

## Deutscher Alterssurvey (DEAS): Kurzbeschreibung des SUF DEAS 2020/21, Version 1.1

Stuth, Stefan

Veröffentlichungsversion / Published Version

Arbeitspapier / working paper

### Empfohlene Zitierung / Suggested Citation:

Stuth, S. (2022). *Deutscher Alterssurvey (DEAS): Kurzbeschreibung des SUF DEAS 2020/21, Version 1.1*. Berlin: Deutsches Zentrum für Altersfragen. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-79742-5>

### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-SA Lizenz (Namensnennung-Nicht-kommerziell-Weitergabe unter gleichen Bedingungen) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-nc-sa/4.0/deed.de>

### Terms of use:

This document is made available under a CC BY-NC-SA Licence (Attribution-NonCommercial-ShareAlike). For more information see: <https://creativecommons.org/licenses/by-nc-sa/4.0>

---

Juni 2022

Deutscher Alterssurvey (DEAS):  
Kurzbeschreibung des SUF DEAS2020/21, Version 1.1

*Stefan Stuth*

---

Zitationsempfehlung: Stuth, Stefan (2022): Deutscher Alterssurvey (DEAS): Kurzbeschreibung des SUF DEAS2020/21, Version 1.1. Berlin: Deutsches Zentrum für Altersfragen.

Deutsches Zentrum für Altersfragen (DZA)  
Forschungsdatenzentrum (FDZ-DZA)  
Manfred-von-Richthofen-Straße 2  
12101 Berlin  
Telefon +49 (0)30 – 26 07 40-0  
Telefax +49 (0)30 – 26 07 40-33  
E-Mail [fdz@dza.de](mailto:fdz@dza.de)

## INHALTSVERZEICHNIS

1.	Der Deutsche Alterssurvey (DEAS) – eine Studie zur zweiten Lebenshälfte	4
1.1	Stichprobe	5
1.2	Themenbereiche der Hauptbefragung 2020/21	6
1.3	Änderungen und Neuerungen in der Hauptbefragung 2020/21	7
1.3.1	Moduswechsel	7
1.3.2	Befragungsinhalte	7
1.3.3	Neue Berufsklassifikation	8
1.3.4	Flag-Variablen	8
1.3.5	Datenbereitstellung	8
1.3.6	Anonymisierung	9
1.3.7	Regionaldaten	9
1.3.8	Datendokumentation	9
1.3.9	Gewichtung	9
1.3.10	Das komplexe Surveydesign des DEAS	9
2	Ausschöpfung und Passungsprüfung der Hauptbefragung 2020/21	10
3	Datengewichtung der Hauptbefragung 2020/21	10
4	Datenformate	12
5	Systematik der Variablennamen und Missing-Codes im DEAS 2020/21	12
6	Meta-Datensatz zur Teilnahmegeschichte der DEAS-Befragten	13
6.1	Teilnahmeprofil	17
6.2	Verbleibsdaten	17
7	Veränderung von Daten im SUF zur Wahrung der faktischen Anonymität	18
8	Das komplexe Surveydesign des DEAS	18
9	Sonstiges	19
9.1	Dokumentation zum DEAS2021	19
9.2	Regionaldaten	20
	Literatur	20



## 1. DER DEUTSCHE ALTERSSURVEY (DEAS) – EINE STUDIE ZUR ZWEITEN LEBENSHÄLFTE

Der Deutsche Alterssurvey (DEAS) ist eine vom Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) geförderte Langzeitstudie des Deutschen Zentrums für Altersfragen (DZA) zum Wandel der Lebenssituationen und Altersverläufe von Menschen, die sich in der zweiten Lebenshälfte befinden. Grundlage sind bundesweit repräsentative Befragungen im Quer- und Längsschnitt von jeweils mehreren tausend Teilnehmerinnen und Teilnehmern im Alter ab 40 Jahren. Die beteiligten Personen werden umfassend zu ihren objektiven und subjektiven Lebensbedingungen befragt. Hinzu kommen Testverfahren zur Erhebung der kognitiven Leistungsfähigkeit und körperlichen Gesundheit. Aufgrund seines kohorten-sequenziellen Längsschnittdesigns, seiner Themenbreite, seiner nationalen Repräsentativität mit vergleichsweise hoher Fallzahl und der bereits relativ langen Beobachtungsdauer stellt der DEAS eine in Deutschland einzigartige Datenquelle zur interdisziplinären Erforschung der Lebenssituationen und Lebensverläufe sowie der subjektiven Sichtweisen und Lebenspläne älter werdender und alter Menschen dar.

Die Mikrodaten des Deutschen Alterssurveys (DEAS) stehen der wissenschaftlichen Forschung für nicht-gewerbliche Zwecke kostenfrei zur Verfügung. Das Forschungsdatenzentrum gibt Wissenschaftlerinnen und Wissenschaftlern Zugang zu den im Rahmen des Alterssurveys erhobenen Informationen und berät sie bei deren Verwendung. Die anonymisierten DEAS-Datensätze aller abgeschlossenen Befragungswellen und die Dokumentationsmaterialien sind kostenlos über das FDZ-DZA (<https://www.dza.de/forschung/fdz>) erhältlich.

Die erste Erhebungswelle des DEAS wurde im Jahr 1996 durchgeführt, gefolgt von Surveywellen in den Jahren 2002, 2008, 2011, 2014, 2017, 2020 und 2020/21. Die ursprünglich für 2020 vorbereitete turnusmäßige persönlich-mündliche Befragung der Panelteilnehmenden und einer neu gezogenen Basisstichprobe konnte aufgrund der Corona-Pandemie nicht wie geplant stattfinden. Zum Auftakt der siebten DEAS-Welle wurde kurzfristig im Juni und Juli 2020 eine schriftlich-postalische Befragung der Panelteilnehmenden durchgeführt. Im November 2020 startete die Haupterhebung mit einer umfangreichen telefonischen Befragung der Panelteilnehmenden, ergänzt durch einen schriftlichen Drop-off zum Selbstausfüllen. Die Haupterhebung endete im Februar 2021, wobei das letzte Interview am 1. März 2021 durchgeführt wurde. Auf die Befragung einer neu gezogenen Basisstichprobe musste verzichtet werden, da eine Befragung unter Corona-Bedingungen nicht im persönlichen Interview durchgeführt werden konnte und für die Basisstichprobe nur Adressdaten, aber keine Telefonnummern vorlagen.

Die Hauptbefragung 2020/21 folgte inhaltlich den regulären DEAS-Wellen. Methodisch mussten in den Befragungsinstrumenten aber Anpassungen vorgenommen werden, um dem Wechsel im Befragungsmodus (von persönlichen Interviews auf telefonische Interviews) Rechnung zu tragen. Die Feldarbeit, Datenerfassung und Datengewichtung bewerkstelligte – wie schon in den Vorwellen – das Institut für angewandte Sozialwissenschaft GmbH (Infas), Bonn.

## 1.1 Stichprobe

Der DEAS kombiniert quer- und längsschnittliche Stichproben. Seit 1996 werden im Abstand von sechs Jahren jeweils neue Basisstichproben aus der 40- bis 85-jährigen Wohnbevölkerung in Privathaushalten als – nach Altersgruppe, Geschlecht und Landesteil (Ost/West) – geschichtete und geklumpte Zufallsstichproben aus den Einwohnermelderegistern gezogen und durch Interviewerinnen und Interviewer persönlich befragt. Ab Welle 2 (2002) werden zudem die panelbereiten Teilnehmerinnen und Teilnehmer aus den vorangehenden Basiserhebungen erneut interviewt. Ausgenommen davon sind die Befragten der separat gezogenen Ausländerstichprobe 2002, die nicht in den Längsschnitt übernommen wurden. Nach der Erhebung 2008 wurde die Taktung zwischen den Panelwellen von sechs auf drei Jahre verkürzt.

In Welle 1 (1996) beschränkte sich die Auswahl auf Personen mit deutscher Staatsangehörigkeit, in Welle 2 (2002) wurden für die Basiserhebung zwei getrennte Stichproben von Deutschen und Nicht-Deutschen gezogen. Seit 2008 wird die Basisstichprobe aus der Grundgesamtheit der 40- bis 85-jährigen Bevölkerung in Privathaushalten (Ort der Hauptwohnung) ohne Unterscheidung der Nationalität gezogen.

Ursprünglich war für das Jahr 2020 die kombinierte Befragung einer neuen Basisstichprobe mit der Befragung der Panelteilnehmer und -teilnehmerinnen geplant. Die Corona-Pandemie stand der Umsetzung dieser Planung jedoch entgegen, da die Einwohnermeldeämter nur Adressdaten bereitstellen können, aber keine Telefonnummern. Unter Pandemiebedingungen sind persönliche Interviews jedoch nicht ohne Gesundheitsrisiken umzusetzen gewesen. Die Befragung der Basisstichprobe konnte in der vorliegenden Welle nicht entsprechend realisiert werden. Aus diesem Grund richtete sich die Hauptbefragung 2020/21 an alle Panelteilnehmerinnen und -teilnehmer, die zuvor bereits mindestens einmal am DEAS teilgenommen haben. Grundlage für die Längsschnitterhebung waren alle noch erreichbaren panelbereiten Teilnehmerinnen und Teilnehmer der Basisstichproben 1996 bis 2014. Insgesamt liegen gültige Interviews für 5.402 Personen im Alter von 46 bis 100 Jahren vor. Davon haben 4.103 Personen zuletzt an der schriftlichen Kurzbefragung im Sommer 2020 teilgenommen.

Der Datensatz SUF DEAS2021 mit den anonymisierten Daten der Hauptbefragung 2020/21 enthält Informationen zu den Untersuchungsteilnehmerinnen und -teilnehmern aus den folgenden Stichproben:

<b>Stichprobe 2020/21</b>	<b>n</b>
Panelbefragte aus B1996	554
Panelbefragte aus B2002	572
Panelbefragte aus B2008	1.717
Panelbefragte aus B2014	2.559
Befragte insgesamt	5.402

Mit den insgesamt vorhandenen Scientific Use Files (SUF) aller bisherigen Erhebungswellen (20.714 Befragte, 50.256 gültige Interviews) lassen sich individuelle und gesellschaftliche Entwicklungen über den Zeitraum von mehr als 24 Jahren betrachten (siehe Klaus et al. 2017). Das Stichprobendesign des Deutschen Alterssurveys eröffnet dabei verschiedene Analysemöglichkeiten. Es bieten sich die Optionen der querschnittlichen Untersuchung von Lebenssituationen und Lebenszusammenhängen im jeweiligen Erhebungsjahr, zur Analyse des sozialen Wandels zwischen 1996 und 2020/21 (siehe Mahne et al. 2017) sowie der individuellen Entwicklung über drei bis 24 Lebensjahre mit den Paneldaten (Vogel et al. 2019). Eine weitere Perspektive ergibt sich aus der Gegenüberstellung der individuellen Entwicklungen über sechs Jahre in den Zeiträumen 1996-2002, 2002-2008, 2008-2014 und 2014-2020 bzw. 2014-2020/21. So ist der Vergleich der Entwicklung in einem konstanten Altersabschnitt für verschiedene Geburtsjahrgangsgruppen möglich (Kohortensequenzanalyse), beispielsweise der prospektive Vergleich von Übergangsprozessen in den Ruhestand oder der gesundheitlichen Entwicklung der Individuen.

## **1.2 Themenbereiche der Hauptbefragung 2020/21**

Wie in den Vorwellen wurden die Studienteilnehmerinnen und -teilnehmer auch in Welle 8 umfassend zu ihrer Lebenssituation befragt. Es wurden Auskünfte zu folgenden Themenbereichen erhoben:

- Arbeit und Ruhestand
- Partnerschaft, Familie und Generationenbeziehungen
- Soziale Netzwerke und soziale Unterstützung
- Freizeittätigkeiten und bürgerschaftliches Engagement
- Wohnen und Mobilität
- Wirtschaftliche Lage und wirtschaftliches Verhalten
- Subjektives Wohlbefinden
- Gesundheit und Gesundheitsverhalten
- Hilfe- und Pflegebedürftigkeit
- Einstellungen, Normen und Werte
- Soziodemografische Grunddaten.

Die Befragung gliederte sich – wie in den Vorwellen – in zwei Hauptteile. Der Großteil der Informationen wurde durch ein persönliches Telefoninterview mit einem umfangreichen Fragebogen im Haushalt der Zielperson erhoben. Der zweite Teil der Befragung bestand aus einem von den Befragten selbst auszufüllenden schriftlichen Fragebogen („Drop-off“). Hier wurden eher subjektive Vorstellungen – wie Einstellungen, Werthaltungen, Altersbilder, Wohlbefinden –, sensiblere Bereiche (z. B. Details zur materiellen Lage und Gesundheit) und Themen abgefragt, die keine differenzierte Filterführung benötigen. Auch die meisten Items zur Erhebung psychologischer Skalen wurden im Drop-off erhoben. 619 der 4.419 Ausfüllenden machten von der Option Gebrauch, den Drop-off als Online-Fragebogen auszufüllen.

## 1.3 Änderungen und Neuerungen in der Hauptbefragung 2020/21

### 1.3.1 Moduswechsel

Der normale Erhebungsmodus des DEAS ist ein persönlich geführtes computerunterstütztes Interview (CAPI). Aufgrund der Corona-Pandemie musste der Erhebungsmodus auf ein Computer Assisted Personal Interview am Telefon (CAPI-per-Phone) umgestellt werden.<sup>1</sup> In diesem mündlichen Interview mit einer durchschnittlichen Dauer von ca. 75 Minuten wurden im Wesentlichen die Fragen analog zur Welle 2017 eingesetzt. In der Hauptbefragung 2017 wurde die Möglichkeit des Proxy-Interviews neu eingeführt: Panelpersonen, die sich aus gesundheitlichen Gründen nicht mehr in der Lage sahen, das Interview eigenständig zu führen, konnten festlegen, dass eine Person ihres Vertrauens an ihrer Stelle die Fragen beantwortete. Da die Umstellung auf Capi-per-Phone diesen Ansatz stark erschwerte, wurden bei der Befragung 2020/21 keine Proxy-Interviews durchgeführt. Aufgrund der pandemiebedingten Umstellung auf CAPI-per-Phone konnten keine körperlichen Funktionstest zur Bestimmung des Atemvolumens (Lungenfunktionstest) und keine Tests zur kognitiven Leistungsfähigkeit (Zahlen-und-Zeichen-Test) durchgeführt werden. Da die InterviewerInnen die Befragten nicht aufsuchen konnten, war es ihnen nicht möglich Angaben zum Wohnumfeld der Befragten zu machen.

Auch wenn das Interview weiterhin mündlich erfolgte, mussten mit dem Moduswechsel die meisten Fragen angepasst werden. Während die Frageformulierungen davon kaum betroffen waren, konnten Hilfsmittel wie Listen, Personenkarten oder ähnliches gar nicht mehr verwendet werden. Weiterhin musste für einige Fragen der Umfang der Antwortkategorien, stark reduziert werden, da die Hilfestellungen wie Listen- oder Personenkarten wegfielen.

### 1.3.2 Befragungsinhalte

Gegenüber der Welle 2017 gab es nur wenige Änderungen an den Befragungsinhalten. Beispiele: Ausgeweitet wurden die Fragen zu dem Pflege- und Unterstützungsbedarf der noch lebenden Eltern und Schwiegereltern. Wenn die Befragten angeben, dass sie andere Personen unterstützen, wird nun nachgefragt, ob die unterstützte Person an Demenz erkrankt ist. Weiterhin gibt es neue Fragen zur Vereinbarkeit von Pflege und Beruf, die die Nutzung der neuen Familienpflegezeit oder des zinslosen Darlehens erfassen.

Als Ergebnis eines Aufrufs des DZA zur Einreichung von Modulen sind drei neue Module zu den Themen altersbezogene Rückzugs- und Aktivierungsnormen, Big-Five-Persönlichkeitsmerkmalen sowie zur bevorzugten und subjektiven Lebenserwartung verfügbar. Für diese neuen Inhalte gilt jedoch eine Sperrfrist. Sie werden erst ab Januar 2023 verfügbar sein. Mehr [Informationen](#) finden sich im Dokument „eingeschränkte Datenverfügbarkeit für eingereichte Module“.

---

<sup>1</sup> Infas nutzt die Bezeichnung Capi-per-Phone um wesentliche Unterschiede zur klassischen CATI Befragung hervorzuheben. Im Gegensatz zu regulären CATI Interviews wurde die Erhebung nicht durch anonyme Telefonstudios durchgeführt, sondern durch die Interviewer und Interviewerinnen, die die Befragung bereits in den Vorjahren persönlich durchführten.

### 1.3.3 Neue Berufsklassifikation

Um die Analyse von Berufen und beruflichen Laufbahnen mit einer heimischen Klassifikation vornehmen zu können (anstelle mit der auf internationale Vergleiche ausgelegten ISCO08 Klassifikation), steht ab sofort die Klassifikation der Berufe (KldB 2010) zur Verfügung. Informationen zu dem aktuellen, den allerersten ausgeübten Beruf oder zum zuletzt ausgeübten Beruf (KldB2010) stehen für die Befragten (und teilweise für deren Partnerinnen und Partner) ab dem Befragungsjahr 2014 zur Verfügung.

### 1.3.4 Flag-Variablen

Im Zuge der Aufbereitung und Prüfung der Befragungsdaten durch das DZA werden auch Konsistenz- und Plausibilitätschecks durchgeführt, deren Ergebnisse ab dem DEAS 2020/21 auch den Datennutzerinnen und Datennutzern zur Verfügung stehen. Flag-Variablen heben die Fälle hervor, die aufgrund der Prüfungen auffällig geworden sind. Diese Prüfungen finden für diejenigen Variablen statt, für die entweder Zusatzinformationen aus anderen Fragen vorliegen, die einen Abgleich ermöglichen oder es können Grenzwerte definiert werden, deren Unter- oder Überschreitung unplausibel erscheinen. In der Instrumentendokumentation liegen die Variablen für die Flags vor und die zu Grunde liegenden Prüfregeln sind detailliert ausgeführt. Die Flag-Variablen starten mit dem flag-Prefix nach dem der Variablenname folgt, auf den sich die Flag-Variable bezieht. Beispiel: Die Variable kc105c, erfasst die Höhe des Abschlages den Renterinnen und Rentner bei vorzeitigem Ausscheiden aus dem Berufsleben in Kauf nehmen müssen. Die Höhe des Abschlages kann einen gesetzlichen Grenzwert nicht überschreiten. Manche Befragte haben aber eine Abschlagshöhe angegeben, die über dieser Grenze liegt. Die Variable flag\_kc105c markiert diese Fälle. Alle Befragten, die unauffällig waren, erhalten den Wert 0, während für Fälle die Werte über dem Grenzwert angegeben haben, der Wert 1 eingetragen wird.

Die Ergebnisse dieser Prüfungen werden zusätzlich blockweise kumuliert, um Personen hervorzuheben, die mehrfach bei den Fragen des jeweiligen Themenblocks auffällig wurden. Die Variablen sehen folgendermaßen aus: flag\_BLOCK-NRcum. Beispiel: Die Variable flag\_kc500cum addiert die (wiederholten) Inkonsistenzen im Gesundheitsblock (kc501-kc556a) zusammen.

Die Variable flag\_kcgesamt summiert die Anzahl aller inkonsistenten oder unplausiblen Angaben im gesamten mündlichen Interview.

Nähere Informationen zu Design, Inhalt und Ablauf der Erhebung 2020/21 finden sich im Methodenbericht von infas (Schiel et al. 2021). Die Instrumente und Methodenberichte zu allen Erhebungswellen des DEAS stehen auf den Seiten des FDZ-DZA zum Download zur Verfügung: <https://www.dza.de/forschung/fdz>.

### 1.3.5 Datenbereitstellung

Mit der Herausgabe des DEAS 2020/21 ändert sich der Modus der Datenbereitstellung. Nutzerinnen und Nutzer erhalten automatisch das vollständige Datenpaket mit allen Wellen von 1996 bis zur aktuellsten Erhebung. Damit umgeht das FDZ das Problem, dem sich insbesondere Datennutzerinnen und Datennutzer gegenübersehen, die noch keine Erfahrung mit den DEAS-Daten haben: Der DEAS

erhebt bestimmte Daten sparsam, um die Fragenlast für die Befragten nicht unnötig zu erhöhen. So werden Informationen zum Beispiel für ausgeübten Beruf nur dann erfasst, wenn dieser sich im Vergleich zur letzten Welle oder Teilnahme verändert hat. Hat sich nichts geändert erhalten sparsam erhobene Variablen den Code für überfiltert. Mit der Auslieferung aller Wellen haben die Datennutzerinnen und Datennutzer nun die Möglichkeit diese Lücken zu füllen, ohne sich schon bei Antragstellung für spezifische Erhebungsjahre entscheiden oder ggfs. später weitere Datensätze nachbestellen zu müssen.

### *1.3.6 Anonymisierung*

Um die Anonymität der Befragten zu gewährleisten, prüft das FDZ regelmäßig, ob bestehende Schutzmaßnahmen ausreichend sind. In diesem Zuge sind nur noch folgende Regionalmerkmale in den SUF verfügbar: Bundesland, BBSR Kreistyp und die „primary sampling units“ (PSU). Die PSUs bestehen in der Mehrheit aus Gemeinden, die in wenigen Fällen auf die Kreisebene vergrößert wurden. Die PSUs sind maskiert und lassen keinen Rückschluss mehr auf die Original-Erhebungseinheit zu.

### *1.3.7 Regionaldaten*

Aufgrund der hohen Anforderungen an die Wahrung der Anonymität der Befragten, ist es künftig nicht mehr möglich, regionale Zusatzinformationen an die Daten der SUFs anzuspielen. Analysen mit einem regionalen Fokus müssen in Zukunft vor Ort (am DZA) am Gastwissenschaftlerarbeitsplatz erfolgen. Mittelfristig prüft das FDZ aber die Möglichkeit, in Zukunft gröbere Regionalindikatoren auf der Ebene der Raumordnungsregionen bereitzustellen, damit Datennutzer und Datennutzerinnen, nach Absprache mit dem FDZ, weitere raumordnungsregionenbasierte Informationen zuspieren können.

### *1.3.8 Datendokumentation*

Die ausführliche Dokumentation der generierten Variablen befindet sich ab sofort nicht mehr im Kurzbericht, sondern in der Variablendokumentation.

### *1.3.9 Gewichtung*

Es haben sich geringfügige Anpassungen in den Ausfallmodellen ergeben, die als Grundlage für die Gewichtungsfaktoren dienen. Diese sind online im Dokument „Hinweise zu Neuerungen in der Gewichtung des DEAS“ unter <https://www.dza.de/forschung/fdz/deutscher-alterssurvey/deas-dokumentation> zu finden.

### *1.3.10 Das komplexe Surveydesign des DEAS*

Ab sofort stehen im Meta-Datensatz zusätzliche Merkmale zur Verfügung, mit denen das komplexe Surveydesign des DEAS in den Analysen berücksichtigt werden kann (siehe Kapitel 8).

## 2 AUSSCHÖPFUNG UND PASSUNGSPRÜFUNG DER HAUPTBEFRAGUNG 2020/21

Aus einer Bruttoeinsatzstichprobe von 8.379 Adressen wurden 5.418 CAPI-Interviews realisiert. Das entspricht einer Ausschöpfung von 66 Prozent. Darüber hinaus wurden 4.425 Drop-offs ausgefüllt, was eine Rücklaufquote von 82 Prozent bedeutet (vgl. infas-Methodenbericht: Schiel et al. 2021).

In zwei Prüfschritten wurde anhand von Geburtsjahr/Alter und Geschlecht geprüft, ob (1) die 2020/21 befragte Person mit der jeweiligen befragten Person aus den Vorwellen übereinstimmt und ob (2) die mündlich befragte Person mit der Person identisch ist, die den schriftlichen Zusatzfragebogen ausgefüllt hat. Von den 5418 Fällen wurden 11 Fälle durch Infas als nicht auswertbar ausgeschlossen. Fünf weitere Fälle brachen ihr Interview ab und setzten es bis zum Ende des Befragungszeitraumes auch nicht fort. Diese Fälle wurden ebenfalls gelöscht. Kein Fall musste im Rahmen der Panelpassungsprüfung gelöscht werden, da es sich bei den 2020/21 befragten Personen mit hoher Wahrscheinlichkeit um die eigentlichen Panelpersonen handelte. Nach Abschluss der Prüfungen liegen 5402 gültige Interviews vor.

Fünf mündlich befragte Personen stimmten nicht mit den Personen überein, die den schriftlichen Zusatzfragebogen ausgefüllt haben. Die vorhandenen Drop-off-Informationen wurden für diese Fälle komplett gelöscht. Alle Variablen wurden auf -7 gesetzt (gelöscht in Datenaufbereitung). Die Fälle stammen aus dem Paper Pencil. Ein Drop-off Fragebogen wurde sehr früh abgebrochen. Die wenigen vorhandenen DO-Informationen wurden für diesen Fall komplett gelöscht. Auch in diesen Fällen wurden alle Variablen auf -7 gesetzt (gelöscht in Datenaufbereitung). Der Fall stammt aus dem CAWI. Am Ende der Drop-off & CAPI-per-Phone Passungsprüfung verbleiben 4.419 Drop-offs (619 CAWI und 3.800 Paper Pencil).

## 3 DATENGEWICHTUNG DER HAUPTBEFRAGUNG 2020/21

In enger Abstimmung mit dem DEAS-Projektteam des DZA hat infas die Datengewichtung vorgenommen. Das gewählte Vorgehen entspricht im Grundsatz dem Verfahren, das auch in anderen Panelstudien angewendet wird.<sup>2</sup> Grundlage der *längsschnittlichen Gewichtung* sind Ausfallmodelle (logistische Regression) zur Bestimmung der Teilnahmewahrscheinlichkeit von einer Erhebungswelle zur darauffolgenden Erhebungswelle (ohne Personen, die zwischen den beiden Wellen verstorben sind) mit einem begrenzten Satz von Prädiktoren (siehe Schiel et al. 2021: Kapitel 8.1). Das Längsschnittgewicht ergibt sich aus der Multiplikation des Querschnittgewichts der Ausgangswelle mit dem Kehrwert der Teilnahmewahrscheinlichkeit an der Folgewelle. Dieses Verfahren wurde für jedes Erhebungsjahr seit 2002 angewendet. Im Gewichtungssatz für den aktuellen Scientific Use File zum DEAS 2020/21 enthalten die Variablen *ls20\_21* und *ls17\_21* dieses Längsschnittgewicht.

---

<sup>2</sup> Siehe Hammon et al. 2016; Bethmann 2013; European Central Bank 2016

Grundsätzlich hat jeder Datennutzer die Möglichkeit, durch Kombination der wellenspezifischen Teilnahmewahrscheinlichkeiten Vergleiche über mehrere Wellen hinweg im balanced panel Design vornehmen (für Details hierzu siehe Methodenbericht: Kapitel 7). Die längsschnittlichen GewichtungsvARIABLEN und die Variablen der reziproken Teilnahmewahrscheinlichkeiten aller Einzelwellen des DEAS werden in einer gesonderten Datei bereitgestellt und stehen registrierten Datennutzerinnen und Datennutzern zum Download zur Verfügung.

Zusätzlich wird für jede Erhebungswelle ab 2002 eine *integrierte Querschnittsgewichtung* vorgenommen. Dazu werden jeweils alle Teilstichproben einer Welle zu einer gemeinsamen Querschnittstichprobe integriert. Bei den Teilstichproben handelt es sich um (a) Panelfälle mit Teilnahme in der Vorwelle (Wiederteilnehmende), (b) Panelfälle ohne Teilnahme in der Vorwelle (Rückkehrer nach temporärem Ausfall) und (c) neue Basisfälle (Erstbefragte, sofern im Erhebungsjahr eine neue Stichprobe hinzukommt). Ausgangsgewicht für die Integration ist für die Wiederteilnehmenden aus der Vorwelle das Längsschnittgewicht der Betrachtungswelle. Für die Rückkehrer nach temporärem Ausfall (Panelbefragte, die die Vorwelle übersprungen haben) wird mittels zweier Ausfallmodelle (Wahrscheinlichkeit der Nichtteilnahme, Wahrscheinlichkeit der Rückkehr) ein Hilfsgewicht berechnet (für Details hierzu siehe Schiel et al. 2021: Kapitel 7.2). Die Ausgangsgewichte der Teilstichproben werden in einem zweistufigen Verfahren durch Konvexkombination (composite weighting) zu einem integrierten Querschnittgewicht zusammengefügt, indem das jeweilige Ausgangsgewicht mit dem relativen Anteil der Teilstichprobe an der Gesamtstichprobe multipliziert wird. Das Querschnittgewicht im DEAS 2020/21 wertet die Kurzbefragung 2020 als Vorwelle, um die temporären Ausfälle zu bestimmen.

Die integrierten Querschnittgewichte der Welle 2020/21 dienen als Startgewichte für eine Anpassung an die Merkmalsverteilung im Mikrozensus des Jahres 2020. Die Poststratifizierung erfolgt mittels „iterative proportional fitting“ nach den Merkmalen Altersgruppe\*Geschlecht\*Landesteil. Wegen der relativ geringen Anzahl von Befragten im Alter ab 91 Jahren empfehlen wir, das poststratifizierte Querschnittgewicht nicht auf diese Gruppe anzuwenden. Im Scientific Use File zum DEAS 2020/21 enthält die Variable *qsps\_21\_u91* dieses integrierte Querschnittgewicht. Ein poststratifiziertes Querschnittgewicht ohne Altersbegrenzung sowie die integrierten Querschnittgewichte ohne Poststratifizierung sind in einer gesonderten Datei für registrierte Datennutzer verfügbar.

Das Verwenden von Gewichten liegt in der Verantwortung der einzelnen Nutzerinnen und Nutzer. Die Datengewichtung bildet eine von verschiedenen Möglichkeiten des Umgangs mit unterschiedlichen Ziehungswahrscheinlichkeiten und dem Problem der selektiven Teilnahme und Panelmortalität. Der Einsatz der Querschnittgewichtung ist vor allem für allgemeine deskriptive Darstellungen und Analysen vorgesehen, die sich auf einen einzelnen Zeitpunkt beziehen.

Es haben sich geringfügige Anpassungen in den Ausfallmodellen ergeben, die als Grundlage für die Gewichtungsfaktoren dienen. Diese sind online im Dokument „Hinweise zu Neuerungen in der Gewichtung des DEAS“ unter <https://www.dza.de/forschung/fdz/deutscher-alterssurvey/deas-dokumentation> zu finden. Die Änderungen in der Gewichtung haben sich aus zwei Gründen ergeben. Zum einen muss die Kurzbefragung aus dem Sommer 2020 in das Gewichtungsschema integriert werden. Die Kurzbefragung 2020 enthält aber nicht alle Merkmale, die normalerweise in den

Ausfallmodellen der regulären Erhebungen berücksichtigt werden. Das Hinweisdokument stellt die Abweichungen dar. Zum anderen hatte das DEAS-Team in Zusammenarbeit mit dem FDZ geprüft, ob es sinnvoll ist, die Ausfallmodelle geringfügig zu erweitern, um systematische Verzerrungen aufgrund von omitted-variable-bias zu vermeiden. Die Anpassungen, auch wenn nur geringfügig, finden für die Erhebung 2020/21 und für alle folgenden Erhebungen Anwendung.

#### 4 DATENFORMATE

Die Scientific Use Files (SUF) der Jahre 1996 bis 2020/21, der Gewichtungssatz sowie der SUF Meta (siehe Abschnitt 6) sind sowohl im SPSS-Format als auch im Stata-Datenformat verfügbar. Sollte es versionsbedingte Probleme mit dem Einlesen der Daten geben, dann wenden Sie sich bitte direkt an die Beratung des FDZ-DZA ([fdz@dza.de](mailto:fdz@dza.de)). Alle Daten und die dazugehörigen Dokumentationen sind in Deutsch und Englisch verfügbar.

#### 5 SYSTEMATIK DER VARIABLENNAMEN UND MISSING-CODES IM DEAS 2020/21

Auch in der siebten Welle des DEAS wird die 2008 eingeführte Systematik zur Benennung der Variablen verwendet. In den ersten beiden Wellen des DEAS wurden die Variablenbezeichnungen noch entsprechend der Verkodungstradition eines Kartensystems vergeben.<sup>3</sup> Um die Bezeichnung der Variablen besser erfassbar zu machen, wurden sie ab der dritten Welle anhand der Fragennummer fortlaufend vergeben. Die Fragennummern in den DEAS-Wellen sind thematisch gruppiert (beispielsweise enthält der 300er Fragenblock alle Angaben zur familiären Situation).

Die Variablen der Befragung des Jahres 2020/21 beginnen mit einem ‚k‘ als Wellenmarkierer, gefolgt von einem ‚c‘ oder ‚d‘ als Kennzeichnung der Befragungsart (c = mündliches Interview; d = schriftlicher oder Onlinefragebogen). Die dann folgende Zahl entspricht der jeweiligen Fragennummer. Bei Mehrfachantworten werden die Antwortoptionen durch eine fortlaufende numerische Endung unterschieden (z.B. kc427\_1, kc427\_2, etc.). Offene Angaben sind durch die Endung ‚o‘ gekennzeichnet (kc103b\_5o).

Einen vollständigen Überblick über alle jemals in den DEAS-Datensätzen verwendeten Variablennamen befindet sich in der Variablenkorrespondenzliste auf der [Dokumentationsseite](#) des FDZ-DZA. Auf der

---

<sup>3</sup> Zum Beispiel befindet sich die Frage 101 zum Erwerbsstatus aus dem persönlichen Interview der zweiten Welle im Jahr 2002 (bezeichnet mit ‚w‘) auf der Karte 14 an *zweiter Stelle* und lautet daher w14\_12.

Dokumentationsseite steht auch das Codebuch mit der Grundauszählung aller im SUF DEAS 2020/21 vorhandenen Variablen inklusive der Labels.

Die Vergabe der Missing-Codes folgt der ab Welle 4 neu entwickelten Systematik. Diese erleichtert das automatische Definieren fehlender Werte und vereinheitlicht die Ausprägungen und die Ursachen für das Vorliegen nicht-valider Angaben.<sup>4</sup>

Die folgende Tabelle beschreibt die im Datensatz vorkommenden Codes für fehlende Werte.

Wert SPSS	Wert Stata	Label [Erklärung]
-1	.a	verweigert
-2	.b	weiß nicht
-3	.c	überfiltert (Frage)
-4	.d	überfiltert (Stichprobe) <i>[wenn sich die Frage nur an Panel- oder Basisbefragte richtet; entfällt in der Welle 2020/21]</i>
-5	.e	kein Drop-off vorhanden
-6	.f	keine Angabe ( <i>Drop-off</i> )
-7	.g	gelöscht in Datenaufbereitung

Im Allgemeinen wurden die Missing-Codes im Datensatz als fehlende Werte (user-defined-system missings) formatiert. Die Ausnahme sind die Wellen 1 bis 3 (1996, 2002 und 2008), die noch nicht der allgemeinen Systematik der Missingbenennung folgen. Für diese Wellen empfehlen wir vor der Analyse, Häufigkeitsauszählungen der analysierten Variablen vorzunehmen, um Missingcodes im validen Wertebereich (z.B. 8, 9, oder 98) vorab zu identifizieren und als Missings zu deklarieren.

## 6 META-DATENSATZ ZUR TEILNAHMEHISTORIE DER DEAS-BEFRAGTEN

Zur verbesserten Übersicht, zum einfacheren Zusammenspielen von Analysedatensätzen und für Verbleibsanalysen stellt das FDZ einen Meta-Datensatz zur Verfügung, der alle Personen enthält, die jemals Befragungsteilnehmer und -teilnehmerinnen im DEAS waren. Im Einzelnen finden sich folgende

---

<sup>4</sup> In den Daten der Wellen 1 bis 3 waren fehlende Werte entweder als 0, -1, -2 oder als 7,8,9 bzw. 97, 98, 99 etc. codiert – je nach Ausprägung der validen Werte. Eine einheitliche Missingdefinition über alle Variablen hinweg war damit nicht möglich.

Variablen im Meta-Datensatz, die teilweise auch in den wellenspezifischen SUFs enthalten sind (Stand: Januar 2021):

Variable	Label	Erläuterung
fallnum	Fallnummer (anonymisierte Befragten-ID)	zum Zusammenspielen der Informationen aus allen SUFs (1996-2014)
stich	Stichprobenzugehörigkeit	Angabe der Basisstichprobe, aus der die befragte Person stammt
part_96	Teilnahme in 1996	ja; nein
part_02	Teilnahme in 2002	ja – Basis 2002; ja – Ausl. 2002; ja – Panel; nein
part_08	Teilnahme in 2008	ja – Basis 2008; ja – Panel; nein
part_11	Teilnahme in 2011	ja – Panel; nein
part_14	Teilnahme in 2014	ja – Basis 2014; ja – Panel; nein
part_17	Teilnahme in 2017	ja – Panel; nein
part_20	Teilnahme an Kurzbefragung 2020	ja – Panel; nein
part_21	Teilnahme in 2020/21	ja – Panel; nein
part	Teilnahmeprofil, 1996-2020/21	zeigt an, in welchen Erhebungsjahren die Person befragt wurde
lastpart	Letzte Teilnahme	Jahr der letzten Befragung
n_obs	Anzahl der Teilnahmen	Anzahl individueller Befragungsteilnahmen (1-8)
entry	DEAS-Eintrittsjahr	analog zu stich: Kalenderjahr der Erstbefragung
yob	Geburtsjahr	Geburtsjahr (evtl. aktualisiert im Zuge der Datenbereinigung)
gender	Geschlecht	Geschlecht
natdeutsch	Deutsche Staatsangehörigkeit	Merkmal beim ersten Interview: deutsch; nicht-deutsch
migrat	Migrationshintergrund	ohne MH; eingewandert; MH ohne eigene Einwanderung
bildung4	Bildungsniveau, 4-stufig	ermittelt aus höchstem Schul- und Berufsbildungsabschluss: niedrig, mittel, gehoben, hoch

<b>Variable</b>	<b>Label</b>	<b>Erläuterung</b>
isced	Bildungsniveau nach ISCED, 3-stufig	niedrig (ISCED 0-2), mittel (ISCED 3-4), hoch (ISCED 5-6)
bland_96 bis bland_21	Bundesland	gemäß Wohnort zum Befragungszeitpunkt
westost_96 bis westost_21	West-Ost-Zugehörigkeit des Wohnorts	mit Aufteilung Berlins
westost_21neu	West-Ost-Zugehörigkeit des Wohnorts	Berlin=Ost
bbsr_kreistyp96 bis bbsr_kreistyp21	Kreistyp	Siedlungsstruktureller Kreistyp des Wohnorts nach BBSR/INKAR
rlc2002_kat bis rlc2021_kat	Rücklaufcode	Gründe der Nichtteilnahme an Welle 2, 3, 4, 5, 6, 7 (kurz) oder 7 (regulär) (oder Code für Teilnahme)
vitalstatus_last	Letzter bekannter Vitalstatus	gibt an, ob die Person noch lebt oder inzwischen verstorben ist
vitalstatus_source	Jahr der Information über den Vitalstatus	Kalenderjahr, aus dem die Information über den Vitalstatus stammt
surv_birth	Überlebensdauer seit Geburt	Monate seit Geburt bis Tod oder Beobachtungsende
surv1996	Überlebensdauer seit Befragung 1996	Monate seit Befragung 1996 bis Tod oder Beobachtungsende
surv2002	Überlebensdauer seit Befragung 2002	Monate seit Befragung 2002 bis Tod oder Beobachtungsende
surv2008	Überlebensdauer seit Befragung 2008	Monate seit Befragung 2008 bis Tod oder Beobachtungsende
surv2011	Überlebensdauer seit Befragung 2011	Monate seit Befragung 2011 bis Tod oder Beobachtungsende
surv2014	Überlebensdauer seit Befragung 2014	Monate seit Befragung 2014 bis Tod oder Beobachtungsende
surv2017	Überlebensdauer seit Befragung 2017	Monate seit Befragung 2017 bis Tod oder Beobachtungsende
surv2021	Überlebensdauer seit Befragung 2021	Monate seit Befragung 2021 bis Tod oder Beobachtungsende
psu	Primary Sampling Unit	Die PSUs sind die zufällig und größenproportional gezogenen 290 Gemeinden aus denen die Befragten

Variable	Label	Erläuterung
		zufällig und stratifiziert gezogen wurden. In den Fällen wo die Gemeinden sehr schwach mit Befragten besetzt sind (N<10) wurde die Variable für den SUF vergrößert. Die Befragten wurden dem übergeordneten Kreis zugeordnet, zusammen mit Befragten aus den Gemeinden, die ebenfalls dem entsprechenden Kreis angehören. Von den ursprünglichen 290 PUSs bleiben nach der Vergrößerung 266 PSUs übrig. Die vergrößerte überschneidungsfreie PSU-Variable bezieht sich auf das Jahr der Erstbefragung und ist maskiert. D.h. die PSUs sind durch systemfreie laufende Ziffern voneinander differenziert, die keinen Rückschluss mehr auf die Originalerhebungseinheit zulassen.
ssu	Secondary Sampling Unit	Die ssu identifiziert die Befragten eindeutig mit Hilfe einer unveränderlichen Personenidentifikationsnummer. Sie entspricht der Variable fallnum.
strata_gemeinde	Schichtungsvariable Gemeindeebene	Die PSU gehört zu den alten oder neuen Bundesländern
strata	Schichtungsvariable Individualebene	Ordnet die SSUs bzw. die Befragten einer der 12 Stratifizierungszellen zu, die sich aus der Kombination von 3 Altersgruppen, Geschlecht und alte oder neue Bundesländer ergeben. Die Zuordnung zu einer der 12 Zellen bezieht sich auf den Zeitpunkt der Erstbefragung.

Vor allem das Erstellen eines Längsschnittdatensatzes wird durch Verwendung des Meta-Datensatzes erleichtert. Die Verbleibsdaten und Drop-out-Informationen eröffnen zudem die Möglichkeit von Mortalitäts- und Ausfallanalysen. Bei Interesse an einer differenzierten Analyse der Ausfallgründe sind auf Nachfrage beim FDZ auch noch detailliertere Ausfallcodes verfügbar.

Das komplexe Surveydesign des DEAS kann nun mit Hilfe der Variablen *psu*, *ssu*, *strata\_gemeinde* und *strata* für Analysen berücksichtigt werden (siehe 8. Das Komplexe Surveydesign des DEAS).

## 6.1 Teilnahmeprofil

In der Variable *part* sind die Teilnahmen über alle bisherigen DEAS-Erhebungsjahre zusammengefasst. Allen Personen, die jemals DEAS-Befragte waren, ist ein achtstelliger Wert zugewiesen: bspw. bedeutet die Ausprägung 11111111, dass die Befragungsperson an allen acht DEAS-Erhebungen teilgenommen hat, Personen mit der Ausprägung 01001111 haben erstmals in 2002 teilgenommen, waren in 2008 und 2011 nicht dabei, haben aber in 2014, 2017, 2020 kurz und 2020/21 wieder das Frageprogramm durchlaufen. Mithilfe des Teilnahmeprofiles kann das Analysesample sehr differenziert ausgewählt und zusammengestellt werden. Soll eine Auswahl nur über die Anzahl der individuellen Teilnahmehäufigkeit erfolgen, kann dies mittels der Variable *n\_obs* erfolgen. Auf Grundlage der vorselektierten Beobachtungen lassen sich in einem weiteren Schritt die inhaltlichen Informationen aus den wellenspezifischen SUFs über die Identifizierungsvariable *fallnum* hinzuspielen. Bei Fragen zur Nutzung des Meta-Datensatzes wenden Sie sich bitte an die Beratung des FDZ: [fdz@dza.de](mailto:fdz@dza.de).

## 6.2 Verbleibsdaten

Bei den Kontaktaufnahmen zu den Befragten anlässlich der Wiederholungsbefragungen sowie mehrerer Panelpflegeaktionen und spezieller Verbleibsrecherchen über die Einwohnermeldeämter wurde vom Feldforschungsinstitut jeweils der Vitalstatus erfasst und protokolliert. Bei bekannt gewordenen Todesfällen wurde dabei möglichst auch das Sterbedatum ermittelt. Die Variable *vitalstatus\_last* weist den letzten bekannten Vitalstatus aller jemals befragten DEAS-Teilnehmerinnen und –teilnehmer aus (0=lebend, 1=tot). Wenn Befragte nur einmal interviewt wurden und niemals eine Panelbereitschaft abgegeben haben, bezieht sich die Angabe auf das Datum des ersten und einzigen Interviews, ansonsten auf den letzten Informationsstand. Die Variable *vitalstatus\_source* enthält das Kalenderjahr, in dem Informationen über den letzten bekannten Vitalstatus vorliegen. Die Variable *surv\_birth* enthält die seit der Geburt verstrichenen Monate bis zum Sterbedatum bzw. bis zum letzten bekannten Überlebensdatum. Die Variablen *surv\_birth*, *surv1996*, *surv2002*, *surv2008*, *surv2011* und *surv...* zur Überlebensdauer seit Geburt und seit der jeweiligen Befragung bilden zusammen mit der Variable *vitalstatus\_last* (Sterbe- oder Zensierungsindikator) die Grundlage für Ereignisanalysen zum Sterblichkeitsrisiko.

Die Variablen *rlc2002\_kat* bis *rlc2021\_kat* enthalten die kategorisierten Rücklaufcodes anlässlich der Kontaktaufnahmen für die DEAS-Erhebungen 2002 bis 2021. Sie geben Auskunft darüber, ob die Zielperson befragt werden konnte oder welcher Grund maßgeblich für das Nicht-Zustandekommen eines gültigen Interviews war. Interviews, die nachträglich storniert wurden, enthalten den Code 6 für einen Ausfall aus sonstigen Gründen.

## 7 VERÄNDERUNG VON DATEN IM SUF ZUR WAHRUNG DER FAKTISCHEN ANONYMITÄT

Um die faktische Anonymität der Untersuchungsteilnehmerinnen und -teilnehmer zu wahren, wurden einzelne Variablen nicht in den Scientific Use File (SUF) DEAS2021 aufgenommen und bei einer Reihe von Variablen Vergrößerungen der Ausprägungen vorgenommen.

- Gelöscht wurden unter anderem alle Variablen zum räumlichen Kontext, die unter Umständen eine Bestimmung der Wohngemeinde der Befragten ermöglichen würden. Entfernt wurden auch die Angaben zum Geburtstag und -monat der Befragten.
- Einzelne Variablen mit offenen Angaben wurden gelöscht oder es wurden offene Angaben verallgemeinert, die sehr spezifische Merkmale enthielten. Solche Veränderungen bei offenen Angaben wurden durch das Voranstellen des Zeichens ‚#‘ kenntlich gemacht.
- Seltene familiendemografische Merkmale wurden vergrößert.
- Namensangaben die zu Behelfszwecken für die Erfassung z.B. der sozialen Netzwerke erfasst wurden, wurden ebenfalls gelöscht
- Seltene Jahresangaben wurden zu Kategorien zusammengefasst. Dies geschah durch Rekodierung des Werts in der vorhandenen numerischen Variable auf Werte der Mitte des zusammengefassten Bereichs oder auf den Modalwert (mit entsprechender Kennzeichnung durch Wertelabels).
- Vereinzelt wurden Maximalwerte zu einer Maximalkategorie zusammengefasst (z.B. Haushaltsgröße).

Am DZA sind speziell gesicherte PC-Arbeitsplätze für Gastnutzer eingerichtet, die es erlauben, Auswertungen mit den vollständigen DEAS-Datensätzen vorzunehmen, z. B. zur kleinräumigen Analyse auf Kreisebene oder unter Einbezug von Kontextmerkmalen der unmittelbaren Nachbarschaft. Interessenten wenden sich dazu bitte an die Beratung des FDZ: [fdz@dza.de](mailto:fdz@dza.de).

## 8 DAS KOMPLEXE SURVEYDESIGN DES DEAS

Der DEAS beruht nicht auf einer einfachen, sondern auf einer mehrfach geschichteten und geklumpten Zufallsstichprobe. Um sowohl zufälligen Verzerrungen (z.B. könnten Männer in einfachen Zufallsstichproben zufällig überproportional häufig gesampelt werden), als auch systematische Verzerrungen vorzubeugen (weil zum Beispiel Hochaltrige aus gesundheitlichen Gründen seltener an der Befragung teilnehmen als jüngere Befragte), wird die Population vorab in Schichten (Strata) eingeteilt. Im DEAS entstehen diese Schichten aus einer Kombination von Geschlecht, drei Altersgruppen und der Zugehörigkeit zu den alten oder den neuen Bundesländern. Personen die zur ältesten Altersgruppe und den neuen Bundesländern gehören werden zudem oversampled. Die zufällige Ziehung der Befragten findet innerhalb dieser Schichten statt. Da der DEAS als persönliches Interview durchgeführt wird, ist es logistisch und kostentechnisch nahezu unmöglich die gesamte Bundesrepublik abzudecken. Die Reise- und Organisationskosten für die Durchführung der persönlichen Interviews wären unverhältnismäßig hoch. Aus diesem Grund wird der DEAS geklumpt. Aus den 14627

Gemeinden die im Jahr 1996 zur Auswahl standen, wurde ein größen-proportionales Sample aus den alten Bundesländern und ein größen-proportionales Sample aus den neuen Bundesländern gezogen. Größen-proportional heißt, dass größere Gemeinden eine höhere Ziehungswahrscheinlichkeit hatten als kleinere Gemeinden. Im Ergebnis wurde die stratifizierte Individualstichprobe jeweils aus 90 Gemeinden in den neuen Bundesländern und 200 Gemeinden aus den alten Bundesländern gezogen.

Wird dieses komplexe Stichprobendesign nicht berücksichtigt (insbesondere die Klumpung), werden die Befragten als voneinander unabhängige Beobachtungseinheiten gewertet. Befragte die jedoch aus den gleichen Gemeinden stammen, können einander aber ähnlich sein (z.B. Befragte aus armen/reichen Gemeinden sind sich in ihrem Einkommen wahrscheinlich sehr ähnlich). Wird die Klumpung nicht berücksichtigt, wird die Varianz unterschätzt und im Ergebnis werden signifikante Effekte/Zusammenhänge gefunden, die bei der korrekten Berücksichtigung des Designs, nicht gefunden werden würden.

Aus diesem Grund werden im DEAS zusätzliche Variablen bereitgestellt, um den statistischen Analyseprogrammen die Komplexität des Surveydesigns mitteilen zu können. In Stata kann dies mit der svy-Suite geschehen. Im Metadatensatz werden dafür die folgenden Merkmale bereitgestellt: psu (Primary Sampling Unit), ssu (Secondary Sampling Unit), strata\_gemeinde (Stratifizierung der PSUs) und strata (Stratifizierung der SSUs).

Der svyset Befehl in Stata, mit dem die Surveydesignmerkmale des DEAS definiert werden sollten, sieht folgendermaßen aus:

```
svyset psu, strata(strata_psu) vce(linearized) singleunit(scaled) || ssu [pw=weight], strata(strata_ssu)
```

Das passende Gewicht (weight) muss gemäß dem Analysedesign selbständig gewählt oder erstellt werden. Alle folgenden Analysen, die den Suffix -svy- beinhalten, berücksichtigen das komplexe Surveydesign des DEAS, das mit Hilfe des svyset-Befehls definiert wurde.

## 9 SONSTIGES

### 9.1 Dokumentation zum DEAS2021

Auf der [Website](#) des FDZ-DZA befinden sich im Bereich der DEAS-Dokumentation zur Ansicht und zum Download diverse Dokumentationsmaterialien zu allen verfügbaren DEAS-Wellen - das vollständige Erhebungsinstrument, der infas-Methodenbericht, das Codebuch, diverse Arbeitshilfen (etwa zu Einzelmerkmalen), eine Variablenkorrespondenzliste in Tabellenform, die ausführliche Variablendokumentation sowie eine Erläuterung zur Datengewichtung. Die meisten dieser Dokumente sind zusätzlich in Englisch verfügbar.

Ebenfalls auf der FDZ-DZA-Seite sind alle uns bekannten Publikationen auf Basis des DEAS gelistet - als komplettes PDF-Dokument oder zum komfortablen Durchsuchen mit Filteroptionen.

## 9.2 Regionaldaten

Den Befragungsdaten können Merkmale des Regionalkontexts auf der Ebene der Landkreise und kreisfreien Städte oder auf der Ebene von Nachbarschaften hinzugespielt und für Analysen an einem besonders geschützten Nutzerarbeitsplatz im DZA zur Verfügung gestellt werden. Interessierte wenden sich bitte direkt an die Beratung des FDZ-DZA.

## LITERATUR

- Bundesamt für Bau-, Stadt- und Raumforschung (BBSR) im Bundesamt für Bauwesen und Raumordnung (2012). Raumabgrenzungen und Raumtypen des BBSR. Bonn: BBSR. URL: (<https://www.inkar.de/>) [zuletzt gesehen am 01.05.22].
- European Central Bank (2016). The Household Finance and Consumption Survey (HFCS): Methodological report for the second wave, ECB Statistical Paper Series No. 17). Frankfurt a. M.: European Central Bank. URL: <https://www.ecb.europa.eu/pub/research/statistics-papers/html/index.en.html> [zuletzt gesehen am 01.06.22].
- Hammon, A., Zinn, S., Aßmann, C., & Würbach, A. (2016). Samples, weights, and nonresponse: the adult cohort of the National Educational Panel Study (Wave 2 to 6) (NEPS Survey Paper No. 7). Bamberg, Germany: Leibniz Institute for Educational Trajectories, National Educational Panel Study. URL: <https://www.neps-data.de/Datenzentrum/Publikationen/NEPS-Survey-Papers> [zuletzt gesehen am 01.06.22].
- Klaus, D., Engstler, H., Mahne, K., Wolff, J. K., Simonson, J., Wurm, S., & Tesch-Römer, C. (2017). Cohort Profile: The German Ageing Survey (DEAS). International Journal of Epidemiology, 46(4), 1105-1105g. doi: <https://doi.org/10.1093/ije/dyw326>
- Mahne, K., Wolff, J. K., Simonson, J., & Tesch-Römer, C. (Eds.) (2017). Altern im Wandel. Zwei Jahrzehnte Deutscher Alterssurvey (DEAS). Wiesbaden: Springer VS. Open Access: doi: <https://doi.org/10.1007/978-3-658-12502-8>.
- Schiel, S., Ruiz Marcos, J., Schulz, S., Dickmann, C. & Fahrenholz, M. (2021). Methodenbericht. Deutscher Alterssurvey (DEAS): Durchführung der 7. Erhebungswelle 2020/21. Bonn: infas. URL: <https://www.dza.de/forschung/fdz/fdz-publikationen> [zuletzt gesehen 01.06.22].
- Trapmann, M. (2014). In: Bethmann, A., Fuchs, B., Wurdack, A., Beste, J., Trappmann, M., Müller, G., Eggs, J., Gundert, S., Wenzig, C., Gebhardt, D. (Hrsg.) (2013). User guide "Panel Study Labour Market and Social Security" (PASS) \* wave 6. (FDZ-Datenreport, 07/2013 (en)), Nürnberg: IAB

(S. 81-99). URL: <https://fdz.iab.de/187/section.aspx/Publikation/k131218302> [zuletzt gesehen am 01.06.22].

Vogel, C., Wettstein, M., & Tesch-Römer, C. (Eds.) (2019). Frauen und Männer in der zweiten Lebenshälfte. Älterwerden im sozialen Wandel. Wiesbaden: Springer VS. Open Access doi: <https://doi.org/10.1007/978-3-658-25079-9>