

Methodenbericht

Deutscher Alterssurvey (DEAS) Schriftliche Panelbefragung 2020

Autoren: Stefan Schiel, Jonathan Ruiz Marcos, Moritz Fahrenholz

infas Institut für angewandte
Sozialwissenschaft GmbH

Friedrich-Wilhelm-Straße 18
D-53113 Bonn
Tel. +49 (0)228/38 22-0
Fax +49 (0)228/31 00 71
info@infas.de
www.infas.de

Bericht an

Deutsches Zentrum für Altersfragen (DZA)
Manfred-von-Richthofen-Straße 2
12101 Berlin

Vorgelegt von

infas Institut für angewandte
Sozialwissenschaft GmbH
Friedrich-Wilhelm-Straße 18
53113 Bonn

Kontakt

Stefan Schiel
Senior-Projektleiter Sozialforschung

Tel.: 0228/3822-424

Fax: 0228/310071

E-Mail: s.schiel@infas.de

Autoren

Stefan Schiel, Jonathan Ruiz Marcos, Moritz Fahrenholz

Projekt

7348
Bonn, Oktober 2020
SI, RJ

© infas Institut für angewandte Sozialwissenschaft GmbH
Der Inhalt dieses Berichts darf ganz oder teilweise
nur mit unserer schriftlichen Genehmigung veröffentlicht,
vervielfältigt, gedruckt oder in Informations- und
Dokumentationssystemen (information storage and
retrieval systems) gespeichert, verarbeitet oder ausgegeben
werden.

infas ist zertifiziert
nach ISO 20252 für die Markt-,
Meinungs- und Sozialforschung

ISO 20252



infas ist Mitglied im
Arbeitskreis Deutscher Markt- und
Sozialforschungsinstitute e.V.
(ADM) und ESOMAR

ADM

ESOMAR
member

Zusammenfassung

Mit dem Forschungsvorhaben „Deutscher Alterssurvey – Die zweite Lebenshälfte“ (DEAS) werden seit 1996 die Lebenssituation, Einstellungen und Bedürfnisse älterer Menschen ab einem Alter von 40 Jahren erforscht. Die für den Frühling und Sommer 2020 geplante siebte Erhebungswelle als kombinierte Basis- und Panelerhebung konnte aufgrund der aktuellen Corona-Pandemie nicht, wie ursprünglich geplant, stattfinden und wurde daher auf einen späteren Zeitpunkt verschoben.

Um die Veränderungen in verschiedenen Lebensbereichen in Folge der durch die Corona-Pandemie erfolgten Kontaktbeschränkungen zwischen März und Mai 2020 zu messen, hat das Deutsche Zentrum für Altersfragen (DZA) infas mit einer schriftlichen Befragung der Panelteilnehmerinnen und -teilnehmer im Juni 2020 beauftragt.

Da die Ergebnisse der Befragung möglichst zeitnah ausgewertet werden sollten, war eine vergleichsweise kurze Feldzeit von 5 Wochen vorgesehen. Im Zeitraum vom 08.06.2020 bis zum 22.07.2020 haben über die Hälfte aller Panelteilnehmerinnen und -teilnehmer einen ausgefüllten Fragebogen an infas zurückgesendet. Nach Prüfung der Daten liegen dem DZA 4.823 Fragebögen für die Auswertung vor. Dies entspricht einer Ausschöpfungsquote von 55,1 Prozent.

Im Anschreiben zum schriftlichen Fragebogen wurde auch die für den Herbst/Winter 2020/2021 geplante telefonische Panelbefragung angekündigt. Um die Erreichbarkeit der Personen der Panelstichprobe zu verbessern, wurde um Mitteilung der aktuellen Telefonnummer gebeten. Insgesamt haben 1.235 Personen eine neue Rufnummer mitgeteilt, so dass die Rufnummerndichte in der Panelstichprobe auf 93 Prozent angehoben werden konnte.

Der vorliegende Methodenbericht dokumentiert das Erhebungsdesign, die Stichprobe, die Durchführung der Erhebung und das Feldergebnis sowie die Datenprüfung und -gewichtung.

Inhaltsverzeichnis

1	Gegenstand	6
2	Erhebungsdesign	7
3	Panelstichprobe	8
4	Erhebungsunterlagen	9
4.1	Der schriftliche Fragebogen	9
4.2	Mitteilungskarte	9
4.3	Anschreiben	9
4.4	Datenschutzerklärung	9
4.5	Begleitschreiben des Ministeriums	10
4.6	Erinnerungsaktion	10
5	Feldzeit	11
6	Ergebnis	12
6.1	Fragebogenrücklauf	12
6.2	Stichprobenausschöpfung	13
6.3	Panelbereite Zielpersonen (Stand Ende September 2020)	14
6.4	Mitteilung neuer Telefonnummern	15
7	Datenerfassung und -prüfung	16
7.1	Rücklauferfassung	16
7.2	Datenerfassung	16
7.3	Datenprüfung	16
8	Gewichtung	18
8.1	Längsschnittgewichtung	18
8.2	Integration der Stichproben in einen gemeinsamen Querschnitt	19
8.3	Poststratifizierte Querschnittgewichtung	22
8.4	Kennziffern der Gewichtungsfaktoren	22
	Anhang	25

Tabellenverzeichnis

Tabelle 1	Zusammenfassung des Studiendesigns	7
Tabelle 2	Entwicklung der Panelstichprobe zwischen Panelpflege 2019 und Feldstart der schriftlichen Befragung im Juni 2020	8
Tabelle 3	Fragebogenrücklauf – Ersteinsendung	12
Tabelle 4	Fragebogenrücklauf – Erste und zweite Rücksendung	12
Tabelle 5	Bruttoausschöpfung (zum Feldende am 22.07.2020)	13
Tabelle 6	Bereinigte Stichprobenausschöpfung	14
Tabelle 7	Umfang Panelstichprobe (Stand Ende September 2020)	15
Tabelle 8	Kennziffern der Gewichtung	23
Tabelle 9	Ausfallmodellierung Panelstichprobe: Interview in 2020, Basis: Fälle mit Interview in 2017 (logistische Regression)	25
Tabelle 10	Ausfallmodellierung Panelstichprobe temporäre Ausfälle: PAPI-Interview in 2020, Basis: Fälle mit Nichtteilnahme in 2017 (logistische Regression)	26
Tabelle 11	Verteilungsvergleich ungewichtete und kalibrierte Stichprobe	27

1 Gegenstand

Das Deutsche Zentrum für Altersfragen (DZA) hat infas mit der Vorbereitung und Durchführung der 7. Welle des Deutschen Alterssurveys als kombinierte Basis- und Panelerhebung für das Jahr 2020 beauftragt. Aufgrund der aktuellen Corona-Pandemie kann die Befragung im Jahr 2020 jedoch nicht, wie ursprünglich geplant, mit persönlich-mündlichen CAPI-Interviews stattfinden. Für den Herbst/Winter 2020/21 ist eine telefonische Befragung der Panelstichprobe als Alternative zum Face-to-Face-Interview vorgesehen. Sofern es die Situation bzgl. der Corona-Pandemie 2021 zulässt, soll im nächsten Jahr die Befragung der Basisstichprobe als persönlich-mündliches CAPI-Interview stattfinden.

Der Beginn der ursprünglich für das Jahr 2020 geplanten Erhebung war den Panelteilnehmerinnen und -teilnehmern im Vorfeld für das Frühjahr 2020 angekündigt worden, zuletzt im Rahmen der Panelpflege 2019. Kurz bevor die Ankündigungsschreiben zum Start der anstehenden Befragung im April 2020 verschickt werden sollten, wurde die Face-to-Face-Befragung im März aufgrund der aktuellen Entwicklung der Corona-Pandemie in Abstimmung mit dem DZA ausgesetzt.

Die Panelteilnehmerinnen und -teilnehmer sind im Juni über die geänderte Planung aufgrund der Pandemie informiert worden. In dem Anschreiben wurde ihnen mitgeteilt, dass die geplante Befragung auf den Herbst verschoben und in diesem Jahr telefonisch durchgeführt wird. Dem Anschreiben lag eine Mitteilungskarte bei, die für eine Übermittlung einer Rufnummer für das Telefoninterview genutzt werden konnte sowie für eine mögliche Änderung oder Korrektur der Adresse. Zusätzlich wurden die Panelteilnehmerinnen und -teilnehmer gebeten, einen 20-seitigen Fragebogen auszufüllen.

Mit dem Fragebogen sollte die Veränderung des Alltags, der Freizeitgestaltung und anderer Lebensbereiche in Folge der Kontaktbeschränkungen im Rahmen der Corona-Pandemie im März, April und Mai 2020 gemessen werden.

2 Erhebungsdesign

Mit der schriftlichen Befragung zu den Auswirkungen der Corona-Pandemie war zugleich eine Pflege der Panelstichprobe verbunden, nachdem der geplante Start der Hauptbefragung von Ende April auf den Herbst 2020 verschoben werden musste. Aus zwei Gründen sollte die schriftliche Befragung möglichst zeitnah starten. Zum einen sollten die Zielpersonen über die Verschiebung der Hauptbefragung informiert werden. Zum anderen sollten die Lebensumstände und Veränderungen im Alltag der Panelteilnehmerinnen und -teilnehmer seit Mitte März (Beginn der verschärften Kontaktbeschränkungen) mit möglichst geringem zeitlichen Abstand ermittelt werden. Um möglichst zügig mit der Erhebung starten zu können wurde auf die Durchführung eines Pretest verzichtet. In Tabelle 1 ist das Studiendesign der schriftlichen Befragung zusammengefasst dargestellt.

Tabelle 1 Zusammenfassung des Studiendesigns

Einsatzstichprobe	<ul style="list-style-type: none"> – Nach der Panelpflege 2019 noch panelbereite Zielpersonen aus den 4 Panelstichproben P1996, P2002, P2008, P2014 – Umfang der Bruttostichprobe: 8.754 Adressen
Erhebungsmethode	Schriftlich-postalische Befragung
Erhebungsinstrument	20-seitiger Papierfragebogen, Selbstausfüller
Erhebungsmethode	Deutsch
Erhebungsmaterialien	<ul style="list-style-type: none"> – Gemeinsames Anschreiben von infas und DZA – Gemeinsame Datenschutzerklärung von infas und DZA – Begleitschreiben des Bundesministeriums für Familie, Senioren, Frauen und Jugend (BMFSFJ) – Rückumschlag (groß) für den Fragebogen – Mitteilungskarte zur Adressaktualisierung und Mitteilung einer Telefonnummer, mit Rückumschlag (klein) – Kombiniertes Dank- und Erinnerungsschreiben
Rückversicherungsmöglichkeit	<ul style="list-style-type: none"> – Kostenfrei Hotline – Studienspezifische E-Mail-Adresse – Studienankündigung über infas-Website
Erhebungszeitraum	08.06.2020 bis 22.07.2020
Rücklauf	<p>Insgesamt: n=5.388 erfasste Fragebögen, davon</p> <ul style="list-style-type: none"> – n=4.855 Ersteinsendungen – n=543 doppelte Einsendungen <p>Eingang nach Feldende: n=95, davon</p> <ul style="list-style-type: none"> – n=80 Ersteinsendungen – n=15 doppelte Einsendungen
Stichprobenausschöpfung	<ul style="list-style-type: none"> – Bruttoausschöpfung: 55,4 Prozent (innerhalb der Feldzeit) – Bruttoausschöpfung: 56,3 Prozent (inkl. Rücklauf nach Feldende)
Datenerfassung	<ul style="list-style-type: none"> – Erfassung der Mitteilungskarten – Manuelle Datenerfassung aller ausgefüllten Fragebögen (Einfachfassung)
Datenaufbereitung und -prüfung	Rücklaufkontrolle, formale Datenprüfung, Erstellung eines Stata- und SPSS-Datensatzes inklusive Labelung
Datenlieferung	<ul style="list-style-type: none"> – Befragungsdatensatz – Quer- und Längsschnittgewichte – Versand der ausgefüllten Fragebögen an das DZA

3 Panelstichprobe

Die Bruttostichprobe der schriftlichen Befragung setzt sich ausschließlich aus panelbereiten Zielpersonen der vorangegangenen DEAS-Befragungen zusammen. In Vorbereitung der für 2020 geplanten Haupterhebung wurde im Jahr 2019 eine Panelpflege durchgeführt, bei der um Mitteilung von Adressänderungen gebeten wurde. Bei unzustellbaren Anschreiben wurden jeweils die aktuellen Adressen der Zielpersonen bei Einwohnermeldeämtern recherchiert. Zwischen dem im Panelpflegebericht dokumentierten Stand im Januar 2020 und dem Anschreibenversand im Juni 2020 ist eine Person ins Ausland verzogen. Die Einsatzstichprobe für die schriftliche Befragung umfasste n=8.754 Personen.

Tabelle 2 Entwicklung der Panelstichprobe zwischen Panelpflege 2019 und Feldstart der schriftlichen Befragung im Juni 2020

Spalten%	Panelstichprobe 1996		Panelstichprobe 2002		Panelstichprobe 2008		Panelstichprobe 2014		Gesamt	
	abs.	%	abs.	%	abs.	%	abs.	%	abs.	%
Panelpflege Stand Januar 2020	1.013	100,0	996	100,0	2.757	100,0	3.989	100,0	8.755	100,0
Ins Ausland verzogen	-	-	-	-	-	-	1	0,0	1	0,0
Einsatzstichprobe Juni 2020	1.013	100,0	996	100,0	2.757	100,0	3.988	100,0	8.754	100,0

Quelle: infas-Panelstichprobendatei

4 Erhebungsunterlagen

4.1 Der schriftliche Fragebogen

Der schriftliche Fragebogen wurde vom Deutschen Zentrum für Altersfragen (DZA) entworfen und in einer gemeinsamen Endredaktion von infas und DZA finalisiert. Das Fragenprogramm wurde von infas in ein professionelles Layout überführt, das sich am Layout des Drop-Off-Fragebogens der Hauptbefragung orientiert. Es wurde für die schriftliche Befragung überarbeitet und etwas moderner gestaltet. Die gelbe Farbe, die den Zielpersonen seit vielen Wellen als Erkennungszeichen des schriftlichen Drop-off im Alterssurvey bekannt ist wurde auf der Titelseite und für die Ausfüllhinweise auf der Seite 2 weiterhin als Hintergrundfarbe genutzt. Auf den Folgeseiten wird gelb als Hintergrund für die Fragennummern und Filter eingesetzt. Die Farbe Rot aus dem DZA-Logo findet sich als Trennlinie zwischen den Fragen und bei den Ausfüllhinweisen und Filtern wieder. Die Texte sind weiterhin in größerer Punktgröße gesetzt, um den Menschen mit Sehproblemen das Lesen zu erleichtern. Der insgesamt 20-seitige Fragebogen umfasst 18,5 Seiten mit inhaltlichen Fragen. Für den Rückversand des ausgefüllten Fragebogens lag ein C4 Großumschlag mit Aufdruck „Porto zahlt Empfänger“ bei, der an infas adressiert war.

4.2 Mitteilungskarte

Zur Vorbereitung der telefonischen Befragung ab Herbst 2020 lag dem Anschreiben neben dem Fragebogen auch eine Karte zur Mitteilung von Telefonnummern und Adressänderungen bei. Auf der Karte waren die verschiedenen Möglichkeiten für die Rückmeldung an infas aufgeführt (per Karte, Telefon, E-Mail oder Online-Adressportal). Das Formular auf der Rückseite konnte zur Mitteilung von Telefonnummern und Adressänderungen genutzt werden. Für den Rückversand lag ein an infas adressierter Rückumschlag mit Aufdruck „Porto zahlt Empfänger“ bei.

4.3 Anschreiben

Dem Fragebogen und der Mitteilungskarte lag außerdem ein von infas und dem DZA unterzeichnetes Anschreiben bei. Mit dem Anschreiben wurden die Panelpersonen über die Verschiebung der regelmäßigen Befragung auf den Herbst und über den Wechsel der Durchführungsart (per Telefon statt persönlich-mündlich) informiert. Zudem wurden sie gebeten, den Fragebogen auszufüllen. Bei der Entwicklung des Anschreibens wurde besonderen Wert darauf gelegt den Moduswechsel zu erläutern und die Verschiebung der regelmäßigen Befragung in den Herbst anzukündigen. Gleichfalls sollte das Anschreiben über die Zielsetzung der schriftlichen Befragung aufklären und zum Ausfüllen motivieren. Der Inhalt des Anschreibens wurde eng zwischen infas und DZA abgestimmt.

4.4 Datenschutzerklärung

Den Unterlagen lag eine gemeinsame Datenschutzerklärung von infas und DZA bei.

4.5 Begleitschreiben des Ministeriums

Um die Wichtigkeit der Erhebung und das öffentliche Interesse zu unterstreichen lag – wie in den vorangegangenen Befragungen – wieder ein Begleitschreiben des Bundesministeriums für Familie, Senioren, Frauen und Jugend (BMFSFJ) bei.

4.6 Erinnerungsaktion

Als ausschöpfungssteigernde Maßnahme wurde nach rund 2 Wochen an alle Panelpersonen, die in der Zwischenzeit nicht als verstorben oder ins Ausland verzogen gemeldet wurden bzw. ihre Panelbereitschaft zurückgezogen haben, ein kombiniertes Dank- und Erinnerungsschreiben gesendet (n=8.727). Die Zielpersonen wurden gebeten den Fragebogen möglichst bis zum 07.07.2020 an infas zurückzusenden. Allen, die den Bogen bereits ausgefüllt zurückgeschickt hatten, wurde gleichfalls gedankt. Im Anschreiben wurde erneut die telefonische Durchführung der auf den Herbst verschobenen regelmäßigen Befragung angekündigt.

Dem Anschreiben lagen – analog zum Erstversand – ein Fragebogen, eine Datenschutzerklärung, eine Mitteilungskarte und jeweils ein großer (für den Fragebogen) und ein kleiner (für die Mitteilungskarten) an infas adressierter Rückumschlag bei. Lediglich das Begleitschreiben des Ministeriums wurde nicht erneut beigelegt.

5 Feldzeit

Mit der schriftlichen Befragung der Panelteilnehmerinnen und -teilnehmer sollten möglichst rasch Daten vorliegen, die die Erfahrungen und unmittelbaren Veränderungen in den ersten Wochen der Corona-Pandemie seit Mitte März abbilden. Vor diesem Hintergrund wurde eine kurze Feldzeit von rund 5 Wochen mit einer Erinnerungsaktion geplant. Die Erhebungsunterlagen wurden am 08.06.2020 (Kalenderwoche 24) an die Panelteilnehmerinnen und -teilnehmer gesendet. Zur Erinnerung an das Ausfüllen des Fragebogens wurde in der Mitte der vorgesehenen Feldzeit am 22.06.2020 (Kalenderwoche 26) ein kombiniertes Dank- und Erinnerungsschreiben verschickt. Da nach dem ursprünglich vorgesehenen Feldende am 10.07. (Kalenderwoche 28) immer noch ausgefüllte Fragebögen bei infas eingingen, verständigten sich das DZA und infas auf eine Verlängerung bis zum 22.07.2020 (Mitte Kalenderwoche 30).

6 Ergebnis

6.1 Fragebogenrücklauf

Der Eingang der Fragebögen wurde bei infas täglich (montags bis freitags) erfasst. Somit war eine tagesaktuelle Bewertung der Rückläufe durch die infas-Projektleitung möglich. Bis zum ursprünglich für den 10.07.2020 geplanten Feldende erreichten infas n=4.704 ausgefüllte Fragebögen. Da in der letzten Feldwoche täglich noch mehrere Fragebögen bei infas eintrafen, wurde die Feldzeit in Absprache mit dem DZA entsprechend verlängert. Aufgrund der feldbegleitenden Erfassung der Fragebögen konnten alle bis zum 22.07.2020 eingegangenen Fragebögen noch rechtzeitig erfasst und formal geprüft werden, ohne dass der geplante Termin für die Übergabe der Befragungsdaten an das DZA am 24.07.2020 gefährdet war. Zwischen dem 10.07. und dem 22.07.2020 erreichten infas weitere 141 Fragebögen. Somit konnten insgesamt n=4.845 Fragebögen erfasst und für die Auswertung berücksichtigt werden (siehe Tabelle 3).

Tabelle 3 Fragebogenrücklauf – Ersteinsendung

	Erste Rücksendung
bis zum 10.07.2020	4.704
bis zum 22.07.2020	141
Rücklauf insgesamt	4.845

Quelle: infas-Rücklaufdatenbank

Doppelte Einsendungen

Das Erhebungsdesign sah ein kombiniertes Dank- und Erinnerungsschreiben mit erneutem Fragebogenversand bereits zwei Wochen nach Versand der Erstan-schreiben vor. Im Zuge dieser Erinnerungsaktion haben rund 11 Prozent aller Personen, die bereits einen Fragebogen eingeschickt hatten, ein zweites Mal den Fragebogen ausgefüllt und zurück gesendet. Insgesamt n=543 bis zum 22.07.2020 eingegangene Zweiteinsendungen wurden ebenfalls erfasst und mit den Befragungsdaten an das DZA ausgeliefert (siehe Tabelle 4).¹

Tabelle 4 Fragebogenrücklauf – Erste und zweite Rücksendung

	Erste Rücksendung	Zweite Rücksendung	Gesamt
bis zum 10.07.2020	4.704	513	5.217
bis zum 22.07.2020	141	30	171
Rücklauf insgesamt	4.845	543	5.388

Quelle: infas-Rücklaufdatenbank

¹ Nach dem 22.07.2020 erreichten infas bis zur Berichtslegung weitere 95 Fragebögen, die nicht mehr erfasst wurden (davon 80 Erst- und 15 Zweitrücksendungen).

6.2 Stichprobenausschöpfung

Es wurden alle verbleibenden 8.754 Panelteilnehmerinnen und -teilnehmer angeschrieben und um die Beantwortung und Rücksendung des schriftlichen Fragebogens gebeten. Weit mehr als die Hälfte der Panelteilnehmerinnen und -teilnehmer schickte einen Fragebogen zurück. Bis zum Feldende am 22.07.2020 lagen insgesamt 4.857 Fragebögen vor (55,5 Prozent, siehe Tabelle 5). Abzüglich der 12 vollständig leeren Fragebögen wurden 4.845 ausgefüllte Fragebögen (55,4 Prozent) an das DZA ausgeliefert. Nach eingehender Prüfung durch das DZA wurden 22 Fragebögen von der weiteren Auswertung ausgeschlossen. Somit verbleiben n=4.823 Fragebögen für die inhaltliche Auswertung durch das DZA (55,1 Prozent, siehe Tabelle 5).

Tabelle 5 Bruttoausschöpfung (zum Feldende am 22.07.2020)

	<i>abs.</i>	<i>%</i>
Bruttostichprobe	8.754	100,0
Nicht mehr Zielgruppe	221	2,5
Zielperson verstorben	221	2,5
Keine Rückmeldung	3.093	35,3
Nonresponse – nicht erreicht	315	3,6
Erhebungsunterlagen unzustellbar zurück	314	3,6
Zielperson in Feldzeit nicht zu erreichen	1	0,0
Nonresponse – nicht befragbar	38	0,4
Zielperson (lt. Auskunft) nicht befragbar/dauerhaft krank/behindert	38	0,4
Nonresponse – verweigert	135	1,5
Verweigert: keine Zeit/dauert zu lange	1	0,0
Verweigert: Teilnahme nur persönlich bzw. nur telefonisch	1	0,0
Verweigert: krank	1	0,0
Grundsätzliche Verweigerung, Panelbereitschaft zurückgezogen	121	1,4
Verweigert: nicht in dieser Welle (temporärer Ausfall)	6	0,1
Verweigert: sonstige Gründe	5	0,1
Nonresponse – Sonstige Rückmeldungen der Zielpersonen	95	1,1
Rückgesendete Fragebögen	4.857	55,5
Fragebogen ausgefüllt, auswertbar	4.823	55,1
Fragebogen ausgefüllt, nicht auswertbar	22	0,3
Fragebogen leer zurück	12	0,1

Quelle: Rücklaufdatenbank, eigene Berechnungen

Die folgende Tabelle 6 weist die Ausschöpfung der um die neutralen Ausfälle bereinigten Stichprobe aus. Für 221 Personen ist bekannt, dass sie zum Befragungszeitpunkt verstorben waren. Abzüglich dieser Personen ergibt sich, eine bereinigte Rücksendequote von 56,9 Prozent. Alleine auf die auswertbaren Fragebögen bezogen ergibt sich eine bereinigte Ausschöpfungsquote von 56,5 Prozent (siehe Tabelle 6). Vor dem Hintergrund der erstmaligen rein schriftlichen Befragung der Panelstichprobe sowie der kurzen Feldzeit und der damit einher-

gehenden einmaligen Erinnerungsaktion ist der Rücklauf erfreulich hoch und die Ausschöpfung als sehr gut einzuschätzen.

Tabelle 6 Bereinigte Stichprobenausschöpfung

	<i>abs.</i>	%
Bruttostichprobe	8.754	100,0
Stichprobenneutrale Ausfälle, Zielperson verstorben	221	2,5
Bereinigte Stichprobe	8.533	100,0
Rücklauf insgesamt, davon	4.857	56,9
Ausgefüllte Fragebögen, nicht auswertbar	22	0,3
Fragebogen leer zurück	12	0,1
Auswertbare Fragebögen	4.823	56,5

Quelle: Rücklaufdatenbank, eigene Berechnungen

Nach dem Feldende gingen weitere 80 ausgefüllte Fragebogen bei infas ein (nicht in Tabelle 6 ausgewiesen). Zählt man diese Fragebögen zum Rücklauf dazu, steigt die Rücksendequote bezogen auf die Bruttostichprobe nochmal um 0,9 Prozentpunkte auf 56,4 Prozent. Bezogen auf die bereinigte Bruttostichprobe erhöht sich die Rücklaufquote von 56,9 auf 57,9 Prozent.

6.3 Panelbereite Zielpersonen (Stand Ende September 2020)

Während der Feldzeit wurden 221 Personen als verstorben gemeldet und 121 Personen zogen ihre Panelbereitschaft zurück. Nach dem Feldende wurde für jeweils 2 weitere Personen zurückgemeldet, dass sie verstorben bzw. ins Ausland verzogen sind. Darüber hinaus haben weitere 26 Personen nach dem Feldende ihre Panelbereitschaft zurückgezogen. Zuvor hatten diese 26 Personen noch an der schriftlichen Befragung teilgenommen und einen Fragebogen ausgefüllt. Insgesamt sind somit 223 verstorbene und 2 ins Ausland verzogene Personen sowie 147 Rücknahmen der Panelbereitschaft zu verzeichnen.

Nach Prüfung der Befragungsdaten durch das DZA wurde ein weiterer Fall aus dem Panel ausgeschlossen. Insgesamt reduziert sich die Panelstichprobe somit um 373 Fälle, so dass noch 8.381 panelbereite Zielpersonen – Stand Ende September 2020 – für die geplante telefonische Erhebung im Herbst 2020 verbleiben.

Tabelle 7 Umfang Panelstichprobe (Stand Ende September 2020)

Spalten%	Panelstichprobe 1996		Panelstichprobe 2002		Panelstichprobe 2008		Panelstichprobe 2014		Gesamt	
	abs.	%	abs.	%	abs.	%	abs.	%	abs.	%
Einsatzstichprobe Juni 2020	1.013	100,0	996	100,0	2.757	100,0	3.988	100,0	8.754	100,0
Zielperson verstorben	39	3,9	43	4,3	70	2,5	71	1,8	223	2,6
Zielperson ins Ausland verzogen	-	-	-	-	2	0,1	-	-	2	0,0
Zielperson zieht Panelbereitschaft zurück	16	1,6	16	1,6	45	1,6	70	1,8	147	1,7
Nach Prüfung realisierter Fälle vom DZA ausgeschlossen	-	-	-	-	1	0,0	-	-	1	0,0
Verbleibende Personen für Panelbefragung 2020	958	94,6	937	94,1	2.639	95,7	3.847	96,5	8.381	95,7

Quelle: infas-Panelstichprobendatei

6.4 Mitteilung neuer Telefonnummern

Im November 2020 soll die regelmäßig stattfindende persönliche Befragung der Panelstichprobe telefonisch durchgeführt werden. Vor Durchführung der schriftlichen Befragung im Juni 2020 lag für rund 90 Prozent der Panelstichprobe eine Telefonnummer vor.

Um möglichst viele Panelteilnehmerinnen und -teilnehmer im Herbst/Winter telefonisch erreichen und befragen zu können, wurden die Zielpersonen im Rahmen der schriftlichen Befragung gebeten ihre Telefonnummer der infas-Hotline, per E-Mail, online über das infas-Adresstool oder durch den Rückversand der ausgefüllten Mitteilungskarte (s. Abschnitt 4.2) anzugeben. Die Mitteilungskarte lag sowohl dem Erstanschreiben als auch dem kombinierten Dank- und Erinnerungsanschreiben bei. Bis Ende September meldeten sich 1.331 Personen bei infas und teilten (mindestens) eine aktuelle Telefonnummer mit. Die überwiegende Mehrheit der Panelpersonen (73,1 Prozent) kontaktierte für die Mitteilung der Telefonnummer den infas-Rücklauf über die Hotline, die studien-spezifische E-Mail-Adresse oder die Mitteilungskarte. Rund ein Viertel aller Rückmeldungen (26,9 Prozent) erfolgten über das Online-Adresstool.

Alle mitgeteilten Rufnummern wurden auf fehlerhafte Angaben geprüft und mit den bereits vorliegenden Kontaktdaten abgeglichen. In 96 Fällen (7,2 Prozent) waren die mitgeteilten Rufnummern fehlerhaft bzw. bereits bekannt. Im Saldo haben im Rahmen der Rückmeldeaktion erfreulicherweise 1.235 Personen mindestens eine (zusätzliche) neue Telefonnummer mitgeteilt. Dies kann als positives Zeichen für die Ankündigung der telefonischen Befragung gedeutet werden.

Nach Abschluss der schriftlichen Befragung liegt für 7.777 Fälle der verbleibenden Bruttostichprobe (n=8.378) mindestens eine Telefonnummer vor. Damit wurde die Telefonnummerndichte auf gut 93 Prozent angehoben und die Aktualität der vorliegenden Rufnummern verbessert. Die Grundlage für die telefonische Kontaktierung der Zielpersonen ist damit erfreulich gut.

7 Datenerfassung und -prüfung

7.1 Rücklauferfassung

Über jeden eingehenden Fragebogen wurde doppelt Buch geführt. Der erste Verarbeitungsschritt nach Eingang der Fragebögen bei infas stellte die Eingangskontrolle dar. Bei der Rücklaufkontrolle der schriftlichen Fragebögen wurde nach der Eingangskontrolle und Sichtung des Fragebogens ein Rücklaufcode, der den Status des Bogens (z.B. ausgefüllt, leer, verweigert) wiedergibt, in die Datei eingetragen.

Durch das Verfahren der Rücklauferfassung wurde der Eingang bei infas protokolliert. Im Anschluss daran wurden die Antworten in den Fragebögen komplett erfasst. Die Eingangsverbuchung erfolgt über einen direkten Abgleich der auf dem Fragebogen aufgedruckten personenspezifischen Kennzeichnung (Fallnummer) mit der Bruttostichprobe. Dadurch wurde sichergestellt, dass nur Fragebögen mit in der Bruttostichprobe vorhandenen Fallnummern erfasst wurden.

7.2 Datenerfassung

Sämtliche Fragebögen mit mindestens einer beantworteten Frage wurden manuell von professionellen Codiererinnen und Codierern erfasst. Basis der Erfassung waren die zwischen infas und dem DZA abgestimmten Erfassungsregeln. Die Regeln basieren auf dem Regelwerk, das seit mehreren Wellen bereits für den schriftlichen Drop-off zum Einsatz gekommen ist.

Die Befragungsdaten wurden in einen elektronischen Datensatz (Rohdatensatz) überführt. Die Datenerfassung war am 22.07.2020 abgeschlossen. Im Anschluss konnte die formale Datenprüfung abgeschlossen werden.

Die Angaben auf den zurückgesendeten Mitteilungskarten wurden anhand der Fallnummer im Stichprobenverwaltungstool (infas Sample-Management-System, iSMS) eingepflegt.

7.3 Datenprüfung

Der Einsatz des papierbasierten Selbstausfüllers hatte den Vorteil, dass die Erhebung in vergleichsweise kurzer Zeit parallel bei allen Panelteilnehmerinnen und -teilnehmern stattfinden konnte. Das Bearbeiten eines papierbasierten Selbstausfüllers ist jedoch im Vergleich zu einem elektronischen Instrument deutlich fehleranfälliger – anders als beispielsweise bei einem Online-Fragebogen, bei dem sich der Befragte nicht um Filterführungen kümmern muss und automatisch auf dem richtigen Weg durch den Fragebogen geleitet wird. Die Zielpersonen sind beim Ausfüllen des Papierfragebogens auf sich allein gestellt. Es ist im Gegensatz zu einem Face-to-Face-Interview auch kein Interviewer anwesend, der beim Ausfüllen unterstützend eingreifen oder Rückfragen beantworten kann.

Vor diesem Hintergrund war bei der Entwicklung des Fragebogens darauf geachtet worden, dass möglichst wenige Filter verwendet wurden und der Bogen möglichst leicht auszufüllen war.

In Abstimmung mit dem DZA hat infas lediglich eine formale Prüfung der Daten vorgenommen. Die inhaltlich Prüfung und Bereinigung der Befragungsdaten übernimmt das DZA. Auf der Basis der am 24.07.2020 ans DZA übermittelten Befragungsdaten hat das DZA entschieden, welche Fälle vor der Berechnung der GewichtungsvARIABLEN aus dem Datensatz ausgeschlossen werden sollen. Dies betraf n=22 Fälle. Es verbleiben somit n=4.823 Fälle. Sie stellen die Basis für die Berechnung von Gewichten dar.

Die ausgefüllten Fragebögen, inkl. der 543 Zweiteinsendungen (siehe Kapitel 6.1), wurden am 07.08.2020 ans DZA nach Berlin geschickt.

8 Gewichtung

Die Gewichtung umfasst eine Längsschnittgewichtung für die Panelteilnehmerinnen und -teilnehmer, die auch an der Befragung 2017 teilgenommen haben und eine Querschnittsgewichtung für die Teilnehmerinnen und Teilnehmer der schriftlichen Panelbefragung. Die letztgenannte Gruppe enthält auch Personen, die 2017 kein Interview gegeben haben. Das Querschnittsgewicht wurde an bekannten Bevölkerungsverteilungen kalibriert.

8.1 Längsschnittgewichtung

Zur Berechnung der Längsschnittgewichte wurden Ausfallmodelle (logistische Regression) zur Bestimmung der Teilnahmewahrscheinlichkeit von der Erhebungswelle 2017 zur aktuellen Erhebungswelle berechnet.² Die Grundgesamtheit des Modells sind die Teilnehmerinnen und Teilnehmer der Erhebungswelle 2017 abzüglich Personen, die zwischen den beiden Befragungen verstorben sind. Berechnet wird hierüber die individuelle Wahrscheinlichkeit einer Teilnahme in der Folgewelle. Die Prädiktoren³ des Modells sind:

- Landesteil: West, Ost
- Kreistypen nach BBSR gruppiert: kreisfreie Großstädte, städtische Kreise, ländliche Kreise/dünn besiedelte ländliche Kreise
- Geburtsjahrgang gruppiert: 1941 und früher, 1942-1950, 1951-1959, 1960 und später
- Geschlecht: männlich, weiblich
- Bildungsabschluss ISCED gruppiert: niedrig, mittel, hoch
- Netzwerkgröße gruppiert: bis 2 Personen, 3 bis 5 Personen, 6 und mehr Personen
- Äquivalenzeinkommen gruppiert: unteres Terzil, mittleres Terzil/keine Angabe, oberes Terzil
- Subjektiver Gesundheitsstatus gruppiert: sehr gut/gut, mittel, schlecht/sehr schlecht

Berücksichtigt wurden die Ausprägungen einer Variablen in der Befragung 2017. Das Geschlecht und das Geburtsjahr sind konstant, Bildung bezieht sich auf die Erstbefragung eines Falls.

² Die Ergebnisse des Ausfallmodells sind im Anhang dargestellt.

³ Die Auswahl der Prädiktoren erfolgte in Abstimmung mit dem DZA.

Um die Längsschnittgewichte zu berechnen, wurde das Querschnittgewicht der Ausgangswelle (Startgewicht = integriertes Querschnittgewicht) mit der reziproken Bleibewahrscheinlichkeit ($\frac{1}{\text{Teilnahmewahrscheinlichkeit}}$) multipliziert:

$$ls_{17_20} = qs_{17} \cdot \frac{1}{p_{\text{Teilnahme_2020}}}$$

Dieses Längsschnittgewicht passt die Verteilungen der Panelfälle an die Verteilungen in der Ausgangswelle an. Somit kann die Veränderung von Verteilungen zwischen der Ausgangswelle und der Folgewelle betrachtet werden.

Hinweise zur Erstellung und Benutzung der Gewichte: Die an das DZA übergebenen Längsschnittgewichte beziehen sich auf den Betrachtungszeitraum von der Welle 2017 zur schriftlichen Panelbefragung 2020. Um einen Vergleich zwischen diesen beiden Wellen durchzuführen, muss die Verteilungsausählung für die Ausgangswelle mit dem Querschnittgewicht der Ausgangswelle und die Verteilungsausählung für die Folgewelle mit dem Längsschnittgewicht gewichtet werden. Längere Zeithorizonte (Beispielsweise 3 Wellen) können erschlossen werden, indem Gewichte auf folgende Weise gebildet werden: Querschnittgewicht der Ausgangswelle (t1) * reziproke Teilnahmewahrscheinlichkeit Folgewelle (t2) * reziproke Teilnahmewahrscheinlichkeit Folgewelle (t3). Auf diese Weise kann jedes denkbare Längsschnittgewicht erstellt werden. Voraussetzung für einen Vergleich über mehrere Wellen ist, dass die befragte Person in jeder der betrachteten Erhebungswelle an der Befragung teilgenommen haben muss.

8.2 Integration der Stichproben in einen gemeinsamen Querschnitt

Für die Bildung von Querschnittgewichten müssen die Teilstichproben der Erhebungswelle in eine gemeinsame Querschnittstichprobe integriert werden. Diese Teilstichproben sind:

- Panelfälle mit Teilnahme in der Vorwelle 2017
- Panelfälle ohne Teilnahme in der Vorwelle 2017

Das Ausgangsgewicht für die Integration ist bei den wieder teilnehmenden Panelfällen das aktuelle Längsschnittgewicht. Für die temporären Ausfälle (d.h. keine Teilnahme 2017) wurde ein Hilfsgewicht erstellt. Basis für dieses Hilfsgewicht war das Querschnittgewicht der Welle, in dem der Fall das letzte Mal teilgenommen hat. Dieses Gewicht wurde anhand der Ergebnisse zweier Ausfallmodelle adjustiert:

1. Die individuelle Wahrscheinlichkeit einer Nichtteilnahme in der Folgewelle nach der letzten Teilnahme (1-Teilnahmewahrscheinlichkeit) wurde in den Ausfallmodellen der Längsschnittgewichtung berechnet.

2. Zur Bestimmung der individuellen Wahrscheinlichkeit, in der Betrachtungswelle teilzunehmen, unter der Voraussetzung in der Vorwelle (bzw. in den Vorwellen) nicht teilgenommen zu haben (Rückkehrwahrscheinlichkeit), wurde eine separate logistische Regressionen gerechnet.⁴

Grundgesamtheit der jeweiligen Modelle sind alle Panelfälle ohne Teilnahme in der Ausgangswelle 2017. Prädikatoren eines jeden Modells sind:

- Landesteil: West, Ost.
- Geschlecht: männlich, weiblich.
- Geburtsjahrgang gruppiert: 1941 und früher, 1942-1950, 1951-1959, 1960 und später

Der Landesteil bezieht sich auf die erste Erhebungswelle eines Falls, das Geschlecht ist konstant. Das Alter ist das Alter zum Zeitpunkt der Ausgangswelle des Modells.

Zur Bildung des Hilfsgewichts werden folgende Elemente multiplikativ verknüpft: Querschnittgewicht der Welle, in welcher der Fall das letzte Mal teilgenommen hat, reziproke Wahrscheinlichkeit der Nichtteilnahme in der Folgewelle nach der letzten Teilnahme (1-Teilnahmewahrscheinlichkeit aus den Ausfallmodellen zur Berechnung der Längsschnittgewichte), reziproke Rückkehrwahrscheinlichkeit in der Betrachtungswelle:

$$hw_{2020} = qs_{t1} \cdot \frac{1}{1 - p_{\text{Teilnahme}_t2}} \cdot \frac{1}{p_{\text{Teilnahme}_{2020}}}$$

wobei t1 jeweils die Welle bezeichnet, in der ein/e Befragte/r das letzte Mal an der Befragung teilgenommen hat und t2 jeweils die Folgewelle.

Die Zusammenführung der Teilstichproben in eine gemeinsame Stichprobe muss berücksichtigen, dass die Auswahlgesamtheiten jeder der beiden Teilstichproben überlappend waren. Befragte konnten also sowohl in der einen als auch in der anderen Stichprobe ausgewählt werden. Die Auswahlwahrscheinlichkeiten, auf deren Basis das Designgewicht berechnet wird, müssen dementsprechend angepasst werden.

Dies kann auf im Wesentlichen zwei unterschiedliche Arten erfolgen, zum einen über die fallweise Bestimmung der Wahrscheinlichkeiten, jeweils in den beiden Stichproben enthalten zu sein. Dies setzt allerdings weitreichende Informationen z.B. über die Mobilität (Zuzüge und Fortzüge) voraus und ist, insbesondere auch für die temporären Ausfälle, schwierig bzw. gar nicht möglich. Alternativ

⁴ Die Ergebnisse der Modelle sind im Anhang dargestellt.

kann die Integration auch über eine Konvexkombination (composite weighting) erfolgen, was für den Alterssurvey die zweckmäßige Vorgehensweise darstellt.⁵

Ist die Voraussetzung erfüllt, dass zwei Stichproben vorliegen, die erwartungstreue Schätzungen für den wahren Parameter erlauben, dann ist jede Konvexkombination, d.h. die Berechnung von Konvexgewichten, die sich zur Gesamtheit summieren, der erwartungstreuen Schätzer ebenfalls erwartungstreu. Bei der Konvexgewichtung werden die ursprünglichen Ausgangsgewichte jeder Stichprobe mit einem beliebigen Faktor multipliziert, wobei die Gewichte der Stichprobe 1 mit einem beliebigen Faktor α , die Gewichte der Stichprobe 2 mit $1-\alpha$ multipliziert wird. Als Faktor kann dabei auch die Fallzahl verwendet werden:

$$\alpha = \frac{n_1}{n_1 + n_2} \text{ und } 1 - \alpha = \frac{n_2}{n_1 + n_2}.$$

Dies entspricht der einfachen Regel, die ursprünglichen Gewichte mit dem Anteil der jeweiligen Teilstichprobe am Gesamtstichprobenumfang zu multiplizieren.

Eine Konvexkombination der Gewichte führt also zu erwartungstreuen Schätzern, allerdings kann unter Umständen die Varianz der Gewichte höher sein als bei der Bestimmung der Auswahlwahrscheinlichkeiten.

Die Konvexkombination erfolgte, indem die Panelteilnehmer mit Wiederteilnahme mit solchen mit temporärem Ausfall zusammengefügt wurden. Dabei wurden die ursprünglichen Gewichte jeweils mit dem Anteil der jeweiligen Teilstichprobe am Gesamtstichprobenumfang multipliziert.

Das so entstandene Querschnittsgewicht berücksichtigt die Auswahl- und Teilnahmewahrscheinlichkeiten über die verschiedenen Stichproben hinweg. Es dient als Startgewicht für eine Anpassung an bekannte Sollverteilungen (poststratifiziertes Querschnittsgewicht). Zudem stellt es ein hilfsweises Querschnittsgewicht dar, um Längsschnittgewichte zu berechnen.

⁵ Lohr, Sharon L. (2011). Alternative survey sample designs: Sampling with multiple overlapping frames. *Statistics Canada - Survey Methodology*, 37:197–213.

Brick, J. M.; Dipko, S.; Presser S. Tucker C.; Yuan Y. (2006). Nonresponse Bias in a Dual Frame Sample of Cell and Landline Numbers. *Public Opinion Quarterly*, 70:780 – 793.

Xia, K.; Pedlow, S.; Davern M. (2010): Dual-Frame Weights (Landline and Cell) for the 2009 Minnesota Health Access Survey. *American Statistical Association - Proceedings of the Survey Research Methods Section*, S. 3912 – 3922.

Spieß, M. & Rendtel, U. (2000): Combining an ongoing panel with a new cross-sectional sample. *DIW-Discussion Papers* 198. Berlin.

Rendtel, U. (1999): The Application of the Convex Weighting Estimator to Household Panel Surveys. Mimeo, Frankfurt.

8.3 Poststratifizierte Querschnittgewichtung

Das Querschnittgewicht wurde mittels "iterative proportional fitting" (IPF) an die bekannten Verteilungen Geburtsjahr x Geschlecht x Landesteil laut Mikrozensus (2019) angepasst.⁶ Als Startgewicht diente das Gewicht welches aus der Querschnittsintegration entstanden ist. Zusätzlich wurde ein poststratifiziertes Gewicht erstellt in dem Teilnehmer/innen mit einem Geburtsjahr vor 1930 nicht berücksichtigt wurden, aufgrund der zu geringen Fallzahl pro Geburtsjahrgang.

8.4 Kennziffern der Gewichtungsfaktoren

Zur Beurteilung der Längsschnittgewichte wird im Folgenden neben dem Mittelwert und der Standardabweichung auch das Effektivitätsmaß angegeben (E).⁷ Das Effektivitätsmaß E basiert auf der Varianz des Gewichtungsfaktors. Das Effektivitätsmaß gibt in Prozent der realisierten Fallzahl an, wie groß bei Verwendung des Gewichts die effektive Fallzahl (n') bei einem passiven Merkmal ist, das mit den aktiven Merkmalen nicht korreliert (Worst-case-Szenario).⁸ Die effektive Fallzahl entspricht der Anzahl Befragter, die bei einer uneingeschränkten Zufallsauswahl, gegeben die Varianz des Merkmals in der Stichprobe, den gleichen Stichprobenfehler produziert hätte. Das Effektivitätsmaß drückt das Verhältnis von n zu n' aus und wird berechnet über:

$$E = \frac{n'}{n} \text{ wobei: } n' = \frac{(\sum_i g_i)^2}{\sum_i g_i^2}$$

⁶ In den bisherigen DEAS-Wellen wurde Berlin anhand der Postleitzahlen in Ost- und Westberlin aufgeteilt und jeweils getrennt den Landesteilen Ostdeutschland bzw. Westdeutschland bei der Gewichtung zugeordnet. Für die an dieser Stelle berichtete Gewichtung wurde Berlin auf Wunsch des DZA vollständig dem Landesteil „Ost“ zugeordnet. Ein Verteilungsvergleich der ungewichteten und gewichteten Stichprobe ist im Anhang dargestellt.

⁷ Vergleiche zum Effektivitätsmaß und Designeffekt (E = 1/Designeffekt) auch:

Kish, L. (1965). *Survey Sampling*. New York: Wiley.

Kish, L. (1987). Weighting in Deft2. *The Survey Statistician*, June 1987.

Little, R. J.A., S. Lewitzky, S. Heeringa, J. Lepkowski & R.C. Kessler (1997): Assessment of weighting methodology for the National Comorbidity Survey. *American Journal of Epidemiology*, 146, 439–449.

Rösch, G. (1994): Kriterien der Gewichtung einer nationalen Bevölkerungsstichprobe: Gabler, S., J. H. P. Hoffmeyer-Zlotnik & D. Krebs (Hg.): *Gewichtung in der Umfragepraxis*, Westdeutscher Verlag, 7-26.

⁸ Aktive Merkmale sind Merkmale, die bei der Kalibrierung eingesetzt wurden. Alle übrigen erhobenen Informationen sind passive Merkmale.

Die folgende Tabelle stellt die Kennziffern der ausgelieferten Gewichte dar:

Tabelle 8 Kennziffern der Gewichtung

Variable	Bezeichnung	Fallzahl	Mittelwert	Standardabweichung	Min.	Max.	Effektivitätsmaß (%)
ls17_20	Längsschnittgewicht Panel 2017-2020	4.228	1	1,212	0,059	20,8	40,5
qs_20	Querschnittgewicht integrierte Stichprobe 2020	4.823	1	2,199	0,035	71,6	17,1
qsps_20	Querschnittgewicht integrierte Stichproben 2020 – Poststratifiziert	4.823	1	1,628	0,020	24,6	27,4
qsps_20_u90	Querschnittgewicht integrierte Stichproben 2020 – Poststratifiziert – ab 1930 geboren	4.763	1	1,627	0,020	24,7	27,8

Quelle: infas – eigene Berechnung

Bei der Bewertung der dargestellten Kennziffern ist das Ziel der Gewichtung zu berücksichtigen. Das Hauptziel der Gewichtung ist der Ausgleich von Verzerrungen, die durch das Stichprobendesign (mit ungleichen Auswahlwahrscheinlichkeiten) und durch selektive Teilnahme bzw. Nichtteilnahme (ungleiche Realisierungswahrscheinlichkeiten) entstehen. Unter Verwendung der Gewichte lassen sich Populationswerte aus der Stichprobe erwartungstreu⁹ schätzen. Weisen die Gewichte allerdings eine starke Streuung auf, so kann dies zu einer großen Varianz der Schätzfunktionen führen. Hier tritt der in der Statistik immer wieder anzutreffende Trade-off zwischen Bias und Varianz auf. Die Gewichtung reduziert den Bias, eine allzu große Steigerung der Varianz durch die Gewichtung sollte allerdings auch vermieden werden.

Bei Längsschnitterhebungen tritt häufig ein kontinuierliches Absinken der Effektivität auf, da sich die Selektivitätseffekte über die Wellen kumulieren. Selten lässt sich beobachten, dass die Selektivitäten (beispielsweise durch die verstärkte Teilnahme von temporären Ausfällen) wieder zurückgehen und die Effektivität steigt. Dieser Effekt verstärkt sich, da aufgrund der Corona-Pandemie die eigentlich geplante Auffrischung der Stichprobe in Form einer neuen Basisstichprobe im Sommer 2020 nicht stattfinden konnte und die hier berichtete Befragung der Panelstichprobe ausschließlich schriftlich stattfand (Moduswechsel).

Unter Berücksichtigung dieser Umstände sinken die Effektivitätsmaße insgesamt in einem erwartbaren Rahmen ab. Wesentlich ist auch, dass die effektive Fallzahl in der Regel noch so hoch bleibt, dass die Konfidenzintervalle in einer akzeptablen Größenordnung bleiben. Hierbei ist auch noch einmal darauf hinzuweisen, dass es sich bei den effektiven Fallzahlen um ein Worst-Case-Szenario

⁹ Ein Schätzer heißt erwartungstreu, wenn sein Erwartungswert gleich dem wahren Wert des zu schätzenden Parameters ist. Ist eine Schätzfunktion nicht erwartungstreu, spricht man davon, dass der Schätzer verzerrt ist. Das Ausmaß der Abweichung seines Erwartungswerts vom wahren Wert heißt Verzerrung oder Bias. Die Verzerrung drückt den systematischen Fehler des Schätzers aus.

handelt (gilt nur für Merkmale, die überhaupt nicht mit den aktiven Gewichtsmerkmalen korrelieren), d.h. bei der überwiegenden Zahl der Merkmale wird die effektive Fallzahl (deutlich) höher sein.

Anhang

**Tabelle 9 Ausfallmodellierung Panelstichprobe: Interview in 2020,
Basis: Fälle mit Interview in 2017 (logistische Regression)**

Abhängige Variable: Y = 1: Teilnahme, Y = 0: Keine Teilnahme	Odds Ratio	p-Wert
Landesteil		
West	Referenz	
Ost	0,95	0,43
Kreistyp (BBSR)		
Kreisfreie Großstadt	Referenz	
Städtische Kreise	0,96	0,61
Ländliche Kreise	0,83	0,01
Geburtsjahrgang		
Vor 1942	Referenz	
1942-1950	1,41	0,00
1951-1959	1,01	0,90
1960 und später	0,68	0,00
Geschlecht		
Männer	Referenz	
Frauen	1,22	0,00
Bildung (ISCED)		
Niedrig (ISCED 1-2)	Referenz	
Mittel (ISCED 3-4)	1,80	0,00
Hoch (ISCED 5-6)	2,42	0,00
Netzwerkgröße		
Bis 2 Personen	Referenz	
3 bis 5 Personen	1,22	0,01
6 und mehr Personen	1,39	0,00
Einkommen		
Unteres Terzil	Referenz	
Mittleres Terzil/keine Angabe	1,17	0,02
Oberes Terzil	1,41	0,00
Subjektiver Gesundheitszustand		
Sehr gut/gut	Referenz	
Mittel	0,84	0,00
Schlecht/ sehr schlecht	0,61	0,00
Konstante	0,75	0,06
n		6.452
Pseudo-R²		0,04
LR Chi²		301,94
Prob > Chi²		0,00

P-Werte basieren auf einfachen Standardfehlern.

Hinweis zur Fallzahl: CAPI-Interview Erhebung 2017, ohne Verstorbene

Quelle: infas-eigene Berechnung

**Tabelle 10 Ausfallmodellierung Panelstichprobe temporäre Ausfälle:
PAPI-Interview in 2020, Basis: Fälle mit Nichtteilnahme in 2017
(logistische Regression)**

Abhängige Variable: Y = 1: Teilnahme, Y = 0: Keine Teilnahme	Odds Ratio	p-Wert
Geburtsjahrgang		
Vor 1942	0,45	0,00
1942-1950	Referenz	
1951-1959	1,11	0,41
1960 und später	1,47	0,00
Geschlecht		
Männer	Referenz	
Frauen	1,05	0,59
Landesteil		
West	Referenz	
Ost	0,73	0,00
Konstante	0,07	0,00
n		11.043
Pseudo-R2		0,03
LR Chi²		130,16
Prob > Chi²		0,00

P-Werte basieren auf einfachen Standardfehlern.
Quelle: infas-eigene Berechnung

Tabelle 11 Verteilungsvergleich ungewichtete und kalibrierte Stichprobe

Spalten%	ungewichtet		gewichtet	bekannte Soll-Verteilung*
	abs.	%	%	%
Gesamt	4.823	100,0	100,0	100,0
Geburtsjahr x Gebiet x Geschlecht				
Nach 1964 West männlich	140	2,9	11,6	11,6
Nach 1964 West weiblich	193	4,0	11,7	11,7
Nach 1964 Ost männlich	55	1,1	2,6	2,6
Nach 1964 Ost weiblich	77	1,6	2,5	2,5
1955-1964 West männlich	400	8,3	11,7	11,7
1955-1964 West weiblich	494	10,2	11,9	11,9
1955-1964 Ost männlich	148	3,1	3,0	3,0
1955-1964 Ost weiblich	224	4,6	3,0	3,0
1945-1954 West männlich	494	10,2	7,6	7,6
1945-1954 West weiblich	479	9,9	8,3	8,3
1945-1954 Ost männlich	217	4,5	2,1	2,1
1945-1954 Ost weiblich	242	5,0	2,3	2,3
1935-1944 West männlich	481	10,0	5,7	5,7
1935-1944 West weiblich	392	8,1	7,1	7,1
1935-1944 Ost männlich	263	5,5	1,7	1,6
1935-1944 Ost weiblich	241	5,0	2,2	2,2
1930-1934 West männlich	86	1,78	1,0	1,0
1930-1934 West weiblich	63	1,31	1,6	1,6
1930-1934 Ost männlich	43	0,89	0,3	0,3
1930-1934 Ost weiblich	31	0,64	0,5	0,5
Vor 1930 West männlich	27	0,56	0,4	0,4
Vor 1930 West weiblich	15	0,31	0,9	0,8
Vor 1930 Ost männlich	10	0,21	0,1	0,1
Vor 1930 Ost weiblich	8	0,17	0,2	0,2

* Mikrozensus 2019, Zahlen sind gerundet, daher Abweichungen zur Summe 100 möglich
Quelle: infas-eigene Berechnung