation ins long-Format

```
reshape long evn einsvn zweivn dreivn viervn fuefnv woevni , i(id2) j(positionvig)

sort id2 ausw_

** Neue Variablennamen vergeben
* von Infratest generierte ID
rename fuefnv infratest_vig_id
label var infratest_vig_id „Von Infratest generierte ID“

*Vignetten, die nicht bewertet wurden, d. h. bei denen „weiter ohne Einschätzung“ angegeben wurde
rename woevni vig_missing

* Variable gibt die Position in der Reihenfolge an, in der der Betrieb die Vignette bekommen hat
rename ausw_ reihenfolge
label var reihenfolge „Position in der Reihenfolge, in der Betrieb Vignette bekommen hat“

* Variable gibt die Bewertung der Vignette, d. h. wie wahrscheinlich die/der Bewerberin/Bewerber
* eingeladen wird: von 1 (sehr unwahrscheinlich) bis 10 (sehr wahrscheinlich)
rename evni bewertung
label var bewertung „Bewertung der Vignette“

* Modulspezifische ID der Vignette über Sets hinweg, wird für Match mit Setup-Daten benötigt
rename viervni id_vignr
label var id_vignr „Modulspezifische Vignetten ID über Sets hinweg“
```

```
save "$output\Vignetten_Gesamt_long.dta", replace

* Getrennte Datensätze für Modul A und B
drop if modul==2
save "$output\Vignetten_modula_long.dta", replace

use "$output\Vignetten_Gesamt_long.dta", clear

drop if modul==1
save "$output\Vignetten_modulB_long.dta", replace

* Modul A
* setup/vignettendatensatz laden
use "$input\wzb_setup_modula_gwa.dta", clear

merge 1:m id_vignr using "$output\Vignetten_modula_long.dta"
sort id2 zweivn reihenfolge

tab _merge
drop _merge modul einsvn

* Doppelungen von Variablen

* dreivn, set_vignr (vignettendatensatz) und positionvig (reshape befehl) geben jeweils die ID der Vignette
* im Set an
```

```
* zweivn und deck (vignettendatensatz) geben jeweils die ID des Vignettenset an

* test ob dreivn, set_vignr und positionvig jeweils gleiche Werte haben
* es dürfen jeweils keine Fälle angezeigt werden, dann sind Variablen tatsächlich gleich
list id2 id_vignr if dreivn != positionvig
list id2 id_vignr if dreivn != set_vignr
list id2 id_vignr if positionvig != set_vignr

* wenn Tests erfolgreich, dann:
drop positionvig dreivn

* Test ob zweivn und deck immer die gleichen Wert haben
list id2 id_vignr if zweivn !=deck

*Codierungsfehler in Befragungsdaten bei Beobachtungen id2 6023 und 15745
* id2 6023, infratest_vig_id:1464229, zweivn==16 statt richtig 46;
* id2 15745, infratest_vig_id:1334164, zweivn==. statt richtig 33
* --> da infratest_vig_id korrekt, kann zweivn gelöscht werden
drop zweivn

* Test ob vig_missing der Info entspricht, dass Vignette nicht bewertet wurde
list id2 id_vignr if vig_missing==1 & bewertung !=.

*wenn Tests erfolgreich, dann:
drop vig_missing
```

```
sort id2 reihenfolge
order id2 reihenfolge id_vignr deck set_vignr infratest_vig_id bewertung

save "$output\Merge_modulA.dta"

*Modul B
* setup/vignettendatensatz laden
use "$input\wzb_setup_modulB_gwa.dta", clear

merge 1:m id_vignr using "$output\Vignetten_modulB_long.dta"
sort id2 zweivn reihenfolge

tab _merge
drop _merge modul einsvn

* Doppelungen von Variablen
* dreivn, set_vignr (vignettendatensatz) und positionvig (reshape befehl) geben jeweils die ID der Vignette
* im Set an
* zweivn und deck (vignettendatensatz) geben jeweils die ID das Vignettenset an
* test ob dreivn, set_vignr und positionvig jeweils gleiche Werte haben
* es dürfen jeweils keine Fälle angezeigt werden, dann sind Variablen tatsächlich gleich
list id2 id_vignr if dreivn != positionvig
list id2 id_vignr if dreivn != set_vignr
```

```
list id2 id_vignr if positionvig != set_vignr

* wenn Tests erfolgreich, dann:
drop positionvig dreivn

* Test ob zweivn und deck immer die gleichen Wert haben
list id2 id_vignr if zweivn !=deck

* wenn Tests erfolgreich, dann:
drop zweivn

* Test ob vig_missing der Info entspricht, dass Vignette nicht bewertet wurde
list id2 id_vignr if vig_missing==1 & bewertung !=.

* wenn Tests erfolgreich, dann:
drop vig_missing

sort id2 reihenfolge
order id2 reihenfolge id_vignr deck set_vignr infratest_vig_id bewertung

save "$output\Merge_modulB.dta"
```

Sonderedition zum 50. Jahrgang mit Beiträgen aus fünf Jahrzehnten

Fünf Jahrzehnte lang begleitet die BWP den Fachdiskurs zur Berufsbildung in Wissenschaft, Politik und Praxis.

Aus rund 3.500 Beiträgen, die seit der ersten Ausgabe erschienen sind, hat Georg Hanf, ehemaliger Arbeitsbereichsleiter im BIBB, für jedes Jahrzehnt je einen Beitrag aus Wissenschaft, Politik und Praxis ausgewählt.

Die Beiträge behandeln im Wesentlichen zwei übergreifende Themen:

- die Qualität der Ausbildung und
- die Gleichwertigkeit von beruflicher und allgemeiner Bildung.

Das Fazit von Georg Hanf: »Der Rückblick auf 50 Jahre BWP verdeutlicht: Es dauerte oft lange, bis Erkenntnisse und Vorschläge, die in der Zeitschrift erstmals vorgebracht wurden, Eingang in die reguläre Praxis fanden. Manches harrt weiter der Erfüllung.«

www.bwp-zeitschrift.de/se-2021



Bundesinstitut für Berufsbildung (Hrsg.)
Berufsbildung in Wissenschaft und Praxis
Eine Beitragsauswahl aus fünf Jahrzehnten
Zusammengestellt von Georg Hanf
Sonderedition 2021, 98 Seiten
ISSN 0341-4515

Die Sonderedition liegt als digitale Version vor, die Beiträge stehen einzeln zum Download zur Verfügung.



Bundesinstitut für Berufsbildung
Robert-Schuman-Platz 3
53175 Bonn

Telefon (0228) 1 07-0

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