



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

FDZ-DATENREPORT

Dokumentation zu Arbeitsmarktdaten

13|2021 DE Panel „Betriebe in der Covid-19 Krise“ – 20/21

Eine Längsschnittstudie in deutschen Betrieben – Welle 1 - 14

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

Eine Längsschnittstudie in deutschen Betrieben – Welle 1 - 14

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	14
4.1	Datenorganisation	14
4.2	Datenprüfung und Datenbereinigung.....	15
4.3	Gewichtung	17
4.4	Verlinkung mit administrativen Betriebsdaten (BHP)	18
4.5	Anonymisierung.....	19
5	Verwendungshinweise und Datennutzung	20
5.1	Vorbereitungen zur Datenanalyse.....	20
5.2	Stata-Auswertungsbeispiel.....	20
5.3	Datenzugang und Zusatzmaterial	22

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	13
Tabelle 7: Fortgeschriebene Variablen	15
Tabelle 8: Klassifikation fehlender Werte	16

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. Bislang sind vierzehn Wellen (Zeitraum: 03.08.2020 – 18.06.2021) für externe Forscherinnen und Forscher verfügbar.

Details zu verschiedenen Zugangswegen, den Voraussetzungen der Nutzung sowie zur Antragstellung des Datensatzes können der FDZ-Internetseite unter <http://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. So far, fourteen waves (period: 08/03/2020 - 06/18/2021) 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 <http://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 werden 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 resultiert in einem aus 20 geplanten Wellen bestehendem Panel, das einen Zeitraum vom 03. August 2020 bis voraussichtlich Ende 2021 erfasst. In jeder Welle wurden 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 Forschungskoooperation mit der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) und dem Exzellenzcluster ECONtribute der Universitäten Bonn und Köln. In bislang drei 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 der ersten vierzehn Wellen der Betriebsbefragung (Zeitraum: 03. August 2020 bis 18. Juni 2021). 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. Weitere Wellen der Betriebsbefragung werden bis zum Ende der Studie (voraussichtlich Dezember 2021) in regelmäßigen Abständen über das Forschungsdatenzentrum veröffentlicht.

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

2 Datenbeschreibung

2.1 Kurzbeschreibung

Tabelle 1: Inhaltliche Charakteristika

Kategorie	Beschreibung
Aktuelle Datenversion	BeCovid 2020/2021
Themen/Merkmalgruppen	Für jede Welle erhoben: Betriebsstruktur Personalentwicklung Liquiditätsreserven Ausmaß und Art der wirtschaftlichen Auswirkungen durch die Corona Krise Kurzarbeit (ab Welle 3) In einzelnen Wellen erhoben: Arbeitsicherheit und Gesundheitsschutz Ausbildung und Ausbildungsprämie Homeoffice Weiterbildung Ex- und Importe Auswirkungen des Lockdown Digitalisierung Atypische Beschäftigung Arbeitsorganisation Betriebliche Maßnahmen für psychische Gesundheit
Untersuchungseinheit	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). Stichtag für Wellen 1 bis 4: 30. November 2019 Stichtag ab Welle 5: 28. Februar 2020
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 18. Juni 2021 (vierwöchiger 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. • deren Erstbefragung höchstens sieben Wellen zurückliegt. • 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; Pohlan, Laura; Robelski, Swantje; Roth, Duncan; Schierholz, Malte; Sommer, Sabine; Stegmaier, Jens; Tisch, Anita; Umkehrer, Matthias; Aminian, Armin (2021): "Panel „Betriebe in der Covid-19 Krise“ (IAB BeCovid) – Welle 01-14". Forschungsdatenzentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB). DOI: 10.5164/IAB.BeCovid_W01-W14.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 Pohlan, Swantje Robelski, Duncan Roth, Malte Schierholz, Sabine Sommer, Jens Stegmaier, Anita Tisch, Matthias Umkehrer, Armin Aminian (2021): Panel zu „Betriebe in der Covid-19 Krise – 20/21“. Eine Längsschnittstudie in deutschen Betrieben – Welle 1-14. FDZ-Datenreport, 13/2021 (de), Nürnberg. DOI: 10.5164/IAB.FDZD.2113.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_v2.dta	32, 33	2020	1.723	100%	0%
BeCovid_W2_v2.dta	35, 36	2020	1.556	53,92%	46,08%
BeCovid_W3_v2.dta	38, 39	2020	1.733	40,33%	59,67%
BeCovid_W4_v2.dta	41, 42	2020	1.794	38,97%	61,03%
BeCovid_W5_v2.dta	44, 45	2020	1.760	31,38%	68,62%
BeCovid_W6_v2.dta	47, 48	2020	2.001	38,98%	61,02%
BeCovid_W7_v2.dta	50, 51	2020	1.511	33,09%	66,91%
BeCovid_W8_v2.dta	3, 4	2021	1.999	54,23%	45,77%
BeCovid_W9_v2.dta	6, 7	2021	2.000	34,45%	65,55%
BeCovid_W10_v1.dta	9, 10, 11	2021	2.002	33,72%	66,28%
BeCovid_W11_v1.dta	12, 13, 14	2021	2.009	37,33%	62,67%
BeCovid_W12_v1.dta	16, 17	2021	2.001	30,03%	69,97%
BeCovid_W13_v1.dta	19, 20, 21	2021	2.001	35,52%	64,48%
BeCovid_W14_v1.dta	23, 24	2021	2.001	39,78%	60,22%

3 Datenerhebung

3.1 Inhalt und Methode

Die inhaltliche und methodische Konzeption der Panelbefragung „Betriebe in der Covid-19 Krise“ wird gemeinsam vom zuständigen Projektteam des IAB und verschiedenen Kooperationspartnern, insbesondere der BAuA, erarbeitet. Die Durchführung der Befragung erfolgt durch Kantar Public. Im Fokus der Befragung stehen die wirtschaftlichen und strukturellen Auswirkungen der Corona-Krise auf Betriebe in Deutschland, sowie deren Veränderung im zeitlichen Verlauf der Pandemie. Bislang wurden im Zeitraum zwischen dem 03. August 2020 und dem 18. Juni 2021 insgesamt 14 Einzelwellen erhoben. Tabelle 5 enthält Informationen zu Beginn und Dauer der jeweiligen Feldzeiten. In der Regel startete in einem drei- bis vierwöchigen Intervall die Erhebung einer neuen Welle. Die aktive Feldzeit pro Einzelwelle betrug zwei Wochen. Interviews wurden ausschließlich telefonisch (CATI) geführt.

Die Befragung erfolgt anhand eines rotierenden Paneldesigns. Hierzu werden Betriebe nach maximal sechs Teilnahmen, oder spätestens sieben Wellen nach ihrem Erstinterview, aus dem Panel rotiert. Dabei ist nicht relevant, wie häufig der Betrieb tatsächlich befragt wurde. So rotieren beispielsweise auch Betriebe aus dem Panel, die im Zeitraum von sieben Wellen nach ihrem Erstinterview lediglich zwei Mal erfolgreich befragt wurden. Des Weiteren werden 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 werden durch neu rekrutierte Betriebe ersetzt.

Die Teilnahme an der Befragung erfolgt 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 wird. Des Weiteren enthält das Schreiben Informationen zur Zielsetzung der Studie, als auch Hinweise zum Datenschutz.

Der Fragebogen enthält 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“ 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/de/FDZ_Establishment_Data/iab_becovid/IABBeCovid_W01-W14.aspx).

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 werden¹.

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	Universität Heidelberg / 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)
Führung/Leistungsbeurteilung, Tests, Homeoffice, KuG+	12	19.04. – 30.04.2021	Universität Heidelberg
Arbeitsorganisation	13	10.05. – 26.05.2021	Rheinische Friedrich-Wilhelms-Universität Bonn
Betriebliche Maßnahmen für psychische Gesundheit (+Homeoffice)	14	07.05 – 18.06.2021	BAuA

¹ Die Fragebögen sind über die Internetseite des FDZ (https://fdz.iab.de/de/FDZ_Establishment_Data/iab_becovid/IABBeCovid_W01-W14.aspx) verfügbar.

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 gilt²:

- 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.

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

² Eine detaillierte Beschreibung der Betriebsnummernlogik gibt Fritsch/Brix 2004, sowie die Internetseite der Bundesagentur für Arbeit (https://www.arbeitsagentur.de/datei/dok_ba015200.pdf)

Stichprobe zu gelangen. Diese Disproportionalität wird im Rahmen der Gewichtung wieder ausgeglichen.

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 = 30551) zusätzlich von der Ziehung ausgeschlossen. Ab der fünften 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.

Aufgrund der langen Gesamtfeldzeit wurde die Stichprobe in Tranchen aufgeteilt. Diese wurden zu unterschiedlichen Zeitpunkten gezogen und sukzessiv an Kantar übermittelt. Dort erfolgte für alle Betriebe eine Telefonnummernrecherche. Insgesamt konnten 83,49% der gezogenen Betriebe erfolgreich eine Telefonnummer zugeordnet werden. Anschließend wurden sowohl doppelte Telefonnummern (2,34%) als auch jene, die sich bereits auf einer Kontakt-Sperrliste des Erhebungsinstituts befinden (6,65%), aussortiert. Zusätzlich wurden im Erstinterview anhand eines Screenings erneut Betriebe des öffentlichen Dienstes aus der Stichprobe entfernt (0,014%). Insgesamt waren ca. 75,41% 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.
- **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*), den Wellenidentifikator (*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	Einziges 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	

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) und 26100 (Welle 14). Der Fehler wurde während der Feldphase frühzeitig erkannt und behoben.
.h	Alle Beschäftigte	Betrifft ausschließlich home07000 und home09000 in Welle 4 und home04000, und home15000 in Welle 4, 8, 10, 11 und 12. 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 „ausrek“ 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.
.v	Betrieb bildet derzeit nicht aus	Betrifft pe05001 ab Welle 6 und ausb03001 in Welle 7. Wird entsprechend der Ausführung im Fragebogen vergeben.

Missing-Code	Bezeichnung	Anmerkungen
.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 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.

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. Die Hochrechnungsfaktoren erzielen 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 und August 2020 Kurzarbeitergeld abgerechnet hat. Die abschließende Kalibrierung erfolgte wie zuvor anhand der Schichtungsmerkmale und der Region.

Der Hochrechnungsfaktor *hrf_kal_refined* berücksichtigt, dass Betriebe, die im Frühjahr/Sommer 2020 Kurzarbeit abgerechnet haben, häufiger zur Teilnahme an der

Befragung bereit sind. Diese Erkenntnis war ausschlaggebend für die Erstellung dieses verbesserten Hochrechnungsfaktors. Jedoch besteht die Annahme, dass die Korrelation zwischen abgerechneter Kurzarbeit und Teilnahmebereitschaft zwar in den ersten Wellen (August/September 2020) besonders stark war, in späteren Wellen jedoch nachlässt. Daher ist davon auszugehen, dass die vorgenommene Nonresponse-Korrektur in den ersten Wellen besser gelingt als bei Wellen, die erst mehrere Monate später durchgeführt wurden. Dies führt möglicherweise, wenn weit auseinanderliegende Wellen verglichen werden sollen, zu mangelhafter Vergleichbarkeit der Ergebnisse im Zeitverlauf. Diese Einschränkung trifft für *hrf_kal* nicht zu.

Aufgrund der oben genannten Ausführungen dienen *hrf_kal* und *hrf_kal_refined* lediglich als vorläufige Hochrechnungsfaktoren. Endgültige Hochrechnungsfaktoren werden zum Ende der Erhebung angestrebt.

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) zur Verfügung. Dies umfasst unter anderem allgemeine Informationen zur Betriebs- und Beschäftigtenstruktur. Seit dem 2. Quartal 2021 sind Querschnittsdaten des BHP aus dem Jahr 2019 (Stichtag: 30. Juni) verfügbar. Ausführliche Informationen und Unterlagen zum Inhalt und Entstehungsprozess dieser Daten finden sich auf der Internetseite des Forschungsdatenzentrums:

https://fdz.iab.de/de/FDZ_Establishment_Data/Establishment_History_Panel.aspx.

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*“ können beide Quellen zu einem Datensatz zusammengespielt werden (siehe Datenbeispiel 1). Dabei ist folgendes zu beachten:

- Für 0.2 Prozent 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 „betriebeinheit01100“.
- Betriebe, welche nach dem Stichtag des BHP 2019 neu gegründet wurden, können nicht verlinkt werden.
- Betriebe, deren Betriebsnummer sich zwischen dem Stichtag des BHP und dem Erhebungszeitpunkt der Studie geändert hat³, können ebenfalls nicht verlinkt werden.

Insgesamt haben 92.85 Prozent 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,

³ Eine detaillierte Beschreibung der Betriebsnummernlogik gibt Fritsch/Brixy 2004, sowie die Internetseite der Bundesagentur für Arbeit (https://www.arbeitsagentur.de/datei/dok_ba015200.pdf)

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 BHP-Ergänzung wird auf begründeten Antrag zur Verfügung gestellt. Die enthaltenen Merkmale finden sich unter http://doku.iab.de/fdz/access/BHP_Variablen.pdf.

Datenbeispiel 1: Verlinkung mit administrativen Betriebsdaten

```
use {orig}\BeCovid_W1_v2.dta, clear
```

```
merge 1:1 betnr using {orig}\BeCovid_W01-W14_v1_bhp_7519_v2_2019.dta, keep(1 3)
```

4.5 Anonymisierung

Aus datenschutzrechtlichen Gründen sind sowohl die Daten des Panels zu „Betrieben in der Covid-19 Krise“, als auch des BHP standardmäßig schwach anonymisiert. Besonders sensible Merkmale, welche eine Rückführung auf den befragten Betrieb ermöglichen, sind nur in vergrößerter Form verfügbar. Dies betrifft die Angaben zum Arbeitsort (Bundesland) und zum Wirtschaftszweig (3-Steller). Im Bedarfsfall können diese Informationen auch im Original zur Verfügung gestellt werden. Dies muss unter expliziter Angabe des Analyseziels im Antrag auf Datenzugang vermerkt werden. Diese aus datenschutzrechtlicher Sicht besonders sensiblen Merkmale sind:

Administrative Betriebsdaten:

- Arbeitsort - Kreis (ao_kreis)
- Wirtschaftszweig - Wirtschaftsunterklasse (w08_5)
- Erstes Auftreten Betriebsnummer (grd_dat)
- Letztes Auftreten Betriebsnummer (lzt_dat)

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 [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“. 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 Anteil 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 = 14
```

```
*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'_v2.dta", clear
```

```
}
```

```

else{

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

}

*Setup der Gewichtung

svyset betnr [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",angle(40)) title("Einstellungen  
während der Corona-Krise") subtitle("Betriebe in Deutschland (2020/21)")
```

```
graph export ${log}/Einstellungen_w2_bis_w14.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 das Jahr 2019.

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 13|2021

Veröffentlichungsdatum

27. September 2021

Herausgeber

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

Rechte

Nachdruck - auch auszugsweise - nur mit Genehmigung des FDZ gestattet

Bezugsmöglichkeit dieses Dokuments

http://doku.iab.de/fdz/reporte/2021/DR_13-21.pdf

Dokumentation Version

BeCovid_W01-W14_DE_v1_dok1, DOI: 10.5164/IAB.FDZD.2113.de.v1

Datensatz Version

BeCovid W01-W14 v1, DOI: 10.5164/IAB.BeCovid_W01-W14.de.en.v1

Bezugsmöglichkeit aller Veröffentlichungen der Reihe „FDZ-Datenreport“

https://fdz.iab.de/de/FDZ_Publications/FDZ_Publication_Series/FDZ-Datenreporte.aspx

Website

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

Rückfragen zum Inhalt

Patrick Gleiser

Telefon: (911) 179 7987

E-Mail: patrick.gleiser@iab.de