



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

FDZ-DATENREPORT

Dokumentation zu Arbeitsmarktdaten

09|2022 DE Panel „Betriebe in der Covid-19 Krise“ –
20/21/22

Eine Längsschnittstudie in deutschen Betrieben – Wellen 1 - 24

Nils Backhaus, Lutz Bellmann, Patrick Gleiser, Sophie Hensgen, Christian Kagerl, Theresa Koch, Corinna König, Eva Kleifgen, Moritz Kuhn, Ute Leber, Michael Moritz, Laura Pohlan, Swantje Robelski, Duncan Roth, Malte Schierholz, Sabine Sommer, Jens Stegmaier, Anita Tisch, Matthias Umkehrer, Armin Aminian



Bundesagentur für Arbeit

Panel „Betriebe in der Covid-19 Krise“ – 20/21/22

Eine Längsschnittstudie in deutschen Betrieben – Wellen 1 - 24

Nils Backhaus, Lutz Bellmann, Patrick Gleiser, Sophie Hensgen, Christian Kagerl, Theresa Koch, Corinna König, Eva Kleifgen, Moritz Kuhn, Ute Leber, Michael Moritz, Laura Pohlan, Swantje Robelski, Duncan Roth, Malte Schierholz, Sabine Sommer, Jens Stegmaier, Anita Tisch, Matthias Umkehrer, Armin Aminian

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

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

Inhaltsverzeichnis

1	Einleitung.....	5
2	Datenbeschreibung.....	7
2.1	Kurzbeschreibung.....	7
2.2	Mengengerüst	9
3	Datenerhebung.....	10
3.1	Inhalt und Methode.....	10
3.2	Grundgesamtheit und Stichprobenziehung	12
4	Datenaufbereitung	15
4.1	Datenorganisation	15
4.2	Datenprüfung und Datenbereinigung.....	16
4.3	Gewichtung	18
4.4	Verlinkung mit administrativen Betriebsdaten (BHP)	19
5	Verwendungshinweise und Datennutzung	22
5.1	Vorbereitungen zur Datenanalyse.....	22
5.2	Stata-Auswertungsbeispiel.....	22
5.3	Datenzugang und Zusatzmaterial	25

Tabellenverzeichnis

Tabelle 1: Inhaltliche Charakteristika	7
Tabelle 2: Methodische Charakteristika	8
Tabelle 3: Datenzugang	8
Tabelle 4: Mengengerüst	9
Tabelle 5: Themenübersicht des flexiblen Moduls	11
Tabelle 6: Schichtungsmerkmale der Bruttostichproben	14
Tabelle 7: Fortgeschriebene Variablen	16
Tabelle 8: Klassifikation fehlender Werte	17

Zusammenfassung

Dieser Datenreport beschreibt die Daten der im Rahmen des Projekts „Betriebe in der Covid-19-Krise“ durchgeführten Betriebsbefragung. In regelmäßigen Abständen ist dabei eine repräsentative Stichprobe von 1.500 bis 2.000 Betrieben aus verschiedenen Wirtschaftszweigen und Betriebsgrößen befragt worden.

Durch ein hochfrequentes, rotierendes Panel-Design bietet die Befragung eine Datenbasis, welche die vielseitigen Auswirkungen der Pandemie auf Betriebe in Deutschland in einem zeitlich engen Takt erfasst. Insgesamt stehen vierundzwanzig Wellen (Zeitraum: 03.08.2020 – 30.06.2022) für externe Forscherinnen und Forscher zur Verfügung.

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

Abstract

This data report describes the data from an establishment survey conducted as part of the "Establishments in the Covid-19 Crisis" project. A representative sample of 1,500 to 2,000 establishments of different sizes and from various economic sectors was surveyed at regular intervals.

By utilizing a high-frequency, rotating panel design, the survey provides a database that closely tracks the consequences of the pandemic. A total of twenty-four waves (period: 08/03/2020 - 06/30/2022) are available to external researchers.

Details on data access, conditions of use and how to request the data set can be found on the FDZ website at <https://fdz.iab.de/>.

Keywords

Betriebsbefragung, Panel Datensatz, Datendokumentation, Covid-19-Pandemie

Danksagung

Wir danken unseren Kooperationspartnern, insbesondere der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA), und dem Exzellenzcluster ECONtribute der Universitäten Bonn und Köln, für die finanzielle und inhaltlich-konzeptionelle Beteiligung an der Studie sowie Kantar Public für die Umsetzung des Projekts und die Durchführung der Erhebung.

1 Einleitung

Die Corona-Krise stellt Betriebe vor erhebliche Herausforderungen. Um das Virus einzudämmen, veranlassen Behörden Restriktionen und zeitlich begrenzte „Lockdowns“ für große Teile des gesellschaftlichen Lebens. Infolgedessen müssen viele Betriebe ihre Geschäftstätigkeit zeitweise reduzieren oder ihre Arbeitsabläufe vielseitig ändern, um die dynamischen Anforderungen des Arbeits- und Gesundheitsschutzes zu erfüllen. Zusätzlich bedingt die Krise Lieferengpässe sowie eine reduzierte Nachfrage nach bestimmten Waren und Dienstleistungen. Betriebe sind von diesen Herausforderungen jedoch in sehr unterschiedlichem Maße betroffen. Viele berührt die Krise nur am Rande, während andere ihre Geschäftstätigkeit weitgehend einstellen müssen. Des Weiteren stehen Betrieben unterschiedliche Konzepte und Strategien offen, um den behördlichen Restriktionen gerecht zu werden, beispielsweise eine Ausweitung der Home-Office Nutzung oder die Etablierung umfassender Hygienekonzepte. Das Panel „Betriebe in der Covid-19-Krise“ untersucht die wirtschaftlichen und betriebsorganisatorischen Auswirkungen auf Betriebe in Deutschland.

Im Rahmen eines hochfrequenten, rotierenden Paneldesigns wurden Betriebe aus verschiedenen Wirtschaftsbereichen und Größenklassen zu den Auswirkungen der Corona-Krise in Bezug auf verschiedene Teilbereiche in einem drei- bis vierwöchigen Rhythmus telefonisch befragt. Die Erhebung resultierte in einem aus 24 Wellen bestehendem Panel, das einen Zeitraum vom 03. August 2020 bis zum 30. Juni 2022 erfasst. In jeder Welle wurden bislang zwischen 1.500 und 2.000 Betriebe befragt (Erst- und Panelbefragte).

Die Daten der Betriebsbefragung ermöglichen repräsentative Quer- und Längsschnittanalysen zu den vielfältigen Auswirkungen der Corona-Krise auf Betriebe in Deutschland. Das hochfrequente Intervall der Befragung ermöglicht es zudem, alle Phasen der Pandemie in einem engen zeitlichen Takt abzubilden. Um das Analysepotenzial weiter zu erhöhen, können die Daten mit administrativen Daten der Bundesagentur für Arbeit verknüpft werden.

Die Datensätze entstanden im Projekt „Betriebe in der Covid-19 Krise“. Teile des Projekts erfolgten im Rahmen einer Forschungskooperation mit der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) und dem Exzellenzcluster ECONtribute der Universitäten Bonn und Köln. In bislang vier Wellen war die BAuA sowohl an der inhaltlichen Konzeption des Erhebungsinstruments als auch der Finanzierung beteiligt. Das Exzellenzcluster beteiligte sich finanziell und inhaltlich-konzeptionell vor allem an Welle 13.

Dieser Datenreport dokumentiert die Daten aller vierundzwanzig Wellen der Betriebsbefragung (Zeitraum: 03. August 2020 bis 30. Juni 2022). Er enthält Informationen über die Inhalte der Befragung, Erhebungsmethode, Stichprobenziehung, Aufbereitung, Anonymisierung, Gewichtung und Datenorganisation. Fragebögen, sowie eine Variablenübersicht sind separat über das Forschungsdatenzentrum der Bundesagentur für Arbeit verfügbar.

Die Datensätze sind für externe Wissenschaftler und Wissenschaftlerinnen über das Forschungsdatenzentrum erhältlich. Verschiedene Möglichkeiten zum Datenzugang können der offiziellen Webseite entnommen werden (<https://fdz.iab.de/>).

2 Datenbeschreibung

2.1 Kurzbeschreibung

Tabelle 1: Inhaltliche Charakteristika

Kategorie	Beschreibung
Aktuelle Datenversion	BeCovid 2020/2021/2022
Themen/Merkmalgruppen	<p>Für jede Welle erhoben: Betriebsstruktur Personalentwicklung Liquiditätsreserven Ausmaß und Art der wirtschaftlichen Auswirkungen durch die Corona Krise Kurzarbeit (ab Welle 3)</p> <p>In einzelnen Wellen erhoben: Arbeitsicherheit und Gesundheitsschutz Impfangebote in Betrieben Ausbildung und Ausbildungsprämie Homeoffice Weiterbildung Ex- und Importe Lieferengpässe Auswirkungen des Lockdowns Digitalisierung Atypische Beschäftigung Arbeitsorganisation Betriebliche Maßnahmen für psychische Gesundheit Beschäftigungsentwicklung Betriebsrat und Betriebsleitung Frauen in Führungspositionen Maßnahmen zur Vereinbarung von Beruf & Familie Rekrutierung von Personal Auswirkungen des Kriegs in der Ukraine</p>
Untersuchungseinheit	<p>Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten zum Stichtag. Ausgeschlossene Wirtschaftssektoren: „Öffentliche Verwaltung, Verteidigung, Sozialversicherung“ (WZ 2008:84), „private Haushalte“ (WZ 2008: 97+98) und „exterritoriale Organisationen“ (WZ 2008: 99).</p> <p>Stichtag für Wellen 1 bis 4: 30. November 2019 Stichtag ab Welle 5: 28. Februar 2020 Stichtag ab Welle 21: 30. April 2021</p>
Fallzahlen	1.500 bis 2.000 Betriebe pro Welle (Erst- und Wiederbefragte)
Zeitraum	03. August 2020 bis 18. Dezember 2020 (dreiwöchiger Rhythmus) 18. Januar 2021 bis 26. Mai 2021 (dreiwöchiger Rhythmus) 7. Juni 2021 bis 14. Dezember 2021 (vierwöchiger Rhythmus) 13. Januar 2022 bis 30. Juni 2022 (sechs- bis neunwöchiger Rhythmus)
Zeitbezug	Befragungszeitpunkt
Regionale Gliederung	Ost/West
Gebietsstand	Zum jeweiligen Stichtag

Tabelle 2: Methodische Charakteristika

Kategorie	Beschreibung
Bruttostichproben	<p>Wiederholerstichprobe: Betriebe, die:</p> <ul style="list-style-type: none"> • in der vorigen Befragung weiteren Teilnahmen zugestimmt haben. • zwischen einzelnen Wellen temporär ausfallen. • weniger als sechsmal erfolgreich befragt wurden. <ul style="list-style-type: none"> ○ ausgeschlossen sind Großbetriebe ab Welle 19. • deren Erstbefragung höchstens sieben Wellen zurückliegt. <ul style="list-style-type: none"> ○ Ausgeschlossen sind Großbetriebe ab Welle 19. • deren letztes Interview nicht mehr als vier Wellen zurückliegt. <p>Auffrischerstichprobe: Betriebe, die:</p> <ul style="list-style-type: none"> • zum Befragungszeitpunkt nicht an der IAB-Stellenerhebung oder am IAB-Betriebspanel teilnehmen. • noch nicht an der Befragung teilgenommen haben oder bislang nicht kontaktiert wurden. • nach einer ersten Kontaktaufnahme die Befragung nicht verweigerten.
Erhebungsmodus	Telefonische Interviews (CATI)
An der Erhebung beteiligte Institutionen	Institut für Arbeitsmarkt- und Berufsforschung (IAB), Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA), Kantar Public
Dateiformat/-größe	STATA (ca. 500 KB pro Welle)
Dateiorganisation	nach Welle

Tabelle 3: Datenzugang

Kategorie	Beschreibung
Zugangswege	Gastaufenthalt, anschließende Datenfernverarbeitung
Anonymisierungsgrad	Schwach anonymisiert
Sensible Merkmale	<p>Befragungsdaten: keine</p> <p>Administrative Betriebsdaten: Arbeitsort - Kreis (ao_kreis), Wirtschaftsunterklasse (w08_5), Erstes Auftreten der Betriebsnummer (grd_dat), Letztes Auftreten der Betriebsnummer (lzt_dat)</p>
Vorgaben zum Zitieren der Daten und Datendokumentation	<p>Daten: "Backhaus, Nils; Bellmann, Lutz; Gleiser, Patrick; Hensgen, Sophie; Kagerl, Christian; Koch, Theresa; König, Corinna; Kleifgen, Eva; Kuhn, Moritz; Leber, Ute; Moritz, Michael; Pohlen, Laura; Robelski, Swantje; Roth, Duncan; Schierholz, Malte; Sommer, Sabine; Stegmaier, Jens; Tisch, Anita; Umkehrer, Matthias; Aminian, Armin (2022): "Panel „Betriebe in der Covid-19 Krise“ (IAB BeCovid) – Wellen 01-24". Forschungsdatenzentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB). DOI: 10.5164/IAB.BeCovid_W01-W24.de.en.v1</p> <p>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: Nils Backhaus, Lutz Bellmann, Patrick Gleiser, Sophie Hensgen, Christian Kagerl, Theresa Koch, Corinna König, Eva Kleifgen, Moritz Kuhn, Ute Leber, Michael Moritz, Laura Pohlen, Swantje Robelski, Duncan Roth, Malte Schierholz, Sabine Sommer, Jens Stegmaier, Anita Tisch, Matthias Umkehrer, Armin Aminian (2022): Panel zu „Betriebe in der Covid-19 Krise – 20/21/22“. Eine Längsschnittstudie in deutschen Betrieben – Wellen 1-24. FDZ-Datenreport, 09/2022 (de), Nürnberg. DOI: 10.5164/IAB.FDZD.2209.de.v1</p>

2.2 Mengengerüst

Die Betriebsbefragung setzt sich in jeder Welle zu Teilen aus Erst- und Wiederbefragten zusammen. Tabelle 4 zeigt die Fallzahlen sowie die Zusammensetzung der bereitgestellten Datensätze. In jeder Welle wurde eine Mindestzahl von 1.500 Betrieben erfasst. Der Anteil an Erstbefragten ist nach wenigen Wellen auf ein angestrebtes Niveau von ca. 30-40 Prozent gesunken, und dient, unter anderem, zur Auffrischung der Stichprobe. In Welle 8 stieg der Anteil an Erstbefragten erneut auf 54 Prozent, da erstmalig ein großer Teil der Panelbetriebe aus der Stichprobe rotiert wurde.

Tabelle 4: Mengengerüst

Datensatz	Kalenderwoche	Jahr	Nettostichprob enumfang	Anteil Erstbefragte	Anteil Wiederbefragte
BeCovid_W1_v3dta	32, 33	2020	1.723	100%	0%
BeCovid_W2_v3.dta	35, 36	2020	1.556	53,92%	46,08%
BeCovid_W3_v3.dta	38, 39	2020	1.733	40,33%	59,67%
BeCovid_W4_v3.dta	41, 42	2020	1.794	38,97%	61,03%
BeCovid_W5_v3.dta	44, 45	2020	1.760	31,38%	68,62%
BeCovid_W6_v3.dta	47, 48	2020	2.001	38,98%	61,02%
BeCovid_W7_v3.dta	50, 51	2020	1.511	33,09%	66,91%
BeCovid_W8_v3.dta	3, 4	2021	1.999	54,23%	45,77%
BeCovid_W9_v3.dta	6, 7	2021	2.000	34,45%	65,55%
BeCovid_W10_v2.dta	9, 10, 11	2021	2.002	33,72%	66,28%
BeCovid_W11_v2.dta	12, 13, 14	2021	2.009	37,33%	62,67%
BeCovid_W12_v2.dta	16, 17	2021	2.001	30,03%	69,97%
BeCovid_W13_v2.dta	19, 20, 21	2021	2.001	35,52%	64,48%
BeCovid_W14_v2.dta	23, 24	2021	2.001	39,78%	60,22%
BeCovid_W15_v1.dta	27, 28, 29	2021	2.002	39,86%	60,14%
BeCovid_W16_v1.dta	31, 32, 33	2021	1.903	45,23%	54,77%
BeCovid_W17_v1.dta	36, 37, 38	2021	2.010	40,15%	59,85%
BeCovid_W18_v1.dta	40, 41, 42	2021	2.001	38,98%	61,01%
BeCovid_W19_v1.dta	44, 45, 46	2021	2.000	39,75%	60,25%
BeCovid_W20_v1.dta	48, 49, 50	2021	2.000	40,00%	60,00%
BeCovid_W21_v1.dta	2, 3, 4	2022	2.002	35,51%	64,49%
BeCovid_W22_v1.dta	9, 10, 11	2022	1.999	33,72%	66,28%
BeCovid_W23_v1.dta	18, 19, 20	2022	1.978	33,52%	66,48%
BeCovid_W24_v1.dta	24, 25, 26	2022	2.000	38,85%	61,15%

3 Datenerhebung

3.1 Inhalt und Methode

Die inhaltliche und methodische Konzeption der Panelbefragung „Betriebe in der Covid-19 Krise“ wurde gemeinsam vom zuständigen Projektteam des IAB und verschiedenen Kooperationspartnern, insbesondere der BAuA, erarbeitet. Die Durchführung der Befragung erfolgte durch Kantar Public. Im Fokus der Befragung standen die wirtschaftlichen und strukturellen Auswirkungen der Corona-Krise auf Betriebe in Deutschland, sowie deren Veränderung im zeitlichen Verlauf der Pandemie. Insgesamt wurden im Zeitraum zwischen dem 03. August 2020 und dem 30. Juni 2022 insgesamt 24 Einzelwellen erhoben. Tabelle 5 enthält Informationen zu Beginn und Dauer der jeweiligen Feldzeiten. Die ersten zwanzig Wellen wurden in einem drei- bis vierwöchigen Intervall erhoben. Ab der einundzwanzigsten Welle betrug das Intervall zwischen sechs und neun Wochen. Die aktive Feldzeit pro Einzelwelle betrug in der Regel zwei Wochen. Interviews wurden ausschließlich telefonisch (CATI) geführt.

Die Befragung erfolgte anhand eines rotierenden Paneldesigns. Hierzu wurden Betriebe nach maximal sechs Teilnahmen, oder spätestens sieben Wellen nach ihrem Erstinterview, aus dem Panel rotiert. Dabei war nicht relevant, wie häufig der Betrieb tatsächlich befragt wurde. So rotierten beispielsweise auch Betriebe aus dem Panel, die im Zeitraum von sieben Wellen nach ihrem Erstinterview lediglich zwei Mal erfolgreich wiederbefragt wurden. Des Weiteren wurden besonders schwer kontaktierbare Betriebe, mit denen in vier aufeinanderfolgende Wellen kein Interview zustande kam, durch andere Betriebe ersetzt. Ausfälle aufgrund der Rotation oder Non-Response wurden durch neu rekrutierte Betriebe ersetzt.

Die Teilnahme an der Befragung erfolgte zu jedem Zeitpunkt auf freiwilliger Basis. Für die Befragung selektierte Betriebe erhielten einige Wochen zuvor ein Anschreiben, in dem der bevorstehende Anruf eines Interviewers angekündigt wurde. Des Weiteren enthielt das Schreiben Informationen zur Zielsetzung der Studie, als auch Hinweise zum Datenschutz.

Die Fragebögen enthalten drei Arten von Fragen: „Strukturfragen“, „Panelfragen“ und „Modulfragen“. *Strukturfragen* werden üblicherweise einmalig bei der ersten Teilnahme eines Betriebs gestellt. Sie erheben grundsätzliche Merkmale der Betriebsorganisation und ermöglichen damit eine differenzierte Datenanalyse für verschiedene Subpopulationen. Zu den Merkmalen zählen beispielsweise Angaben über die Zahl der Beschäftigten vor Beginn der Pandemie, Teilaspekte der Ausbildungsaktivität und zur Organisationsstruktur. Bei einigen Strukturangaben handelt es sich um zeitlich konstante Merkmale. Daher werden sie für Wiederbefragte kein weiteres Mal gestellt, sondern automatisch anhand der Angaben aus der Erstbefragung in den Daten fortgeschrieben. Detailliertere Informationen zum Fortschreibungsprozess sind in Abschnitt [4.1](#) enthalten. Für Betriebe, die einer Verknüpfung ihrer Angaben mit administrativen Daten der Bundesagentur für Arbeit zugestimmt haben (92,85% der befragten Betriebe), ist zusätzlich ein umfassenderes Spektrum an Betriebsstrukturinformationen aus den Daten des Betriebs-Historik-Panels (BHP) verfügbar.

Diese können dem Panel anhand des Betriebsidentifikators „betnr_becovid“ zugespielt werden (siehe Abschnitt [4.4](#)).

Panelfragen werden sowohl für Erst- als auch Wiederbefragte in regelmäßigen Abständen erhoben und beziehen sich häufig auf die Auswirkungen der Covid-19-Krise. Sie gliedern sich in die Themenkomplexe „Betroffenheit und wirtschaftliche Auswirkungen“, „Personalentwicklung“ und „Kurzarbeit“. Durch das hochfrequente Befragungsintervall ermöglichen diese Angaben engmaschige Zeitreihenanalysen zu einigen relevanten Indikatoren der Covid-19-Krise. Eine Übersicht zu den Erhebungszeitpunkten einzelner Panelfragen bietet die Variablenübersicht welche separat über die Internetseite des Forschungsdatenzentrums verfügbar ist:

(<https://fdz.iab.de/betriebsdaten/panel-betriebe-in-der-covid-19-krise-iab-becovid-welle-01-24/>).

Bei den *Modulfragen* handelt es sich um einen flexiblen Fragenteil, dessen Inhalt sich in jeder Welle ändert. In der Regel beziehen sich die Modulfragen einer Welle auf konkrete und vom Rest der Erhebung gesonderte Themenschwerpunkte. Tabelle 5 gibt einen Überblick der einbezogenen Themen, sowie Informationen zum Befragungszeitpunkt. Der Inhalt des Fragenprogramms kann dem Fragebogen der jeweiligen Welle entnommen werden. Die Fragebögen sind für externe Wissenschaftlerinnen und Wissenschaftler über das Forschungsdatenzentrum erhältlich.

Tabelle 5: Themenübersicht des flexiblen Moduls

Thema	Welle	Feldzeit	Beteiligte Institutionen
Kurzarbeit, Krisenbetroffenheit	1	03.08 – 18.08.2020	-
Arbeitssicherheit, Gesundheitsschutz	2	25.08 – 04.09.2020	BAuA
Ausbildung, Ausbildungsprämie	3	14.09 - 25.09.2020	-
Homeoffice	4	05.10 – 19.10.2020	BAuA
Weiterbildung	5	26.10 – 09.11.2020	FB: Arbeitsförderung und Erwerbstätigkeit (IAB)
Zukunftstrends, Exporte, Lockdown	6	16.11 – 27.11.2020	FB: Migration, Integration und internationale Arbeitsmarktforschung (IAB)
Ausbildungsjahr 2021/22, Ausbildungsprämie (II), Rekrutierung	7	07.12 – 18.12.2020	Bundesagentur für Arbeit
Kurzarbeitergeld, Lockdown, Mindestlohn, Homeoffice	8	18.01 – 29.01.2021	Mindestlohnkommission
Digitalisierung	9	08.02 – 19.02.2021	Ruprecht-Karls-Universität Heidelberg; Luxembourg Institute of Socio-Economic Research (LISER)
Ausbildungsprämie III, Homeoffice	10	01.03 – 12.03.2021	Forschungsdatenzentrum (IAB)
Vignette atypische Beschäftigung (Einstellungsverhalten), Corona-Tests, Homeoffice	11	22.03. – 07.04.2021	FB: Bildung, Qualifizierung und Erwerbsverläufe (IAB); FB: Betriebe und Beschäftigung (IAB)

Thema	Welle	Feldzeit	Beteiligte Institutionen
Führung/Leistungsbeurteilung, Corona-Tests, Homeoffice, Kurzarbeit	12	19.04. – 30.04.2021	Ruprecht-Karls-Universität Heidelberg
Arbeitsorganisation	13	10.05. – 26.05.2021	Rheinische Friedrich-Wilhelms-Universität Bonn
Betriebliche Maßnahmen für psychische Gesundheit (+Homeoffice)	14	07.06. – 18.06.2021	BAuA
Finanzielle Lage, Beschäftigungsentwicklungen und Interessensvertretung innerhalb des Betriebs	15	05.07. – 20.07.2021	Friedrich-Alexander-Universität Erlangen-Nürnberg; Europäische Akademie der Arbeit (Goethe Universität Frankfurt am Main)
Schutzmaßnahmen in Betrieben, Impfungen, Importe und Lieferketten	16	02.08. – 17.08.2021	BAuA
Ausbildungsverhältnisse, Ausbildungsprämien	17	06.09. – 20.09.2021	Bundesinstitut für Berufsbildung (BIBB)
Frauen in Führungspositionen, Maßnahmen zur Vereinbarkeit von Beruf und Familie	18	04.10. – 19.10.2021	
Homeoffice	19	02.11. – 16.11.2021	
Import, Lieferengpässe	20	01.12. – 14.12.2021	FB: Migration, Integration und internationale Arbeitsmarktforschung (IAB)
Rekrutierung bei Ausbildungsplätzen und Übernahmen, Gesundheitsschutz (3G Überprüfung)	21	13.01. – 28.01.2022	
Rekrutierung von Personal	22	01.03. – 16.03.2022	
Auswirkungen des Kriegs in der Ukraine (Energieprobleme, Beschäftigung von Geflüchteten, Importe) Vignette Rekrutierung bei Ausbildungsplätzen	23	02.05. – 17.05.2022	FB: Bildung, Qualifizierung und Erwerbsverläufe (IAB)
Homeoffice, Ausbildungsprämie IV, Auswirkungen des Kriegs in der Ukraine (Importe, Exporte, Vorleistungen, Lieferengpässe)	24	13.06. – 30.06.2022	

3.2 Grundgesamtheit und Stichprobenziehung

Die Grundgesamtheit des Panels umfasst alle Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten. Ein Betrieb entspricht der Definition des Betriebsnummernkonzepts der Bundesagentur für Arbeit (Bundesagentur für Arbeit 2013). Demnach ist ein Betrieb eine regional und wirtschaftlich abgegrenzte Einheit mit Beschäftigten, für die im Rahmen des Meldeverfahrens an die Sozialversicherungsträger nach bestimmten Regeln eine eigene Betriebsnummer vergeben wird. Dabei gilt:

- Niederlassungen eines Arbeitgebers in verschiedenen Gemeinden erhalten grundsätzlich eine eigene Betriebsnummer.
- Niederlassungen eines Arbeitgebers innerhalb einer Gemeinde werden zu einem Betrieb mit einer Betriebsnummer zusammengefasst, sofern diese dieselbe wirtschaftliche Tätigkeit ausüben.
- Niederlassungen eines Arbeitgebers innerhalb einer Gemeinde mit unterschiedlichen wirtschaftlichen Tätigkeiten erhalten eine eigene Betriebsnummer.

Für eine detaillierte Beschreibung der Betriebsnummernlogik sei auf Fritz/Brixy 2004 oder die Internetseite der Bundesagentur für Arbeit (https://www.arbeitsagentur.de/datei/dok_ba015200.pdf) verwiesen.

Basis der Stichprobenziehung ist die Betriebsdatei der Bundesagentur für Arbeit zum jeweiligen Stichtag. Sie enthält alle Betriebe, die im Rahmen des oben genannten Verfahrens eine Betriebsnummer erhalten. Aus forschungspraktischen Gründen wurden folgende Wirtschaftszweige von der Ziehung ausgeschlossen:

- Öffentliche Verwaltung, Verteidigung; Sozialversicherung (WZ 2008: 84)
- Private Haushalte (WZ 2008: 97 + 98)
- Exterritoriale Organisationen und Körperschaften (WZ 2008: 99)
- Betriebe mit fehlenden Angaben bei der Anzahl der sozialversicherungspflichtig Beschäftigten bzw. des Wirtschaftszweigs (notwendige Informationen zur Stichprobenziehung)

Aus dieser Grundgesamtheit wurde eine nach der Betriebsgröße und dem Wirtschaftszweig geschichtete Stichprobe gezogen. Die Ausprägungen der Schichtungsmerkmale wurden so gewählt, dass über den Gesamtzeitraum der Studie ausreichend Betriebe pro Schicht verfügbar sind. Tabelle 6 beschreibt die Schichtungsmerkmale. Die Schichtung entspricht einer Kreuzmatrix beider Merkmale, wobei Betriebe mit mehr als 250 Beschäftigten nicht zusätzlich nach dem Wirtschaftszweig aufgeteilt wurden. Es ergeben sich insgesamt 16 Schichten. Der Umfang einzelner Schichten ergibt sich aus einer beschäftigtenoptimalen Allokation in die einzelnen Schichten (Ausnahme: Großbetriebe mit mehr als 250 Beschäftigten, deren Anzahl für eine beschäftigtenoptimale Allokation nicht ausgereicht hätte). Demnach erfolgte die Ziehung der Stichprobe disproportional anhand der Betriebsgröße. Je größer ein Betrieb, desto höher dessen Wahrscheinlichkeit in die Stichprobe zu gelangen. Diese Disproportionalität wird im Rahmen der Gewichtung wieder ausgeglichen.

Aufgrund der langen Gesamtfeldzeit wurde die Stichprobe in Tranchen aufgeteilt. Diese wurden zu unterschiedlichen Zeitpunkten gezogen und sukzessiv an das Erhebungsinstitut Kantar übermittelt. Durch dieses Verfahren sollte unter anderem verhindert werden, dass inzwischen geschlossene Betriebe in die Stichprobe gelangen. Für die ersten vier Wellen des Panels basiert die Stichprobenziehung auf der Betriebsdatei zum Stichtag 30. November 2019. Dabei zählen insgesamt 2.111.169 Betriebe zur oben genannten Grundgesamtheit. Davon wurden Betriebe, die aufgrund einer schriftlichen Verweigerung nicht mehr für IAB-Befragungen kontaktiert werden dürfen, oder bereits an einer anderen IAB-Betriebsbefragung teilnehmen (N = 30.551) zusätzlich von der Ziehung ausgeschlossen. Von der fünften bis zur zwanzigsten Welle basiert die Stichprobe auf der Betriebsdatei zum Stichtag 28. Februar 2020, in der 2.096.633 Betriebe zur Grundgesamtheit gehören. Davon waren zum Zeitpunkt der Ziehung für Welle fünf 2.055.409 Betriebe verfügbar. Ab der einundzwanzigsten Welle basiert die Stichprobe auf der Betriebsdatei zum Stichtag 30. Juni 2021, in der 2.100.079 Betriebe zur Grundgesamtheit gehören. Davon waren zum Zeitpunkt der Ziehung für Welle 2.079.534

Betriebe verfügbar. Für detaillierte Informationen zur Stichprobenziehung sei auf den Methodenbericht (im Erscheinen) verwiesen.

Für alle Betriebe erfolgte über Kantar eine Telefonnummernrecherche. Dabei konnten insgesamt 84,32% der gezogenen Betriebe erfolgreich eine Telefonnummer zugeordnet werden. Anschließend wurden sowohl doppelte Telefonnummern (3,24%) als auch jene, die sich bereits auf einer Kontakt-Sperrliste des Erhebungsinstituts befinden (5,41%), aussortiert. Zusätzlich wurden im Erstinterview anhand eines Screenings erneut Betriebe des öffentlichen Dienstes aus der Stichprobe entfernt. Insgesamt waren 75,23% der gezogenen Betriebe für ein Interview kontaktierbar.

Durch den Panelcharakter setzen sich einzelne Wellen aus zwei Teilstichproben zusammen:

- **Wiederholerstichprobe:** Umfasst alle Betriebe, die in ihrem letzten gültigen Interview einer weiteren Teilnahme zugestimmt haben. Von dieser Regel ausgeschlossen sind Betriebe, die bereits sechsmal an der Befragung teilgenommen haben, deren Erstbefragung bereits sieben Wellen zurückliegt, oder dessen letztes gültiges Interview mehr als vier Wellen zurückliegt, da diese aus dem Panel rotieren.
 - Ab Welle 19 wurden Großbetriebe mit Wiederbefragungsbereitschaft unbegrenzt in der Wiederholerstichprobe behalten. Des Weiteren wurde ab Welle 21 jede Welle ein Teil der wiederbefragungsbereiten Großbetriebe die bereits ausrotiert waren erneut in die Wiederholerstichprobe aufgenommen. Diese Änderung wurden vorgenommen, da bereits Ende 2021 anhand der Felddaten eine zu geringe Menge verfügbarer Großbetriebe für neue Erstbefragungen prognostiziert wurde. Damit auch bis Ende der Studie genügend Großbetriebe an der Befragung teilnehmen sollten durch diese Maßnahmen als Kompensation mehr Großbetriebe in der Wiederholungsstichprobe gehalten werden.
- **Auffrischerstichprobe:** Umfasst ausschließlich Betriebe, die bislang zu keinem Zeitpunkt der Studie befragt wurden, unabhängig von der Anzahl vorheriger Kontaktversuche. Pro Welle ist ein Minimum von 500 Erstbefragten angestrebt.

Tabelle 6: Schichtungsmerkmale der Bruttostichproben

Merkmal	Kategorien
Betriebsgröße (nach Anzahl Beschäftigte)	1-9, 10-49, 50-249, 250+
Wirtschaftszweig (nach Klassifikation der Wirtschaftszweige 2008)	Bergbau und Gewinnung von Steinen und Erden, Energie und Wasserversorgung, Baugewerbe (WZ 2008: 05-09,35-39,41-43); Verarbeitendes Gewerbe, Land- und Forstwirtschaft, Fischerei (WZ 2008: 01-03,10-33); Handel, Verkehr, Information und Kommunikation (WZ 2008: 45-47,49-53,58-63); Gastgewerbe, Finanzen und Versicherungen, unternehmensnahe Dienstleistungen, Erziehung und Unterricht sowie sonstige Dienstleistungen (WZ 2008: 55-56,64-66,68-82,85,90-93,95-96); Gesundheits- und Sozialwesen, Organisationen ohne Erwerbszweck, öffentliche Verwaltung (WZ 2008: 86-88,94)

4 Datenaufbereitung

4.1 Datenorganisation

Das Panel beinhaltet für alle befragten Betriebe, die mindestens einmal teilgenommen haben, Strukturinformationen. Einige davon werden pro Welle nur für Erstbefragte erhoben. Bei Wiederbefragten werden diese Informationen hingegen stets anhand der Angaben im ersten Interview fortgeschrieben. Eine Liste aller fortgeschriebenen Merkmale zeigt Tabelle 7. Fortgeschriebene Variablen sind zusätzlich durch eine Kennzeichnung im Variablennamen identifizierbar (s.u.).

Die Variablennamen des Panels bestehen aus einem zweiteiligen Code. Der erste Teil besteht aus einer anführenden Buchstabenkette, welcher die Variable einem der inhaltlichen Themenkomplexe zuordnet. Beispielsweise erhalten alle Variablen zum Thema „Personalentwicklung“ das Präfix „pe“. Der zweite Teil besteht aus einem fünfstelligen Zahlenschlüssel. Die ersten beiden Ziffern dienen schlicht als Zähler um die verschiedenen Variablen eines Themenkomplexes zu unterscheiden. Die dritte Ziffer gibt an, ob es sich um ein fortgeschriebenes (1) oder erhobenes Merkmal handelt (0). Die letzten beiden Ziffern dienen als Versionsnummer, beginnend bei 0. Ändert sich zwischen einzelnen Wellen die Erhebungsmethodik eines Merkmals, beispielsweise durch Unterschiede in der Fragenformulierung oder der Filtersetzung, erhöht sich diese Ziffer um 1.

BEISPIEL: WA04001

WA = WIRTSCHAFTLICHE AUSWIRKUNGEN

04 = FRAGE 04 AUS DIESEM THEMENGEBIET

0 = NICHT FORTGESCHRIEBEN

01 = DIE ERHEBUNGSMETHODIK WURDE EINMALIG GEÄNDERT

Variablen, die nicht aus dem Fragebogen stammen, sind mit sprechenden Namen versehen. Hierbei handelt es sich einerseits um das Betriebsidentifikationsmerkmal (*betnr_becovid*), den Wellen-identifikator (*welle*), den Erhebungszeitpunkt (*kalenderwoche*, *jahr*), um die Gewichtungsfaktoren (*hrf_kal*, *hrf_kal_refined*), einen Schichtungsindikator (*strata_id*) sowie einen Indikator für den Wirtschaftszweig (*WZ_abschn*) laut Abschnitt der offiziellen Klassifikation der Wirtschaftszweige 2008 (Statistisches Bundesamt 2009). Zusätzlich sind die Schichtungsmerkmale (*grkl_admin*, *WZ_ziehung*) als auch der Standort (*region*) differenziert nach Ost- und Westdeutschland auf Basis der Angaben in der Betriebsdatei enthalten. Die klassierten Variablen *grkl_fb_aktuell* und *grkl_fb_feb2020* wurden anhand der Angaben in *pe01100* und *pe02000* generiert.

Tabelle 7: Fortgeschriebene Variablen

Variablenname	Variablenlabel	Anmerkungen
betriebeinheit01100	Kontaktierung der richtigen Betriebseinheit	
standort01100	Einziger Standort oder mehrere Betriebsstätten?	
standort02100	Zentrale oder Niederlassung?	
pe01100	Zahl der Beschäftigten laut Auskunftsperson im Februar 2020	
pe04101	Ausbildungsaktiver Betrieb	Statt einer Fortschreibung in pe04101 wurde anhand der Informationen aus pe02000 (Anzahl Beschäftigte), pe04100 (Ausbildungsaktiv ja/nein) und pe05000 (Anzahl Auszubildende aktuell) für alle Befragten die Indikatorvariable „ausbildungsbetrieb“ generiert, welche den Status zur Ausbildungsaktivität zum jeweiligen Befragungszeitpunkt widerspiegelt.
pe09100	Ausbildungsberechtigter Betrieb	Erst ab Welle 8 kontinuierlich erhoben.
struk02100	Qualifikationsstruktur: Beschäftigte für einfache Tätigkeiten	
struk03100	Zahl der Beschäftigten für einfache Tätigkeiten	
struk04101	Exportbetrieb	Erst ab Welle 2 einheitlich erhoben und entsprechend fortgeschrieben.
struk05100	Ausländisches Eigentum	
struk06100	Betriebsrat	
wa01100	Richtung der wirtschaftlichen Auswirkungen bis Ende Mai 2020	
wa02100	Stärke der negativen wirtschaftlichen Auswirkungen bis Ende Mai 2020	
wa06100	Stärke der positiven wirtschaftlichen Auswirkungen bis Ende Mai 2020	
wa26100	Entwicklung Geschäftsvolumen 2020 im Vergleich zu 2019	Erst ab Welle 13 kontinuierlich erhoben.

4.2 Datenprüfung und Datenbereinigung

Ein Teil der Datenprüfung erfolgt automatisch über die vom Befragungsinstitut verwendete Software während der telefonischen Befragung. Diese stellt sicher, dass alle Angaben ausschließlich in den zuvor definierten Wertebereichen liegen. Des Weiteren wurden für einige Angaben Plausibilisierungsroutinen implementiert. Beispielsweise nennen erstbefragte Betriebe stets die Anzahl an Beschäftigten vor der Krise (Februar 2020) als auch zum aktuellen Zeitpunkt der Befragung. Liegen zwischen beiden Angaben zu hohe Abweichungen, muss der Betrieb die Validität in einer Prüffrage erneut bestätigen oder seine Angaben korrigieren.

Während des Datenerhebungsprozess wurden verschiedene Typen von Antwortausfällen mit -7/9997 „trifft nicht zu“, -8/9998 „weiß nicht“, -9/9999 „keine Angabe“ und weiteren Zahlencodes gekennzeichnet. Im Zuge der Aufbereitung wurden diese in entsprechende Missing-Kategorien umgewandelt. Einen Überblick zeigt Tabelle 8. Die Zuordnung in Missingcodes gegenüber der Verwendung von Zahlencodes soll vor allem Fehlern im Auswertungsprozess vorbeugen. Zusätzlich erhöht die inhaltlich differenzierte Kennzeichnung von Antwortausfällen das Analysepotenzial der Daten. Es ermöglicht

beispielsweise Beobachtungen mit „Weiß nicht“ oder anderen Ausfallgründen, separiert in Auswertungen einzuschließen.

Tabelle 8: Klassifikation fehlender Werte

Missing-Code	Bezeichnung	Anmerkungen
.a	Weiß nicht	Wird entsprechend der Ausführung im Fragebogen vergeben.
.b	Keine Angabe	Wird entsprechend der Ausführung im Fragebogen vergeben.
.c	Trifft nicht zu	Betrifft wa05000/1. Wird entsprechend der Ausführung im Fragebogen vergeben.
.z	Frage nicht gestellt	Frage wurde aufgrund der Filterführung nicht gestellt.
.f	Filterfehler, Frage wurde nicht gestellt	Frage wurde bei wenigen Betrieben fälschlicherweise nicht gestellt. Betrifft wa11001, wa11021 kug15000 (Welle 9), struk09000 (Welle 10), wa26100 (Welle 14), wa26101 (Welle 16) und home31000 sowie home32000 (Welle 19). Der Fehler wurde während der Feldphase frühzeitig erkannt und behoben.
.h	Alle Beschäftigte	Betrifft ausschließlich home07000 und home09000 in Welle 4 und home04000/home04001/home04002 und home15000/home15001 in Welle 4, 8, 10, 11, 12, 15, 19 und 24. Wird vergeben, sofern der Betrieb keine konkrete Zahl, sondern mit „Alle Beschäftigte“ antwortet. Entspricht demnach der aktuellen Beschäftigtenzahl (pe02000).
.k	Kommt nicht vor	Betrifft ausschließlich verhalten01000a- verhalten01000h und verhalten02000a – verhalten02000h in Welle 6. Wird vergeben, sofern ein Betrieb angibt, dass das jeweilige Ereignis im Betrieb nicht vorkommt.
.l	Betrieb hat wegen Lockdownmaßnahmen (Nov bis 15.12) schließen müssen	Betrifft wa05000/1b in Welle 7 und 8, und korrigiert die Filterführung nach Angaben im Fragebogen. Wird vergeben, sofern ein Betrieb in wa11000 „ja“ (Betrieb wurde teilweise/vollständig geschlossen) angegeben hat.
.l	Leih-/Zeitarbeit vollkommen unrealistisch	Betrifft atypisch08000_* in Welle 11. Wird entsprechend der Ausführung im Fragebogen vergeben.
.r	Für unseren Betrieb nicht relevant	Betrifft Fragen des Themenkomplex „auslrek“ in Welle 7.
.t	Befragung nach verschärftem Lockdown ab 16.12.2020, daher nicht gestellt	Betrifft wa11000, wa14000 und wa15000. Wird für alle Betriebe vergeben, die in Welle 7 erst nach dem 16.12.2020 befragt wurden, da ab diesem Zeitpunkt die zugehörige Frage aus inhaltlichen Gründen nicht mehr gestellt wurde.

Missing-Code	Bezeichnung	Anmerkungen
.v	Betrieb bildet derzeit nicht aus	Betrifft pe05001 ab Welle 6 und ausb03001 in Welle 7. Wird entsprechend der Ausführung im Fragebogen vergeben.
.g	Es gibt keine vereinbarte Arbeitszeit	Betrifft struk07000, struk09000, struk10000, struk12000 in Welle 10. Wird entsprechend der Ausführung im Fragebogen vergeben.
.t	Befragung vor Einfügen der Frage, daher nicht gestellt	Betrifft kug36000, kug37000, kug38000 in Welle 12. Wird für alle Betriebe vergeben, die in Welle 12 vor dem 21.04.2021 befragt wurden, da die Frage erst danach gestellt wurde.
.x	Nicht möglich	Betrifft org06000* in Welle 13. Wird entsprechend der Ausführung im Fragebogen vergeben.
.p	Gab es bereits vor der Pandemie	Betrifft psych02000* in Welle 14. Wird entsprechend der Ausführung im Fragebogen vergeben.
.g	Betrieb verweist auf inkonsequente Umsetzung (manchmal wird geprüft, manchmal nicht)	Betrifft dreig01000 in Welle 21. Wird entsprechend der Ausführung im Fragebogen vergeben.
.h	weniger als 1 Tag pro Woche	Betrifft home42000 in Welle 24. Wird entsprechend der Ausführung im Fragebogen vergeben.

4.3 Gewichtung

Das Panel enthält zwei vorläufige Hochrechnungsfaktoren, *hrf_kal* und *hrf_kal_refined*. *Hrf_kal* wurde zu Projektbeginn als einfaches Verfahren entwickelt, und bis März 2021 für mehrere Publikationen verwendet. Aufgrund einer deutlich höheren Teilnahmebereitschaft von Betrieben, die bei der Bundesagentur für Arbeit Kurzarbeitergeld abgerechnet haben, wurde der Hochrechnungsfaktor *hrf_kal_refined* entwickelt um diesen Effekt herauszurechnen und weitere Schwächen von *hrf_kal* zu beheben. Für die Analyse zu Fragestellungen mit Bezug zu Kurzarbeit ist daher der Hochrechnungsfaktor *hrf_kal_refined* empfohlen. Ansonsten erzielen die Hochrechnungsfaktoren mit Ausnahme von Analysen zur Kurzarbeit weitgehend ähnliche Ergebnisse.

Für *hrf_kal* wurde zunächst die Inklusionswahrscheinlichkeit eines Betriebs in der Stichprobe getrennt nach Schichten als Anteil der kontaktierten Betriebe an allen verfügbaren Betrieben berechnet. Zur anschließenden Nonresponse-Korrektur diente neben den Schichtungsmerkmalen (vgl. Tabelle 6) auch die Region (Ost/West). Abschließend erfolgte eine Kalibrierung anhand derselben Variablen.

Bei *hrf_kal_refined* wurden die verschiedenen Auswahlsschritte (Stichprobenziehung, nicht recherchierbare Telefonnummern, Sperrungen von Betrieben, doppelt vorhandene Telefonnummern, Nicht-Kontaktierung aufgrund der kurzen Felddauer, Nicht-Kontaktierbarkeit und die Kooperationsbereitschaft von Betrieben) schrittweise mit einer Serie von logistischen Regressionen modelliert, wobei zusätzlich auf eine verbesserte Nonresponse-Korrektur Wert gelegt wurde. Verwendete Kovariablen waren zusätzlich zu den in Tabelle 6 genannten Schichtungsmerkmalen der Wirtschaftszweig (17 Kategorien), die Region (Ost/West), die Zahl der sozialversicherungspflichtig Beschäftigten, das Alter seit

Betriebsgründung, das Vorliegen einer Meldelücke in den administrativen Daten im Jahr 2019, der Ziehungszeitpunkt der (Auffrischungs-)Stichprobe, die Gemeindegröße am Betriebsstandort und – besonders wichtig – die Zahl der Monate, die ein Betrieb zwischen April 2020 und Juni 2021 Kurzarbeitergeld abgerechnet hat. Die abschließende Kalibrierung erfolgte wie zuvor anhand der Schichtungsmerkmale und der Region.

Der Hochrechnungsfaktor *hrf_kal_refined* berücksichtigt, dass Betriebe die Kurzarbeit abgerechnet haben häufiger zur Teilnahme an der Befragung bereit sind. Diese Erkenntnis war ausschlaggebend für die Erstellung dieses verbesserten Hochrechnungsfaktors. Die Korrelation zwischen abgerechneter Kurzarbeit und Teilnahmebereitschaft ist über den Gesamtzeitverlauf der Studie jedoch schwankend, und sinkt gegen Ende der Studie ab. Daher gelinkt die vorgenommene Korrektur in den ersten Wellen zumeist besser als in späteren Wellen. Dies führt beim Vergleich von weit auseinanderliegenden Wellen eventuell zu mangelhafter Vergleichbarkeit der Ergebnisse im Zeitverlauf. Diese Einschränkung trifft für *hrf_kal* nicht zu.

Für weitere Details zur Gewichtung sei auf den Methodenbericht (im Erscheinen) verwiesen.

4.4 Verlinkung mit administrativen Betriebsdaten (BHP)

Für Betriebe, die einer Verlinkung durch administrative Daten der Bundesagentur für Arbeit zugestimmt haben, stehen zusätzlich Merkmale des Betriebs-Historik-Panels (BHP) aus den Jahren 2019 und 2020 zur Verfügung. Dies umfasst unter anderem allgemeine Informationen zur Betriebs- und Beschäftigtenstruktur. Ausführliche Informationen und Unterlagen zum Inhalt und Entstehungsprozess dieser Daten finden sich auf der Internetseite des Forschungsdatenzentrums: (<https://fdz.iab.de/unsere-datenprodukte/betriebsdaten/bhp/>).

Die BHP Daten aller verlinkungsbereiten Betriebe sind bei Datenzugang zum Panel „Betriebe in der Covid-19 Krise“ standardmäßig in einer separaten Datei enthalten. Über die Schlüsselvariable „*betnr_becovid*“ können beide Quellen zu einem Datensatz zusammengespield werden (siehe Datenbeispiel 1). Dabei ist folgendes zu beachten:

- Für 0,2% der befragten Betriebe ist eine Verlinkung nicht sinnvoll, da diese nicht für ihren eigenen Betrieb, sondern eine übergeordnete Einheit antworten. Die im Survey erhobenen Daten beziehen sich demnach nicht auf die selbe Untersuchungseinheit. Einen Indikator für diese Betriebe bildet die Variable „*betriebseinheit01100*“.
- Betriebe, welche nach dem Stichtag des BHP 2019 bzw. 2020 neu gegründet wurden, können nicht verlinkt werden.
- Betriebe, deren Betriebsnummer sich zwischen dem Stichtag des BHP und dem Erhebungszeitpunkt der Studie geändert hat können ebenfalls nicht verlinkt werden.

Insgesamt haben 92,57% der Betriebe einer Verlinkung mit administrativen Daten zugestimmt. Die Verlinkungsbereitschaft wurde stets zu Beginn des ersten Interviews erhoben, wobei im Bedarfsfall am Ende des Interviews erneut nachgefragt wurde. Betriebe, die zunächst eine Verlinkung ablehnten, wurden bei wiederholter Teilnahme erneut zu ihrer

Bereitschaft befragt. Betriebe, die einer Verlinkung zugestimmt haben, konnten diese jederzeit widerrufen. In den Datensätzen aller Einzelwellen ist stets der finale Status eines Betriebes zur Verlinkungsbereitschaft vermerkt und geltend.

Die Angaben aus den BHP-Daten können von den Daten aus der BeCovid-Befragung abweichen. Die Erforderlichkeit des BHPs bzw. jedes sensiblen Merkmales muss bei der Beantragung gesondert begründet werden. Es werden alle unten aufgeführten Merkmale bereitgestellt.

Basismerkmale

- Betriebsidentifikator (betnr_becovid)
- Jahr (jahr_bhp)
- Jahr des 1. Auftretens der Betriebsnummer (grd_jahr)
- Jahr des letzten Auftretens der Betriebsnummer (lzt_jahr)
- Bundesland des Arbeitsorts (ao_bula)
- Anzahl der Beschäftigten insgesamt (az_ges)
- Anzahl der Vollzeitbeschäftigten (az_ges_vz)
- Anzahl der geringfügig Beschäftigten (az_gf)
- Mittelwert imp. Bruttotagesentgelt Vollzeitbeschäftigte (te_imp_mw)

Thematische Variablenblöcke

- Struktur der Beschäftigten allgemein (az_ges az_f az_reg az_gf az_azubi az_atz az_vz 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_med te_p25 te_p75 te_imp_mw 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)

Sensible Merkmale BHP

Zusätzlich können folgende sensible Merkmale aus dem BHP einzeln beantragt werden:

- Kreis des Arbeitsorts (ao_kreis)
 - Erstes Auftreten Betriebsnummer (grd_dat)
 - Letztes Auftreten Betriebsnummer (lzt_dat)
 - Klassifikation der Wirtschaftszweige 2008 – 3-Steller (w08_3)
 - Klassifikation der Wirtschaftszweige 2008 – 5-Steller (w08_5)
-

Datenbeispiel 1: Verlinkung mit administrativen Betriebsdaten (2019)

```
use ${orig}\BeCovid_W1_v3.dta, clear
```

```
merge 1:1 betnr_becovid using ${orig}\BeCovid_W01-W24_v1_bhp_7520_v1_2019.dta, keep(1  
3)
```

5 Verwendungshinweise und Datennutzung

5.1 Vorbereitungen zur Datenanalyse

Für eine repräsentative Analyse der Daten empfiehlt sich die Nutzung des Gewichts „*hrf_kal_refined*“. In Stata können die Daten über folgende Code-Zeile vorbereitet werden:

Datenbeispiel 2: Vorbereitung zur Datenanalyse

```
svyset betnr_becovid [pw= hrf_kal_refined], strata(stratum_id)
```

Alle Einzelwellen können zu einem Paneldatensatz zusammengespielt werden. Als stetiges Identifikationsmerkmal eines Betriebs zwischen den Einzelwellen dient die Variable „*betnr_becovid*“. Es gilt zu beachten, dass lediglich ca. 60% pro Welle Panelbefragte sind (siehe Tabelle 4). Des Weiteren ist nicht gewährleistet, dass ein Betrieb kontinuierlich an jeder Folgewelle teilnimmt.

5.2 Stata-Auswertungsbeispiel

Durch das hochfrequente Paneldesign eignen sich die Befragungsdaten besonders für Zeitreihenanalysen. Im Datenbeispiel 3 wird eine Zeitreihe mit dem Anteil der Betriebe, die Arbeitskräfte einstellen, erzeugt und das Vorgehen anhand der Software „Stata“ demonstriert.

Datenbeispiel 3: Zeitreihe des Anteils an Betrieben die Arbeitskräfte eingestellt haben

```
//Erstellung einer Zeitreihe zur Einstellungsaktivität von Betrieben
```

```
*Installiere Ado estout für Befehl estpost (nur einmalig notwendig)
```

```
fdzinstall estout
```

```
*Local für maximale Welle
```

```
local weakt = 24
```

```
*Schätzung des Anteils an Betrieben, die in den letzten drei Wochen Arbeitskräfte eingestellt haben pro Welle (Merkmal liegt erst ab Welle 2 vor).
```

```
forvalues x = 2(1)`weakt' {
```

```
*Datensatz einlesen
```

```
if(`x' <= 9){
```

```
use "$orig\BeCovid_W`x'_v3.dta", clear
```

```

}

if(`x' <= 14 & `x' >= 10){

use "$orig\BeCovid_W`x'_v2.dta", clear

}

if(`x' >= 15){

use "$orig\BeCovid_W`x'_v1.dta", clear

}

*Variablenamen anpassen

if(`x' == 15){

rename hire07001 hire07000

}

*Setup der Gewichtung

svyset betnr_becovid [pw = hrf_kal_refined], strata(stratum_id)

*Schätzung

di "Ungewichtete Fallzahlen in Welle `x'"

tab hire07000, matcell(nobsW`x') //ungewichtete Fallzahlen

di "Gewichtete Fallzahlen in Welle `x'"

estpost svy: tab hire07000, ci percent //gewichtete Fallzahlen

*Initialisiere Ergebnismatrix

mat W`x' = `x' //Welle

mat eeW`x' = e(b) //Anteil

mat elbW`x' = e(lb) //Lower bound

mat eubW`x' = e(ub) //Upper bound

*Zusammenfügen der einzelnen Vektoren

mat resW`x' = W`x'', eeW`x'[1,1]', elbW`x'[1,1]', eubW`x'[1,1]', nobsW`x'[1,1]'

*Spaltennamen benennen

mat colnames resW`x' = wave share lowerbound upperbound nobs

```



```

}

*Erstelle Ergebnismatrix für alle Wellen für Datenschutzprüfung

local mn = "resW2"

forv x = 3(1)`weakt' {

local mn = "`mn'" + "\resW`x'"

}

mat resMat = `mn'

di "Ergebnismatrix (nobs=ungewichtete Anzahl an Beobachtungen):"

matrix list resMat

*Erstelle Variablen aus Ergebnismatrix

svmat resMat, names(col)

*Generiere Line-Plot (Anteil an Betrieben, die in den letzten drei Wochen Arbeitskräfte
eingestellt haben pro Welle)

gr tw (line share wave, lp(dash)) (rcap lowerbound upperbound wave), ylab(0(10)30)
legend(order(1 "in den letzten 3 Wochen eingestellt")) graphregion(fcolor(white)
lcolor(white)) xtitle("") ytitle("Anteil der Betriebe (gewichtet) in %")

xlabel(2 "25.08 - 04.09" 3 "14.9 - 25.09" 4 "05.10 - 16.10" 5 "26.10 - 09.11" 6 "16.11
- 27.11" 7 "07.12 - 18.12" 8 "18.01 - 29.01" 9 "08.02 - 19.02." 10 "01.03 - 15.03" 11
"22.03 - 08.04" 12 "19.04 - 30.04" 13 "10.05 - 27.05" 14 "07.06 - 18.06" 15 "05.07 -
20.07" 16 "02.08 - 17.08" 17 "06.09 - 20.09" 18 "04.10 - 19.10" 19 "02.11 - 16.11" 20
"01.12 - 14.12" 21 "13.01 - 28.01" 22 "01.03 - 16.03" 23 "02.05 - 17.05" 24 "13.06 -
30.06", angle(40)) title("Einstellungen während der Corona-Krise") ///

    subtitle("Betriebe in Deutschland (2020/08 - 2022/06)")

graph export ${log}/Einstellungen_w2_bis_w24.png, replace

```

5.3 Datenzugang und Zusatzmaterial

Das für externe Forscherinnen und Forscher zur Verfügung gestellte Datenprodukt enthält neben allen bislang erhobenen Einzelwellen des Panels „Betriebe in der Covid-19 Krise“ auch die Betriebsdaten aus dem Betriebs-Historik-Panel für die Jahre 2019 und 2020.

Die schwach anonymisierten Daten können ausschließlich im Rahmen von Gastaufenthalten am FDZ und anschließender Datenfernverarbeitung ausgewertet werden.

Für die Datennutzung ist zunächst ein Antrag beim FDZ einzureichen. Details zur Beantragung des Datensatzes und Datenverarbeitungsmöglichkeiten können den Internetseiten des FDZ entnommen werden.

Literatur

Fritsch, Michael, und Udo Brixy (2004): The Establishment File of the German Social Insurance Statistics. In: Schmollers Jahrbuch 124. 183 – 190.

Ganzer, Andreas; Schmidlein, Lisa; Stegmaier, Jens; Wolter, Stefanie (2020): Betriebs-Historik-Panel 1975-2019. (FDZ-Datenreport, 16/2020 (de)), Nürnberg, 113 S.

Statistisches Bundesamt (Hg) (2009): Klassifikation der Wirtschaftszweige. 2008, Wiesbaden.
URL: <https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/klassifikation-wz-2008.html> (Zugriff: 06.04.2021).

Impressum

FDZ-Datenreport 09|2022

Veröffentlichungsdatum

18. Oktober 2022

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/2022/DR_09-22.pdf

Dokumentation Version

BeCovid_W01-W24_DE_v1_dok1, DOI: 10.5164/IAB.FDZD.2209.de.v1

Datensatz Version

BeCovid W01-W24 v1, DOI: 10.5164/IAB.BeCovid_W01-W24.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

Patrick Gleiser

Telefon: (911) 179 7987

E-Mail: patrick.gleiser@iab.de