



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

# FDZ-DATENREPORT

Dokumentation zu Arbeitsmarktdaten

---

**09|2025 DE** IAB-Stellenerhebung: Betriebsbefragung zu  
Stellenangebot und Besetzungsprozessen, Welle 2000 bis  
2022 mit Folgequartalen von 2006 bis 2023

André Diegmann, Nicole Gürtzgen, Tim Kovalenko, Alexander Kubis, André Pirralha, Laura Pohlan, Martin Popp, Franka Vetter



**Bundesagentur für Arbeit**

# IAB-Stellenerhebung: Betriebsbefragung zu Stellenangebot und Besetzungsprozessen, Welle 2000 bis 2022 mit Folgequartalen von 2006 bis 2023

André Diegmann (IAB, IWH), Nicole Gürtzgen (IAB), Tim Kovalenko (IAB), Alexander Kubis (IAB), André Pirralha (IAB), Laura Pohlan (IAB), Martin Popp (IAB), Franka Vetter (IAB)

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

FDZ-Datenreporte (FDZ data reports) describe FDZ data in detail. As a result, this series of reports has a dual function: on the one hand, those using the reports can ascertain whether the data offered is suitable for their research task; on the other, the data can be used to prepare evaluations.

# Inhaltsverzeichnis:

<b>1</b>	<b>Änderungen am Datensatz gegenüber früheren Versionen .....</b>	<b>6</b>
<b>2</b>	<b>Beschreibung des Datensatzes .....</b>	<b>7</b>
2.1	Kurzbeschreibung.....	7
2.2	Überblick.....	8
2.3	Grundgesamtheit und Stichprobenziehung der Hauptbefragung.....	9
2.4	Hochrechnungsverfahren der Hauptbefragung .....	11
2.5	Stichprobenziehung und Hochrechnung der Quartalsbefragung .....	12
<b>3</b>	<b>Wellenspezifische Besonderheiten .....</b>	<b>13</b>
3.1	Befragungswelle 2022.....	13
3.1.1	Verringerte Bereinigung der Daten.....	13
3.1.2	Veränderte Codierung von Missings .....	13
3.2	Befragungswelle 2021.....	13
3.2.1	Zusätzliche Quartalsbefragung .....	13
3.3	Befragungswelle 2020.....	14
3.3.1	Methodische Experimente .....	14
3.3.2	Erstmalige Abfrage des Einverständnisses zur Verlinkung von administrativen Daten .....	15
3.3.3	Revision der Daten und Gewichte aufgrund von Fehlern in der Datenaufbereitung.....	15
3.4	Befragungswelle 2019.....	15
3.4.1	Zusätzliche Quartalsbefragungen .....	15
3.4.2	Experimental-Daten der Befragungswelle 2019 .....	16
3.4.3	Revision aufgrund von abweichendem Verhalten von Interviewerinnen bzw. Interviewern .....	16
3.5	Befragungswelle 2016.....	17
3.5.1	Vignettenstudie zur Wirkung der Mindestlohnhöhe .....	17
<b>4</b>	<b>Verwendungshinweis für Datennutzerinnen und Datennutzer .....</b>	<b>18</b>
4.1	Struktur des Datensatzes.....	18
4.2	Verwendung der Gewichte.....	19
4.3	Fragebogenteile mit möglichen Mehrfachantworten.....	20
4.4	Klassifikationen von Wirtschaftszweigen und Berufen .....	21
4.5	Verlinkung zu anderen Daten .....	23
4.5.1	BHP-Verlinkung.....	23
4.5.2	Weitere Verlinkungsmöglichkeiten.....	26
4.6	STATA-Auswertungsbeispiel.....	27

4.6.1	Zeitreihe der offenen Stellen .....	27
4.6.2	Auswertung einer Frage mit Mehrfachantworten .....	28

# Tabellenverzeichnis:

Tabelle 1: Inhaltliche Charakteristika .....	7
Tabelle 2: Methodische Charakteristika .....	7
Tabelle 3: Datenzugang .....	7
Tabelle 4: Schichtungsmerkmale der Bruttostichprobe 2000 - 2022.....	10
Tabelle 5: Stichprobenentwicklung 2000-2022 .....	11
Tabelle 6: Codierung von Missings .....	13
Tabelle 7: Technische-Variablen ohne Variablennamen-Präfixe .....	18
Tabelle 8: Dateinamen und Variablennamen-Präfixe .....	19
Tabelle 9: Zu verwendendes Hochrechnungsgewicht nach Modul und Welle .....	20
Tabelle 10: Beispiel für die Umbuchung einer freien Textfeld-Antwort.....	21
Tabelle 11: Berufsklassifikationen nach Befragungswelle .....	22
Tabelle 12: Ergänzung der KldB2010 für Angaben ohne nähere Tätigkeitsbeschreibung (o.n.T) .....	22
Tabelle 13: Dateinamen der zuspieldbaren BHP-Datensätze .....	23

## Zusammenfassung:

Die IAB-Stellenerhebung ist eine quartalsweise durchgeführte und repräsentative Betriebsbefragung über das gesamtwirtschaftliche Stellenangebot sowie Einstellungsprozesse in Deutschland. Die Erhebung ermittelt die Gesamtzahl aller offenen Stellen am Arbeitsmarkt, einschließlich jener Stellen, die nicht der Bundesagentur für Arbeit (BA) gemeldet werden. Das erste Modul des Fragebogens enthält Informationen zur Zahl und Struktur offener Stellen, dem erwarteten künftigen Arbeitskräftebedarf, zur wirtschaftlichen Lage und zur Entwicklung der befragten Betriebe. Das zweite Modul erfragt die betriebliche Einschätzung und Nutzung aktueller arbeitsmarktpolitischer Instrumente sowie den betrieblichen Umgang mit am Arbeitsmarkt benachteiligten Personen. Das dritte Modul enthält Fragen zum letzten Fall einer Neueinstellung und zum letzten Fall eines gescheiterten Rekrutierungsversuchs. Das Forschungsdatenzentrum der Bundesagentur für Arbeit stellt die Datensätze der Befragungswellen ab 2000 mit allen Modulen einschließlich der Quartalsbefragungen für externe Wissenschaftlerinnen und Wissenschaftler bereit.

## Abstract:

The IAB Job Vacancy Survey is a quarterly and representative establishment survey on labour demand and recruitment processes in Germany. The survey identifies the overall stock of vacancies in the German labour market, including those vacancies that are not reported to the Federal Employment Agency (FEA). The first module of the questionnaire collects information about the number and structure of vacancies, future personnel requirements, about the current economic situation and the expected development of participating establishments. The second module enquires about employer attitudes and firm use of current labour market instruments as well as the employer handling of people disadvantaged in the labour market. The third module asks for information about the last new hire and the last case of a failed recruitment effort. The Research Data Centre of the Federal Employment Agency offers the data sets of the survey waves from 2000 onwards.

## Keywords:

Betriebsbefragung, offene Stellen, Stellenbesetzungsprozesse

# 1 Änderungen am Datensatz gegenüber früheren Versionen

Im Vergleich zu früheren Versionen des Datensatzes (vgl. Börschlein et al., 2024) wurde der Zeitraum der bereitgestellten Daten um das 4. Quartal 2022 und die anschließenden Folgequartale erweitert. Im Rahmen des Moduls 2 der Befragung im 4. Quartal 2022 wurden Daten zu Beschäftigungschancen arbeitsloser Personen (siehe auch Kubis/Popp, 2024), zu arbeitsrelevanten Erfahrungen mit Geflüchteten aus der Ukraine und zu gestiegenen Energiekosten und deren mögliche Auswirkungen erhoben.

Im 2. Quartal 2023 wurden zudem Fragen zur Nutzung von Dienstleistungen des Arbeitgeber-Service der Bundesagentur für Arbeit gestellt (siehe hierzu Hentschke/Kubis, 2024).

Die Angaben zu offenen Stellen sowie Suchdauern werden seit dem 1. Quartal 2023 im Rahmen des vierteljährlichen IAB-Arbeitskräftemonitors für die Fachöffentlichkeit aufbereitet (siehe z.B. Gürtzgen/Kubis/Popp, 2023).

Die Befragungswelle 2022 weist Besonderheiten auf, die bei der Nutzung der Daten zu beachten sind. Diese werden im Abschnitt 3 detailliert erläutert.

## 2 Beschreibung des Datensatzes

### 2.1 Kurzbeschreibung

**Tabelle 1: Inhaltliche Charakteristika**

Kategorie	Beschreibung
Themen/Merkmalgruppen	Beschäftigung, Personalbedarf, Aktivitätshemmnisse, abgebrochene Personalsuche, Berufe, für die steigender Bedarf erwartet wird und mit erwarteten Engpässen in den kommenden 3 Jahren, geförderte Beschäftigung, geringfügige Beschäftigung, Teilzeitbeschäftigung, Reform des Arbeitsmarktes, Beschäftigungschancen (langzeit-)arbeitsloser Personen, Weiterbildung der Belegschaft, Ein-Euro-Jobs (Erfahrungen der Betriebe), Arbeitsmarktpartizipation älterer Beschäftigter, arbeitsrelevante Erfahrungen mit Geflüchteten und Menschen mit Schwerbehinderung, Umgang mit der COVID-19 Krise, Umgang mit den Auswirkungen des Ukraine-Kriegs, Erfahrung mit ukrainischen Geflüchteten, Mindestlohnbetroffenheit, Arbeitgeber-Service, detaillierte Informationen zur letzten Einstellung, detaillierte Informationen zur letzten abgebrochenen Personalsuche
Untersuchungseinheit	Betriebe
Fallzahlen	5.150 bis 21.115 Betriebe
Zeitraum	IV.2000 bis III.2023
Zeitbezug	quartalsweise
Regionale Gliederung	Ost-/Westdeutschland, Bundesländer, weitere Merkmale siehe unter "sensible Merkmale"

**Tabelle 2: Methodische Charakteristika**

Kategorie	Beschreibung
Erhebungsdesign	Repräsentative Stichprobe der Betriebe, geschichtet nach Betriebsgröße, Branche und Ost-/Westdeutschland
An der Erhebung beteiligte Institutionen	Auftraggeber: Forschungsbereich AMPI des IAB Durchführung: Economics & Data ED23 GmbH (vormals Economix Research & Consulting), München
Frequenz der Datensammlung	Jährliche schriftliche bzw. web und quartalsweise telefonische Befragung
Dateiformat/-größe	STATA; alle Datensätze zusammen ca. 315 MB
Dateiorganisation	nach Welle

**Tabelle 3: Datenzugang**

Kategorie	Beschreibung
Datenzugang	Datenfernverarbeitung, Gastaufenthalt
Anonymisierungsgrad	Schwach anonymisiert
sensible Merkmale	Wirtschaftszweigklassifikationen: WZ03 5-Steller (2005–2009) WZ08 5-Steller (ab 2010), Kreisnummer (ab 2007), Regierungsbezirk (ab 2005), Betriebs-Historik-Panel: Kreis des Arbeitsorts (ao_kreis), Klassifikation der Wirtschaftszweige 1993 – 5-Steller (w93_5), Klassifikation der Wirtschaftszweige 2003 – 5-Steller (w03_5), Klassifikation der Wirtschaftszweige 2008 – 5-Steller (w08_5)

Kategorie	Beschreibung
Vorgaben zum Zitieren der Daten und Datendokumentation	<p>Daten:  "Gürtzgen, Nicole; Kubis, Alexander; Diegmann, André; Kovalenko, Tim; Pirralha, André; Pohlan, Laura; Popp, Martin; Vetter, Franka (2025): "IAB-Stellenerhebung – Version 0022 v1".  Forschungsdatenzentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB). DOI: 10.5164/IAB.IABSE0022.de.en.v1  Der Datenzugang erfolgte über einen Gastaufenthalt am  Forschungsdatenzentrum der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (FDZ) und/oder mittels kontrollierter Datenfernverarbeitung beim FDZ.“</p> <p>Datendokumentation:  - Diegmann, André; Gürtzgen, Nicole; Kovalenko, Tim; Kubis, Alexander; Pirralha, André; Pohlan, Laura; Popp, Martin; Vetter, Franka (2025): IAB-Stellenerhebung: Betriebsbefragung zu Stellenangebot und Besetzungsprozessen, Wellen 2000 bis 2022 mit Folgequartalen von 2006 bis 2023. FDZ-Datenreport, 09/2025 (de), Nürnberg.  DOI:10.5164/IAB.FDZD.2509.de.v1  - Diegmann, André; Gürtzgen, Nicole; Kovalenko, Tim; Kubis, Alexander; Pirralha, André; Pohlan, Laura; Popp, Martin; Vetter, Franka (2025): The IAB Job Vacancy Survey: Establishment survey on labour demand and recruitment processes, waves 2000 to 2022 and subsequent quarters 2006 to 2023. FDZ-Datenreport, 09/2025 (en), Nuremberg. DOI:10.5164/IAB.FDZD.2509.en.v1</p>

Details zu verschiedenen Zugangswegen, zu den Voraussetzungen der Nutzung sowie zur Antragstellung des Datensatzes können der FDZ-Internetseite unter <https://fdz.iab.de/> entnommen werden.

## 2.2 Überblick

Die IAB-Stellenerhebung ist eine Betriebsbefragung, die repräsentativ das gesamtwirtschaftliche Stellenangebot in Deutschland abbildet (siehe auch Bossler et al., 2020). Sie ermittelt die Gesamtzahl aller offenen Stellen am Arbeitsmarkt, einschließlich jener Stellen, die nicht der Bundesagentur für Arbeit (BA) gemeldet werden. Allein diese Gesamtzahl erlaubt es, valide und unverzerrte Aussagen über die gesamtwirtschaftliche Arbeitsnachfrage zu treffen. Denn die Meldequote, also der Anteil der von der BA gemeldeten an allen offenen Stellen, unterliegt im Zeitverlauf starken Schwankungen und unterscheidet sich systematisch zwischen Wirtschaftszweigen, Regionen und Berufen.

Die Erhebung wird seit 1989 im vierten Quartal jedes Jahres schriftlich mit einem mehrteiligen Fragebogen durchgeführt. Seit 2002 können Betriebe an der Befragung auch online teilnehmen. Der Webfragebogen ist mit dem schriftlichen Fragebogen so identisch wie möglich. Das Modul 1, ehemals Hauptfragebogen genannt, enthält Informationen zur Zahl und Struktur offener Stellen, zum erwarteten künftigen Arbeitskräftebedarf, zur wirtschaftlichen Lage und zur Entwicklung der befragten Betriebe. In Modul 2, ehemals Sonderfragebogen genannt, wird seit 2000 im SGBII-Kontext die betriebliche Einschätzung und Nutzung aktueller arbeitsmarktpolitischer Instrumente erfragt. Zudem werden Daten zum Umgang der Betriebe mit am Arbeitsmarkt benachteiligten Personen erhoben. Das Modul 3, ehemals Zusatzfragebogen genannt, enthält Fragen zum letzten Fall einer Neueinstellung sowie zum letzten Fall eines gescheiterten Rekrutierungsversuchs. Diese regelmäßige und detaillierte Abbildung von Stellenbesetzungsprozessen ist weltweit ein Alleinstellungsmerkmal der IAB-Stellenerhebung. Sie liefert damit Informationen zu Merkmalen der besetzten Stelle, der eingestellten Person, zu Such- und Besetzungswegen einschließlich des Einschaltungsgrades der Arbeitsvermittlung, zu Such- und Besetzungszeiten, Zahlen zu Bewerberinnen und Bewerbern, Besetzungsschwierigkeiten sowie über gemachte Kompromisse.

Das Modul 3 wird nur von Betrieben beantwortet, die im letzten Jahr Personen neu eingestellt oder dies zumindest versucht haben.

Modul 1 und 3 sind von Welle zu Welle im Wesentlichen gleich, abgesehen von redaktionellen Änderungen und der Verlagerung von Schwerpunkten. Modul 2 ändert sich hingegen mit jeder Befragungswelle. Seit 2000 werden in diesem Modul die betriebliche Einschätzung und Nutzung aktueller arbeitsmarktpolitischer Instrumente erhoben. Von 2005 bis 2013 thematisierte das Modul hauptsächlich die betriebliche Nutzung und Einschätzung von Ein-Euro-Jobs. Es richtete sich in diesem Zeitraum deshalb nur an die Wirtschaftszweige im öffentlichen Sektor, in welchen die überwiegende Zahl der Ein-Euro-Jobs angeboten und durchgeführt wurde.

Seit 2005 finden im 1., 2. und 3. Quartal des jeweiligen Folgejahres einer schriftlichen Befragungswelle kurze telefonische Nachbefragungen zu den Kernfragen unter den teilnehmenden Betrieben der Hauptbefragung statt. Diese telefonischen Quartalsbefragungen ergänzen die schriftliche Hauptbefragung aus dem 4. Quartal. Erfragt und aktualisiert werden die wichtigsten Eckdaten aus dem Modul 1. Hierunter fällt u.a. die Zahl der Beschäftigten eines Betriebes, seine Einschätzung der Beschäftigungsentwicklung und die Zahl der offenen Stellen. In der Befragungswelle 2019, die mit der Hauptbefragung im 4. Quartal 2019 begann, fanden zusätzliche Nachbefragungen im 4. Quartal 2020 und im 1. Quartal 2021 statt. In der Befragungswelle 2021, beginnend mit der Hauptbefragung im 4. Quartal 2021, fand eine zusätzliche Nachbefragung im 4. Quartal 2022 statt.

Im Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit werden die Befragungswellen ab 2000 mit allen Modulen einschließlich der Quartalsbefragungen für externe Wissenschaftlerinnen und Wissenschaftler angeboten. Seit September 2015 stehen die Variablen- und Wertelabels des Datensatzes wie auch die vollständige Dokumentation einschließlich der Codepläne und Datenreports auch in englischer Sprache zur Verfügung.

### 2.3 Grundgesamtheit und Stichprobenziehung der Hauptbefragung

Grundgesamtheit der Befragung im 4. Quartal eines jeweiligen Jahres sind sämtliche Betriebe mit mindestens einem oder einer sozialversicherungspflichtig Beschäftigten im Juni (bis Welle 2004) bzw. Dezember (ab Welle 2005) des Vorjahres unter Ausschluss der Privathaushalte. Ein „Betrieb“ bezieht sich dabei auf eine wirtschaftliche Einheit nach dem Betriebsnummernkonzept des Betriebsnummern-Service der Bundesagentur für Arbeit (Bundesagentur für Arbeit, 2020).

Aus dieser Grundgesamtheit wird jedes Jahr erneut eine disproportional geschichtete Zufallsbruttostichprobe gezogen. Geschichtet wird nach der Region, Betriebsgrößenklasse sowie nach Wirtschaftszweig, sodass eine dreidimensionale Stichprobenmatrix entsteht. Tabelle 4 fasst die Schichtungsmerkmale zusammen. In den Wellen 2005 bis 2013 wurde darüber hinaus eine Ergänzungsstichprobe im öffentlichen Sektor für die Erhebung der Ein-Euro-Jobs nach gleichem Schema gezogen. Nur in der Welle 2005 wurde außerdem aufgrund des geringen Rücklaufs nach Beginn der Feldphase eine weitere Ergänzungsstichprobe zur Erhöhung der Fallzahlen gezogen. In beiden Fällen wurden vor der Ziehung der Ergänzungsstichproben die bereits gezogenen Betriebe aus der Ziehungsgesamtheit entfernt.

**Tabelle 4: Schichtungsmerkmale der Bruttostichprobe 2000 - 2022**

<b>Merkmal</b>	<b>Unterteilungen</b>
Region	2000–2003: Westdeutschland mit West-Berlin, Ostdeutschland mit Ost-Berlin ab 2004: Westdeutschland, Ostdeutschland mit Berlin
Betriebsgröße (Anzahl aller Beschäftigter)	2000–2003: 1–9, 10–19, 20–49, 50–199, 200–499, 500+ 2004: 1–9, 10–19, 20–49, 50–99, 100–199, 200–499, 500–999, 1000+ 2005–2013: 1–9, 10–19, 20–49, 50–199, 200–499, 500–999, 1000+ ab 2014: 1–9, 10–19, 20–49, 50–249, 250–499, 500–999, 1000+
Wirtschaftszweig	2000–2002: 22 Wirtschaftszweige auf Basis der Klassifikation 1973 2003: 14 Wirtschaftszweige auf Basis der Klassifikation 1973 2004–2009: 28 Wirtschaftszweige auf Basis der Klassifikation 2003 2010–2014: 23 Wirtschaftszweige auf Basis der Klassifikation 2008 ab 2015: 24 Wirtschaftszweige auf Basis der Klassifikation 2008

Um von der Bruttostichprobe zur bereinigten Versandstichprobe zu gelangen, werden all jene Betriebe ausgeschlossen, die bei vergangenen Befragungen des IAB geäußert haben, nicht mehr zur Teilnahme aufgefordert werden zu wollen. Diese bildet die Basis für die in Tabelle 5 angegebene Rücklaufquote.

Sämtliche Betriebe der Versandstichprobe werden Anfang Oktober schriftlich per Post kontaktiert und zur Teilnahme aufgefordert. Die Sendung umfasst ein Anschreiben, eine Datenschutzerklärung, den Fragebogen sowie ein Blatt mit Ausfüllhinweisen und Antworten auf häufig gestellte Fragen. Wenige Wochen später wird eine zweite Postsendung versandt, deren Anschreiben an die Teilnahme erinnert und ansonsten den gleichen Umfang und Inhalt hat. Weiterhin ist auf dem Anschreiben vermerkt, telefonische Rückfragen an eine ganztägig besetzte Hotline richten zu können. Die Feldphase dauert in der Regel bis Anfang Januar, wobei der größte Teil des Rücklaufes zwischen Mitte Oktober und Anfang Dezember erfolgt.

Tabelle 5: Stichprobenentwicklung 2000-2022

Welle	Adressbestand Stichtag	Durchschn. Auswahlsatz	Brutto- stichprobe	Versand- stichprobe	Netto- stichprobe	Rücklaufquote
2000	31.12.1999	1,3 %	28.486	28.266	7.578	26,8 %
2001	30.06.2000	1,3 %	27.994	27.827	7.347	26,4 %
2002	30.06.2001	1,3 %	27.418	27.147	5.773	21,3 %
2003	30.06.2002	1,8 %	37.789	37.421	7.310	19,5 %
2004	30.06.2003	2,7 %	56.926	56.699	11.707	20,6 %
2005	31.12.2004	3,9 %	78.032	69.702	11.742	16,8 %
2006	31.12.2005	3,9 %	75.290	69.231	13.537	19,6 %
2007	31.12.2006	3,8 %	75.128	73.635	14.381	19,5 %
2008	31.12.2007	3,8 %	77.543	75.035	13.652	18,2 %
2009	31.12.2008	3,8 %	77.537	74.998	15.288	20,4 %
2010	31.12.2009	3,7 %	77.739	75.000	15.124	20,2 %
2011	31.12.2010	3,6 %	77.685	74.660	15.139	20,3 %
2012	31.12.2011	3,6 %	79.181	75.006	13.807	18,4 %
2013	31.12.2012	3,6 %	77.214	75.486	14.019	18,6 %
2014	31.12.2013	3,6 %	76.767	75.073	12.750	17,0 %
2015	31.12.2014	3,9 %	85.556	83.639	12.852	15,4 %
2016	31.12.2015	4,0 %	110.140	85.000	11.588	13,6 %
2017	31.12.2016	5,1 %	120.102	109.988	14.614	13,3 %
2018	31.12.2017	5,2 %	134.848	109.988	14.528	13,2 %
2019	31.12.2018	5,1 %	114.613	109.986	13.906	12,6 %
2020	31.12.2019	7,1 %	159.979	155.004	20.115	13,0 %
2021	31.12.2020	5,1 %	126.418	109.991	12.815	11,7 %
2022	31.12.2021	6,2 %	146.065	134.991	14.078	10,4 %

## 2.4 Hochrechnungsverfahren der Hauptbefragung

Einzelheiten und Hintergründe zum verwendeten Hochrechnungsverfahren liefert der IAB-Forschungsbericht Nr. 4/2016 (Brenzel et al., 2016). Zur Berechnung der Betriebsgewichte werden folgende Schritte durchgeführt:

1. Berechnung der Designgewichte für jede Zelle der Stichprobenmatrix als Inverse der jeweiligen Auswahlwahrscheinlichkeit.
2. Multiplikation der Designgewichte mit dem Gewicht aus einer Non-Response-Modellierung. Dies erfolgt auf Basis eines logistischen Regressionsmodells, bei dem die Betriebsgrößenklasse, der Wirtschaftszweig, das durchschnittliche Tagesentgelt sowie das Durchschnittsalter der Beschäftigten aus den administrativen Daten der Bundesagentur für Arbeit entnommen und als Hilfsvariablen verwendet werden. Das Non-Response-Gewicht ist die Inverse der geschätzten Antwortwahrscheinlichkeit.
3. Kalibrierung der Gewichte aus Schritt 2 als Startgewichte eines generalisierten Regressionsschätzers (GREG) mit dem Ziel, Eckwerte der Ankervariablen „Zahl der Betriebe“ und „Anzahl der sozialversicherungspflichtig Beschäftigten“ in allen Zellen der Stichprobenmatrix exakt zu erreichen.

Da die Ankervariable „Anzahl der sozialversicherungspflichtig Beschäftigten“ am aktuellen Rand durch Fortschreibung geschätzt werden muss (siehe hierzu Hutter, 2020), wird die Kalibrierung nach jeweils 18 Monaten mit der tatsächlichen Anzahl der sozialversicherungspflichtig Beschäftigten wiederholt und die Hochrechnungsgewichte somit revidiert. Der FDZ-Datensatz enthält ausschließlich die in dieser Weise revidierten Hochrechnungsgewichte; daher kann jede neue Welle frühestens 18 Monate nach Abschluss der Erhebungsphase im FDZ eingestellt werden. Zuvor veröffentlichte Hochrechnungsergebnisse des IAB in Presseinformationen, IAB-Forumsartikeln und IAB-Kurzberichten basieren hingegen in der Regel auf den vorläufigen Hochrechnungsgewichten und können deshalb mit den im FDZ zur Verfügung gestellten Hochrechnungsgewichten nicht exakt repliziert werden; die Unterschiede bewegen sich jedoch in der Regel innerhalb des jeweiligen 95%-Konfidenzintervalls.

Für die rückwirkende Anwendung des neuen Hochrechnungsverfahrens auf die Wellen 2000 bis 2013 wurden die sich über die Jahre verändernden Betriebsgrößenklassen der ursprünglichen Stichprobenmatrizen (vgl. Tabelle 4) auf eine einheitliche Abgrenzung von sechs Betriebsgrößenklassen (1–9, 10–19, 20–49, 50–249, 250–499, 500+) umkodiert. Die Wirtschaftszweige wurden von 2000 bis 2009 einheitlich auf 16 Ausprägungen der Klassifikation der Wirtschaftszweige, Ausgabe 2003, umkodiert. Von 2010 bis 2014 existieren bereits in der Stichprobenmatrix 23 Ausprägungen der Klassifikation der Wirtschaftszweige, Ausgabe 2008. Seit 2015 werden analog zur Stichprobenmatrix 24 Ausprägungen der Wirtschaftszweige, Ausgabe 2008, für die Hochrechnung verwendet.

## 2.5 Stichprobenziehung und Hochrechnung der Quartalsbefragung

Aus den teilnehmenden Betrieben an der schriftlichen Hauptbefragung im vierten Quartal werden jeweils zufällig Betriebe für eine kurze telefonische Nachbefragung in den drei Folgequartalen gezogen. Die angestrebte Zielgröße beträgt 9.000 Betriebe in jedem Quartal. Die Nettostichprobe der schriftlichen Hauptbefragung stellt damit die Bruttostichprobe für die telefonischen Nachbefragungen in den Folgequartalen dar.

Für jedes der Folgequartale wird ein eigenes Betriebsgewicht berechnet, welches sich aus dem Betriebsgewicht der Hauptbefragung im 4. Quartal des Vorjahres und dem Ergebnis einer weiteren Non-Response-Modellierung für das jeweilige Folgequartal zusammensetzt.

## 3 Wellenspezifische Besonderheiten

### 3.1 Befragungswelle 2022

#### 3.1.1 Verringerte Bereinigung der Daten

In vergangenen Wellen gab es für einige Variablen Plausibilitätskontrollen und Anpassungen, um Verzerrungen, etwa durch implausible Angaben, zu korrigieren. Ab Welle 2022 obliegen die Einschätzung und Entscheidung über die Plausibilität von Werten und die eventuelle Bereinigung der Daten zum Großteil den Nutzenden. Lediglich für die Hochrechnung relevante Variablen zur Beschäftigung und zum Personalbedarf werden weiterhin korrigiert.

#### 3.1.2 Veränderte Codierung von Missings

Ab Welle 2022 werden fehlende Werte (Missings) nach dem Grund des Fehlens codiert. Die Codierung ist Tabelle 6 zu entnehmen. Zu beachten ist, dass in Befragungswellen bis 2022 weiterhin nicht nach dem Grund des Fehlens differenziert wird.

**Tabelle 6: Codierung von Missings**

Codierung	Grund
.a	Keine Angabe
.b	Vorzeitiges Beenden des Fragebogens
.k	Trifft nicht zu (Filter)
.l	Trifft nicht zu (Designbedingt)
.u	Inkonsistenter Wert
.v	Unvollständige Angabe bei generierten Variablen
.w	Sonstige Missings
.z	Keine Teilnahme in Quartalsbefragung

### 3.2 Befragungswelle 2021

#### 3.2.1 Zusätzliche Quartalsbefragung

Ähnlich wie die COVID-19-Pandemie in der Befragungswelle 2019 (siehe Abschnitt 3.4.1), hat der russische Angriffskrieg gegen die Ukraine die Notwendigkeit aktueller und kurzfristiger Informationen zur Situation der Betriebe hervorgerufen. Vor diesem Hintergrund wurde die Befragungswelle 2021 um eine zusätzliche Nachbefragung im 4. Quartal 2022 verlängert. Eine Teilstichprobe der befragten Betriebe aus der Hauptbefragung vom 4. Quartal 2021 wurde erneut telefonisch befragt. Die Befragungswelle 2021 stellt somit eine Panelbefragung aus einer Messung vor dem Beginn des Ukraine-Kriegs sowie weiteren Messungen im Verlauf des Jahres 2022 nach dem Beginn des Kriegs dar. Die Nachbefragungen im 3. und 4. Quartal 2022 wurden um zusätzliche Fragen zum betrieblichen Umgang mit den Auswirkungen des Ukraine-Kriegs und zu arbeitsrelevanten Erfahrungen mit ukrainischen Geflüchteten erweitert. Die Befragungswelle 2021 wird inklusive der zusätzlichen Quartalsbefragung auch als „Ukraine-Panel“ bezeichnet.

Für das „Ukraine-Panel“ werden – analog zum regulären Gewichtungsverfahren – eigenständige Hochrechnungsgewichte bereitgestellt (siehe Abschnitt 4.2), die spezifisch gelabelt sind. Die Gewichte haben das Suffix `_ukraine` und tragen dementsprechend die Variablenamen `q421_greggew_ukraine`, `q122_greggew_ukraine`, `q222_greggew_ukraine`, `q322_greggew_ukraine` bzw. `q422_greggew_ukraine`. Im 4. Quartal 2021 und im 1., 2., und 3. Quartal 2022 entsprechen die Gewichte mit Suffix `_ukraine` den regulären Hochrechnungsgewichten ohne Suffix. Da es sich bei der Nachbefragung im 4. Quartal 2022 um eine außerordentliche Zusatzbefragung handelt, steht für dieses Quartal der Welle 2021 kein reguläres Gewicht, sondern ausschließlich das Gewicht mit Suffix `_ukraine` zur Verfügung. Da im 4. Quartal 2022 zeitgleich auch die reguläre Befragung der Welle 2022 im Feld war, sollte zur Identifikation der zusätzlichen Quartalsbefragung der Wellenidentifikator (`jahr`) genutzt werden. Dies ist in Tabelle 8 verdeutlicht.

Die vom IAB berichtete Auswertung zu offenen Stellen im 4. Quartal 2022 basiert auf der Befragungswelle 2022. Die zusätzliche Nachbefragung der Welle 2021 wurde hierfür nicht genutzt.

## 3.3 Befragungswelle 2020

### 3.3.1 Methodische Experimente

Die Welle 2020 wurde zum Test verschiedener Möglichkeiten zur Weiterentwicklung des Erhebungsdesigns mit Hilfe von methodischen Experimenten genutzt. Dabei wurden im 4. Quartal 2020 drei alternative Erhebungsmodi (sequentieller Web-First-Modus, reine Webbefragung, reine Papierbefragung) neben dem bisherigen Standarddesign (gleichzeitige Web- und Papiererhebung) eingesetzt. Mit 110.000 Betrieben wurde der Großteil der Bruttostichprobe mit dem bisherigen Standarddesign kontaktiert. Die alternativen Erhebungsmodi wurden jeweils für ca. 15.000 Betriebe der Bruttostichprobe eingesetzt. Ein weiteres methodisches Experiment wurde im Rahmen der alternativen Erhebungsmodi durchgeführt, um die Wirkung eines zusätzlichen Erinnerungsschreibens zu testen, indem ein solches Schreiben an eine Untergruppe von Betrieben versandt wurde. Dieses zusätzliche Erinnerungsschreiben wurde als „pre-due-date reminder“ konzipiert und stellte eine freundliche Erinnerung an die ablaufende erste Teilnahmefrist dar. Des Weiteren wurde im sequentiellen und im reinen Webbefragungsmodus – ausschließlich im Onlinefragebogen – eine alternative Fragenreihenfolge einzelner Fragen des letzten Falls einer Neueinstellung experimentell getestet. Die Frageformulierung wurde dabei nicht angepasst. Die Beobachtungen aus allen Experimentalgruppen sind Bestandteil der regulären Stichprobe, daher gibt es für alle Beobachtungen Hochrechnungsgewichte. Die Identifikation der Experimentalgruppen kann mit Hilfe der Variable `experiment2020` erfolgen.

Im Folgequartal, d.h. im 1. Quartal 2021, wurden die methodischen Experimente fortgesetzt und die Fortführung der selbst-administrierten Erhebungsmodi in den Nachbefragungen getestet. Dazu wurden die teilnehmenden Betriebe der vier Erhebungsmodi (gleichzeitige Web- und Papierbefragung, sequentieller Web-First-Modus, reine Web- und reine Papierbefragung) aus dem 4. Quartal 2020 jeweils in zwei Gruppen geteilt. Die eine Hälfte wurde, wie in den vorherigen Jahren üblich, telefonisch kontaktiert, und bei der anderen Hälfte wurde das im vierten Quartal genutzte Erhebungsdesign erneut eingesetzt. Ob die Betriebe telefonisch, schriftlich oder online an der

Erhebung teilgenommen haben, kann mit Hilfe der Variable `q121_modus` identifiziert werden. Im 2. und 3. Quartal 2020 wurde die Nachbefragung wieder ausschließlich telefonisch durchgeführt.

### 3.3.2 Erstmalige Abfrage des Einverständnisses zur Verlinkung von administrativen Daten

Seit der Befragungswelle 2020 wird das Einverständnis zur Verlinkung von Befragungsdaten und administrativen Daten explizit im Fragebogen von den Betrieben eingeholt. Für alle Betriebe, die diese Zustimmung nicht explizit erteilt haben, besteht keine Möglichkeit zur Datenverknüpfung mit administrativen Daten (für weitere Details zur Datenverknüpfung siehe Abschnitt 4.5).

### 3.3.3 Revision der Daten und Gewichte aufgrund von Fehlern in der Datenaufbereitung

Im Verlauf unserer Datenqualitätsprüfungen wurden Fehler beim Datenexport aus dem Onlinefragebogen bei einzelnen Variablen in der Erhebungswelle 2020 identifiziert. Da einige dieser Fehler auch hochrechnungsrelevanten Variablen betrafen, wurde im Zuge der Fehlerkorrekturen auch eine Neuberechnung der Hochrechnungsgewichte vorgenommen. Nach sorgfältiger Prüfung wurde beschlossen, auf Revisionen der betroffenen Publikationen zu verzichten, da die auf Basis der korrigierten Daten erzielten Ergebnisse sich nur marginal von den publizierten Ergebnissen unterscheiden. Alle nun im FDZ veröffentlichten Daten enthalten bereits die Korrekturen und die revidierte Hochrechnung. Frühere Wellen sind von dem Fehler beim Datenexport aus dem Onlinefragebogen nicht betroffen.

## 3.4 Befragungswelle 2019

### 3.4.1 Zusätzliche Quartalsbefragungen

Die COVID-19-Krise hat den Bedarf an kurzfristiger und valider Information zum Stellenangebot sowie zum betrieblichen Management im Allgemeinen deutlich gesteigert. Um diesem Informationsbedürfnis nachzukommen, wurde die Befragungswelle 2019 um zwei zusätzliche Quartalsbefragungen im 4. Quartal 2020 und im 1. Quartal 2021 verlängert. In diesen beiden Quartalen wurde eine Teilstichprobe der befragten Betriebe aus dem 4. Quartal 2019, wie in den regulären Quartalsbefragungen, erneut befragt. Die Befragungswelle 2019 stellt somit eine Panelbefragung aus einer Messung vor Ausbruch der Pandemie im 4. Quartal 2019 sowie weiteren Messungen im Verlauf der Pandemie dar. Beginnend mit dem 2. Quartal 2020 wurden alle Quartalsbefragungen um zusätzliche Fragen zum betrieblichen Umgang mit der COVID-19-Pandemie erweitert. Die Befragungswelle 2019, um die beiden zusätzlichen Quartalsbefragungen erweitert, wird auch als „COVID-Panel“ bezeichnet.

Ähnlich wie beim „Ukraine-Panel“ werden für das COVID-Panel – analog zum regulären Gewichtungsverfahren – eigenständige Hochrechnungsgewichte bereitgestellt (siehe Abschnitt 4.2). Die Hochrechnungsgewichte für das COVID-Panel sind spezifisch gelabelt. Die Gewichte haben das Suffix `_covid` und tragen dementsprechend die Variablennamen `q419_greggew_covid`, `q120_greggew_covid`, `q220_greggew_covid`, `q320_greggew_covid`, `q420_greggew_covid` bzw. `q121_greggew_covid`. Im 1., 2., und 3. Quartal 2020 entsprechen die Gewichte mit Suffix `_covid` den regulären Hochrechnungsgewichten ohne Suffix. Da es sich bei den Quartalsbefragungen im 4. Quartal 2020 und im 1. Quartal 2021 um außerordentliche Zusatzbefragungen handelt, stehen für diese beiden

Quartale der Welle 2019 keine regulären Gewichte, sondern ausschließlich Gewichte mit Suffix `_covid` zur Verfügung. Da im 4. Quartal 2020 und im 1. Quartal 2021 zeitgleich auch die reguläre Befragung der Welle 2020 im Feld war, sollte zur Identifikation der zusätzlichen Quartalsbefragungen der Wellenidentifikator (`jahr`) genutzt werden. Dies ist in Tabelle 8 verdeutlicht.

Die vom IAB berichteten Auswertungen zu offenen Stellen im 4. Quartal 2020 und im 1. Quartal 2021 basieren auf der Befragungswelle 2020. Die zusätzlichen Nachbefragungen der Welle 2019 wurden hierfür nicht genutzt.

### 3.4.2 Experimental-Daten der Befragungswelle 2019

Im 4. Quartal 2019 fand neben der regulären Hauptbefragung ein Survey-Experiment (N=3.883) statt. In dem Experiment wurden die Effekte von zusätzlichen erklärenden Hinweisen zu Fragen in Modul 3 auf das Antwortverhalten von Betrieben untersucht (für nähere Informationen zum Experimentaldesign und zu den Ergebnissen siehe Kufner/Sakshaug/Zins, 2021). Für das Experiment wurde eine separate Stichprobe gezogen und - anders als in der Hauptbefragung - eine reine Webbefragung durchgeführt. Die reguläre Hauptbefragung im 4. Quartal und alle nachfolgenden Quartalsbefragungen waren nicht Teil des Experiments und wurden vom Experiment nicht beeinträchtigt oder beeinflusst. Die Haupterhebung steht für sich und bildet die zentrale Grundlage für inhaltliche Analysen.

Die Daten dieser Experimentalbefragung sind Teil der Jahreswelle 2019, die an das FDZ geliefert werden. Diese Beobachtungen können über die Indikatorvariable `experiment2019` eindeutig identifiziert werden. Da für diese Experimentalbeobachtungen keine Hochrechnungsgewichte zur Verfügung stehen, ist eine Nutzung für gewichtete Analysen nicht möglich (`q419_greggew == .`). Die Nutzung der Experimentaldaten kann allerdings für explorative und ungewichtete Zwecke einen Mehrwert bringen, etwa für die Untersuchung von variierendem Antwortverhalten von Betrieben, und steht Nutzerinnen und Nutzern im FDZ daher zur Verfügung. Es sollte beachtet werden, dass mit Ausnahme der im Rahmen des Survey-Experiments durchgeführten Analysen keine zusätzliche Qualitätsprüfung der Experimentaldaten stattfand. Auch wurde nicht untersucht, inwiefern Abweichungen im Erhebungsdesign einen Einfluss auf mögliche Selektionseffekte oder Messeffekte und damit auf die Vergleichbarkeit mit der Haupterhebung haben. Eine Verlinkung mit administrativen Daten ist nicht möglich (siehe Abschnitt 4.5). Daher empfehlen wir standardmäßig den Ausschluss dieser Experimentaldaten (`drop if experiment2019 == 1`) für Analysen inhaltlicher Art.

### 3.4.3 Revision aufgrund von abweichendem Verhalten von Interviewerinnen bzw. Interviewern

Im 2. Quartal in 2021 wurden bei routinemäßigen Prüfungen der Datenqualität auffällige Datenmuster und Zeitreihen identifiziert. Im Rahmen einer ausführlichen Untersuchung konnte festgestellt werden, dass durch eine Interviewerin bzw. einen Interviewer systematisch Interviews manipuliert wurden. Diese Person hat bei einer Filterfrage gehäuft angegeben, dass die befragten Betriebe über keine Vakanzen verfügen. Durch diese verfälschte Angabe mussten keine weiteren Fragen zur Struktur der Vakanzen erhoben werden, wodurch die Befragungsdauer stark verkürzt werden konnte. Im Laufe der Untersuchung konnte noch eine weitere Interviewerin beziehungsweise ein weiterer Interviewer identifiziert werden, die/der im 1. Quartal 2020

abweichendes Verhalten gezeigt hat. Alle Interviews der beiden Personen wurden auch retrospektiv aus der gesamten Befragungswelle 2019 ausgeschlossen. In den früheren Befragungswellen waren keine Interviews der beiden Personen enthalten.

Im Rahmen einer Revision wurden neue Hochrechnungsgewichte berechnet. Zudem wurden alle IAB-Publikationen sowie die auf der Homepage veröffentlichten Zahlen revidiert, die Beobachtungen der Interviewfälschung enthielten. In den nun im FDZ veröffentlichten Daten sind die Beobachtungen der beiden Interviewerinnen beziehungsweise Interviewer ausgeschlossen und die revidierten Hochrechnungsgewichte enthalten. Zuvor im FDZ veröffentlichte Daten sind von der Interviewfälschung nicht betroffen. Weitere Informationen zur Entdeckung des abweichenden Verhaltens der Interviewerinnen bzw. Interviewer, der Revision der Daten sowie zur verbesserten Qualitätssicherung der Telefoninterviews sind im FDZ-Methodenreport von Bossler et al. (2022) festgehalten.

## 3.5 Befragungswelle 2016

### 3.5.1 Vignettenstudie zur Wirkung der Mindestlohnhöhe

Im 2. Quartal 2017 wurde die Befragung für eine Vignettenstudie genutzt. Dabei wurde die betriebliche Reaktion auf einen Anstieg oder ein Absinken des gesetzlichen Mindestlohns untersucht. Die beiden dazugehörigen Fragebogenvarianten sind in den jeweiligen Codeplänen dokumentiert. Die entsprechenden Daten sind Teil des Datensatzes und die über die Fragebögen variierende Höhe des genannten fiktiven Mindestlohns kann der Variable `q217_mindestlohn` entnommen werden.

# 4 Verwendungshinweis für Datennutzerinnen und Datennutzer

## 4.1 Struktur des Datensatzes

Die Daten jeder Befragungswelle sind in jeweils einer Datei im STATA-Format abgelegt. Sie enthält die Informationen des 1., 2. und 3. Moduls sowie ab 2005 die Angaben aus den telefonischen Folgebefragungen in den Folgequartalen. Die Daten werden also nicht separat für Kalenderjahre, sondern das 4. Quartal eines Erhebungsjahrs wird zusammen mit den – im Regelfall drei – Folgequartalen bereitgestellt. Diese Unterteilung ist deshalb sinnvoll, weil sich die teilnehmenden Betriebe an den Folgebefragungen aus den antwortenden Betrieben an der Hauptbefragung im 4. Quartal des Ursprungsjahrs rekrutieren. So entsteht in jeder Befragungswelle ein unbalanciertes Panel mit vier Beobachtungszeitpunkten bzw. sechs und fünf Beobachtungszeitpunkten in den Befragungswellen 2019 und 2021.

Bis auf wenige Ausnahmen (Tabelle 7) enthalten alle Variablennamen jeweils ein Präfix aus der Nummer des jeweiligen Quartals sowie der Jahreszahl minus 2000. Die Variable mit dem Betriebsgewicht `greggew` (vgl. Abschnitt 3.2) heißt im 3. Quartal 2012 demzufolge `q312_greggew`. Tabelle 8 zeigt die Aufteilung der Quartale auf die Datendateien sowie die jeweils zugehörigen Präfixe der Variablennamen. Die Bedeutung der einzelnen Variablennamen und deren Ausprägungen sind der Variablen-tabelle sowie den Codeplänen der jeweiligen Befragungswelle auf der Webseite des FDZ unter [„Betriebs- und Unternehmensdaten: IAB-Stellenerhebung – Version 0022 v1“](#) zu entnehmen. Aus Platzgründen wird in den Codeplänen auf das jeweilige Präfix verzichtet.

**Tabelle 7: Technische-Variablen ohne Variablennamen-Präfixe**

Variablenname	Bedeutung
<code>key</code>	Pseudonymisierter Betriebsidentifikator
<code>jahr</code>	Befragungswelle
<code>hf</code>	Modul 1 (Hauptfragebogen) wurde beantwortet
<code>sf</code>	Modul 2 (Sonderfragebogen) wurde beantwortet
<code>zf</code>	Modul 3 (Zusatzfragebogen) wurde beantwortet
<code>fragebogen</code>	Kombination aus <code>hf</code> , <code>zf</code> und <code>sf</code>
<code>gebiet04</code>	Region (West-/Ostdeutschland mit Berlin)
<code>bl</code> , <code>bl_agg</code>	Bundesland (NUTS 1), Bundesland (aggregiert)
<code>wz03_16/wz03_28</code>	2000–2009: 16/28 Wirtschaftszweige auf Basis der Klassifikation 2003
<code>wz08_23</code>	seit 2010: 23 Wirtschaftszweige auf Basis der Klassifikation 2008
<code>wz08_24</code>	seit 2015: 24 Wirtschaftszweige auf Basis der Klassifikation 2008
<code>wz*_3st</code>	2000-2004: Original-Wirtschaftszweigklassifikationen aus der BA-Adressdatei (3-Steller)
<code>wz*_5st</code>	seit 2005: Original-Wirtschaftszweigklassifikationen aus der BA-Adressdatei (5-Steller) <sup>1)</sup>
<code>kreis</code>	seit 2007: Kreisschlüssel (NUTS 3) <sup>1)</sup>
<code>rbezirk</code>	seit 2005: Regierungsbezirk (NUTS 2) <sup>1)</sup>
<code>gesamtausfueller</code>	seit 2013: Es antwortete die übergeordnete Einheit, nicht der angeschriebene Betrieb
<code>erh_form</code>	Erhebungsform (schriftliche oder mündliche Befragung)
<code>antwort</code>	seit 2011: Antwortart (in Papierform oder online)

**1)** Sensibles Merkmal: Wird externen Nutzerinnen und Nutzern nur auf Antrag zur Verfügung gestellt.

**Tabelle 8: Dateinamen und Variablennamen-Präfixe**

Dateiname	Quartal	Variablennamen-Präfix	Welle (jahr)
iabse_00	4. Quartal 2000	q40_	2000
iabse_01	4. Quartal 2001	q41_	2001
iabse_02	4. Quartal 2002	q42_	2002
iabse_03	4. Quartal 2003	q43_	2003
iabse_04	4. Quartal 2004	q44_	2004
iabse_05	4. Quartal 2005 (Schriftliche Haupterhebung)	q45_	2004
	1. Quartal 2006 (Telefonische Folgebefragung)	q16_	2004
	2. Quartal 2006 (Telefonische Folgebefragung)	q26_	2004
	3. Quartal 2006 (Telefonische Folgebefragung)	q36_	2004
	(...)		
iabse_19	4. Quartal 2019 (Schriftliche Haupterhebung)	q419_	2019
	1. Quartal 2020 (Telefonische Folgebefragung)	q120_	2019
	2. Quartal 2020 (Telefonische Folgebefragung)	q220_	2019
	3. Quartal 2020 (Telefonische Folgebefragung)	q320_	2019
	4. Quartal 2020 (Telefonische Folgebefragung – Zusatz)	q420_	2019
	1. Quartal 2021 (Telefonische Folgebefragung – Zusatz)	q121_	2019
iabse_20	4. Quartal 2020 (Schriftliche Hauptbefragung)	q420_	2020
	1. Quartal 2021 (Telefonische Folgebefragung)	q121_	2020
	2. Quartal 2021 (Telefonische Folgebefragung)	q221_	2020
	3. Quartal 2021 (Telefonische Folgebefragung)	q321_	2020
iabse_21	4. Quartal 2021 (Schriftliche Haupterhebung)	q421_	2021
	1. Quartal 2022 (Telefonische Folgebefragung)	q122_	2021
	2. Quartal 2022 (Telefonische Folgebefragung)	q222_	2021
	3. Quartal 2022 (Telefonische Folgebefragung)	q322_	2021
	4. Quartal 2022 (Telefonische Folgebefragung – Zusatz)	q422_	2021
iabse_22	4. Quartal 2022 (Schriftliche Haupterhebung)	q422_	2022
	1. Quartal 2023 (Telefonische Folgebefragung)	q123_	2022
	2. Quartal 2023 (Telefonische Folgebefragung)	q223_	2022
	3. Quartal 2023 (Telefonische Folgebefragung)	q323_	2022

## 4.2 Verwendung der Gewichte

Bei deskriptiven Auswertungen sind die entsprechenden Hochrechnungsgewichte zu verwenden, um von der realisierten Stichprobe auf die zugrundeliegende Grundgesamtheit schließen zu können.<sup>1</sup> Bei Regressionsanalysen wird üblicherweise auf die Verwendung von Hochrechnungsgewichten verzichtet, wenn sämtliche Schichtungsvariablen (vgl. Tabelle 4) in die Schätzgleichung mit einfließen (vgl. Bossler/Geis/Stegmaier, 2018; Winship/Radbill, 1994).

Für Variablen des Moduls 1 sowie des Moduls 2, soweit sie sich nicht auf den zuletzt eingestellten Ein-Euro-Jobber beziehen, ist das Betriebsgewicht `greggew` zu verwenden. Für Fragen des Moduls 2 zur letzten Neueinstellung ist das Neueinstellungsgewicht (Betriebsgewicht multipliziert mit der Zahl der Neueinstellungen, seit 2004 Zahl der sozialversicherungspflichtigen

<sup>1</sup> Bei Verwendung von Hochrechnungs- oder Gewichtungsfaktoren in deskriptiven Analysen sind auch immer analoge ungewichtete Ergebnisse auszuweisen. Die gewichteten und die dazugehörigen ungewichteten Ergebnistabellen müssen dabei immer direkt untereinander stehen, weil dies die Datenschutzprüfung erleichtert und somit beschleunigt.

Neueinstellungen), für Fragen zum letzten Suchabbruch das Abbruchgewicht (Betriebsgewicht multipliziert mit der Zahl der Suchabbrüche) zu verwenden. Tabelle 9 fasst die zu verwendenden Gewichte zusammen. Für die telefonischen Nachbefragungen in den Folgequartalen steht für jedes Quartal ein eigenes Betriebsgewicht `greggew` zur Verfügung, welches durch das quartalspezifische Präfix vom Betriebsgewicht der Hauptbefragung unterschieden wird. Wie in den Abschnitten 3.1.1 und 3.3.1 beschrieben, stehen für das Ukraine-Panel und das COVID-Panel eigene Hochrechnungsgewichte zur Verfügung.

Bei diesen Gewichten handelt es sich dem Grunde nach um Stichprobengewichte (sampling weights), das heißt, in STATA ist nach Möglichkeit die Gewichtsart `pweights` anzugeben und solche Auswertungsbefehle zu verwenden, die die Angabe dieser Gewichtsart unterstützen (`svyset`, `svy:tabulate`, `regress`). Für die Befehle `tabulate` (ohne `svy:`) und `summarize` kann hilfsweise auch `iweight` verwendet werden; allerdings können dann keine Standardfehler, etwa bei  $\chi^2$ -Tests, interpretiert werden. Soll bei deskriptiven Auswertungen auf statistische Signifikanz zwischen Teilgruppen getestet werden, so darf dies wegen fehlender Gewichtungsmöglichkeit nicht mit dem `ttest`-Befehl geschehen; stattdessen ist hilfsweise eine univariate Regression mit der Gruppenvariable als Regressor unter Angabe der `pweights`-Gewichtung durchzuführen und die Signifikanz des Unterschiedes am P-Wert der Gruppenvariable abzulesen (vgl. STATA-Beispiel 1).

**Tabelle 9: Zu verwendendes Hochrechnungsgewicht nach Modul und Welle**

Module	Zu verwendendes Gewicht	
Modul 1, telefonische Folgebefragungen		<code>greggew</code>
Telefonische Folgebefragung – COVID-Panel		<code>greggew_covid</code>
Telefonische Folgebefragung – Ukraine-Panel		<code>greggew_ukraine</code>
Der letzte Fall einer Einstellung	2000-2003: ab 2004:	<code>greggew*f10</code> <code>greggew*f11_04</code>
Der letzte Fall einer abgebrochenen Personalsuche		<code>greggew*f422</code>
Modul 2 (außer Fragen zum letzten Ein-Euro-Jobber)		<code>greggew</code>
Angaben zur zuletzt in einem Ein-Euro-Job eingestellten Person		<code>greggew*p29</code>

### 4.3 Fragebogenteile mit möglichen Mehrfachantworten

An vielen Stellen im Fragebogen sind Mehrfachantworten möglich. Da *à priori* nicht klar ist, ob ein nicht angekreuztes Feld als „nein“ oder als fehlender Wert zu interpretieren ist, sind angekreuzte Felder im Datensatz mit dem Wert 1, nicht angekreuzte Felder als fehlender Wert (.) codiert. In der Praxis wird bei Auswertungen üblicherweise so vorgegangen, sämtliche nicht angekreuzte Felder dann in 0 („nein“) umzucodieren, wenn mindestens ein Feld innerhalb der Item-Batterie angekreuzt wurde (vgl. STATA-Beispiel 2).

Bei einigen Fragen können Antwortende „Sonstiges“ ankreuzen und nähere Angaben zu Sonstigem in einem freien Textfeld machen. Der Inhalt dieser freien Textfelder wird zur Wahrung des Datenschutzes nicht an externe Nutzerinnen und Nutzer zur Verfügung dargestellt, da über die Formulierung Rückschlüsse auf einzelne Betriebe möglich sein können. Häufiger vorkommende

Antworten werden jedoch kategorisiert in einer entsprechenden Variablen mit der Endung *u* (Umbuchung) zur Verfügung gestellt, deren Ausprägungen den Wertelabels im (Test-)Datensatz zu entnehmen sind.

Sofern eine Antwort im freien Textfeld bereits als angebotene Antwortmöglichkeit im Fragebogen existiert, wird sie entsprechend umbucht. Bei Fragen mit Antwortmöglichkeit „Sonstiges“ existiert jede Antwortvariable deshalb zweimal, einmal mit dem ursprünglichen Inhalt und einmal inklusive dieser Umbuchungen. Die Variablen „inklusive Umbuchungen“ erhalten das Suffix *iu*. Wird beispielsweise beim letzten Fall einer Neueinstellung als beschrittener Suchweg nur „sonstiger Suchweg“ angekreuzt und im freien Textfeld die Antwort „wurde uns empfohlen“ eingetragen, so enthalten die Variablen die in Tabelle 10 dargestellten Werte.

**Tabelle 10: Beispiel für die Umbuchung einer freien Textfeld-Antwort**

Variablenname	Bezeichnung	Wert
zf97	Suchweg: über eigene Mitarbeiter/persönliche Kontakte	. (wurde nicht angekreuzt)
zf9n	Sonstiger Suchweg	1 (wurde angekreuzt)
zf9txt*	Sonstiger Suchweg: freies Textfeld	„wurde uns empfohlen“
zf9u	Umbuchung in...	11 (persönliche Kontakte)
zf97iu	Suchweg: über eigene Mitarbeiter/persönliche Kontakte (inklusive Umbuchungen)	1 (wurde angekreuzt oder durch Umbuchungen generiert)
zf9niu	Sonstiger Suchweg	0 (wurde nicht angekreuzt oder wurde umbucht)

\* nicht im FDZ-Datensatz enthalten

## 4.4 Klassifikationen von Wirtschaftszweigen und Berufen

Der Wirtschaftszweig eines Betriebes wird nicht abgefragt, sondern im Rahmen der Ziehung der Bruttostichprobe aus der Adressendatei der Bundesagentur für Arbeit entnommen. Wie Tabelle 4 zeigt, erfolgte dies bis 2003 anhand der Wirtschaftszweig-Klassifikation 1973 (WZ73), von 2004 bis 2009 anhand der Klassifikation WZ2003 und ab 2010 anhand der Klassifikation WZ2008. Im Datensatz wird hingegen von 2000 bis 2009 durchgängig die Klassifikation WZ2003 sowohl mit 16 als auch 28 Zweigen zur Verfügung gestellt (die Wellen 2000 bis 2003 entsprechend umcodiert), ab 2010 die Klassifikation WZ2008. Die in der Adressendatei der Bundesagentur für Arbeit vorliegenden Original 5-Steller der Klassifikationen 2003 und 2008 sind ab dem Jahr 2005 verfügbar und werden als sensible Merkmale zur Wahrung des Datenschutzes externen Nutzerinnen und Nutzern nur auf Antrag zur Verfügung gestellt.

An mehreren Stellen wird im Fragebogen die Angabe von Berufen verlangt. Da die aufwändige Berufscodierung den Antwortenden nicht zuzumuten ist, wird stattdessen nach der Berufsbezeichnung im Klartext gefragt. Diese Freitextantwort wird zur Wahrung des Datenschutzes nicht zur Verfügung gestellt. Die Berufsbezeichnungen werden vom Befragungsinstitut in die entsprechenden Codes zweier Klassifikationen codiert, wie Tabelle 11 zeigt. In einigen Fällen waren die Angaben zum beruflichen Tätigkeitsfeld zu ungenau, um eine passende Codierung zu finden; hier mussten außerhalb der offiziellen Codierung eigene Codes vergeben werden, die in Tabelle 12 aufgelistet sind.

**Tabelle 11: Berufsklassifikationen nach Befragungswelle**

Welle	Berufscodierung	Variablen
<b>Klassifikation 1 (deutsches System)</b>		
2000-2011	Klassifikation der Berufe, Ausgabe 1992 (KldB92), 3-Steller	c[1-5], bkz[1-3], f418_c[1-5], f423_c[1-5], f424c_[1-5], code, codeab, c430-c432
seit 2012	Klassifikation der Berufe, Ausgabe 2010 – Erste Fassung (KldB2010), 5-Steller	kb10_*
<b>Klassifikation 2 (internationales System)</b>		
2000-2004	Keine	
2005-2011	International Standard Classification of Occupations, 1988 Edition (ISCO-88)	isco*, iscl-3, f418isc*, f423isc*, f424isc*, i430-i432
seit 2012	International Standard Classification of Occupations, 2008 Edition (ISCO-08)	isco08_*

**Tabelle 12: Ergänzung der KldB2010 für Angaben ohne nähere Tätigkeitsbeschreibung (o.n.T)**

Kategorie	Code	Bezeichnung
Technische Mitarbeiter o.n.T.	99001	Produktionsarbeitskraft, Produktionshelfer, Gehilfe
	99002	Werksfachkraft, technische Fachkraft, Geselle
	99003	Techniker, Vorarbeiter, Meister
	99004	Ingenieur, Diplomingenieur
	99005	Technische Mitarbeiter o.n.A.
Kaufmännische Mitarbeiter o.n.T.	99101	Kaufm. Gehilfen
	99102	Kaufmann, Disponent
	99103	Betriebswirt
	99104	Diplomkaufmann
	99105	Kaufmännische Mitarbeiter o.n.A.
Dienstleistungskräfte o.n.T.	99201	Servicepersonal, Supporter, Terminleger
	99202	Servicefachkraft, Berater, Inspektor, nichtwissenschaftl. Mitarbeiter
	99203	Fachberater, Personaldienstleister
	99204	Senior Consultant
	99205	Dienstleistungskräfte o.n.A.
Führungskräfte o.n.T.	99301	(Anforderungsniveau 1 kommt bei Führungskräften nicht vor)
	99302	(Anforderungsniveau 2 kommt bei Führungskräften nicht vor)
	99303	Teamleiter, Leitungspersonal, Standortchef, Abteilungsleiter
	99304	Direktor, Manager, Geschäftsführer
	99305	Führungskräfte o.n.A.
IT-Kräfte o.n.T.	99401	IT-Mitarbeiter (Anforderungsniveau 1)
	99402	IT-Fachkraft
	99403	IT-Spezialist
	99404	IT-Experte
	99405	IT-Beruf, IT-Mitarbeiter o.n.A.
Lehrkräfte o.n.T.	99501	Schulassistent
	99502	Lehrkraft
	99503	Lehrer
	99504	Dozent, Lehrkraft mit Hochschulabschluss
	99505	Lehrer o.n.A.
Arbeitskräfte in Ausbildung o.n.T.	99611	Azubi
	99621	Praktikant / Volontär
	99691	Sonstige Arbeitskräfte in Ausbildung
Arbeitskräfte o.n.T.	99901	Helfer, Aushilfe, Anlernkraft, ungelernete Mitarbeiter
	99902	Fachkraft, Arbeiter, Assistent
	99903	Spezialist
	99904	Experte
	99905	Arbeitskräfte o.n.A.

## 4.5 Verlinkung zu anderen Daten

### 4.5.1 BHP-Verlinkung

Ab dem 4. Quartal 2010 steht für alle verknüpfbaren Betriebe ergänzend das komplette Variablenspektrum des Betriebs-Historik-Panels (BHP), [Version BHP 7523 v1](#), zur Verfügung. Eine Ausnahme stellt das BHP-Erweiterungsmodul zu Insolvenzen dar, das nicht mit der IAB-Stellenerhebung verknüpft werden kann. Ab der Version IAB-SE 0022 v1 wird standardmäßig eine Basisbetriebsdatei bereitgestellt. Sensible Merkmale sowie weitere Betriebsmerkmale in Variablenblöcken und Erweiterungsmodule müssen gesondert beantragt und begründet werden.

In den Befragungswellen 2010-2019 haben die Betriebe durch die Teilnahme an der Befragung automatisch ihr Einverständnis zur Verlinkung der Befragungsdaten mit administrativen Daten der Bundesagentur für Arbeit gegeben. Darauf wurden die Betriebe in den Erklärungen zum Datenschutz hingewiesen. Beginnend mit der Befragungswelle 2020 wird das Einverständnis zur Verlinkung von Befragungsdaten und administrativen Daten explizit im Fragebogen abgefragt. Basierend auf wissenschaftlicher Forschung zur Erzielung einer hohen Zustimmungsbereitschaft (siehe Sakshaug/Vicari, 2018) wird diese Frage am Anfang des Fragebogens gestellt, um das Analysepotential der verlinkten Daten möglichst groß zu halten. Für alle Betriebe ohne explizite Zustimmung besteht jedoch keine Möglichkeit zur Datenverknüpfung mit administrativen Daten. Dementsprechend wird für diese Betriebe auch keine Verlinkung der IAB-Stellenerhebung mit dem Betriebs-Historik-Panel zur Verfügung gestellt.

**Tabelle 13: Dateinamen der zuspieldbaren BHP-Datensätze**

Dateiname	Inhalt
iabse_0022_v1_bhp_7523_m06_v1_Basis	BHP-Basisbetriebsdatei
iabse_0022_v1_bhp_7523_m06_v1_1975 ... iabse_0022_v1_bhp_7523_m06_v1_2023	BHP-Variablenblöcke in „Jahresscheiben“
iabse_0022_v1_bhp_7523_m06_inflow_v1 iabse_0022_v1_bhp_7523_m06_outflow_1	BHP-Erweiterung „Beschäftigtenströme“
iabse_0022_v1_bhp_7523_m06_entry_v1 iabse_0022_v1_bhp_7523_m06_exit_v1	BHP-Erweiterung „Betriebsdynamik“

Bei der Verwendung ist folgendes zu beachten:

1. Die Verlinkung erfolgt ausschließlich mittels der (über alle Jahre eindeutigen) Variable `key`. Die im BHP ursprünglich enthaltene Variable `betnr` ist im BHP-Erweiterungsdatensatz der IAB-Stellenerhebung nicht enthalten.
2. Für eine geringe Anzahl von antwortenden Betrieben war eine Verlinkung nicht möglich, weil sie für mehr als eine Betriebsstätte geantwortet haben; sie sind im BHP-Datensatz nicht enthalten. Weiterhin sind bei einigen `keys` aus dem gleichen Grund die zugespielten BHP-Daten nicht sinnvoll interpretierbar; bei ihnen hat die Variable `gesamtausfueller` den Wert 1.
3. Die Datenschutzerklärung, die an alle zur Teilnahme aufgeforderten Betriebe versendet wird, war in den Wellen 2010-2019 so abgefasst, dass mit der Teilnahme an der Befragung die

Zustimmung zur Datenverlinkung erteilt wird. Mit Ausnahme der in Nummer 2 genannten Betriebe ist somit für sämtliche Betriebe dieser Befragungswellen eine Verlinkung möglich. Seit der Welle 2020 wird die explizite Zustimmung der Betriebe im Fragebogen eingeholt.

4. In den BHP-Datensätzen bezeichnet die Variable `jahr` den Zeitpunkt, auf den sich die jeweiligen BHP-Variablen beziehen. Die Befragungswelle, in welcher der jeweilige Betrieb befragt wurde, ist anhand der zusätzlichen Variable `SE_jahr` zu erkennen.
5. Es ist vereinzelt mit Abweichungen, etwa bei der Betriebsgröße zwischen Befragungsdaten und administrativen Daten, zu rechnen, etwa wenn die Angaben sich auf eine größere oder kleinere Einheit (z. B. Unternehmen versus Betriebsstätte) beziehen als in der für die Adressziehung verwendete Betriebsdatei der Bundesagentur für Arbeit vermerkt ist.
6. Sofern die sensiblen Merkmale `kreis` (IAB-SE) / `ao_kreis` (BHP) oder `wz*_5st` (IAB-SE) / `w*_5` (BHP) bestellt wurden, ist zu beachten, dass sich die entsprechenden Werte in der IAB-Stellenerhebung auf den Zeitpunkt der Stichprobenziehung (vgl. Tabelle 5) beziehen, also auf das Vorjahr der Feldphase. Die Kreisnummern in der IAB-Stellenerhebung sind zum Ziehungszeitpunkt gültig, während die Kreisnummern in den BHP-Daten den Gebietsstand zum 31.12.2023 widerspiegeln.

Folgende Basismerkmale werden in der BHP-Basisbetriebsdatei standardmäßig bereitgestellt:

#### Basismerkmale

- Ordnungsnummer (key)
- Welle der IAB-Stellenerhebung (`SE_jahr`)
- Welle des BHPs (`jahr`)
- Klassifikation der Wirtschaftszweige 1973 – 3-Steller (`w73_3`)
- Klassifikation der Wirtschaftszweige 1993 – 3-Steller (`w93_3`)
- Klassifikation der Wirtschaftszweige 2003 – 3-Steller (`w03_3`)
- Klassifikation der Wirtschaftszweige 2008 – 3-Steller (`w08_3`)
- Zeitkonsistente Klassifikationen der Wirtschaftszweige, 3-Steller (`w73_3_gen`, `w93_3_gen`, `w08_3_gen`) und Information zur Art der Vervollständigung (`group_w73_3`, `group_w93_3`, `group_w08_3`)
- Jahr des ersten Auftretens der Betriebsnummer (`grd_jahr`)
- Jahr des letzten Auftretens der Betriebsnummer (`lzt_jahr`)
- Anzahl der Beschäftigten insgesamt (`az_ges`)
- Anzahl der Vollzeitbeschäftigten (`az_vz`)
- Anzahl der geringfügig Beschäftigten (`az_gf`)
- Mittelwert imp. Bruttotagesentgelt Vollzeitbeschäftigte (`te_imp_mw`)
- Bundesland des Arbeitsorts (`ao_bula`)

Folgende sensible Basismerkmale können auf begründeten Antrag geliefert werden:

#### Sensible Basismerkmale

- Klassifikation der Wirtschaftszweige 1993 – 5-Steller (`w93_5`),
- Klassifikation der Wirtschaftszweige 2003 – 5-Steller (`w03_5`),
- Klassifikation der Wirtschaftszweige 2008 – 5-Steller (`w08_5`),

- Kreis des Arbeitsorts (ao\_kreis)

Des Weiteren können auf begründeten Antrag hin weitere Betriebsmerkmale in Variablenblöcken in „Jahresscheiben“ bereitgestellt werden. Dies gilt auch für die Erweiterungsmodule „Beschäftigungsströme“ (Eintritte/Austritte) und „Betriebsdynamik“ (Gründungen/Schließungen). Die Begründung der Notwendigkeit für das Erreichen des Forschungsziels muss dabei separat für jeden Variablenblock bzw. jedes Erweiterungsmodul erfolgen. Die Variablenblöcke bzw. Erweiterungsmodule enthalten folgende Merkmale:

#### BHP-Variablenblöcke in „Jahresscheiben“

- Struktur der Beschäftigten allgemein (az\_f az\_reg az\_azubi az\_atz az\_tz az\_f\_vz az\_f\_tz az\_reg\_vz)
- Struktur der Beschäftigten nach Schul- und Berufsbildung (az\_gq az\_mq az\_hq az\_gq\_vz az\_mq\_vz az\_hq\_vz)
- Altersstruktur der Beschäftigten (az\_15\_19 az\_15\_19\_vz az\_20\_24 az\_20\_24\_vz az\_25\_29 az\_25\_29\_vz az\_30\_34 az\_30\_34\_vz az\_35\_39 az\_35\_39\_vz az\_40\_44 az\_40\_44\_vz az\_45\_49 az\_45\_49\_vz az\_50\_54 az\_50\_54\_vz az\_55\_59 az\_55\_59\_vz az\_60\_64 az\_60\_64\_vz az\_ab65 az\_ab65\_vz alter\_mw alter\_mw\_vz)
- Aktivitäten Forschung und Entwicklung (az\_ingnat)
- Anzahl atypisch Beschäftigter (az\_leih az\_bfr)
- Struktur der Beschäftigten nach Staatsangehörigkeit (az\_d az\_d\_vz az\_eu)
- Struktur der Beschäftigten nach Blossfeld Berufsqualifikation (az\_bf\_agr az\_bf\_emb az\_bf\_edi az\_bf\_evb az\_bf\_qmb az\_bf\_qdi az\_bf\_qvb az\_bf\_tec az\_bf\_semi az\_bf\_ing az\_bf\_prof az\_bf\_man)
- Struktur der Beschäftigten nach Tätigkeitsniveau (az\_niv1 az\_niv2 az\_niv3 az\_niv4)
- Gehaltsstruktur der Vollzeitbeschäftigten (az\_zens te\_imp\_med te\_imp\_p25 te\_imp\_p75 te\_imp\_mw\_f te\_imp\_med\_f te\_imp\_med\_m te\_imp\_mw\_gq te\_imp\_med\_gq te\_imp\_mw\_mq te\_imp\_med\_mq te\_imp\_mw\_hq te\_imp\_med\_hq te\_imp\_med\_uq te\_imp\_mw\_d te\_imp\_med\_d te\_imp\_med\_a te\_imp\_sum)

#### BHP-Erweiterungsmodule

- Beschäftigtenströme (ein\_ges ein\_f ein\_reg ein\_gf ein\_azubi ein\_vz ein\_tz ein\_f\_vz ein\_f\_tz ein\_reg\_vz ein\_wdr ein\_wdr\_f ein\_bw ein\_bw\_f ein\_bf\_agr ein\_bf\_emb ein\_bf\_edi ein\_bf\_evb ein\_bf\_qmb ein\_bf\_qdi ein\_bf\_qvb ein\_bf\_tec ein\_bf\_semi ein\_bf\_ing ein\_bf\_prof ein\_bf\_man ein\_15\_19 ein\_20\_24 ein\_25\_29 ein\_30\_34 ein\_35\_39 ein\_40\_44 ein\_45\_49 ein\_50\_54 ein\_55\_59 ein\_60\_64 ein\_ab65 aus\_ges aus\_f aus\_reg aus\_gf aus\_azubi aus\_vz aus\_tz aus\_f\_vz aus\_f\_tz aus\_reg\_vz aus\_temp aus\_temp\_f aus\_bw aus\_bw\_f aus\_bf\_agr aus\_bf\_emb aus\_bf\_edi aus\_bf\_evb aus\_bf\_qmb aus\_bf\_qdi aus\_bf\_qvb aus\_bf\_tec aus\_bf\_semi aus\_bf\_ing aus\_bf\_prof aus\_bf\_man aus\_15\_19 aus\_20\_24 aus\_25\_29 aus\_30\_34 aus\_35\_39 aus\_40\_44 aus\_45\_49 aus\_50\_54 aus\_55\_59 aus\_60\_64 aus\_ab65 aus\_senio\_1 aus\_senio\_2 aus\_senio\_3)
- Betriebsdynamik (eintritt besch inflow besch\_vor status\_vor austritt besch outflow besch\_nach status\_nach)

#### 4.5.2 Weitere Verlinkungsmöglichkeiten

Es besteht die Möglichkeit zur Verlinkung des letzten Falls einer Neueinstellung aus der IAB-Stellenerhebung mit den administrativen Prozessdaten der eingestellten Personen ([IABSE-ADIAB](#)). Die Verlinkung steht derzeit für den Zeitraum 2010-2020 zur Verfügung. Die Personendaten folgen in Umfang und Logik der „Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB)“. Weitere Informationen finden sich im FDZ-Datenreport 01/2023 (Lochner, 2023). Die Daten der Stellenerhebung können über einen eindeutigen Schlüssel zugespielt werden. Die in der Stellenerhebung angegebenen eingestellten Personen werden in den Prozessdaten mithilfe eines „Matching-Algorithmus“ identifiziert, der in einem separaten Methodenbericht beschrieben ist (Lochner 2019). Für den zusätzlichen Erhalt der IABSE-ADIAB-Daten müssen diese bei der Datenbestellung ausdrücklich ausgewählt und deren Notwendigkeit begründet werden.

Es besteht außerdem die Möglichkeit, [betriebsspezifische Lohneffekte \(AKM-Effekte\)](#) an die IAB-Stellenerhebung heranzuspielen. Weitere Informationen finden sich im FDZ-Methodenreport 03/2025 (Lochner/Wolter, 2025). Die Verlinkung steht für alle Wellen der IAB-Stellenerhebung zu Verfügung. Add-ons können als Ergänzung zum Projekt formlos per E-Mail an [iab.fdz@iab.de](mailto:iab.fdz@iab.de) beantragt werden.<sup>2</sup>

---

<sup>2</sup> Die entsprechende Datei und die Variablennamen lauten: `iabse_0022_v1_akm_estab_8523_v1.dta` (`feff_1985_1992` `feff_1993_2000` `feff_2001_2008` `feff_2009_2016` `feff_2017_2023`).

## 4.6 STATA-Auswertungsbeispiel

### 4.6.1 Zeitreihe der offenen Stellen

```
/* Ausgeben einer Zeitreihe der offenen Stellen von 2000 bis zu aktuellen Rand, mit
Konfidenzintervallen */

capture log close
log using ${log}/05a_STATABeispiell_de,text replace
set more off

// Schleife über jede Befragungswelle
forvalues Welle=2000/2022 {
local WelleKurz =stroofreal(`Welle'-2000,"%02.0f")
quietly use ${orig}/iabse_`WelleKurz',clear
quietly label language de

/*Ausschluss der Experimentaldaten */
if `Welle' ==2019 {
drop if experiment2019 ==1 & jahr == 2019
}

// Anzahl der Quartale in dieser Welle ermitteln
if `Welle' <2005 {
local Quartale 4
}
else {
local Quartale 4 1 2 3
}

// Schleife über jedes Quartal
foreach Quartal of local Quartale {
// Quartalsspezifisches Variablenpräfix bilden
if `Quartal' ==4 {
local Jahr =`Welle'
}
else {
local Jahr =`Welle' +1
}
local JahrKurz =`Jahr' -2000
local Praefix q`Quartal'`JahrKurz'

quietly {
// Keine Angabe heißt keine offenen Stellen
replace `Praefix'_f20 =0 if `Praefix'_f20 ==.
replace `Praefix'_f40 =0 if `Praefix'_f40 ==.

// Summe aus sofort und später zu besetzenden Stellen
generate `Praefix'_Stellen =`Praefix'_f20 +`Praefix'_f40

// Hochgerechnete Zahl der offenen Stellen berechnen inkl. 95%-
// Konfidenzintervall durch Berechnung des gewichteten Mittel-
// werts multipliziert mit der hochgerechneten Anzahl Betriebe
mean `Praefix'_Stellen [pweight =`Praefix'_greggew]
matrix Ergebnis =r(table)
quietly summarize `Praefix'_greggew
local Betriebe =r(sum)
local b =Ergebnis[1,1] *`Betriebe'
local ll =Ergebnis[5,1] *`Betriebe'
local ul =Ergebnis[6,1] *`Betriebe'
} // quietly

// Ergebniszeile ausgeben
display as text "Offene Stellen im " as result `Quartal' as text ".
Quartal " as result `Jahr' as text ": "
display as result %7.0f `b' as text ", 95% Konfidenzintervall: [" as
result %7.0f `ll' "; " %7.0f `ul' as text "]"
}
}
log close
```

## 4.6.2 Auswertung einer Frage mit Mehrfachantworten

```
/* Anteil der schwierigen Neueinstellungen und Gründe auswerten */

capture log close
log using "${log}/05b_STATABeispiel2_de.log",text replace
set more off

use ${orig}/iabse_22,clear

/*Ausschluss der Experimentaldaten */
capture drop if experiment2019 ==1

label language de

/* Neueinstellungsgewicht bilden */
generate gregneu = q422_greggew*q422_f11_04

/* Gewicht für alle folgenden Auswertungen */
svyset [pweight=gregneu]

/* Anteil der Neueinstellungen mit Schwierigkeiten */
svy: tabulate q422_zf8

/* Falls ein Grund angegeben, alle nicht angekreuzten auf "nein" setzen,
sonst auf Missing belassen */
generate byte beantwortet_zf8grund = q422_zf8a==1 | q422_zf8b1==1 | ///
q422_zf8c1==1 | q422_zf8e==1 | q422_zf8f==1
foreach var of varlist q422_zf8a q422_zf8b1 q422_zf8c1 q422_zf8e q422_zf8f {
    replace `var'=0 if `var'==. & beantwortet_zf8grund==1
}

/* Sicherstellen, dass nur bei Schwierigkeiten ein Grund genannt wurde */
assert beantwortet_zf8grund == 0 if q422_zf8 != 1

/* Gründe nacheinander ausgeben */
foreach var of varlist q422_zf8a q422_zf8b1 q422_zf8c1 q422_zf8e q422_zf8f {
    svy: tabulate `var' if beantwortet_zf8grund
}
log close
```

# Literatur

- Börschlein, Erik-Benjamin; Diegmann, André; Gürtzgen, Nicole; Kubis, Alexander; Pirralha, André; Pohlman, Laura; Popp, Martin; Vetter, Franka (2024): IAB-Stellenerhebung: Betriebsbefragung zu Stellenangebot und Besetzungsprozessen, Welle 2000 bis 2021 mit Folgequartalen von 2006 bis 2022. (FDZ-Datenreport 06/2024 (de)), Nürnberg, 27 S. DOI:10.5164/IAB.FDZD.2406.de.v1
- Bossler, Mario; Geis, Gregor; Stegmaier, Jens (2018): Comparing survey data with an official administrative population: Assessing sample-selectivity in the IAB Establishment Panel. In: *Quality & Quantity*, Jg. 52, H. 2, S. 899-920.
- Bossler, Mario; Gürtzgen, Nicole; Kubis, Alexander; Kufner, Benjamin; Olbrich, Lukas; Schwanhäuser, Silvia (2022): Revision and new data quality concept due to deviant interviewer behavior in the IAB Job Vacancy Survey. (FDZ-Methodenreport, 05/2022 (en)), Nürnberg, 19 S.
- Bossler, Mario; Gürtzgen, Nicole; Kubis, Alexander; Kufner, Benjamin; Lochner, Benjamin (2020). The IAB Job Vacancy Survey: design and research potential. In: *Journal for Labour Market Research*, Jg. 54, H. 1, S. 1-12.
- Bundesagentur für Arbeit (2020): Steckbrief Betriebsnummervergabe. <https://www.informationsportal.de/wp-content/uploads/Steckbrief-Betriebsnummer.pdf> [2021-06-08]
- Brenzel, Hanna; Czepek, Judith; Kiesel, Hans; Kriechel, Ben; Kubis, Alexander; Moczall, Andreas; Rebien, Martina; Röttger, Christof; Szameitat, Jörg; Warning, Anja; Weber, Enzo (2016): Revision der IAB-Stellenerhebung. Hintergründe, Methode und Ergebnisse. (IAB-Forschungsbericht, 04/2016), Nürnberg.
- Gürtzgen, Nicole; Kubis, Alexander; Popp, Martin (2023): IAB-Monitor Arbeitskräftebedarf 1/2023: Verhältnis von Arbeitslosen zu offenen Stellen ist mit 1,5 nach wie vor niedrig. In: *IAB-Forum*, 01.06.2023. DOI: 10.48720/IAB.FOO.20230601.01
- Hentschke, Anna; Kubis, Alexander (2024): Positive Bewertungen bei Kontakt, aber viele Betriebe kennen die Dienstleistungen des Arbeitgeber-Service der Bundesagentur für Arbeit nicht. In: *IAB-Forum* H. 24.04.2024. DOI: 10.48720/IAB.FOO.20240424.01
- Hutter, Christian (2020): A new indicator for nowcasting employment subject to social security contributions in Germany. In: *Journal for Labour Market Research*, Jg. 54, H. 1, Art. 4, S. 1-10.
- Kubis, Alexander; Popp, Martin (2024): Persönliche Empfehlungen tragen dazu bei, betriebliche Vorbehalte gegenüber Langzeitarbeitslosen abzubauen. In: *IAB-Forum* H. 21.08.2024. DOI: 10.48720/IAB.FOO.20240821.01
- Kufner, Benjamin; Sakshaug, Joseph; Zins, Stefan (2021): More Clarification, Less Item Nonresponse in Establishment Surveys? A Split-Ballot Experiment. In: *Survey research methods*, Jg. 15, H. 2, S. 195-206.
- Lochner, Benjamin (2019): A simple algorithm to link “last hires” from the Job Vacancy Survey to administrative records. (FDZ-Methodenreport 06/2019 (en)), Nürnberg, 16 S. DOI:10.5164/IAB.FDZM.1906.en.v1

- Lochner, Benjamin (2023): IAB-Stellenerhebung verknüpft mit administrativen Daten des IAB (IABSE-ADIAB) 1975-2020. (FDZ-Datenreport, 01/2023(de)), Nürnberg. DOI: 10.5164/IAB.FDZD.2301de.v1
- Lochner, Benjamin; Wolter, Stefanie (2025): AKM Effects for German labour market data 1985-2023. (FDZ-Methodenreport 03/2025), Nürnberg, 12 S. DOI:10.5164/IAB.FDZM.2503.en.v1
- Sakshaug, Joseph; Vicari, Basha (2018): Obtaining record linkage consent from establishments. The impact of question placement on consent rates and bias. In: Journal of Survey Statistics and Methodology, Jg. 6, H. 1, S. 46-71. DOI:10.1093/jssam/smx009
- Winship, Christopher; Radbill, Larry (1994): Sampling Weights and Regression Analysis. In: Sociological Methods & Research, Jg. 23, H. 2, S. 230-257.

# Impressum

## FDZ-Datenreport 09|2025

### Veröffentlichungsdatum

26. September 2025

### Herausgeber

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

### Nutzungsrechte

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

### Bezugsmöglichkeit dieses Dokuments

[https://doku.iab.de/fdz/reporte/2025/DR\\_09-25.pdf](https://doku.iab.de/fdz/reporte/2025/DR_09-25.pdf)

### Dokumentation Version

IABSE0022\_DE\_v1\_dok1, DOI: 10.5164/IAB.FDZD.2509.de.v1

### Datensatz Version

IABSE 0022 v1, DOI: 10.5164/ IAB.IABSE0022.de.en.v1

### Bezugsmöglichkeit aller Veröffentlichungen der Reihe „FDZ-Datenreport“

<https://fdz.iab.de/forschung/publikationen/fdz-datenreporte/>

### Website

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

---

### Rückfragen zum Inhalt

Alexander Kubis  
Telefon: 0911 179-8978  
E-Mail: [alexander.kubis@iab.de](mailto:alexander.kubis@iab.de)