

BIBB/BAuA-Erwerbstätigenbefragung 2018

Methodenbericht

Forschungsprojekt im Auftrag des
Bundesinstituts für Berufsbildung (BIBB) und der
Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA)

Durchgeführt von:
Kantar Public (ehemals: TNS Infratest Sozialforschung)

Miriam Gensicke
t 089 5600 1585
miriam.gensicke@kantarc.com

Nikolai Tschersich
t 089 5600 1954
nikolai.tschersich@kantarc.com

München, Juli 2018

MB_315114735

Inhaltsverzeichnis

1	Ziele und Untersuchungsanlage des Gesamtprojekts	1
2	Erhebungsinstrument	2
3	Grundgesamtheit und Anlage der Stichprobe	4
3.1	Auswahlgrundlage und Stichprobenziehung	4
3.1.1	Auswahl einer zu kontaktierenden Telefonnummer	5
3.1.2	Stichprobenrealisierung	6
3.2	Auswahl von Erwerbstätigen ab 15 Jahren	10
3.3	Rekrutierung für die Nachbefragung (Los 2)	11
4	Durchführung der Erhebung	12
4.1	Interviewerschulung	12
4.2	Interviewereinsatz und Durchführung der Feldarbeit	13
4.3	Maßnahmen zur Qualitätssicherung und Interviewerkontrolle	14
4.4	Ergebnisse der Feldarbeit	16
4.5	Kurzbefragung von Nichtteilnahmebereiten	22
5	Datenprüfung und -aufbereitung	23
6	Vercodung der offenen Angaben zur beruflichen Tätigkeit und zum Wirtschaftszweig	25
6.1	Berufevercodung	26
6.1.1	Das Verfahren bei Kantar Public	26
6.1.2	Vorgehensweise und allgemeine Regeln bei der manuellen Vercodung	27
6.1.3	Besonderheiten der jeweiligen Klassifikationssysteme	29
6.2	Branchenvercodung	31
6.2.1	Das Verfahren bei Kantar Public	31
6.2.2	Vorgehensweise und allgemeine Regeln bei der manuellen Vercodung	31
6.2.3	Spezielle Regeln und Besonderheiten der jeweiligen Klassifikationssysteme	32
7	Gewichtung	34
7.1	Designgewichtung	34
7.2	Ausfallgewichtung über Strukturanpassung	36
7.3	Qualitätsindikatoren der Gewichtung	39

1 Ziele und Untersuchungsanlage des Gesamtprojekts

Das Bundesinstitut für Berufsbildung (BIBB) und die Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) haben im Mai 2017 Kantar Public mit der Durchführung der „BIBB/BAuA-Erwerbstätigenbefragung 2018“ beauftragt. Ziel der Erhebung ist es, die Arbeitsbedingungen und -anforderungen, den Erwerb und die Verwertung beruflicher Kenntnisse sowie die diesbezüglichen Entwicklungen und Veränderungen im Zuge technischen und organisatorischen Wandels der Arbeitswelt zu erfassen. Die Gesamtuntersuchung unterteilte sich in zwei Lose, wobei Los 1 die Erhebung von 20.000 CATI-Interviews mit Erwerbstätigen umfasste und Los 2 eine Nacherhebung bei 8.000 Erwerbstätigen aus Los 1 sowie eine Befragung von 1.000 Nicht-Erwerbstätigen beinhaltete, die im Rahmen von Los 1 rekrutiert wurden. Kantar Public erhielt den Zuschlag für Los 1.

Im Rahmen von Los 1 war vorgesehen, von Oktober 2017 bis Anfang April 2018 rund 20.000 Erwerbstätige ab 15 Jahren telefonisch und computerunterstützt, d. h. per CATI (Computer Assisted Telephone Interview) zu befragen. Die Telefonerhebung erfolgte anhand einer Dual-Frame-Stichprobe, d. h. neben Interviews über Festnetznummern wurden erstmals bei der BIBB/BAuA-Erwerbstätigenbefragung auch Interviews über Mobilfunknummern realisiert. Die durchschnittliche Interviewlänge war mit 40 Minuten geplant. Themen der Befragung waren die Arbeitsanforderungen und Tätigkeiten von Erwerbstätigen, ihre Arbeitsbedingungen, gesundheitliche Belastungen und Beschwerden, sowie der Erwerb und die Verwertung beruflicher Qualifikationen.

Im Anschluss an die „BIBB/BAuA Erwerbstätigenbefragung 2018“ (Los 1) erfolgte eine Nachbefragung (Los 2) von 8.000 Erwerbstätigen sowie 1.000 Nicht-Erwerbstätigen, die im Rahmen von Los 1 identifiziert werden sollen. Im Zuge der Zielpersonenermittlung von Los 1 sollten in ausreichendem Umfang befragungsbereite nichterwerbstätige Personen rekrutiert werden, deren Kontaktdaten in 14-tägigem Abstand an das BIBB für die Durchführung der Los 2-Nachbefragung übermittelt werden sollten.

Der vorliegende Methodenbericht beschreibt zunächst die Entwicklung des Erhebungsinstruments. Anschließend werden die Grundgesamtheit und die Anlage der Stichprobe dargestellt. Die Realisierung in der praktischen Feldarbeit, Interviewereinsatz und -kontrolle werden erläutert. Es folgt ein Abschnitt, in dem das Vorgehen bei der Vercodung der beruflichen Tätigkeit sowie des Wirtschaftszweigs dokumentiert ist. Abschließend werden das Gewichtungsmodell sowie die Ergebnisse der Gewichtung dargestellt.

Das Projekt wurde bei Kantar Public von einem Projektteam bearbeitet. Die Projektleitung lag bei Miriam Gensicke und Nikolai Tschersich. Sie wurden beratend unterstützt durch Dr. Josef Hartmann. Zuständig für die Projektkoordination war Sören Schleypen, für die Fragebogenprogrammierung war es Gerd Döring, für die Stichprobensteuerung Susanne Schober und Christian Stumpf, für die Gewichtung waren es Rita Reimer und Martin Pfister.

2 Erhebungsinstrument

Das Erhebungsinstrument der Befragung wurde vom BIBB und der BAuA auf Basis der letzten Erhebung entwickelt und am 2. Juni 2017 in einem persönlichen Treffen mit der Projektgruppe in den Räumlichkeiten des BIBB diskutiert. Wesentliche Bestandteile des Erhebungsinstruments sind dabei identisch zur letzten Erhebung. An einigen Stellen wurden Fragen herausgenommen und durch neue Fragestellungen ersetzt.

Im Anschluss an die Fragebogenabstimmung wurden die Fragen zur Ermittlung der Zielpersonenanzahl und der Auswahl der Zielperson im Rahmen der Festnetzstichprobe entwickelt und mit dem Auftraggeber abgestimmt. Das Erhebungsinstrument wurde bei Kantar Public in ein Programm für computergestützte telefonische Interviews (CATI) umgesetzt. Als Befragungssoftware wurde dabei NIPO 32 genutzt.

Das BIBB und die BAuA erhielten einen Testzugang zum CATI-Fragebogen, um selbst Testinterviews durchführen zu können. Vor dem Start der Haupterhebung wurden zwei Pretests durchgeführt.

In der Erhebung 2018 gab es gegenüber der Erhebung 2012 in der Stichprobenanlage zwei Änderungen, die eine Neuentwicklung des Intervieweinstiegs erforderlich machten:

- Umstellung auf eine Dual-Frame-Stichprobe
- Ermittlung der Zielperson bei Festnetz auf Haushaltsebene unter allen Personen ab 15 Jahren, die mindestens 10 Stunden pro Woche erwerbstätig sind oder ihre Erwerbstätigkeit für maximal 3 Monate unterbrochen haben (2012: Auswahl unter allen Personen im Haushalt ab 15 Jahren).

Die Ermittlung der Anzahl der erwerbstätigen Personen im Haushalt im Rahmen der Festnetzstichprobe, unter denen die Zufallsauswahl erfolgte, wurde anhand von mehreren Fragen vorgenommen, die im Rahmen der beiden Pretests in verschiedenen Varianten getestet wurden. Mit der ausgewählten Zielperson wurde im Anschluss ein kurzes Screeninginterview geführt, um festzustellen, ob sie erwerbstätig im definierten Sinne der Untersuchung ist.

Die Ermittlung der Zielperson in der Mobilfunkstichprobe erfolgte direkt anhand des kurzen Screeninginterviews, wobei zuvor festgestellt wurde, ob die Person 15 Jahre oder älter sowie der Hauptnutzer des Handys ist. Da es sich hier um eine Personenstichprobe und keine Haushaltsstichprobe handelt, erfolgte hier keine Zufallsauswahl unter mehreren potenziellen Zielpersonen. Das Screeninginterview selbst war in den wesentlichen Bestandteilen identisch zu dem der Erhebungswelle 2012.

Designbedingte Unterschiede zur Stichprobenanlage früherer Erhebungswellen können mittels Designgewichtung aufgehoben werden, so dass diese miteinander vergleichbar sind. Die unterschiedlichen Auswahlwahrscheinlichkeiten sind sowohl für den Dual-Frame-Ansatz als auch für die unterschiedliche Anzahl an Zielpersonen im Haushalt bekannt und damit berechenbar (siehe auch Kapitel 7.1).

Die Feldarbeit des ersten Pretests fand vom 31. Juli bis zum 10. August 2017 statt, der zweite Pretest wurde in der Zeit vom 12. September bis zum 18. September 2017 durchgeführt. Direkt vor Start der Feldarbeit des ersten Pretests fand die Interviewereinweisung statt, die gemeinsam von BIBB, BAuA und Kantar Public durchgeführt wurde. Eine Vertreterin der Projektleitung von Kantar Public war bei der Einweisung im Telefonstudio in Parchim. Jeweils eine Vertreterin von BIBB und BAuA sowie ein Vertreter von Kantar Public waren in den Räumlichkeiten von Kantar in München über Webex nach Parchim zugeschaltet. Vor dem zweiten Pretest wurde keine Einweisung durchgeführt, da dieselben Interviewer wie beim ersten Pretest zum Einsatz kamen. Im Rahmen des ersten Pretests wurden 200 Interviews mit Erwerbstätigen durchgeführt, beim zweiten Pretest waren es 100 Interviews.

Um Probleme mit dem Instrument aufzudecken, nutzte die Auftraggeberseite sowie die Projektleitung während beider Pretests die Möglichkeit, bei den Telefoninterviews mitzuhören. Darüber hinaus gab es Rückmeldungen von den Supervisoren des Telefonstudios zu Problemen oder Auffälligkeiten sowie

Hinweise zu Verbesserungsmöglichkeiten für die Haupterhebung auf Basis der Interviewererfahrungsberichte.

Änderungsbedarf zeigte sich in erster Linie beim Intervieweinstieg sowie aufgrund der zu hohen Interviewdauer. Darüber hinaus gab es einzelne Fragen, bei denen sich Verbesserungsbedarf zeigte. Das methodische Vorgehen sowie die Ergebnisse dieser Pretests sind in jeweils eigenen Berichten dokumentiert.¹

Das Erhebungsinstrument beinhaltet unterschiedlich lange Themenblöcke mit folgenden Fragennummern, die Übersicht 1 zu entnehmen sind:

Übersicht 1: Inhalte des Fragebogens

Frageblock	Inhalt
F100ff:	Aktuelle Berufstätigkeit
F200ff:	Arbeitszeit und Arbeitsort
F300ff:	Ausgeübte Tätigkeiten
F400ff:	Berufliche Anforderungen
F500ff:	Arbeitsverhältnis
F600ff/F700ff:	Arbeitsbedingungen
F900ff:	Berufs- und Lebensziele
F1000ff:	Veränderungen in den letzten zwei Jahren
F1100ff:	Schulbildung
F1200ff:	Berufsausbildung (in bis zu 5 Schleifen)
F1225ff:	Zusammenhang (letzte) Ausbildung und Beschäftigung
F1300ff:	Weiterbildung
F1400ff:	Berufsverlauf
F1450ff:	Arbeitszufriedenheit
F1500ff:	Gesundheit
F1600ff:	Zur Person

Die Informationen, die erhoben wurden, bezogen sich größtenteils auf die befragte Person und ihre aktuelle berufliche Tätigkeit. Hierbei wurden auch einige Informationen zum Betrieb, in dem die Person aktuell tätig ist, erhoben. Retrospektive Informationen wurden zur schulischen und beruflichen Ausbildung erhoben. Für den Zeitraum der letzten 2 Jahre wurden Informationen zur beruflichen Weiterbildung sowie zu Veränderungen im Arbeitsumfeld erfragt.

¹ Kantar Public, 2017: BIBB/BAuA-Erwerbstätigenbefragung 2018 – Bericht Pretest 1. München.
Kantar Public, 2017: BIBB/BAuA-Erwerbstätigenbefragung 2018 – Bericht Pretest 2. München.

3 Grundgesamtheit und Anlage der Stichprobe

Grundgesamtheit der BIBB/BAuA-Erwerbstätigenbefragung 2018 waren deutschsprachige Erwerbstätige in Deutschland ab 15 Jahren mit einer Arbeitszeit von mindestens zehn Stunden pro Woche. Als erwerbstätig wurde definiert, wer zum Zeitpunkt der Befragung eine bezahlte bzw. eine mit einem Einkommen verbundene Tätigkeit ausübte. Diese Definition muss an verschiedenen Stellen präzisiert werden:

- Personen, die ihre Tätigkeit kurzzeitig, d. h. bis zu einer Dauer von maximal drei Monaten, unterbrochen hatten, wurden einbezogen. Hierbei konnte es sich um Mutterschutz oder Elternzeit sowie um Sonderurlaub aus einem anderen Grund handeln.
- Ebenso eingeschlossen wurden Personen, die einer vergüteten Tätigkeit neben einer Ausbildung oder eines Studiums oder im Rahmen eines Referendariats oder einer Facharztausbildung nachgingen.
- Auch mithelfende Familienangehörige wurden befragt. Sie zählen als erwerbstätig, obwohl sie u. a. darüber definiert sind, dass sie ohne feste Vergütung im Betrieb oder Geschäft des anderen Familienmitglieds mitarbeiten.
- Ausländer wurden bei der Untersuchung einbezogen, wenn sie ausreichend deutsch sprachen.
- Dagegen wurden Personen, die der bezahlten Tätigkeit ehrenamtlich nachgehen, nicht befragt.
- Ausgeschlossen wurden auch Personen in einem Beschäftigungsverhältnis, das notwendiger Bestandteil einer Ausbildung ist, sowie Tätigkeiten im Rahmen eines Praktikums.
- Ebenso ausgeschlossen wurden Wehr- oder Freiwilligendienstleistende sowie Personen, die sich in einem freiwilligen sozialen oder ökologischen Jahr befanden.

3.1 Auswahlgrundlage und Stichprobenziehung

Die Entwicklung auf dem Mobilfunkmarkt und das veränderte Kommunikationsverhalten der Bevölkerung haben in den letzten Jahren dazu geführt, dass es etliche Personen gibt, die ausschließlich per Mobiltelefon zu erreichen sind und keinen Festnetzanschluss mehr besitzen (= Mobile-Only). Diese Gruppe kann nicht vernachlässigt werden, da sie mittlerweile einen Anteil von etwa 12 % an allen telefonisch Erreichbaren ausmacht (vgl. z. B. ADM-Forschungsprojekt „Dual-Frame-Ansätze“ 2011/2012 Forschungsbericht). Hinzu kommt, dass es sich dabei um Personen mit spezifischen Merkmalen handelt, so dass die Erwerbstätigenbefragung 2018 als Dual-Frame-Stichprobe konzipiert wurde.² Der Anteil der Interviews über Mobilfunk wurde auf 30 % festgelegt.

Der Auswahlrahmen für die Stichprobenziehung wurde zunächst von dem Institut BIK über ein Random-Digit-Dialing-Verfahren erstellt, das im nächsten Abschnitt näher erläutert wird. Bei diesem Verfahren liegen in der Bruttostichprobe unter den Mobilfunknummern deutlich mehr ungültige bzw. nicht geschaltete Nummern vor als in der Festnetzstichprobe. Dies gilt nach wie vor, obwohl es mittlerweile ein Verfahren gibt (HLR Lookup, siehe ebenfalls nächster Abschnitt), die gezogenen Rufnummern einer Vorabprüfung zu unterziehen. Aufgrund dieses erhöhten Anteils an ungültigen Nummern unter den

² Vor allem jüngere und mobile Personen, zu denen auch ein Teil der Erwerbstätigen zählt, oder Personen, die kürzlich umgezogen sind, sind überdurchschnittlich häufig nur noch mobil erreichbar (Quelle: TNS Infratest Mehrthemenumfrage in 2014).

Mobilfunknummern waren mehr Nummern in der Bruttostichprobe nötig, um die gewünschte Interviewzahl zu realisieren, als dies in der Festnetzstichprobe der Fall war. Das Ziel des Dual-Frame-Ansatzes war es, Mobile-Onlys in ausreichendem Umfang zu erfassen. Aus diesem Grund bezieht sich der Anteil der Mobilfunknummern auf den Anteil der realisierten Interviews aus der Mobilfunkstichprobe. Der Anteil von 30 % Interviews aus der Mobilfunkstichprobe wurde dabei über eine Nettosteuerung realisiert, wobei der benötigte Bruttoansatz höher als beim Festnetzanteil war. Da die über Mobilfunk kontaktierten Personen häufig auch über einen Festnetzanschluss verfügten, war es nicht ausgeschlossen, dass zwar der Kontakt zunächst über Mobilfunk erfolgte, das eigentliche Interview jedoch über eine Festnetznummer stattfand. In diesen Fällen zählten die Interviews zur Mobilfunkstichprobe, da sie über Mobilfunknummern kontaktiert wurden.

3.1.1 Auswahl einer zu kontaktierenden Telefonnummer

Die Stichprobenziehung der Festnetznummern und Mobilfunknummern basiert auf dem hauseigenen Telefon-Master-Sample. Die Dual-Frame-Befragung wurde mit zwei vergleichbaren Auswahlrahmen als kombinierte Festnetz- und Mobilfunkstichprobe durchgeführt. Beide Sampling-Frames bezogen sich auf die gleiche Grundgesamtheit und konnten ex post auf Basis einer Designgewichtung so miteinander verschmolzen werden, dass sie zusammen ein repräsentatives Abbild der Grundgesamtheit bilden. Das Dual-Frame-Verfahren führt zu verzerrungsfreien Stichproben ohne Klumpeneffekte. Sie entsprechen den Anforderungen an Zufallsstichproben, die auf der Wahrscheinlichkeitstheorie aufbauen (probability sampling).

Das Random-Digit-Dialing wurde gemäß dem ADM-Standard (ADM = Arbeitsgemeinschaft Deutscher Markt- und Sozialforschungsinstitute) durchgeführt, eine Erweiterung des sog. „Gabler-Häder-Verfahrens“. Datenbasis war die Auswahlgrundlage für Fest- und Mobilnetznummern der Arbeitsgemeinschaft ADM-Telefonstichproben.

Erstellungsbasis der ADM-Auswahlgrundlage waren die Rufnummernstammlisten der Bundesnetzagentur (BNA) für Mobil- und Festnetz sowie das aktuelle Telefonverzeichnis. Die BNA weist Nummernblöcke aus, die sie den Netzbetreibergesellschaften zugeteilt hat. Der BNA-Bestand ist durch 10er-, 100er-, 1.000er- und 10.000er-Blöcke gekennzeichnet. Im Rahmen eines Abgleichs mit dem aktuellen Telefonbuch wurde ermittelt, ob die generierte Nummer im Telefonbuch eingetragen ist und ob – soweit erkennbar – es sich dabei um einen gewerblichen Anschluss handelt. Wenn Teile in den Einträgen offensichtlich für einen gewerblichen Anschluss sprechen (bspw. „GmbH“ oder „Ltd.“), werden diese Telefonnummern für Personenbefragungen nicht berücksichtigt.

Festnetzstichprobe

Jede Festnetznummer des ADM-Systems wurde auf Gemeindeebene regionalisiert. Die Festnetzstichprobe wurde so gezogen, dass sich eine ausgewogene regionale Verteilung ergab. Eingetragene Rufnummern wurden mit den Informationen des Verzeichnisses direkt verortet. Nichteingetragene Nummern waren nur dann eindeutig verortbar, wenn sich alle eingetragenen Rufnummern des betreffenden Blockes in einer einzigen Gemeinde befanden. War dies nicht der Fall, wurde im Telefon-Master-Sample-System bei den generierten Rufnummern mit mehreren möglichen Gemeindekennziffern eine per Zufall innerhalb des betreffenden Blockes ausgewählt. Diese Zufallsauswahl war mit Bedeutungsgewicht gesteuert; die Häufigkeitsverteilung der Gemeindekennziffern der nichteingetragenen Nummern im jeweiligen Block entsprach der der eingetragenen.

Bei den Festnetznummernblöcken ohne einen einzigen Eintrag war die Gemeindezuordnung komplexer. Es wurden im Falle eines Nummernblockes der Telekom die Gemeindekennziffern der Nachbarblöcke mit Einträgen angesetzt. Bei anderen nationalen und überregionalen Anbietern erfolgte diese Zuordnung pro Ortsnetzvorwahl gemäß der Verteilung der Einträge dieser Gruppe von Netzbetreibergesellschaften.

Im Fall von regionalen Anbietern war die Verteilung der Gemeindekennziffer pro Vorwahl des jeweils einzelnen Anbieters entscheidend.

Die Schichtung der Festnetzstichprobe erfolgte zum einen anhand von Kriterien der amtlichen Gebietseinteilung, zum anderen anhand der BIK-Gemeindetypen. Für die Allokationsrechnung wurde das Cox-Verfahren verwendet. Die allozierte Sollverteilung des Schichtungstableaus wurde haushaltsproportional auf die jeweiligen schichtangehörigen Gemeinden verteilt und so die Ziehungsvorgabe berechnet. Die Ziehung der Rufnummern erfolgte pro Gemeinde per reiner Zufallsauswahl.

Die Stichprobe mit Festnetznummern stellt eine Haushaltsstichprobe dar, bei der ein zweistufiges Auswahlverfahren verwendet wurde. Im Interview wurde mit dem Schwedenschlüssel (Kish-Selection-Grid) zufällig eine Zielpersonenauswahl innerhalb des jeweiligen Haushaltes getroffen (vgl. dazu Abschnitt 3.2).

Mobilfunkstichprobe

Mobilfunknummern wurden mit einfacher Zufallsauswahl gezogen. Die Auswahl der Personen war mit der Selektion der Nummern in der Mobilfunkstichprobe als einfache Zufallswahl bereits abgeschlossen: Der kontaktierte Nutzer des Gerätes war Zielperson. Für die Mobilfunknummernblöcke konnte keine regionale Verortung vorgenommen werden, da die Eintragsdichte mit ca. 0,5 % generell zu niedrig und ein regionaler Rückschluss einer eingetragenen Nummer auf andere Nummern im Nummernblock nicht möglich war. Eine Ex-ante-Schichtung und regionale Filterung entfiel somit für dieses Subsample. Im Telefon-Master-Sample wurde daher eine Ex-post-Schichtung im Verlauf der Feldarbeit nach Abfrage der Postleitzahl und des Wohnorts im Interview verwendet (siehe dazu Abschnitt 3.1.2).

Die Auswahlgrundlage der Festnetz- sowie der Mobilfunkstichprobe enthält lediglich Telefonnummern und qualifizierende Sekundärmerkmale, jedoch keine Namen und Adressen. Die Nummern aller Personen bzw. Betriebe, die auf keinen Fall im Zusammenhang mit einer Befragung angerufen werden wollten, werden in einem zentralen Register gesperrt. Alle Telefonnummern der Bruttostichprobe werden vor dem ersten Anruf mit dieser Liste abgeglichen und gesperrte Nummern erhalten einen entsprechend Ausfallcode und werden nicht kontaktiert.

Dieses erweiterte Gabler-Häder-Verfahren sichert – auch durch die jährliche Aktualisierung – die Inklusion aller vergebenen Rufnummern mit berechenbaren Auswahlwahrscheinlichkeiten. Es garantiert die Selektion von Rufnummern ohne Eintrag und ebenso die Auswahl von Nummern, die in Blöcken ohne Eintrag vorlagen. Dies gilt für Mobilfunk- oder Festnetznummern. Für die Praxis bedeutete dies, dass erstens die wachsende Zahl von Anschlüssen bei den neueren Netzbetreibergesellschaften enthalten war. Zweitens erhielten auch Personen, die über keinen Festnetzanschluss, dafür aber über Mobilfunk verfügen, eine positive Auswahlchance.

3.1.2 Stichprobenrealisierung

Beim vorliegenden Dual-Frame-Design wurde im Vorfeld festgelegt, welches Mischungsverhältnis Festnetz zu Mobilfunk für die Nettofallzahl erreicht werden sollte. Ausschlaggebend bei der Wertung war dabei, ob der erste Kontakt über eine Mobilfunk- bzw. Festnetznummer stattfand, und nicht, ob für das Interview oder Teile des Interviews auf eine andere Nummer umgestiegen wurde. Die Bruttostichproben wurden separat alloziert, getrennt gezogen, aber gemeinsam in der Feldarbeit abgearbeitet. Bei beiden Ziehungen waren – soweit erkennbar – nicht-private Einträge, bereits gezogene sowie generell gesperrte Rufnummern nicht auswählbar.

Die Stichprobenrealisierung erfolgte nach dem Konzept der Nettosteuerung per Sample-Management-System, das zu mikrostratifizierten, ungeklumpten Stichproben führt, die sich bevölkerungsproportional

auf die Mikrozellen aufteilen. Multistratifikation und Aufteilung der Stichprobe auf die Zellen erfolgten vollautomatisch. Das Schichtungstableau der Allokationsrechnung ging als Sollstruktur in die Steuerung ein. Mobilfunk- und Festnetznummern wurden gemeinsam in einem System gesteuert, das gewährleistete, dass in jeder Zelle die erforderliche Zahl von Interviews durchgeführt wurde. Innerhalb jeder Zelle waren die Festnetznummern nach Zufallszahlen sortiert. Mobilfunknummern wurden gemäß ihrem angestrebten Anteil zufällig in die Kontakte insgesamt eingemischt, wobei die Zellenzuordnung im Interview durch Abfrage von Postleitzahl und Wohnort erfolgte. Dieses Verfahren stellte sicher, dass eine Zellenübererfüllung kaum vorkam und die regionalen Strukturen durch Gewichtung gut korrigierbar waren.

Die beiden folgenden Grafiken veranschaulichen das Verfahren:

Abbildung 1: Ablauf einer Dual-Frame-Studie

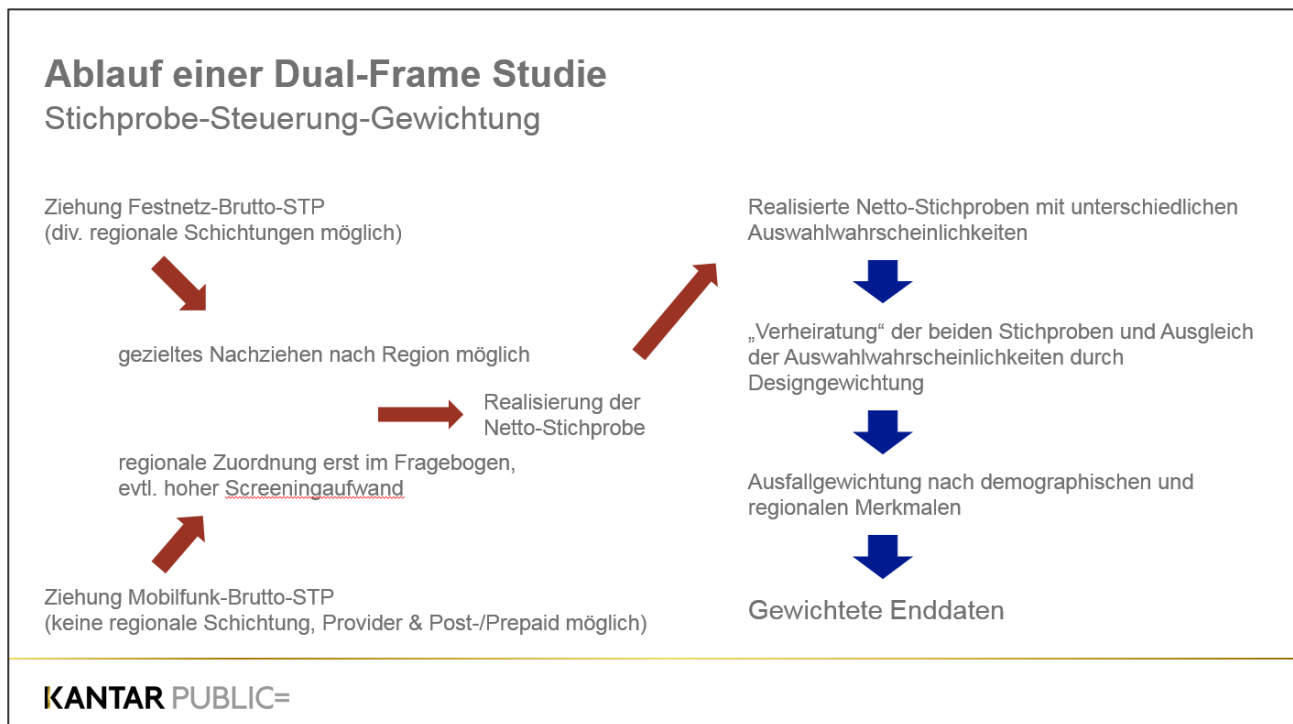
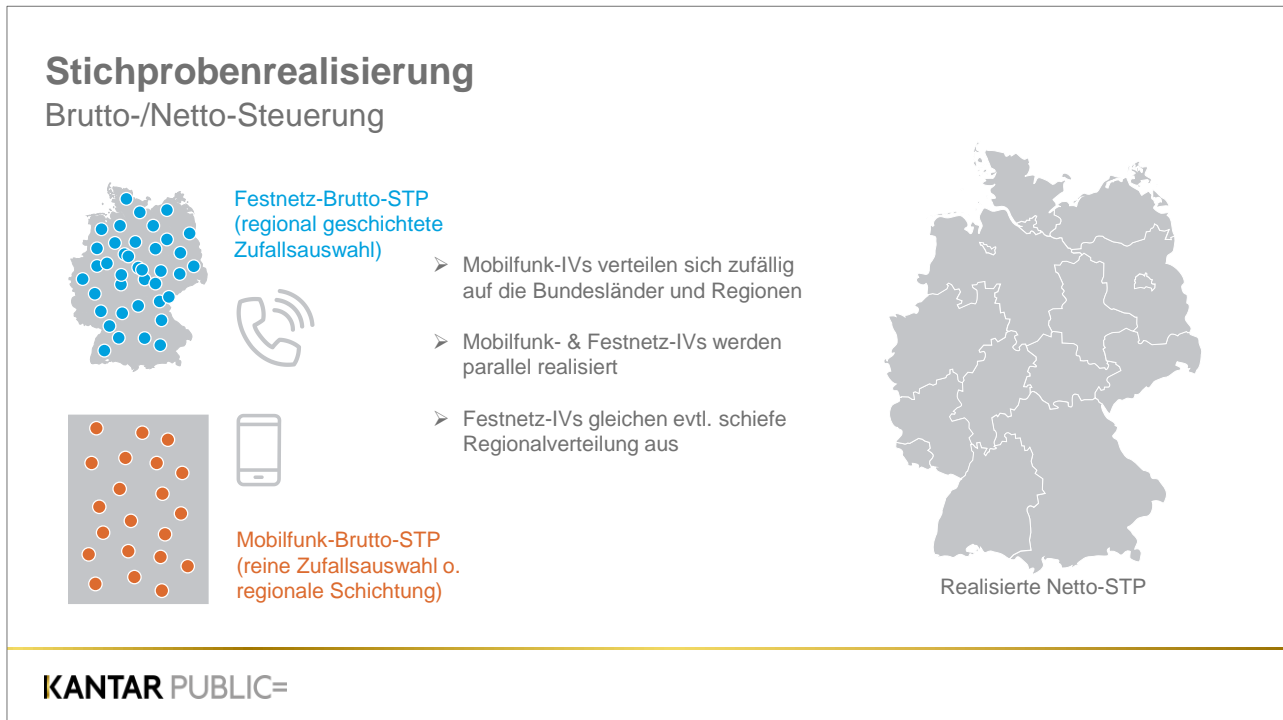


Abbildung 2: Stichprobenrealisierung



Generell gibt es im Mobilfunksample neben den bereits beschriebenen Unterschieden zur Festnetzstichprobe, die sich auf die regionale Auswahl und die Personen- statt die Haushaltsebene beziehen, weitere technische Unterschiede. Dies betrifft den bereits erwähnten höheren Anteil an ungültigen bzw. nicht geschalteten Nummern. Dieser begründet sich darin, dass Mobilfunknummern nur zu einem sehr geringen Anteil in Telefonbüchern registriert sind und deshalb der Anteil an generierten Nummern deutlich höher als bei der Festnetzstichprobe ist. Zudem kann der Autodialer bei Mobilfunknummern seltener erkennen, ob es sich um eine geschaltete Nummer handelt oder nicht. Dies liegt wiederum daran, dass bei nicht geschalteten Mobilfunknummern – je nach Netzbetreiber – auch Bandansagen oder Freizeichen ertönen, die nicht als falsche Nummer erkannt werden können. Der dadurch erhöhte Anteil an Fehlkontakten macht nochmals die benötigte höhere Übersetzung an Telefonnummern bei Mobilfunk deutlich.

Durch das Home-Location-Register-(HLR)-Lookup-Verfahren kann dieser Aufwand mittlerweile reduziert werden. Die folgende Grafik zeigt dies auf:

Abbildung 3: HLR-Lookup

HLR-Lookup

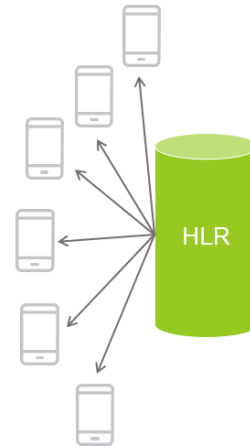
Technische Prüfung von Mobilfunknummern

Problemstellung:

- Das Mobilfunksample der ADM enthält einen hohen Anteil generierter Nummern
 - Die Dialererkennung der ISDN-Codes ist weniger eindeutig als bei Festnetztelefonie
- relativ hoher Anteil an Fehlkontakten mit unsicherem Ergebniscode

Home-Location-Register (HLR)-Lookup:

- Elementarer Bestandteil des Mobilfunknetzes sind Datenbanken, die gewährleisten, dass ein Endgerät im Netz „gefunden“ wird
- In diesen Datenbanken kann abgefragt werden, ob eine Nummer vergeben (valid) oder nicht (invalid) ist



KANTAR PUBLIC=

Seit einiger Zeit ist es Standard bei Kantar, dass lediglich Mobilfunknummern, die diesen Test durchlaufen haben, kontaktiert werden. Im Rahmen des HLR-Lookups werden etwa 40 %-50 % der generierten Mobilfunknummern vorab aussortiert. Dennoch verbleibt ein nennenswerter Anteil ungültiger Nummern in der eingesetzten Bruttostichprobe. Dabei kann ein Teil eindeutig als falsche Telefonnummer identifiziert werden. Bei einem wesentlich größeren Teil der Telefonnummern kann kein Kontakt hergestellt werden, wobei in diesen Fällen eine Unsicherheit besteht, ob – wie eben beschrieben – eine „gültige“ oder „ungültige“ Nummer vorliegt. (vgl. Abschnitt 4.4).

Während der Feldarbeit nicht erreichte Rufnummern wurden im Sample-Management-System zurückgelegt und kamen nach größerem zeitlichem Abstand zu anderen Tageszeiten zur Wiedervorlage. Die an einem bestimmten Tag nicht Erreichten wurden durch solche substituiert, die an anderen Tagen nicht erreicht wurden. Damit entfiel der sogenannte Not-at-home-Bias weitgehend. Festnetznummern, bei denen in sechs Anrufen kein Kontakt zu einer Person im Haushalt hergestellt werden konnte, also immer entweder das Freizeichen ertönte oder ein Anrufbeantworter geschaltet war, wurden angesteuert. Bei Mobilfunknummern war dies nach dem dritten Anruf ohne Kontakt zu einer Person der Fall, da es sich dabei im Regelfall um (noch) nicht genutzte Rufnummern handelte. Telefonnummern, bei denen in einem der ersten sechs bzw. drei Versuchen ein Kontakt zu einer Person hergestellt werden konnte, wurden so lange kontaktiert, bis ein Interview oder ein finaler Ausfallgrund vorlag.

Vom Sample-Management-System wurde je nach Bedarf auch die Gleichverteilung der Interviews auf Befragungstage und Tageszeiten gesteuert. In der vorliegenden Untersuchung war hier insbesondere an eine gleichmäßige Wochenverteilung zu denken, um saisonale Effekte zu vermeiden. Die üblichen Telefonzeiten waren Spätnachmittags bis Abends und am Wochenende am Nachmittag, um Privatpersonen, insbesondere Erwerbstätige, auch erreichen zu können. Um mögliche Einflüsse auf Untersuchungsergebnisse von vornherein auszuschalten, wurde die Stichprobe nach einem Verfahren der dynamischen Repräsentativität bezüglich der Besetzung der Zellen des Multistatifikationsstableaus optimiert, so dass sich für jedes Stundenintervall proportionale Teilstichproben ergaben.

Es kann somit festgehalten werden: Telefon-Master-Sample-Stichproben von Kantar sind über das Bindeglied amtlicher Flächeneinheiten auf die Verteilung der Privathaushalte justiert. Bei der Stichprobenanlage wurde dabei sowohl auf der Ebene der Bruttoziehung (Schichtung à priori) als auch

bei der Stichprobenrealisierung eine optimale Aussteuerung garantiert, so dass eine Gewichtung nach Haushaltszahlen pro Region zu einer äußerst kleinen Faktorenspanne führt und praktisch unterbleiben könnte. Das Telefon-Master-Sample-Stichprobendesign erforderte jedoch zwingend eine Designgewichtung (Schichtung à posteriori), da unterschiedliche (aber berechenbare) Auswahlchancen ausgeglichen werden müssen (vgl. hierzu Abschnitt 7.1).

3.2 Auswahl von Erwerbstätigen ab 15 Jahren

Festnetzstichprobe

Bei der Teilstichprobe mit Festnetznummern wurde auf der ersten Stufe zunächst ein Haushalt ausgewählt. In diesem wurde dann über einen Zufallsmechanismus eine zu befragende Zielperson ausgewählt. Hierfür musste zunächst geklärt werden, ob und wie viele Personen ab 15 Jahren im Haushalt leben, die erwerbstätig im Sinne der Untersuchung sind. Da bei diesem Vorgehen auch Proxy-Informationen, also Informationen über andere Personen durch die Auskunftsperson, erhoben wurden, musste die Abfrage nach der Zahl der Erwerbstätigen im Haushalt so gestaltet werden, dass keine erwerbstätige Person von der Auskunftsperson vergessen wurde. Bei der Ermittlung der Zielperson wurde schrittweise vorgegangen.

Zunächst wurde die Auskunftsperson direkt gefragt, ob sie erwerbstätig ist. Im Anschluss wurde ermittelt, wie viele erwerbstätige Personen ab 15 Jahren insgesamt im Haushalt leben. Durch eine separate Nachfrage wurden Personen in Ausbildung ausgeschlossen. Durch eine weitere Nachfrage wurden Personen mit einer maximalen Unterbrechung der Erwerbstätigkeit berücksichtigt. Schließlich wurde ermittelt, ob die erwerbstätigen Personen mindestens 10 Stunden pro Woche arbeiten. Aus der so ermittelten Zielpersonenanzahl wurde eine Person zufällig mit Hilfe des Schwedenschlüssels (bzw. Kish-Methode) ausgewählt. Um den Einstieg ins Interview zu optimieren, wurde zu Beginn der Haupterhebung der Einstieg modifiziert, die finale Version kam am 14.11.2017 zum Einsatz. Das genaue Vorgehen ist im Fragebogen dokumentiert.

Im Anschluss an die Ermittlung der Zielperson schloss sich ein kurzes Screeninginterview an, über das sichergestellt wurde, dass die Zielperson erwerbstätig im Sinne der Untersuchung ist und die Tätigkeit z. B. nicht ehrenamtlich oder im Rahmen einer Ausbildung ausübt.

Mobilfunkstichprobe

Die Ermittlung der Zielperson war bei der Teilstichprobe mit Mobilfunknummern insofern einfacher, da es sich bereits um eine Personenstichprobe handelt. Bei dem Erstkontakt musste mit der Person am Telefon deshalb zunächst nur geklärt werden, ob sie Hauptnutzer des Handys und mindestens 15 Jahre alt ist. War dies der Fall, wurde anhand des Screeninginterviews ermittelt, ob die Person erwerbstätig im Sinne der Untersuchung ist. Um die regionale Verortung des Interviews vornehmen zu können, war es zudem zwingend, dass die Person uns die Postleitzahl ihres Wohnortes nannte. War sie dazu nicht bereit, wurde das Gespräch an dieser Stelle beendet.

Um später die Auswahlwahrscheinlichkeiten berechnen zu können, wurde in allen Interviews die Zahl der Personen ab 15 Jahren insgesamt sowie die Anzahl der Erwerbstätigen ab 15 Jahren im Haushalt erfasst. Ebenso wurde die Zahl der Telefonnummern (Handy und Festnetz) erhoben, unter denen die Zielpersonen zu erreichen sind.

3.3 Rekrutierung für die Nachbefragung (Los 2)

Es war vorgesehen im Rahmen der BIBB/BAuA-Erwerbstätigenbefragung, Zielpersonen für eine Nachbefragung des BIBB (Los 2) zu ermitteln, bei der 8.000 Erwerbstätige und 1.000 Nicht-Erwerbstätige befragt werden sollten.

Dabei wurden alle Erwerbstätigen, die ein vollständiges Interview gegeben hatten, nach ihrem Einverständnis für eine Wiederbefragung gefragt. Dieses Einverständnis gaben 17.408 Personen, eine ausreichende Basis für die Nachbefragung von 8.000 Erwerbstätigen. Das entspricht einer Wiederbefragungsbereitschaft von 87 %, was als sehr guter Wert bewertet werden kann und für eine Interviewführung spricht, bei der eine gute Vertrauensbasis geschaffen wurde.

Um die 1.000 Interviews mit Nicht-Erwerbstätigen durchführen zu können, sollte von 3.000 Nicht-Erwerbstätigen das Einverständnis für eine Nachbefragung eingeholt werden. Für die Rekrutierung der Nicht-Erwerbstätigen war zunächst nur vorgesehen, die Mobilfunkstichprobe zu nutzen. Im Anschluss an das Screeninginterview wurden die nicht erwerbstätigen Personen gefragt, ob sie bereit wären, an einer weiteren Befragung im Auftrag des BIBB teilzunehmen. Da sich Anfang März abzeichnete, dass die erwünschte Zahl an befragungsbreiten Nicht-Erwerbstätigen rein über die Mobilfunkstichprobe möglicherweise nicht erreicht wird, wurde ab dem 6. März 2018 auch über die Festnetzstichprobe Nicht-Erwerbstätige rekrutiert. Hierbei wurden lediglich Haushalte einbezogen, in denen kein Erwerbstätiger lebte. Insgesamt wurden dem BIBB 3.182 Telefonnummern von befragungsbereiten Nicht-Erwerbstätigen zur Verfügung gestellt, darunter wurden 1.227 über die Festnetzstichprobe rekrutiert.

4 Durchführung der Erhebung

4.1 Interviewerschulung

Vor Start der Haupterhebung wurde eine projektspezifische Interviewerschulung durch die Projektleitung durchgeführt. Die Schulung fand als WebEx-Schulung am 2. Oktober 2017 statt und wurde aufgezeichnet, so dass sie für alle weitere Nachschulungen zur Verfügung stand. Jeder Interviewer und jede Interviewerin, der bzw. die bei der Studie zum Einsatz kam, wurde vorab mit dem Schulungsvideo auf das Interview vorbereitet.

Bei der Schulung wurde auf folgende Punkte eingegangen:

- Vorstellung der Kundenseite und des Projekts: Inhalte und Ziele, Methodische Eckdaten, Besonderheiten der Befragung
- Einstieg, Ermittlung der Zielperson und Möglichkeit zum Wechsel von Mobilfunk auf Festnetz
- Informationsmöglichkeiten für die Befragten
- Überblick über den Fragebogen
- Erläuterungen zu ausgewählten Fragen und Erläuterung des Inhalts schwieriger Fragen
- Erklärung von schwierigen Begriffen
- Wiederholungsbefragungen und Wiederbefragungsbereitschaft

Die Interviewer erhielten zudem schriftliches Schulungsmaterial, das ihnen während des Telefonierens zur Verfügung stand und folgende Materialien enthielt:

- die Schulungspräsentation,
- eine Übersicht mit den ausgewählten Fragen und Erläuterungen,
- eine Übersicht zu den beruflichen Ausbildungsabschlüssen mit Erläuterungen.

Im Anschluss an die Schulung gingen alle Interviewer ein Testinterview durch bis sie mit allen Fragen vertraut waren, bevor sie das erste Interview durchführten. Die Projektverantwortlichen an den verschiedenen Standorten und die Studienleitung standen für Rückfragen zur Verfügung. Dies gewährleistete eine sichere Interviewführung von Beginn an. Alle eingewiesenen Interviewer erhielten in der Projektdatenbank einen Vermerk, dass sie dieses Projekt bearbeiten können.

Wenn in einem Telefonstudio Fragen auftraten, wurden die Antworten auch an die anderen Standorte kommuniziert. Im Verlauf der Feldarbeit wurde an einzelnen Stellen im Erhebungsinstrument Nachschulungsbedarf festgestellt. Dies betraf folgende Aspekte:

- die Erfassung der Postleitzahl des Betriebes,
- die Ermittlung des Ausbildungsberufs sowie
- die korrekte Eingabe, wenn eine Person zwar wiederbefragungsbereit war, jedoch den Namen nicht nennen wollte.

Ebenfalls wurde nach der Modifikation der Einstiegsfragen (vgl. Abschnitt 3.2) eine Schulungsunterlage erstellt, die den Interviewern die vorgenommenen Änderungen und deren Hintergrund erläuterte. Dieses Vorgehen stellte sicher, dass alle eingesetzten Interviewer den gleichen Informationsstand hatten.

Das schriftliche Schulungsmaterial wurde im Vorfeld mit dem Auftraggeber abgestimmt und ist im Anhang dokumentiert.

4.2 Interviewereinsatz und Durchführung der Feldarbeit

Nach Abschluss der beiden Pretests und der Einarbeitung des dort festgestellten Anpassungsbedarfs startete die Feldarbeit der Haupterhebung wie geplant am 2. Oktober 2017. Die letzten Interviews wurden am 5. April 2018 durchgeführt. Insgesamt wurden 20.012 Interviews durchgeführt.

Kantar beschäftigt ausnahmslos Interviewer im Angestelltenverhältnis oder in geringfügiger Beschäftigung. Zum Zeitpunkt der Befragung betrug die durchschnittliche Verweildauer im Interviewerstab rund 42 Monate. Über die gesamte Feldzeit kamen 847 Interviewer zum Einsatz. Die Zahl der eingesetzten Interviewer liegt geringfügig niedriger als bei der letzten Erhebung. Im Durchschnitt befragte jeder Interviewer 23,6 Erwerbstätige.

Tabelle 1: Zahl der Interviews pro Interviewer

Zahl der Interviews	Zahl der Interviewer	Anteil in %
1 bis 10 Interviews	355	42%
11 bis 20 Interviews	186	22%
21 bis 30 Interviews	113	13%
31 bis 50 Interviews	83	10%
51 bis 90 Interviews	77	9%
91 oder mehr Interviews	33	4%
Insgesamt	881	100%

Zur Sicherstellung einer möglichst hohen Erreichbarkeit wurde vorwiegend nachmittags und abends bzw. an den Wochenenden kontaktiert. Die Kontaktzeiten wurden dabei über das Sampling Management System nach Uhrzeit und Wochentag variiert (siehe Abschnitt 4.2).

Um die Interviewer bei der Gewinnung der Zielpersonen zu unterstützen, wurden verschiedene vertrauensbildende Maßnahmen vorgesehen.

■ Informationsschreiben

Zunächst stand den Interviewern ein Informationsschreiben der Auftraggeber zur Verfügung, das sie auf Nachfrage per Email als PDF-Dokument an potentielle Zielpersonen versenden konnten. Zusammen mit diesem Informationsschreiben wurde ein Datenschutzblatt versendet. Das Informationsschreiben und das Datenschutzblatt sind im Anhang dokumentiert.

■ Hinweis auf Homepages

Des Weiteren existiert bereits eine Projekthomepage zur BIBB/BAuA-Erwerbstätigenbefragung, deren Adresse im Informationsschreiben genannt wurde. Auch während der Einstiegsphase des Interviews konnten die Interviewer die potentiellen Befragten auf diese Projekthomepage hinweisen. Sowohl auf der Projekthomepage als auch auf der Homepage von Kantar Public wurde während der Feldphase ein Hinweis auf die laufenden Feldarbeiten aufgenommen, so dass sich die Befragten von der Korrektheit des Anrufs überzeugen konnten.

■ Auftraggeber und Thema der Befragung

Eine weitere wichtige Unterstützung ist es für die Interviewer, wenn sie zu Beginn des Gesprächs den Auftraggeber nennen können, zumal wenn es sich – wie es bei der vorliegenden Studie der Fall ist – um öffentliche und seriöse Institute handelt. Auch das Thema der Befragung – die eigene berufliche Tätigkeit – ist für die Befragten überwiegend ein wichtiges und interessantes Thema.

■ Hinweise zur Studie bei Nachfragen

Der Einleitungstext wurde so kurz wie möglich gehalten, um mit den Zielpersonen möglichst rasch ins Gespräch zu kommen. Für Nachfragen zu verschiedenen Aspekten, wie Zielsetzung der Studie, Länge des Interviews, Datenschutz u. ä. standen den Interviewern Texte zur Verfügung, die sie je nach Bedarf und Nachfrage des Befragten aufrufen konnten (siehe Anhang A – Screeningfragebogen).

Angesichts der Länge des Interviews hatten die Interviewer verschiedene Möglichkeiten, um auf die Befragten einzugehen. So konnten Termine zu jeder Uhrzeit und an allen Wochentagen – außer Sonntags – vereinbart werden. Begonnene Interviews konnten jederzeit unterbrochen werden und zu einem anderen Zeitpunkt fortgeführt werden. Befragte, die über Mobilfunk kontaktiert wurden, hatten die Möglichkeit, - sofern vorhanden – eine Festnetznummer anzugeben und unter dieser das Interview zu führen.

Um die termingerechte Realisierung der Fallzahl von 8.000 Interviews zum Jahreswechsel 2017/2018 sowie die Fallzahl von 20.000 zu Anfang April 2018 sicherzustellen, fand ein monatlicher Jour-Fix mit der CATI-Produktionssteuerung statt, in der die aktuelle Fallzahl und die Interviewerkapazitäten besprochen wurden. Phasenweise wurde der Turnus auf 2 Wochen verkürzt. Die Auftraggeber erhielten regelmäßige Feldstandsberichte, die die Anzahl der Interviews, wesentliche soziodemographische Merkmale der Befragten sowie die Anzahl und Art der Screenout-Fälle enthielten. Insgesamt wurden 33 Feldstandsberichte erstellt.

4.3 Maßnahmen zur Qualitätssicherung und Interviewerkontrolle

Ein wesentlicher Vorteil von CATI-Umfragen ist die Möglichkeit der zeitnahen Qualitätskontrolle während der Feldzeit. Dazu wurden Prüfprogramme im Rahmen der Datenprüfung sowie folgende Monitoring-Techniken bei der Interviewdurchführung eingesetzt:

- (1) Frühzeitige Datenprüfung anhand von Zwischendaten
- (2) Regelmäßig Kurzberichte aus den Telefonstudios zu Schwierigkeiten oder Fragen
- (3) Monitoring von Interviews durch das Qualitätsmanagement

Zu (1): Frühzeitig und regelmäßig wurden Zwischendaten gezogen, die die Projektleitung auf korrekte Filterführung, korrekte Erfassung der Postleitzahl des Wohnorts bei Mobilfunk wie des Betriebs, Plausibilität sowie Verwertbarkeit der offenen Tätigkeitsbezeichnung prüfte. Die Prüfung von Zwischendaten erfolgte insgesamt 12 Mal bei folgenden Interviewanzahlen: 43 Interviews, bei rund 1.000, 1.500, 2.000, 4.000, 6.000, 7.000, 8.000, 10.000, 13.000, 15.000 und zuletzt bei rund 18.000 Interviews. Die Auftraggeber erhielten 3 Mal Zwischendaten (Näheres dazu in Abschnitt 5).

Zu (2): Die Projektleitung stand in engem Kontakt zu den Ansprechpartnern in den Telefonstudios und erhielt regelmäßig Kurzberichte, ob Schwierigkeiten oder Fragen aufgetreten sind.

Im Rahmen dieser ersten beiden Maßnahmen zeigte sich bei drei Aspekten Nachschulungsbedarf für die Interviewer, wofür Handouts erstellt wurden, die an alle eingesetzten Interviewer verteilt wurden.

Zu (3): Grundsätzlich werden für die Gewährleistung einer hohen Qualität bei der Interviewdurchführung die eingesetzten Interviewer durch Monitoringmaßnahmen beobachtet, die auch bei der aktuellen BIBB/BAuA-Erwerbstätigenbefragung zum Einsatz kamen. Das Monitoring (Mithören und -sehen) dient zur Überprüfung der Richtlinien zur korrekten und optimalen Datenerfassung und wird auf dem Qualitätsformular dokumentiert.

Durch ein PC-gestütztes Programm zum Mithören und Mitsehen kann der Supervisor erkennen, in welchem Programmabschnitt sich der Interviewer befindet (Einstieg, Maske xx, etc.) und zu welchen Zeiten und wie lange der Interviewer nicht aktiv ist.

Mit diesem Instrument werden in der Kontaktphase folgende Aspekte geprüft:

- Wird die richtige Zielperson ausgewählt und befragt?
- Wie ist das Argumentations- und Motivationsverhalten?
- Werden Ausfälle richtig eingestuft?
- Werden Terminvereinbarungen korrekt vorgenommen?

Im Rahmen einer Kurz-Qualitätskontrolle werden durch zufälliges Umschalten über einen kurzen Zeitraum Fragen oder Fragebogenteile verfolgt. Darüber hinaus wird spezielles Augenmerk auf von der Projektleitung als wichtig definierte Fragen gelegt. Bei den zudem stichprobenartig durchgeführten Voll-Qualitätskontrollen wird über die Kontaktphase hinaus das vollständige Interview anhand des folgenden Leitfadens beobachtet:

- Verhalten während des Vorlesens der Fragen
- Beherrschung der Fragetechnik (wörtlich, korrekte Eingabe)
- Motivationsfähigkeit, Tempo der Gesprächsführung
- Verhalten bei Nachfragen
- Wahrung der Neutralität
- Verabschiedung höflich, Hinweis auf Wiederholungsanrufe, Datenschutz

Aus den sich hierbei ergebenden Hinweisen erstellt der Supervisor eine Gesamtbeurteilung des Interviewers. Diese drückt sich in seiner Benotung für „Qualität“ aus:

1 = Sehr gutes Interview

2 = Gutes Interview

3 = Leichte Mängel

4 = Interview mit großen Mängeln

Eine Bewertung mit dem Wert 4 erfolgt beispielsweise, wenn eine falsche Zielpersonenauswahl oder eine suggestive Fragestellung erfolgte. Die Note 3 wird z.B. vergeben, wenn Frage- oder Antworttexte verkürzt vorgelesen werden oder Antwortskalen nicht ausreichend häufig vorgelesen wurden. Die Bewertung wird anschließend in das Interviewerinformationssystem eingegeben. Es werden jeweils die aktuellsten zehn Bewertungen gespeichert, um ein ausreichendes Abbild der gelieferten Qualität des Interviewers zu haben. Bei neuen Interviewern werden die ersten drei Einsätze voll kontrolliert. Bei den übrigen Interviewern werden in regelmäßigen Abständen Voll-Qualitätskontrollen und Kurz-Qualitätskontrollen in verschiedenen Studien vorgenommen.

Bei der vorliegenden Untersuchung wurden insgesamt 1.055 Kontrollen durchgeführt, darunter ein Drittel während der Kontaktphase, etwa jede fünfte Kontrolle war eine sog. Kurzkontrolle und in gut der Hälfte der Kontrollen wurde eine sog. Vollkontrolle durchgeführt. Die genaue Anzahl der durchgeführten Kontrollen und das Ergebnis der Gesamtbeurteilung wird in Tabelle 2 dargestellt.

Tabelle 2: Durchgeführte Interviewerkontrollen und ihr Ergebnis

Note	Anzahl Kontrollen	Art der Kontrolle in %			Kontrollen in %
		Kontaktphase	Kurzkontrolle	Vollkontrolle	
1	473	10%	11%	24%	45%
2	472	16%	6%	22%	45%
3	88	4%	2%	2%	8%
4	22	1%	0,5%	0,2%	2%
Insgesamt	1.055	32%	21%	51%	100%

Der Anteil an Kontrollen, bei denen nicht korrektes Interviewerverhalten konstatiert werden musste, ist gering. Die meisten der festgestellten Mängel wurden dabei während der Kontaktphase beobachtet und betreffen damit nicht unbedingt Interviews, sondern auch Ausfälle. Je nach Bewertung erhielten die Interviewer im Anschluss an die Kontrolle ein Lob, Hinweise auf die Punkte, die sie verbessern sollen, eine gezielte Nachschulung oder es erfolgte eine Verwarnung und der Abzug vom Projekt. Interviews mit der Note 4 wurden ausnahmslos als nicht verwertbar aus dem Datensatz entfernt, bei der Note 3 erfolgte eine Nachschulung durch die Supervision.

4.4 Ergebnisse der Feldarbeit

In der Zeit vom 2. Oktober 2017 bis zum 5. April 2018 wurden 20.012 auswertbare Interviews mit Erwerbstätigen durchgeführt, davon 14.008 Interviews, die über Festnetz kontaktiert wurden, und 6.004 Fälle, die über eine Mobilfunknummer kontaktiert wurden. Zum 31.12.2017 lagen 8.136 Interviews vor.

Die Interviewdauer war mit durchschnittlich 40 Minuten geplant, wobei hier die Nettodauer, also nach Identifikation der Zielperson, gemeint ist. Die realisierte durchschnittliche Nettointerviewdauer lag mit 42,2 Minuten jedoch um gut 2 Minuten höher als geplant. Betrachtet man den Median liegt die Interviewdauer bei 40,7 Minuten. Hinzukommt die Zeit, die für die Kontaktphase sowie das Screeninginterview benötigt wird. Die durchschnittliche Gesamtdauer liegt damit bei 46 Minuten, der Median bei 44,4 Minuten. Vergleicht man die Interviewdauern der Festnetz- und der Mobilfunkstichprobe, so dauerten die Interviews in der Mobilfunkstichprobe eineinhalb Minuten länger als in der Festnetzstichprobe. Die folgende Tabelle gibt einen Überblick über die jeweiligen Interviewdauern.

Tabelle 3: Interviewdauer in Minuten

	Gesamtdauer	Kontaktphase	Screening	Nettodauer
Durchschnitt Festnetz	45,5	1,1	2,7	41,7
Median Festnetz	43,8	1,0	1,8	40,2
Durchschnitt Mobil	47,2	1,4	2,6	43,2
Median Mobilfunk	45,5	1,2	1,8	41,8
Durchschnitt Gesamt	46,0	1,2	2,7	42,2
Median Gesamt	44,4	1,0	1,8	40,7

Tabelle 4 gibt einen Überblick über die verschiedenen Ausfallgründe sowie die Ausschöpfungs- und Kooperationsrate nach AAPOR.³ Auf Ebene der Verweigerungen wird dabei unterschieden zwischen der Kontaktperson (KP) und der Zielperson (ZP). Bei einer Verweigerung durch eine Kontaktperson ist i.d.R. die Zielperson noch nicht ermittelt.

3 AAPOR - American Association for Public Opinion Research, Standard Definitions Report (9th edition, 2016)

Tabelle 4: Ausschöpfungs- und Kooperationsrate nach AAPOR

A Bruttoansatz	Festnetz	Mobilfunk	Insgesamt
	2.083.931	577.781	2.661.712
Fax/Modem	4,9%	1,6%	4,2%
ZP: nicht in der Lage/krank	0,0%	0,0%	0,0%
ZP: Sprachprobleme	0,0%	0,1%	0,0%
Kein Privathaushalt	1,1%	0,7%	1,0%
Falsche Telefonnummer	42,4%	11,4%	35,7%
Doppeladresse	0,0%	0,1%	0,1%
System error	0,1%	0,1%	0,1%
Abandoned call (Leitungsproblem ohne Kontakt)	0,1%	3,5%	0,8%
Stratifikation erreicht (Regionalzelle voll)	6,8%	0,3%	5,4%
max. Kontaktanzahl erreicht	24,3%	51,6%	30,2%
Firmen Anrufbeantworter	1,1%	1,6%	1,2%
KP: Sprachprobleme	0,4%	1,9%	0,7%
Summe neutraler Ausfälle (not eligible)	81,2%	72,9%	79,4%
B Bereinigtes Brutto	391.670	156.723	548.393
Teilnehmer nimmt nicht ab	0,7%	0,0%	0,5%
Privater Anrufbeantworter	0,2%	0,1%	0,2%
Anschluss besetzt	0,0%	0,0%	0,0%
Angerufene nicht mehr in Leitung	5,9%	7,6%	6,4%
Summe nicht neutraler Ausfälle (unknown eligibility)	6,9%	7,7%	7,1%
C Verbleibende Adressen (eligible)	364.752	144.594	509.346
Kein Termin während der Feldzeit möglich	2,0%	3,9%	2,5%
KP: ungefährrer Termin konnte nicht realisiert werden	0,7%	0,8%	0,7%
KP: fester Termin konnte nicht realisiert werden	0,0%	0,1%	0,1%
KP: kein Interesse am Thema	29,6%	30,8%	29,9%
KP: keine Auskunft/sonstiger Grund	33,2%	32,7%	33,1%
KP: Absolutverweigerer	0,3%	0,5%	0,4%
KP: wortlos aufgelegt	13,5%	12,3%	13,1%
KP: Verweigerung im Namen der Zielperson	1,7%	1,5%	1,6%
KP: Verweigerung aufgrund von Bedenken wegen Datenschutz	0,2%	0,2%	0,2%
KP: Verweigerung wegen Länge des Interviews	1,3%	1,6%	1,4%
ZP konnte unter Telefonnummer nicht ermittelt werden	0,5%	0,7%	0,6%
ZP: ungefährrer Termin konnte nicht realisiert werden	0,2%	0,1%	0,2%
ZP: fester Termin konnte nicht realisiert werden	0,0%	0,0%	0,0%
ZP: Verweigerung aus Zeitgründen	0,3%	0,3%	0,3%
ZP: Verweigerung per Mail, Brief, über Qualitätsmanagement	0,0%	0,0%	0,0%
ZP: kein Interesse am Thema	0,6%	0,6%	0,6%
ZP: keine Auskunft/sonstiger Grund	0,6%	1,1%	0,7%
ZP: Absolutverweigerer	0,0%	0,0%	0,0%
ZP verweigert aufgrund von Bedenken wegen Datenschutz	0,0%	0,0%	0,0%
ZP verweigert wegen Länge des Interviews	0,5%	0,4%	0,4%
ZP: wortlos aufgelegt	0,2%	0,3%	0,2%
Verweigerung wg. Mithören, d.h. Abbruch nach Mithörfrage	0,0%	0,0%	0,0%
Keine Angabe bei PLZ (Mobilfunk)	0,0%	0,5%	0,1%
keine Angabe zur Berufsbezeichnung	0,0%	0,0%	0,0%
Abbruch im Interview	1,1%	3,9%	1,9%
Summe nicht neutraler Ausfälle (eligible)	85,3%	88,2%	86,1%
Durchgeführte Interviews	49.661	10.738	60.399
Interview mit erwerbstätiger Person	28,2%	55,9%	33,1%
Person unter 15 Jahren	0,0%	9,9%	1,7%
Keine erwerbstätige Person im Haushalt	71,0%	-	58,4%
Person gehört lt. Screeninginterview nicht zur ZG (MF)	-	34,2%	6,1%
Person gehört lt. Screeninginterview nicht zur ZG (FN)	0,8%	-	0,7%
Alle Personen im HH unter 15 Jahren (FN)	0,0%	-	0,0%
Ausschöpfungsrate (in % von B, d.h. bereinigtes Brutto)	12,7%	6,9%	11,0%
Kooperationsrate (in % von C, d.h. verbleibende Adressen)	13,6%	7,4%	11,9%

Betrachtet man Ausschöpfungsübersicht, fällt der hohe Anteil an neutralen Ausfällen auf. Dieser ist begründet in der Tatsache, dass die ausgewählten Telefonnummern zufallsgeneriert sind und es damit entsprechend häufig keine real existierenden bzw. vergebenen Telefonnummern sind. Die

unterschiedlich hohen Anteile von „falschen Telefonnummern“ sowie „maximale Kontaktzahl erreicht“ bei der Festnetzstichprobe und der Mobilfunkstichprobe ergeben sich daraus, dass nicht vergebene Festnetznummern i.d.R. als ungültige Nummern mit einem entsprechenden Ton oder einer Ansage über den Autodialer zu identifizieren sind. Bei der Mobilfunkstichprobe wird zwar ein Teil der ungültigen Nummern über das HLR-Lookup-Verfahren vorab identifiziert, der Großteil der nicht vergebenen Telefonnummern ist jedoch nicht zu identifizieren, da auch hier ein Freizeichen beim Anruf ertönt. Entsprechend ist davon auszugehen, dass im Mobilfunkteil der Stichprobe ein Großteil der Telefonnummern, bei denen trotz maximaler Kontaktzahl niemand erreicht werden konnte, nicht vergebene Telefonnummern sind (vgl. Abschnitt 3.1.1). Bei Festnetznummern, unter denen während der Feldzeit kein Kontakt hergestellt werden konnte, konnte nicht festgestellt werden, ob es sich um einen Privathaushalt handelt. Aus diesem Grund sind diese Ausfälle den neutralen Ausfällen zugeordnet.

Unter den „unknown eligible“ Adressen ist der größte Anteil an Ausfällen „Angerufene nicht mehr in der Leitung“. Hier wurde über den Autodialer ein Kontakt hergestellt, dieser konnte jedoch zu keinem Interviewer durchgestellt werden, da alle Interviewer zu diesem Zeitpunkt im Gespräch waren. In diesen Fällen wird eine Bandansage eingespielt. Personen, die auflegen, bevor ein Interviewer das Gespräch aufnimmt, werden hier eingestuft. Dies ist ein notwendiges und übliches Verfahren, um die Masse an Telefonnummern effizient abarbeiten zu können. Gemessen an den etwas mehr als zweieinhalb Millionen eingesetzten Telefonnummern ist dieser Anteil sehr gering. Nach den ADM Richtlinien sind diese Fälle als Ausfälle zu werten und sollen nicht erneut angerufen werden.

Unter den „eligible“ Adressen, also denjenigen, bei denen geklärt werden konnte, dass die Telefonnummern zu einem Privathaushalt oder einer Privatperson gehört, sind die höchsten Anteile an Ausfällen auf der Ebene der Kontaktperson zu verzeichnen. Bei den Festnetznummern konnte bei den Verweigerungen durch die Kontaktperson i.d.R. keine Information zur Anzahl der Erwerbstätigen im Haushalt und damit auch keine Zielperson ermittelt werden. Bei den Mobilfunknummern konnte bei diesen Ausfällen nicht geklärt werden, ob der Hauptnutzer des Handys am Apparat ist bzw. die Person 15 Jahre oder älter ist. Die Verweigerungen durch die Kontaktpersonen wurden also ausgesprochen, ohne dass die Personen am Telefon Informationen preisgaben. Der Anteil an Verweigerungen durch Zielpersonen ist dagegen relativ gering.

Um festzustellen, ob unter der kontaktierten Telefonnummer eine Zielperson zu erreichen ist, waren bei der Festnetzstichprobe zunächst 3-4 Fragen durch die Kontaktperson zu beantworten. Im Anschluss wurde mit der zufällig ausgewählten Zielperson das Screeninginterview geführt, um sicherzugehen, dass sie zur Zielgruppe gehört. Bei der Mobilfunkstichprobe wurde ein Screeninginterview geführt, sofern die Person Hauptnutzer des Handys und 15 Jahre oder älter war. Da in all diesen Konstellationen Informationen erfragt wurden, können diese Anrufe als Interviews gewertet werden.

Das Feldergebnis kann auf verschiedene Art und Weise berechnet werden, entscheidend ist, dass die Zuordnung der Ausfälle sowie der Weg der Berechnung transparent dargestellt sind. Eine andere Berechnungsweise wurde z. B. bei der BAuA-Arbeitszeitbefragung 2015 vorgenommen.⁴ Da beide Studien hinsichtlich der methodischen Anlage sowie der Zielgruppendefinition nahezu identisch sind und, um einen Vergleich zu ermöglichen, haben wir die bei uns verwendeten Response Codes auch nach dem bei der BAuA-Arbeitszeitbefragung eingesetzten Schema eingeordnet. Zwischen diesen beiden Berechnungen gibt es zwei wesentliche Unterschiede:

- (1) Die Zuordnung der Fälle, bei denen sich im Kontakt- oder Screeninginterview herausstellte, dass die Person/der Haushalt nicht zur Zielgruppe gehört.
- (2) Die Zuordnung der Fälle, in denen die Kontaktperson verweigerte.

Bei der Berechnung der Kooperationsrate und der Ausschöpfung in Tabelle 4 gehen die Verweigerungen der Kontaktpersonen ein, die in der Regel erfolgen, bevor eine Zielperson ausgewählt werden konnte.

4 Siehe „Methodenbericht und Fragebogen zur BAuA-Arbeitszeitbefragung 2015“, S.24ff. Dort wurden nach AAPOR eine Ausschöpfungsrage von 33,3 % und eine Kooperationsrate von 36,1 % ausgewiesen. Download unter: <https://www.baua.de/DE/Angebote/Publikationen/Berichte/F2360-2.html>

Bei dieser Berechnung ergeben sich eine Ausschöpfungsrate von 11 % und eine Kooperationsrate von 12 %.

Die Berechnung der Kooperationsrate und der Ausschöpfung in Tabelle 5 erfolgt dagegen ausschließlich auf Ebene der ausgewählten Zielpersonen. Verweigerungen auf Ebene der Kontaktpersonen, die einen sehr hohen Anteil der Ausfälle darstellen, werden in Tabelle 5 unter „Unbekannte Auswahl (UE) (ZG-Zugehörigkeit ungeklärt)“ subsummiert und fließend damit nicht in die Berechnung der Kooperationsrate und der Ausschöpfung ein. Bei der Berechnung in Tabelle 5 ergeben sich eine Ausschöpfung von 43 % und eine Kooperationsrate von 47 %. Die folgende Tabelle zeigt die Zuordnung und die Berechnung der Quoten im Einzelnen.

Tabelle 5: Ausschöpfungs- und Kooperationsrate entsprechend der BAuA-Arbeitszeiterhebung

Bruttoansatz	2.661.712		
Nicht Zielgruppe (NE)	80,1%		
Fax/Modem	4,2%		
Kein Privathaushalt	1,0%		
Falsche Telefonnummer	35,7%		
Doppeladresse	0,1%		
System error	0,1%		
Abandoned call (Leitungsproblem ohne Kontakt)	0,8%		
Stratifikation erreicht (Regionalzelle voll)	5,4%		
maximale Kontaktanzahl erreicht	30,2%		
Firmen Anrufbeantworter	1,2%		
Person unter 15 Jahren	0,0%		
Keine erwerbstätige Person im Haushalt	1,3%		
Person gehört lt. Screeninginterview nicht zur Zielgruppe (MF)	0,1%		
Person gehört lt. Screeninginterview nicht zur Zielgruppe (FN)	0,0%		
Alle Personen im HH unter 15 Jahren (FN)	0,0%		
Unbekannte Auswahl (UE) (ZG-Zugehörigkeit ungeklärt) absolut	528.657		
Unbekannte Auswahl (UE) (ZG-Zugehörigkeit ungeklärt) in %	18,1%		
Teilnehmer nimmt nicht ab	0,1%		
Privater AB (Wiedervorlage)	0,0%		
Anschluss besetzt	0,0%		
Angerufene nicht mehr in Leitung	1,3%		
KP: Sprachprobleme	0,7%		
Kein Termin während Feldzeit möglich	0,5%		
KP: ungefährender Termin konnte nicht realisiert werden	0,1%		
KP: fester Termin konnte nicht realisiert werden	0,0%		
KP: kein Interesse am Thema	5,7%		
KP: keine Auskunft/sonstiger Grund	6,3%		
KP: Absolutverweigerer	0,1%		
KP: wortlos aufgelegt	2,5%		
KP: Verweigerung im Namen der Zielperson	0,3%		
KP: Verweigerung aufgrund von Bedenken wegen Datenschutz	0,0%		
KP: Verweigerung wegen Länge des Interviews	0,3%		
Zugehörigkeit zur Zielgruppe bekannt absolut	46.978	46.978	
Zugehörigkeit zur Zielgruppe bekannt in %	1,8%	100%	
Nonresponse - nicht befragbar (NR-NA)	0,1%	8,1%	
ZP: nicht in der Lage/krank	0,0%	0,4%	
ZP: Sprachprobleme	0,0%	1,3%	
ZP konnte unter Telefonnummer nicht ermittelt werden	0,1%	6,4%	
Nonresponse - nicht erreicht (NR-NC) absolut	43.173	43.173	
Nonresponse - nicht erreicht (NR-NC) in %	0,0%	2,1%	
ZP: ungefährender Termin konnte nicht realisiert werden	0,0%	1,9%	
ZP: fester Termin konnte nicht realisiert werden	0,0%	0,2%	
Zielperson/Haushalt erreicht und befragbar absolut	42.188	42.188	42.188
Zielperson/Haushalt erreicht und befragbar in %	1,6%	89,8%	100%
Nonresponse - Verweigerung (NR-R)	0,8%	47,2%	52,6%
ZP: Verweigerung aus Zeitgründen	0,1%	3,0%	3,4%
ZP: Verweigerung per Mail, Brief, über Qualitätsmanagement	0,0%	0,0%	0,0%
ZP: kein Interesse am Thema	0,1%	6,5%	7,2%
ZP: keine Auskunft/sonstiger Grund	0,1%	8,1%	9,0%
ZP: Absolutverweigerer	0,0%	0,1%	0,1%
ZP verweigert aufgrund von Bedenken wegen Datenschutz	0,0%	0,1%	0,1%
ZP verweigert wegen Länge des Interviews	0,1%	4,8%	5,4%
ZP: wortlos aufgelegt	0,0%	2,3%	2,6%
Verweigerung wg. Mithören, d.h. Abbruch nach Mithörfrage	0,0%	0,2%	0,2%
Keine Angabe bei PLZ (Mobilfunk)	0,0%	1,4%	1,6%
keine Angabe zur Berufsbezeichnung	0,0%	0,1%	0,1%
Abbruch im Interview	0,4%	20,5%	22,9%
Durchgeführte Interviews	20.012	20.012	20.012
Interview mit erwerbstätiger Person	0,8%	42,6%	47,4%

4.5 Kurzbefragung von Nichtteilnahmebereiten

Allen nichtteilnahmebereiten Zielpersonen wurden gebeten, zumindest einige wenige Angaben zu machen. Ziel dabei war es, Ausfallanalysen erstellen zu können und nach Abschluss der Feldarbeit Hinweise auf soziodemographische Strukturen der Verweigerer zu erhalten. Die Verweigerer wurden gebeten, ihr Geburtsjahr und ihren Schulabschluss zu nennen. Das Geschlecht der Person wurde vom Interviewer eingestuft.

Voraussetzung für die Fragen an die Verweigerer war, dass es vorher bereits gelungen war, die Zielperson zu ermitteln. Für den Fall, dass die Kontaktperson verweigerte, was in den meisten Fällen so war, wurden keine Angaben erhoben.

Insgesamt wurde 11.363 Mal versucht, von den verweigernden Zielpersonen Informationen zu erhalten. Das Anliegen wurde wie folgt eingeleitet: „Auch wenn Sie nicht an der Studie teilnehmen möchten, würden Sie mir dennoch ihr Geburtsjahr und Ihren Schulabschluss nennen?“. Hierzu erklärten sich lediglich 581 Personen bereit (5,1 %).

Auf die Frage nach dem Geburtsjahr gab es 567 Angaben und 14-mal mal „Keine Angabe“. Die Frage nach dem Schulabschluss beantworteten 517 Personen, „keine Angabe“ war die Antwort von 64 Verweigerern. Eine Einschätzung durch die Interviewer zum Geschlecht liegt für 10.834 Personen vor, in 529 Fällen konnte das Geschlecht nicht eingeschätzt werden. Für das Geschlecht gab es mit 54 % nichtteilnahmebereiten Männern praktisch keine Abweichung gegenüber den Zahlen des Mikrozensus. Für die übrigen Merkmale sind die Fallzahlen gering, so dass auch hier nicht von signifikanten Verzerrungen gesprochen werden kann.

5 Datenprüfung und -aufbereitung

Der Prozess der Datenprüfung wird bei computerunterstützten Befragungen bereits in die Entwicklungsphase der Erhebung verlagert. Prinzipiell ist hier das Ziel, ein Erhebungsinstrument zu entwickeln, welches fehlerfreie, konsistente und plausible Daten liefert: Das Fragenprogramm läuft unter Berücksichtigung der richtigen Filterführung auf dem Bildschirm ab; die Antworten werden vom Interviewer direkt eingegeben. In das Programm implementierte Prüfroutinen und integrierte Verfahren der Datenprüfung und -edition stellen die Plausibilität und Konsistenz der Antworten sicher. Sollten sich während des Interviews Unstimmigkeiten ergeben, so können diese sofort mit dem Befragten geklärt und korrigiert werden. Vor Feldstart finden ausgiebige technische Tests des Erhebungsinstruments statt.

Vor dem Hintergrund, dass das Fragenprogramm unter Berücksichtigung der Filterführung auf dem Bildschirm abläuft, ist eine Datenbereinigung bei computerunterstützten Befragungen eigentlich nicht notwendig. Voraussetzung hierfür ist allerdings ein fehlerfreies Erhebungsinstrument. Dennoch kann es sich bei den Auswertungen zeigen, dass bei einem komplexen Erhebungsinstrument gewisse Inkonsistenzen bestehen – dies beispielsweise deshalb, weil man auf eine entsprechende Prüfung und Nachfrage verzichtet hat, um an einer bestimmten Stelle nicht den Gesprächsfluss zu stören, um keinen Interviewabbruch zu provozieren oder weil der Befragte die Inkonsistenz nicht korrigieren wollte oder konnte.

Da trotz intensiven Prüfungen des computergestützten Erhebungsinstruments vor Beginn der Erhebung nicht ausgeschlossen werden kann, dass die Filterführung völlig fehlerfrei abgelaufen ist, wurde eine intensive Prüfung der Filter vorgenommen. Kantar Public hat hierfür aufwendige Prüfprogramme in SPSS erstellt. Diese Programme wurden auf Syntaxbasis programmiert, so dass der komplette Prozess der Prüfung und Fehlerbereinigung nachvollziehbar ist und Änderungen dokumentiert sind. In einem ersten Schritt wurde die korrekte Filterführung geprüft. In einem weiteren Schritt wurden Plausibilitätsprüfungen zwischen einzelnen Fragen durchgeführt.

In allen Fällen, in denen Auffälligkeiten entdeckt wurden, wurde die Notwendigkeit einer Datenbereinigung sowie die Möglichkeiten einer sinnvollen Bereinigung geprüft, dem Auftraggeber Vorschläge unterbreitet und die Bereinigung anschließend in Abstimmung mit ihm vorgenommen. Eine enge Abstimmung bei der Erstellung von Entscheidungsregeln muss auch wegen der Vergleichbarkeit zwischen den verschiedenen Erhebungswellen gewährleistet sein: Die Tests waren deshalb möglichst analog zu denen aus den Vorwellen aufgebaut, so dass die Entscheidungen zwischen den Wellen konsistent ausfallen konnten. An folgenden Variablen wurden zu Beginn der Feldphase Probleme sichtbar, woraufhin Korrekturen während der Feldphase vorgenommen wurden:

- Fragen G1217 H1217 I1217 J1217: bei den Nachfragen innerhalb der Ausbildungsschleifen zwei bis fünf, warum die Ausbildung mit einer Dauer von weniger als einem Jahr so ungewöhnlich kurz war, gab es einen Filterfehler. Bis dieser behoben wurde, wurden die Fragen in allen Schleifen insgesamt 115 Mal fälschlicherweise nicht gestellt. Es wurde ein Sondercode mit 96 mit dem Wertelabel „Frage nicht gestellt“ vergeben.
- Frage F1605D: Bei der Frage nach dem Alter des ältesten Kindes gab es einen Fehler in der Fragebogenkonzeption, so dass für manche Konstellationen die Frage erstmal nicht gestellt wurde. Insgesamt waren 253 Fälle davon betroffen, auch hier wurde der Sondercode 96 mit dem Wertelabel „Frage nicht gestellt“ vergeben.

Zusätzlich wurden bei zwei Fragen Korrekturen der Werte vorgenommen:

- Frage 231 zur Arbeitszeit, die i.d.R. im Durchschnitt von zu Hause aus geleistet wird: hier wurde 12x 97 Stunden und 3x 99 Stunden angegeben. Da die angegebene Arbeitszeit jeweils darunter lag, konnte man davon ausgehen, dass hier jeweils eine Ziffer der Sondercodes ‚997‘ und ‚999‘ zu wenig eingegeben wurde. Entsprechend wurden die Daten in Abstimmung mit den Auftraggebern korrigiert. Im Datensatz ist die korrigierte Variabel „F231“ und die originale Variable ohne Änderungen „F231_orig“ enthalten.

- Frage F1408 zur Anzahl der Jahre, die die Berufstätigkeit unterbrochen wurde: hier gab es in 4 Fällen sehr hohe Angaben zur Unterbrechung – auch im Verhältnis zum Alter der Zielperson. Diese vier Fälle haben wir in Abstimmung mit den Auftraggebern auf „99“ für keine Angabe gesetzt. Auch hier ist im Datensatz die korrigierte Variabel „F1408“ und die originale Variable ohne Änderungen „F1408_orig“ enthalten.

Die Datenprüfung fand mehrfach anhand von verschiedenen (Zwischen-)Datensätzen statt:

- Im Anschluss an jeden Pretest wurde ein Datensatz generiert, der aufbereitet und auf Filterfehler und Plausibilität geprüft wurde
- Jeder Zwischendatensatz wurde vor Auslieferung an den Kunden geprüft (nach 1.000, 10.000 und 15.000 Interviews)
- Intensive Prüfung der endgültigen Befragungsdaten

Direkt nach Ende der Befragung stellt die verwendete Befragungssoftware NIPO die Befragungsdaten als ASCII-Datensatz bereit. Aus technischen Gründen entsprachen die Variablennamen der unaufbereiteten Daten nicht immer denen der Fragebogenvorlagen. Dies kommt beispielsweise bei Verschleifungen vor. Kantar Public hat die Variablen entsprechend den Bezeichnungen in den Fragebogenvorlagen umbenannt, so dass die endgültigen Daten den Variablennamen der Fragebogenvorlage entsprechen. Darüber hinaus wurden Variablen für einfachere inhaltliche Auswertungen gebildet (z. B. Umrechnen des Geburtsdatums in das aktuelle Alter als eigene, zusätzliche Variable). Dies geschah mit einer eigenen SPSS-Syntax, so dass der Auftraggeber diese auch im Nachhinein einfach anpassen kann. Alle Datensätze wurden mit Variablen- und Wertelabels ausgeliefert.

Sämtliche Schritte der Datenprüfung und Datenaufbereitung wurden auf SPSS-Syntaxbasis durchgeführt. Damit ist sichergestellt, dass jede Veränderung, von den Rohdaten bis hin zum endgültigen Datensatz dokumentiert und nachvollziehbar ist.

Folgende Datensätze wurden an das BIBB und die BAuA geliefert:

Los 1:

- 08.11.2017: Zwischendaten mit 1.000 Fällen
- 22.12.2017: Zwischendaten mit 7.000 Fällen
- 20.03.2018: Zwischendaten mit 15.000 Fällen
- 13.04.2018: Finale Daten mit 20.012 Fällen
- 20.04.2018: Gewichtungsfaktoren
- 30.04.2018: Bruttoband mit allen eingesetzten Telefonnummern und dem jeweils dazugehörigen finalen Kontaktergebnis (2.661.712 Fälle)
- 18.05.2018: Datensatz mit Verweigererdatensatz (11.363 Fälle)
- 30.06.2018: Datensatz mit der Berufe- und Branchenvercodung

Los 2:

- Zweiwöchentliche Lieferungen der Adressen für die Folgebefragung
- 09.04.2018: Letzte Lieferung der Los-2-Adressen

6 Vercodung der offenen Angaben zur beruflichen Tätigkeit und zum Wirtschaftszweig

Im Erhebungsinstrument gab es an verschiedenen Stellen offene Textangaben zum Wirtschaftszweig sowie zur beruflichen Tätigkeit der Zielperson und ihrer Eltern zu dem Zeitpunkt als sie 15 Jahre alt war. Ebenso wurden Angaben von bis zu fünf Ausbildungsberufen erhoben. Sämtliche offenen Textangaben wurden im Anschluss an die Datenerhebung in verschiedene Klassifikationssysteme vercodet. Im Folgenden wird die Anzahl der zu vercodenden Angaben dargestellt, nach welchen Systemen die Angaben vercodet wurden und wie hierbei vorgegangen wurde.

Zur Ordnung von Berufsinformationen und, darauf aufbauend, zur Beschreibung von Struktur und Verteilung beruflicher Tätigkeiten sind in Deutschland derzeit parallel zwei aktuelle Klassifikationssysteme vorhanden. Das erste System ist die nationale „Klassifikation der Berufe 2010“ (KldB 2010), beim anderen System handelt es sich um die Internationale Standardklassifikation der Berufe („International Standard Classification of Occupations“) des „International Labour Office“ (ILO) aus dem Jahr 2008 (ISCO 08). Darüber hinaus sind für Auswertungen der BIBB/BAuA-Erwerbstätigenerhebung für beide Systeme auch die älteren Klassifikationssysteme nach KldB 1992 und ISCO 88 von Bedeutung, um Trendanalysen mit älteren Erhebungswellen durchführen zu können.⁵ Demnach wurden die beruflichen Tätigkeiten nach vier Klassifikationssystemen vercodet.

Die Angaben zu den Wirtschaftszweigen können zum einen nach der älteren Wirtschaftszweigklassifikation der Europäischen Union („Nomenclature des statistiques des activités économiques de la Communauté Européenne“: NACE Rev. 1.1) bzw. der deutschen Klassifikation der Wirtschaftszweige 2003 (WZ 2003) eingeteilt werden, zum anderen nach der neusten Wirtschaftszweigklassifikation der Europäischen Union („Nomenclature des statistiques des activités économiques de la Communauté Européenne“: NACE Rev. 2) bzw. der deutschen Klassifikation der Wirtschaftszweige 2008 (WZ 2008). Bei der Wirtschaftszweig-klassifikation nach NACE handelt es sich um das System der Europäischen Union, das System WZ stellt das der Bundesrepublik Deutschland dar. Die beiden Klassifikationssysteme stimmen jeweils bis zur Ebene der Klassen überein, die letzte Ebene der Unterklassen stellt ein nationales Spezifikum dar.

Im Folgenden werden die Mengengerüste der offenen Angaben, die durch Kantar Public vercodet wurden, dargestellt und im Anschluss das genaue Vorgehen bei Kantar Public erläutert.

Übersicht über die beruflichen Angaben:

- KldB 2010 123.016 Angaben
- ISCO 08 84.670 Angaben
- KldB 1992 84.670 Angaben
- ISCO 88 84.670 Angaben

Da bei der KldB 2010 auch die aktuellen Berufe mit vercodet wurden, liegt die Anzahl höher als bei den drei anderen Systemen.

Übersicht über die Angaben zum Wirtschaftszweig:

- WZ 2008 19.719 Angaben
- WZ 2003 19.719 Angaben

⁵ Die Vercodung nach KldB 2010 und ISCO 08 liegt vollständig für alle beruflichen Angaben seit der Erhebungswelle 2012 vor.

6.1 Berufeencoding

6.1.1 Das Verfahren bei Kantar Public

Die Encoding der offenen beruflichen Angaben ist ein komplexes Verfahren. Dies liegt zum einen an den teilweise sehr unspezifischen Angaben der Befragten, zum anderen beinhalten die Klassifikationen sehr viele verfügbare Codes. Während beim KldB 2010 der fünfstellige Code die feinste Untergliederung hat, ist es bei den drei anderen Systemen jeweils ein vierstelliger Code. Die jeweils feinste Gliederung wurde von Kantar Public encoded. Der KldB 2010 hat demnach 1.286 verschiedene Codes, beim KldB 1992 sind es 2.287, der ISCO 08 beinhaltet 436 und der ISCO 88 umfasst 390 Codes.

Da häufig Berufsbezeichnungen angegeben werden, die nicht spezifisch genug sind, um diese eindeutig encode zu können (z. B. nur „Mechaniker“ statt „KFZ-Mechaniker“ oder nur „Lehrer“ statt „Gymnasiallehrer“), wurde im CATI-Programm eine Liste mit zu allgemeinen Berufsangaben hinterlegt. Wenn ein solcher Begriff genannt wurde, kam eine Nachfrage, ob es noch eine genauere Bezeichnung für diesen Beruf gibt. Mit diesem Verfahren war es möglich, typische – für eine Encoding zu allgemein gehaltene – Berufsbezeichnungen im Vorfeld zu identifizieren und die Befragten auffordern zu können, ihre Berufsbezeichnung detaillierter zu nennen.

Bei der Berufeencoding bei Kantar Public gab es zwei Teilschritte. Zuerst wurde eine automatische Encoding auf Basis von elektronisch verfügbaren Verzeichnissen (näheres dazu s. u.) durchgeführt, Begriffe, die in diesen Listen nicht enthalten waren oder anders geschrieben wurden, wurden im Anschluss manuell encoded. Um die Qualität der manuelle Encodings bestmöglich sicherzustellen, hat Kantar Public für die Encoding der beruflichen Angabe eine manuelle Doppelencoding durchgeführt. D. h. jeder Begriff, der nicht automatisch encoded werden konnte, wurde von zwei unterschiedlichen Encodern klassifiziert und bei Abweichungen wurde von einem erfahrenen Encoder entschieden, welcher Code der passendere ist oder ob sogar ein anderer, dritter Code vergeben wird.

Um die oben genannte, sehr umfangreiche Menge an Encodings leisten zu können, hat Kantar Public zehn Personen für dieses Projekt eingesetzt und umfassend im Vorfeld geschult (aufgrund der Doppelencoding waren insgesamt etwa 520.000 offene Berufsangaben manuell zu encode; hinzu kamen noch die Schiedsfälle). Zu Beginn wurde eine einwöchige Einarbeitung durchgeführt. Diese beinhaltete eine Basisschulung und Testencodings zum ersten Klassifikationssystem. Anschließend wurden die Encodings und abweichende Ergebnisse bei der Encoding in der Gruppe diskutiert und geklärt, wie in diesen Fällen zu verfahren ist. War die Encoding eines Systems abgeschlossen und auf das nächste umgestiegen, wurde vorab auf die Besonderheiten des neuen Klassifizierungssystems hingewiesen. Auch hier waren zunächst Testencodings zu leisten, bevor mit den BIBB/BAuA-Daten gearbeitet wurde. Die Schulungsunterlagen zur Basisschulung und der KldB 2010 sind im Anhang E. Nach den ersten zwei Wochen wurde geprüft, an welchen Stellen noch Unklarheiten bestehen und entsprechende Nachschulungen vorgenommen, sofern dies erforderlich schien.

Um die Organisation und das Datenhandling bei der Doppelencoding mit zehn Encodern zu ermöglichen, hat Kantar Public eine Datenbank entwickelt, die eine Steuerung der Zuweisung der noch zu bearbeitenden Fälle ermöglichte. Damit wurde garantiert, dass die doppelt zu encodehenden Fälle von zwei verschiedenen Encodern bearbeitet wurden. Ein bereits bearbeiteter Fall war für den ersten Encoder gesperrt. Jeder Encoder arbeitete unabhängig an dem Datensatz und konnte zudem nach eigener thematischer Auswahl Fälle aufrufen. Zusätzlich ermöglichte die Datenbank z. B. die Möglichkeit, zusätzliche Informationen für die Encoder einzublenden (s. u.) und stellte sicher, dass keine Codes eingegeben werden konnten, die in den Klassifikationssystemen nicht vorgesehen sind.

Das Kantar-Verfahren basiert auf einem mehrstufigen Verfahren:

1. Automatik: Grundlage der automatischen Zuordnung der Codes sind die elektronisch verfügbaren Verzeichnisse der Klassifikationssysteme (Dictionary 1). Wenn ein Begriff mit dem in der Befragung angegebenen Begriff genau übereinstimmt, wird der zugehörige Code verwendet. Die Encoding ist damit abgeschlossen.

2. Kantar Public hat auf Basis der bisher durchgeführten Vercodungsprojekte ein eigenes, umfangreiches Dictionary aufgebaut. Dieses wurde fortlaufend um Begriffe erweitert, bei denen die Vercoder davon ausgegangen sind, dass diese eindeutig zuzuordnen sind (Dictionary 2). Da es in Einzelfällen trotzdem sein kann, dass es (z. B. aufgrund von zusätzlichen Angaben) vielleicht einen passenderen Code gibt, wird dieser Code nicht ohne Kontrolle übernommen, sondern bei der manuellen Vercodung als Vorschlag angezeigt. Die Vercoder waren angehalten, diesen Code noch einmal zu prüfen.
3. Mögliche Umstiege zwischen dem KldB 2010 und dem ISCO 08 bzw. dem KldB 92 werden ebenso als Vorschlag angezeigt.⁶ Diese Umstiegslisten sind ebenso frei zugänglich. Da es aber nicht für jeden KldB 2010-Code einen Umstieg gibt, diese nicht immer eindeutig sind (es gibt manchmal für einen Code mehrere Vorschläge für die anderen Systeme) und unserer Erfahrung nach auch nicht immer passend sind, werden auch diese nicht ohne Kontrolle übernommen, sondern müssen manuell geprüft werden.
4. Die manuelle Vercodung der Leftovers aus Schritt 1 (also den Fällen, die im Rahmen von Schritt 1 nicht vercodet werden konnten) findet unter Berücksichtigung der Vorschläge aus Schritt 2 und 3 statt. Die bei der manuellen Vercodung angewendeten Regeln sind eindeutig formuliert und dokumentiert (siehe Abschnitt 1.1.2).
5. Schiedsverfahren

6.1.2 Vorgehensweise und allgemeine Regeln bei der manuellen Vercodung

Grundsätzlich wird die Angabe vercodet, die als berufliche Tätigkeit genannt wird. Liegen weitere Angaben in anderen Variablen (z. B. Branche, Stellung, Beamtenlaufbahnstufe, für den Beruf erforderliche Ausbildung) vor, können diese zur genaueren Vercodung herangezogen werden. Um die Vercoder durch die zusätzlichen Angaben nicht sofort zu beeinflussen, wird in der Datenbank nur angezeigt, ob zusätzliche Informationen vorliegen – was konkret in diesen Angaben enthalten ist, wird erst angezeigt, wenn die Vercoder diese aufrufen. Hintergrund dieses Vorgehens ist, dass die Angabe vercodet werden soll, die genannt wurde. Sofern dies eindeutig möglich ist, werden keine zusätzlichen Angaben verwendet.

Angaben zu Berufsausbildungen werden so behandelt, als wäre die Ausbildung abgeschlossen, soweit eine Zuordnung zu einem Beruf möglich ist (z. B. Schreinerlehre als Schreiner). Bei Studiengängen wird der Beruf vercodet, der mit höchster Wahrscheinlichkeit mit diesem Abschluss ausgeübt wird.

Typischerweise enthalten die offenen Angaben Nennungen, die so allgemein sind, dass die Zuordnung von Codes schwierig oder nicht eindeutig möglich ist. Um dennoch möglichst viele Codes zuweisen zu können, mussten Entscheidungen für schwierige Fälle getroffen werden. Dies erscheint dann allerdings wenig problematisch, wenn die Entscheidungsregeln dokumentiert und diese Fälle entsprechend in einer Variable „XXX_prob“ gekennzeichnet werden.⁷

Die allgemeinen Regeln werden an dieser Stelle kurz dargestellt:

⁶ Dies gilt nur für den ISCO 08 und den KldB 1992, da hier schon die finalen Ergebnisse des KldB 2010 vorlagen.

⁷ In Abhängigkeit vom Vercodungssystem heißt die Variable bspw. f100_kldb2010_prob, f100_isco08_prob, f100_kldb1992_prob und f100_isco88_prob

Regel 1:

Vercodet wird grundsätzlich der angegebene Begriff, wenn er in der jeweiligen Klassifizierung der Berufe aufgeführt ist.

Regel 2:

Bei mehreren Nennungen in einer Klartextangabe wird die erste vercodet, es sei denn, die weiteren Angaben dienen einer näheren Spezifizierung. Bei Doppelangaben, bei denen die eine Angabe die andere näher spezifiziert, wird die jeweils spezifischere Angabe vercodet. Bei zusammengesetzten Begriffen, die so nicht vercodet werden können, wird, wenn möglich, der übergeordnete Begriff vercodet (z. B. Hellermonteur als Monteur).

Regel 3:

Wo eine Vercodung auf der Basis der offenen Angabe allein nicht möglich ist, werden zusätzliche Informationen – soweit vorhanden – hinzugezogen. Falls dies der Fall ist, wird dies in einer Variable „XXX_info“ (Name vom Klassifizierungssystem abhängig) dokumentiert.⁸

Regel 4:

Ist die offene Angabe so unspezifisch, dass sie mehreren Kennziffern zugeordnet werden kann (z. B. Händler – hier ist nicht klar, ob es sich um Einzel- oder Großhändler handelt), wird sie derjenigen zugeordnet, für die angenommen werden kann, dass sie empirisch am häufigsten vorkommt. Diese Fälle werden in der Variable „XXX_prob“ (Name vom Klassifizierungssystem abhängig) gekennzeichnet (s. u.).

Regel 5:

Wenn die Angabe zur beruflichen Tätigkeit keinen Aufschluss über die für die Tätigkeit erforderliche Qualifikation gibt und keine zusätzlichen Informationen hierzu vorliegen (siehe auch Regel 3), wird der Code der Tätigkeit zugewiesen, für die die geringste Qualifikation notwendig ist. Diese Fälle werden in der Variable „XXX_prob“ (s. u.) gekennzeichnet.

Problematische Angaben, bei denen auf die Regeln 4 oder 5 zurückgegriffen wird, werden in einer Zusatzvariablen „XXX_prob“ als solche gekennzeichnet. Dadurch besteht die Möglichkeit, Auswertungen mit diesen Fällen und ohne sie vorzunehmen und somit Sensitivitätstests durchzuführen. Um darüber hinaus eine Hilfestellung zur Beantwortung der Frage zu geben, inwiefern ein solches Vorgehen inhaltlich notwendig ist, werden im Folgenden die konkreten Sachverhalte, bei denen die Entscheidungsregeln angewendet werden, detailliert dokumentiert.

Die Variable „XXX_prob“ hat damit folgende Ausprägungen:

- XXX_prob:**
- 1: Qualifikationsstufe nicht bekannt, niedrigste vercodet
 - 2: Mehrere Codes möglich, Entscheidung getroffen

Bei der Doppelvercodung kann es vorkommen, dass zwar die beiden vergebenen Codes identisch sind, aber von einem Kodierer ein XXX_prob oder XXX_info Code vergeben wurde, vom Zweitbearbeiter jedoch nicht. In diesen Fällen wird die „sicherere“ Variante übernommen. Das bedeutet, sollten bei einer Bearbeitung Zusatzinformationen verwendet worden sein, so wird der XXX_info Code auch final vergeben. Mit XXX_prob verhält es sich ebenso, wobei hier bei unterschiedlicher Vergabe die Ausprägung 2 („Mehrere Codes möglich, Entscheidung getroffen“) sticht. Wird ein Fall im Rahmen des Schiedsverfahrens in der Drittbearbeitung entschieden, so werden die Ausprägungen von XXX_prob und XXX_info aus dieser Bearbeitung übernommen, unabhängig von den vorangegangenen zwei Bearbeitungen.

Im Unterschied zu den beiden älteren Systemen (KldB 1992 und ISCO 88), in denen für bestimmte Berufe Allgemeincodes vorgesehen waren (z. B. für Angestellter; Ingenieur, Selbstständiger usw.), gibt

⁸ In Abhängigkeit vom Vercodungssystem heißt die Variable bspw. f100_kldb2010_info, f100_isco08_info, f100_kldb1992_info und f100_isco88_info

es diese Codes in den neueren Systemen des KldB 2010 und ISCO 08 nicht mehr. Kantar Public codiert solche Angaben mit Sondercodes im Minusbereich (–10, –11 usw.), die im nächsten Abschnitt erläutert werden.

6.1.3 Besonderheiten der jeweiligen Klassifikationssysteme

KldB 2010 und ISCO 08

Trotz der in eben erläuterten Entscheidungsregeln beim Vorliegen nicht eindeutiger Angaben ist es nicht möglich, allen offenen Angaben einen Code zuzuordnen. Um die vorliegende Information zur beruflichen Tätigkeit dennoch zu verwerten und um dem Forscher Hinweise auf deren Inhalt zu geben, werden in solchen Fällen, wo möglich, Sondercodes vergeben. Die Bedeutung der Sondercodes wird im Folgenden dokumentiert.

- 10: Aushilfe
- 11: Auszubildende, Azubi, Berufsvorbereitung
- 12: Arbeiter, Facharbeiter, Hilfsarbeiter, Produktionsarbeiter, Werker, Industrieangestellter, Industriearbeiter, Industriefacharbeiter, Teilwerker
- 13: Industriemeister, Meister
- 14: Techniker, technischer Angestellter
- 15: Ingenieur, technischer Ingenieur
- 16: Praktikant, Trainee, Volontär
- 17: Freiberufler, Geschäftsmann/-frau, Selbstständiger
- 18: mithelfender Familienangehöriger (sowohl in als auch außerhalb der Landwirtschaft)
- 19: Bundesfreiwilligendienst, Freiwilliges soziales Jahr, Wehrdienst, Zivildienst
- 20: Dienstleister, Dienstleistung
- 21: Zeitarbeit
- 22: Fachkraft

- 30: studentische Angestellte/Hilfskraft, Werkstudent, wissenschaftliche Hilfskraft
- 31: Bachelor
- 32: Diplom, Magister, Master, 1. Staatsexamen, Staatsexamen
- 33: 2. Staatsexamen, Assessor, 2. Staatsprüfung, Referendariat
- 34: Studium, Studium an Hochschule, Universitätsstudium, Fachhochschulabschluss, Fachhochschule, Aufbaustudium, Fachdiplom

- 50: Berater, Consultant
- 51: Einsatzleiter, Leiter, leitender Angestellter
- 52: Betriebsleiter
- 53: Kaufmann/-frau

- 54: Gutachter, Sachverständiger
- 55: Referent

- 60: sozial, Sozialbereich, Sozialwesen
- 61: medizinisch, Medizinbereich, medizinischer Bereich
- 62: Therapeut

- 90: Hauptschulbesuch, Hauptschulabschluss, Berufseinstiegsjahr (BEJ), Berufsgrundbildungsjahr (BGJ), Berufsvorbereitungsjahr
- 91: Realschulbesuch, Realschulabschluss, mittlere Reife, Fachoberschulreife, Berufsgrundschuljahr, Wirtschaftsschule, Wirtschaftsschulabschluss
- 92: Berufsfachschulbesuch, Fachschulreife
- 93: Fachschulbesuch, Fachschulabschluss (nur wenn eine Fachrichtung nicht erkennbar ist; wenn es sich z. B. um eine soziale, medizinische, pflegerische oder erzieherische Fachschule handelt, wird der entsprechende Code vergeben)
- 94: Handelsschulbesuch, Abschluss einer Handelsschule
- 95: Fachoberschulbesuch (FOS), Fachhochschulreife, Berufsfachoberschule, Berufsoberschule
- 96: Gymnasium, Abitur, Hochschulreife, Fachabitur, fachbezogener Hochschulzugang
- 1: keine Zuordnung möglich

Für die automatische Vercodung des KldB 2010 wurde die Liste der „Bundesagentur für Arbeit, 2011: Klassifikation der Berufe 2010 – alphabetisches Verzeichnis der Berufsbenennungen“ verwendet, für den ISCO 08 war Grundlage die Liste „ISCO 08 – gemeinsame deutschsprachige Titel und Erläuterungen auf Basis der englischsprachigen Version von Juni 2009“, Stand der Bearbeitung 2010; erstellt von „STATISTIK AUSTRIA – Die Informationsmanager“.

KldB 1992 und ISCO 88

Eine Besonderheit der älteren Klassifikationssysteme ist, dass es einen direkten Umstieg zwischen dem KldB 1992 und dem ISCO 88 gibt. D. h., dass an dieser Stelle nur der KldB 1992 manuell vercodet wurde und der ISCO 88 durch den Umstieg generiert wurde.

Berufsangaben, die in der Klassifizierung der Berufe von 1992 nicht definiert sind, werden, wo möglich, nach vergleichbaren Berufen vercodet. Teilweise wird dabei nach den nächsten Oberbegriffen vercodet: So erhalten der „Gussschweißer“ oder der „Buntschweißer“ die Kennziffer für „Schweißer“ (2410).

Besonders tritt dieses Problem aufgrund neuer oder sich im Zuge der technologischen Entwicklung ändernder Berufsbilder auf, die in der Klassifizierung der Berufe gar nicht aufgeführt sind, speziell im Zusammenhang mit den neuen Medien, Internet usw. (z. B. Web-Designer, Medienkaufmann, Multimedia...). Hier wird, wo möglich, nach vergleichbaren Berufen vercodet oder nach den nächsten Oberbegriffen, z. B. Medienkaufmann als nichtselbstständiger Kaufmann.

Für die automatische Vercodung des KldB 1992 wurde die Liste „Klassifizierung der Berufe, Ausgabe 1992“ des Statistischen Bundesamtes verwendet.

6.2 Branchenvercodung

6.2.1 Das Verfahren bei Kantar Public

Durch das Verfahren von Kantar Public wird der zweistellige numerische Code für die Abteilungen zugewiesen.

Wie auch schon bei der Berufe-Vercodung beschrieben, erfolgt auch die Branchenvercodung in zwei Schritten: in einem ersten Schritt wird eine automatische Vercodung vorgenommen, wenn der angegebene Begriff genau in der Liste zur Klassifikation der Wirtschaftszweige enthalten ist, in einem zweiten Schritt wurden die übrig gebliebenen Begriffe manuell vercodet.

Grundlage für die Bearbeitung Angaben zum Wirtschaftszweig nach WZ 2008 ist die „Klassifikation der Wirtschaftszweige. Mit Erläuterungen – 2008“ (Statistisches Bundesamt Ausgabe 2009), für den WZ 2003 ist es die „Klassifikation der Wirtschaftszweige. Mit Erläuterungen – 2003“ (Statistisches Bundesamt Ausgabe 2003).

Neben den dort vorgegebenen konkreten Codes werden bei sehr allgemeinen Angaben vier zusätzliche Codes für „Industrie“, „Handwerk“, „Dienstleistung“ und „Produzierendes Gewerbe“ vergeben.

Zudem können bei nicht eindeutigen Angaben zum Wirtschaftszweig weitere Informationen zur Spezifizierung herangezogen werden (z. B. Angabe zu beruflicher Tätigkeit, Stellung im Beruf o. ä.).

6.2.2 Vorgehensweise und allgemeine Regeln bei der manuellen Vercodung

Typischerweise enthalten die offenen Angaben Nennungen, die so allgemein sind, dass die Zuordnung von Codes schwierig oder nicht eindeutig möglich ist. Um dennoch möglichst viele Codes zuweisen zu können, mussten Entscheidungen für schwierige Fälle getroffen werden. Dies erscheint dann allerdings wenig problematisch, wenn die Entscheidungsregeln dokumentiert werden. Kantar Public hat alle bislang aufgetretenen Problemfälle erfasst und festgehalten, wie entschieden wurde. So entstand eine Sammlung sowohl von allgemeinen als auch speziellen Regeln, die genau festlegen, wie in solchen Fällen zu verfahren ist.

Die allgemeinen Regeln werden an dieser Stelle kurz dargestellt:

Regel 1:

Vercodet wird grundsätzlich der angegebene Begriff, wenn er in der „Klassifikation der Wirtschaftszweige“ des Statistischen Bundesamtes aufgeführt ist.

Regel 2:

Bei mehreren Nennungen in einer Klartextangabe wird die erste vercodet, es sei denn, die weiteren Angaben dienen einer näheren Spezifizierung. Bei Doppelangaben, bei denen die eine die andere näher spezifiziert, wird die jeweils spezifischere Angabe vercodet.

Regel 3:

Wo eine Vercodung auf der Basis der offenen Angabe allein nicht möglich ist, werden zusätzliche Informationen – soweit vorhanden – hinzugezogen; diese eventuell zudem in geeigneter Kombination. Falls dies der Fall ist, wird dies in einer Variable „WZ2008_info“ bzw. „WZ2003_info“ dokumentiert.

Regel 4:

Ist die offene Angabe so unspezifisch, dass sie mehreren Kennziffern zugeordnet werden kann, wird sie derjenigen zugeordnet, für die angenommen werden kann, dass sie empirisch am häufigsten vorkommt.

Regel 5:

Bei Angaben zum Wirtschaftszweig, die nicht erkennen lassen, ob Herstellung oder Handel gemeint ist, wird Herstellung vercodet.

Problematische Angaben, bei denen auf Regel 4 zurückgegriffen wird, werden in der Zusatzvariablen „WZ2003_prob“ bzw. „WZ2008_prob“ mit Wert 2 als solche gekennzeichnet. Problematische Angaben, bei denen Regel 5 zur Verwendung kommt, werden in der Zusatzvariablen „WZ2003_prob“ bzw. „WZ2008_prob“ mit Wert 1 gekennzeichnet.

Dadurch besteht die Möglichkeit, Auswertungen mit diesen Fällen und ohne sie durchzuführen und somit Sensitivitätstests durchzuführen.

Die Variable „WZ2003_prob“ bzw. „WZ2008_prob“ hat damit folgende Ausprägungen:

XXX_prob:	1:	Herstellung oder Handel unklar, Herstellung vercodet
	2:	Angabe unspezifisch, wahrscheinlichste vercodet

6.2.3 Spezielle Regeln und Besonderheiten der jeweiligen Klassifikationssysteme

Die Befragten nennen als Branche oftmals den Namen der Firma, nochmals die berufliche Tätigkeit oder eine Bezeichnung, die nicht in das Klassifikationsschema passt bzw. unter mehrere der Wirtschaftszweig-Abteilungen fallen könnte, z. B. wenn der Betrieb eine Mischform darstellt. Grundsätzlich stellt die Vercodung nach NACE bei Letzterem auf den *überwiegenden* Geschäftsbereich ab, der von uns aber oft nicht feststellbar ist.

Vielfach kennen die Befragten ihre Branche offensichtlich auch gar nicht, wie Angaben wie Industrie, Handwerk, IG Metall, IG *Phantasienamen, die den Bereich irgendwie umschreiben*, zeigen.

Ebenso sind allgemeine Angaben wie Elektro, Elektriker, Elektrowerkzeug, Elektrobranche, Elektrobetrieb usw., Metall, Metallbearbeitung, Metallbereich, Metallindustrie, Metallgewerbe nicht eindeutig vercodbar.

Auch Angaben wie z. B. Agrargenossenschaft, Kommune, öffentlicher Dienst oder ein Gewerkschaftsbereich beziehen sich, wie den zugeordneten Berufen zu entnehmen ist, vielfach nicht auf diese Branche, sondern dokumentieren offensichtlich das Zugehörigkeitsgefühl bzw. die Mitgliedschaft der Befragten zu, nicht aber ihr Beschäftigungsverhältnis in dieser Kategorie. Hier einfach die entsprechenden Codes (Kommune = Kommunalverwaltung, Agrargenossenschaft = Interessengemeinschaft, IG Metall oder IG Druck und Papier = Gewerkschaft = Interessenvertretung) zu vergeben, würde die Aussagekraft der Angaben deutlich verfälschen. In solchen Fällen wird – wo möglich – nach der beruflichen Tätigkeit der entsprechende Wirtschaftszweig ermittelt. Lässt die Angabe zur beruflichen Tätigkeit darauf schließen, dass der Befragte tatsächlich in dem fraglichen Bereich tätig ist oder sein könnte, also z. B. in der Kirchenverwaltung oder als Gewerkschaftssekretär, oder keine Berufsangabe vorliegt, wird der entsprechende Code vergeben.

Bei zu allgemeinen Angaben, wie z. B. Bergbau oder Fensterbau, werden Codes vergeben, die spekulativ erscheinen mögen, aber aufgrund von überwiegendem Vorkommen getroffen werden (z. B. Fenster-/Türen Bau = Holzverarbeitung, da Holz hierfür das überwiegende Material darstellt). Diese Codes werden in „WZ2003_prob“ bzw. „WZ2008_prob“ mit 2 gekennzeichnet.

Bei mehreren Angaben (z. B. Metall-Kunststoff) wird in der Regel die erste vercodet, es sei denn, die weiteren Angaben dienen einer näheren Bestimmung der ersten.

Allgemeine Angaben ohne nähere Charakterisierung, die nicht erkennen lassen, ob Herstellung oder Handel gemeint ist, wie z. B. Kraftfahrzeuge/Lebensmittel o. ä., werden dem Bereich zugeordnet, der in

der Wertschöpfungskette am weitesten vorne liegt, also beispielsweise dem Bereich Förderung, Erzeugung oder Herstellung und nicht dem Bereich Handel, Vermietung oder Reparatur. Auf der oben angesprochenen Zusatzvariablen „WZ2003_prob“ bzw. „WZ2008_prob“ wird in diesen Fällen der Wert 1 gesetzt.

Eine Ausnahme von dieser Regel bildet die Angabe Computer, Computerbranche ohne nähere Spezifizierung. Hier wird nicht Computerherstellung vercodet, sondern Erbringung von Dienstleistungen der Informationstechnologie, da diesem Wirtschaftszweig im Vergleich zur Herstellung die größere Reichweite zugeordnet werden kann. Auch hier wird auf der Zusatzvariablen „WZ2003_prob“ bzw. „WZ2008_prob“ der Wert 1 gesetzt.

Bei Metallbranchen wird die allgemeine Angabe Metall, Metallindustrie o.ä. wie *Verarbeitung*, also mit 25 (hier „WZ2003_prob“ bzw. „WZ2008_prob“ Wert 2), vercodet; bei den einzelnen Metallen (wie Aluminium, Stahl o. ä.) werden die analogen Angaben wie *Erzeugung*, also mit 24 (hier „WZ2003_prob“ bzw. „WZ2008_prob“ Wert 1), vercodet, da die zusätzlichen Berufsangaben weit überwiegend diese Einordnung nahelegen.

Außerdem werden je Klassifikationssystem vier weitere Codes für Angaben vergeben, die sonst nicht zuzuordnen wären:

WZ 2008:

201	Industrie
202	Handwerk
203	Dienstleistungen
204	Produzierendes Gewerbe
–1	<i>keine Zuordnung möglich</i>

WZ 2003:

96	Industrie
97	Handwerk
98	Dienstleistungen
100	Produzierendes Gewerbe
–1	<i>keine Zuordnung möglich</i>

7 Gewichtung

Ziel der Untersuchung ist es, eine repräsentative Stichprobe der Grundgesamtheit zu erhalten. Telefonbasierte Befragungen über Festnetz haben zwei Designkomponenten, die die Proportionalität der Abbildung beeinträchtigen:

- die von der Zahl der für Gespräche nutzbaren Anschlüsse abhängige Auswahlwahrscheinlichkeit für Telefonhaushalte und
- die von der Zahl der zur Zielgesamtheit gehörenden Personen im Haushalt abhängende Auswahlwahrscheinlichkeit für Personen.

Eine besondere Herausforderung entsteht, wenn zusätzlich eine Befragung über Mobilfunk durchgeführt wird, da dann eine Haushaltsstichprobe (Festnetz) und eine Personenstichprobe (Mobilfunk) zusammengeführt werden müssen.

In der Mobilfunkstichprobe ist die Auswahl der Personen mit der Selektion der Nummern als einfache Zufallsauswahl bereits abgeschlossen: Der kontaktierte (Haupt-) Nutzer des Gerätes ist die Zielperson. Demgegenüber wird in der Festnetzstichprobe ein zweistufiges Auswahlverfahren angewandt. In der Kontaktphase wird mit dem Schwedenschlüssel (Kish-Selection-Grid) zufällig eine Zielpersonenauswahl innerhalb des jeweiligen Haushaltes vorgenommen.

Darüber hinaus gibt es einen zweiten Ursachenkomplex für Abbildungsprobleme: Einerseits ist bei Zufallsstichproben ein Zufallsfehler unvermeidlich, andererseits sind bei jeder auf Freiwilligkeit basierenden Befragung nonresponse-bedingte Selektivitäten zu konstatieren.

Aufgrund dieser Stichprobenanlage sind die Befragungsdaten stets zu gewichten. Mit einer Designgewichtung werden die unterschiedlichen Auswahlchancen korrigiert. Zusätzlich ist eine Ausfallgewichtung nach demografischen Größen auf Basis amtlicher Werte ein bewährtes Mittel zur Verbesserung der Abbildung der Grundgesamtheit durch die Stichprobe. Beide Verfahren werden im Folgenden beschrieben.

7.1 Designgewichtung

Die designbedingten Unterschiede in den Auswahlwahrscheinlichkeiten sind theoretisch und praktisch leicht mittels Designgewichtung zu korrigieren. Unterschiedliche Auswahlwahrscheinlichkeiten ergeben sich durch den Dual-Frame-Ansatz, durch die unterschiedliche Anzahl an Telefonnummern, unter denen die Zielperson erreichbar ist und durch die unterschiedliche Anzahl an Zielpersonen im Haushalt.

Beim Dual-Frame-Design ist die Inklusionswahrscheinlichkeit insgesamt abhängig von der Wahrscheinlichkeit, per Festnetz oder per Mobilfunk befragt zu werden. Die Wahrscheinlichkeit, für die gleiche Untersuchung in beiden Modes ausgewählt zu werden, ist ausgesprochen klein und wird deshalb an dieser Stelle vernachlässigt.

Jede Telefonnummer hat innerhalb des Modes die gleiche Auswahlchance. Personen, die über mehrere Rufnummern verfügen, haben daher eine der Zahl dieser Anschlüsse entsprechende Mehrfachchance bei der Auswahl. Zu deren Bereinigung wird die realisierte Stichprobe mit der reziproken Zahl der Telefonnummern multipliziert. Gleichzeitig ist die Abhängigkeit der Auswahl von der Anzahl der Zielpersonen im Haushalt zu korrigieren. In jedem der Haushalte der Festnetzstichprobe wird nur eine Zielperson für das Interview per Zufall ausgewählt. Hierzu wird der Schwedenschlüssel verwendet. Dieses Verfahren produziert innerhalb des einzelnen Haushaltes gleiche Auswahlchancen. Auf der

Ebene der Gesamtstichprobe aber ist die Auswahlchance des Einzelnen umgekehrt proportional zur Zahl der potenziellen Zielpersonen im jeweiligen Haushalt. Daher ist die Stichprobe ex-post mathematisch mit dem reziproken Wert (Designgewicht) umzuformen.

Eine weitere denkbare Abhängigkeit der Inklusionswahrscheinlichkeit von der Anzahl der möglichen Nutzer eines Handys kann derzeit vernachlässigt werden. In fast allen Fällen handelt es sich bei Mobilfunktelefonen um persönliche Geräte, die nicht gemeinsam genutzt werden.

Nach Gabler und Ayhan (2007) ergibt sich die Inklusionswahrscheinlichkeit⁹:

$$\pi_i \approx k_i^F \frac{m^F}{M^F} \cdot \frac{1}{z_i} + k_i^C \frac{m^C}{M^C}, i = 1, \dots, N$$

$m^{F/C}$ Anzahl Nummern in der Stichprobe (Festnetz / Mobilfunk)

$M^{F/C}$ Anzahl gültiger Rufnummern insg. (Festnetz / Mobilfunk)

$k_i^{F/C}$ Anzahl der Rufnummern (Festnetz / Mobilfunk) über die der Haushalt bzw. die Person erreicht werden kann

z_i Anzahl Zielpersonen im Haushalt

Folgende Komponenten werden im Interview erhoben: Anzahl der persönlichen Rufnummern in Fest- (k_i^F) und Mobilnetz (k_i^C) und die Anzahl der Erwerbstätigen Personen ab 15 Jahren im Haushalt (z_i). Dies gilt für alle Befragten, unabhängig davon, in welchem Mode sie interviewt wurden.¹⁰ Darüber hinaus ist der unterschiedliche Auswahlatz (m^F/M^F bzw. m^C/M^C) der beiden Netzziehungen zu berücksichtigen. Die Werte m^F und m^C werden unter Verwendung des ersten Designfaktors gezählt. Die Werte für die Parameter M^F und M^C werden im Rahmen einer großen, unabhängigen Grundlagenstudie bestimmt. Hierzu kann auf eine regelmäßig (jährlich) durchgeführte F2F-Zufalls-Bevölkerungsstichprobe in der Größe von rund 20.000 Befragten zurückgegriffen werden, für die vorliegende Studie auf die Daten aus 2017. In dieser Studie wird neben der Demografie unter anderem auch die Haushaltsausstattung und persönliche Ausstattung bezüglich Festnetztelefon und Mobilfunk erhoben. Damit ist man in der Lage, die Größen M^F respektive M^C für die Grundgesamtheit der Erwerbstätigen auf unabhängige Weise zu bestimmen.

Für die vorliegende Studie wurden M^F und M^C anhand der Auswertung der Mehrthemenbefragung mit Unterstützung folgender Basiswerte ermittelt:

Haushalte mit erwerbstätigem Haupteinkommensbezieher (Mikrozensus 2017): 26,173 Mio.

Erwerbstätige Personen 15 Jahre und älter (Mikrozensus 2017): 37,434 Mio.

Daraus resultieren die folgenden Werte:

M^F : 33,172 Mio.

M^C : 40,275 Mio.

Dieses Designgewicht, welches auch „Umwandlungsfaktor“ genannt wird, bildet den Eingangsfaktor der folgenden Ausfallgewichtung. Ab diesem Schritt werden die Festnetz- sowie die Mobilfunkstichprobe als eine gemeinsame Personenstichprobe behandelt.

Da bei der Erwerbstätigenbefragung 2018 ein Dual-Frame-Ansatz gewählt wurde, ist das Verfahren bei der Designgewichtung komplexer und nicht unmittelbar vergleichbar mit der Designgewichtung bei der Erwerbstätigenbefragung 2012, bei der eine reine Festnetzstichprobe gezogen wurde. Im Rahmen der Designgewichtung erfolgte aufgrund des Dual-Frame-Ansatzes ein Ausgleich der beiden Kontaktwege. Durch diesen Ausgleich der ungleichen Auswahlwahrscheinlichkeiten wurde auch der im Vergleich zum Anteil in der erwerbstätigen Bevölkerung zu geringe Anteil an sogenannten „Mobile-Onlys“, also

9 Gabler, S. und Ayhan, Ö. (2007): Gewichtung bei Erhebungen im Festnetz und über Mobilfunk: Ein Dual-Frame-Ansatz, in: Mobilfunktelefonie – Eine Herausforderung für die Umfrageforschung, Gabler S., Häder S. (Hrsg.), GESIS – ZUMA-Nachrichten Spezial Band 13.

10 In der entsprechenden Frage werden zudem explizit Telefonnummern, die nur für Fax oder PC benutzt werden, ausgeschlossen.

Mobilfunkbesitzer, die über keine Festnetznummer verfügen, in der realisierten Stichprobe ausgeglichen.¹¹

7.2 Ausfallgewichtung über Strukturanpassung

Über Designkomponenten hinaus gibt es einen zweiten Ursachenkomplex für Abbildungsprobleme: Fehlende Erreichbarkeit und Nonresponse sind Quellen für mögliche Selektivitäten. Diese werden im Rahmen einer Ausfallgewichtung über eine Strukturanpassung ausgeglichen.

Im Rahmen der Strukturanpassung werden Referenzstatistiken herangezogen, die (per Konvention) als Abbildung der Grundgesamtheit gelten und die interessierenden Strukturmerkmale in vergleichbarer Weise abbilden. Ausfallgewichtungen orientieren sich in diesen Fällen an Vergleichen der Strukturen der realisierten Stichprobe mit den Strukturen der Grundgesamtheit und passen dann die Stichprobenstrukturen denen der Grundgesamtheit mittels Gewichtungsfaktoren an (Kalibrierung).

Jede Art von Ausfallgewichtung basiert auf der Annahme, dass die Strukturmerkmale, für die im Rahmen der Ausfallgewichtung Anpassungen vorgenommen werden, mit den inhaltlich interessierenden Merkmalen korreliert sind und somit über die Gewichtung auch deren Strukturen besser an die in der Grundgesamtheit angepasst werden. Es empfiehlt sich, solche Strukturgewichtungen so behutsam wie möglich vorzunehmen und dabei Merkmale (und deren Kombinationen) zu verwenden, die möglichst hoch mit den Zielmerkmalen der Untersuchung korreliert sind (vgl. z. B. Heyde 2014, S. 58).¹²

Die Berechnung der Gewichtungsfaktoren erfolgt in einem iterativen Randsummenverfahren. Auf Basis der Merkmale oder Merkmalskombinationen wird das Modell als Gewichtung mit mehrdimensionalen Randverteilungen realisiert. Die Gewichtungsfaktoren werden dabei auf ein vorgegebenes geschlossenes Intervall begrenzt, die Faktorenspannbreite darf also ein bestimmtes vorher definiertes Maximum nicht überschreiten. Damit werden extrem hohe oder extrem niedrige Faktoren verhindert. Die Begrenzung der Gewichtungsfaktoren wurde bei der vorliegenden Untersuchung bei Minimum 0.200 und Maximum 4.9 vorgenommen. Die Priorität der Gewichtung gibt an, welcher Gewichtungsrand die anderen Gewichtungsänder dominieren, d. h. welcher Gewichtungsrand bevorzugt wird, wobei sich dadurch die nicht-dominanten Gewichtungsänder in ihrer Anpassung etwas verschlechtern können.

Ein integriertes Verfahren mit iterativem Vorgehen ermöglicht es, Verschiebungen der Strukturen von in einem vorangegangenen Iterationsschritt berücksichtigten Merkmalen in späteren Iterationsschritten, die sich auf diese Merkmale beziehen, wieder zu korrigieren. Dadurch wird die Anpassung an alle Ränder sukzessive solange verbessert, bis ein vorher definiertes Konvergenzkriterium erreicht ist. Als Konvergenzkriterium dient die Vorgabe, dass für jede Randverteilung maximal 0,1 % Abweichung in der Summe der einzelnen Zellen toleriert wird. Die Iteration wird dann beendet, wenn die Anpassung an die einzelnen Randverteilungen nicht mehr verbessert wird.

Während der Gewichtung wird des Weiteren geprüft, ob die Faktoren in einzelnen Zellen zu hoch oder zu niedrig werden oder die Zellenbesetzung gleich Null ist. In diesen Fällen werden Zusammenfassungen mit einer geeigneten Nachbarzelle vorgenommen.

11 Der ungewichtete Anteil an „Mobile-Onlys“ in der Stichprobe beträgt 6,2%. Dieser Anteil steigt im Rahmen der Designgewichtung auf 19,4%.

12 Heyde, C. von der (2014): Techniken und Möglichkeiten der Realisierung von Random-Stichproben. S. 37 – 65 in: ADM Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V.: Stichproben-Verfahren in der Umfrageforschung. Eine Darstellung für die Praxis. 2. Auflage. Wiesbaden: Springer VS.

Im Folgenden wird das Vorgehen kurz genauer dargestellt. Die allgemeine Formel für die Gewichtung lautet (vgl. Deming, Stephan 1940; Cochran 1968): ¹³

$$g_j = \frac{1}{\frac{n_j}{n} * \frac{N}{N_j}} = \frac{n * N_j}{n_j * N}$$

g= Gewicht

j = Schicht, definiert durch das jeweilige Gewichtsmerkmal

n_j = Anzahl (designgewichteter) Fälle in Schicht j in der Stichprobe

n = Anzahl der Fälle in der Stichprobe

N_j = Anzahl der Fälle in Schicht j in der Grundgesamtheit

N = Anzahl der Fälle in der Grundgesamtheit

Diese allgemeine Formel kann im iterativen Randsummenverfahren wiedergefunden werden. Wir illustrieren dies am Beispiel eines iterativen Randsummenverfahrens über zwei Schichten. Hierbei werden zwei Randverteilungen nacheinander iterativ angepasst. Zuerst werden die Gewichte gemäß der Verteilung von Schicht i berechnet. Darauffolgend werden die Gewichte gemäß der Verteilung von Schicht j berechnet, wobei das Ergebnis aus dem 1. Rand die Eingangsverteilung für diesen 2. Rand bedeutet.

$$n_{ij} \Rightarrow m'_{ij} \Rightarrow m''_{ij} \Rightarrow m'''_{ij} \Rightarrow m''''_{ij} \Rightarrow \text{usw.}$$

1. Rand über Schicht i (gemäß Gleichung 52 in Deming, Stephan (1940))

$$m'_{ij} = n_{ij} * \frac{n * N_{i.}}{n_{i.} * N}$$

2. Ev. Kappung von Faktoren höher als bestimmte Grenze

3. Rand über Schicht j (gemäß Gleichung 53 in Deming, Stephan (1940))

$$m''_{ij} = m'_{ij} * \frac{n * N_{.j}}{n_{.j} * N}$$

4. Ev. Kappung von Faktoren höher als bestimmte Grenze

wobei:

n_{ij} = Anzahl der Fälle in Zelle ij (ev. designgewichteter) in der Stichprobe

¹³ Vgl. Cochran, W.G. (1968): The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*. 24. 295-313.
Deming, W.E., Stephan, F.F. (1940): On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Annals of Mathematical Statistics*. 11. 427-444.

m'_{ij} resp. m''_{ij} = Anzahl der Fälle in der Zelle ij nach jeweiliger Randanpassung in der Stichprobe

i bzw. j = Schicht i bzw. j , i und j laufen über die Ausprägungen in der jeweiligen Schicht: i kann zum Beispiel das Alter in 4 Ausprägungen und j die Bildung in 3 Ausprägungen sein.

n = Anzahl der Fälle insgesamt der Stichprobe

n_i = Anzahl der Fälle in Schicht i in der Stichprobe, ungeachtet von Schicht j (analog: $n_{.j}$)

N = Anzahl der Fälle in der Grundgesamtheit

N_i = Anzahl der Fälle in Schicht i in der Grundgesamtheit, ungeachtet von Schicht j (analog: $N_{.j}$)

Nun beginnt die Iteration: das Resultat aus der Anpassung an Schicht j (also Verteilung nach dem 2. Rand) gilt als Eingangsverteilung für eine erneute Berechnung gemäß der Verteilung von Schicht i und wiederum nachfolgend Schicht j usw. Die Iteration wird so lange wiederholt, bis die Anpassungen zufriedenstellend sind bzw. keine Verbesserung der Anpassung zu erkennen ist. Bei der Iteration über weitere Schichten wird jeweils analog verfahren.

Zusammenfassend ist festzuhalten: Im Rahmen der Kalibrierung zur Erstellung der Gewichtungsfaktoren wird systematisch geprüft, ob und in welchem Umfang disproportionale Ausfälle auftreten. Dazu wird die realisierte Nettostichprobe mit den Verteilungen der Grundgesamtheit verglichen. Nach Festlegung der in Bezug auf eine Anpassung relevanten Merkmale wird das Gewichtungsmodell mit mehrdimensionalen Randverteilungen realisiert. Es wird ein iterativer Gewichtungsprozess aufgesetzt, dessen Ergebnis Gewichtungsfaktoren sind, die die realisierte Stichprobe an alle vorgegebenen Sollverteilungen mit vordefinierter Genauigkeit und minimaler Varianz anpassen.

Die Ausfallgewichtung erfolgte nach denselben Strukturmerkmalen wie bei der Gewichtung der Erwerbstätigenbefragung 2012. Aufgrund des geänderten Screeningverfahrens (vgl. Abschnitt 3.2) erfolgte die Ausfallgewichtung 2018 direkt auf die Strukturen der Erwerbstätigen ab 15 Jahren, während 2012 in einem vorgelagerten Schritt auf die Gesamtbevölkerung ab 15 Jahren gewichtet wurde.

Zur Ermittlung des endgültigen Gewichtungsmodells wurden – wie in 2012 – zwei Varianten gerechnet, wobei das Merkmal „Stellung im Beruf“ (StiB) bei Variante 1 nicht berücksichtigt und in Variante 2 berücksichtigt wurde. Die Referenzstrukturen bei der Gewichtung lieferten Sonderauswertungen des Mikrozensus 2017.¹⁴

14 An dieser Stelle möchten wir uns herzlich bei Destatis für die zur Verfügung gestellten Sonderauswertungen und beim BIBB für das Beschaffen der benötigten Zählungen bedanken.

Tabelle 6: Ränderkombinationen der Gewichtungsmodelle (Referenzstruktur: Mikrozensus2017)

Ränderkombinationen ¹⁾	Variante 1	Variante 2
Bundesland x 10er BIK	x	x
Bundesland x Haushaltsgröße	x	x
West/Ost x Haushaltsgröße	x	x
Stellung im Beruf		x
West/Ost x Geschlecht x Nationalität	x	x
Familienstand x Geschlecht x Alter (6 Gruppen)	x	x
West/Ost x Bildung x Alter (15-50 Jahre, 51+) ²⁾	x	x
Bundesland x Geschlecht x Alter (6 Gruppen)	x	x
Geschlecht x Alter (6 Gruppen)	x	x
Bundesland x Geschlecht	x	x
Fallzahl	20.012	20.012

1) Die Merkmalsausprägungen der Variablen sind Tabelle 7 zu entnehmen.

2) Hier wurde eine Dichotomisierung verwendet. Hintergrund ist, dass im Mikrozensus die Angabe zur Schulbildung freiwillig ist, wenn die befragte Person mindestens 51 Jahre alt ist. Dementsprechend ist für diese Befragten ein deutlich höherer Anteil fehlender Angaben festzustellen.

7.3 Qualitätsindikatoren der Gewichtung

Im Folgenden werden die Qualität der beiden Gewichtungsmodelle und die Güte der Stichprobe betrachtet. Zunächst werden die Merkmale, die in die Gewichtung einfließen, in ihrer Verteilung in der Referenzstatistik und in den beiden Gewichtungsmodellen dargestellt sowie die „Strukturinkongruenz“ betrachtet. Die Strukturinkongruenz wird hier gemessen als die Summe der absoluten Abweichungen zwischen Soll- und Stichprobenstruktur nach Gewichtung für die in Tabelle 7 dargestellten Merkmale. Anzumerken ist, dass die Abweichung noch geringer ausfallen würde, wenn man jeweils nur die gültigen Angaben heranziehen und darauf prozentuieren würde. Dies liegt am Vorgehen bei der Entwicklung der Gewichtungsfaktoren, welches darauf ausgelegt ist, die Verteilungen bezüglich der gültigen Angaben anzupassen.

Tabelle 7: Vergleich zwischen Stichproben- und Referenzstruktur (Mikrozensus2017)

		Ist-Struktur gewichtet	
Strukturmerkmal	MZ 2017	Variante 1: ohne StiB	Variante 2: mit StiB
Geschlecht			
Männlich	54,7	54,7	54,7
Weiblich	45,3	45,3	45,3
Familienstand			
Ledig	34,1	34,0	34,0
Verheiratet	55,3	54,9	55,0
Verwitwet	1,5	1,5	1,5
Geschieden	9,1	9,2	9,2
Angabe fehlt		0,4	0,4
Deutsch - Nicht-Deutsch			
Nicht-deutsch	11,2	11,2	11,2
Deutsch	88,8	88,7	88,7
Angabe fehlt		0,1	0,1
Alter			
15 bis 20	1,1	1,1	1,1
21 bis 30	16,5	16,3	16,3
31 bis 40	22,2	22,0	22,0
41 bis 50	25,5	25,2	25,2
51 bis 64	32,7	32,3	32,3
65 und älter	2,0	1,9	1,9
Angabe fehlt		1,0	1,0
Stellung im Beruf			
Arbeiter	18,4	15,9	18,4
Angestellte	65,7	69,5	65,6
Beamte	5,2	5,8	5,2
Selbständig	10,4	8,3	10,3
Mith. Familienangeh.	0,3	0,3	0,3
Angabe fehlt		0,1	0,1
Höchster Schulabschluss			
Maximal Volks-/Hauptschule	25,4	25,4	25,4
Mittlere Reife/POS	34,8	34,7	34,7
Fachhochschulreife, Abitur	39,5	39,5	39,5
Angabe fehlt	0,3	0,4	0,4

		Ist-Struktur gewichtet	
Strukturmerkmal	MZ 2017	Variante 1: ohne StiB	Variante 2: mit StiB
Bundesland, Wohnort			
Schleswig-Holstein	3,4	3,4	3,4
Hamburg	2,4	2,4	2,4
Niedersachsen	9,4	9,4	9,4
Bremen	0,8	0,8	0,8
Nordrhein-Westfalen	20,6	20,6	20,6
Hessen	7,5	7,5	7,5
Rheinland-Pfalz	4,8	4,8	4,8
Baden-Württemberg	13,8	13,8	13,8
Bayern	16,8	16,8	16,8
Saarland	1,2	1,2	1,2
Berlin	4,5	4,5	4,5
Brandenburg	3,1	3,1	3,1
Mecklenburg-Vorpommern	1,9	1,9	1,9
Sachsen	4,9	4,9	4,9
Sachsen-Anhalt	2,6	2,6	2,6
Thüringen	2,6	2,6	2,6

Zur Beurteilung der Qualität der Gewichtungsmodelle werden die folgenden Indikatoren herangezogen:

- **Effektivität:**
Diese wird berechnet als das Verhältnis der Stichprobenvarianz unter Berücksichtigung des aktuellen komplexen Designs zur Stichprobenvarianz einer einfachen Zufallsstichprobe derselben Größe. Die Effektivität wird in der Regel mit der Designgewichtung berechnet und kann damit bei gegebener Stichprobengröße n zur Berechnung der effektiven Stichprobengröße n' herangezogen werden. Sie nimmt im optimalen Fall den Wert 1 bzw. 100% an. Je niedriger ihr Wert ist, umso stärker ist die Varianzzunahme durch die Gewichtung. Für die Folgerungen auf die Grundgesamtheit kann man trotz des tatsächlich realisierten Stichprobenumfangs n nur von der effektiven Stichprobengröße n' ausgehen. Dementsprechend größer sind die Konfidenzintervalle anzusetzen.
Die Effektivität der Stichprobe nach der Anpassungsgewichtung (Kalibration) ergibt zusätzlich ein Maß für die weitere Zunahme der Varianz der Gewichtungsfaktoren, allerdings ohne weitere zusätzliche Reduzierung des effektiven Stichprobenumfangs.
- **Standardabweichung der Gewichtungsfaktoren:**
Dies ist ein Maß für die Variation der Gewichtungsfaktoren. Sie sollte möglichst klein sein.
- **Strukturinkongruenz:**
Hier wird die Summe der absoluten Abweichungen zwischen Soll- und Stichprobenstruktur nach Gewichtung gemessen.

Tabelle 8: Qualitätsindikatoren der Gewichtungsmodelle für die Erwerbstätigenstichprobe (Referenzstruktur: Mikrozensus2017)

Qualitätsindikatoren	Variante 1 - ohne StiB	Variante 2 - mit StiB
Fallzahl ungewichtet	20.012	20.012
Fallzahl gewichtet	20.012	20.012
Effektivität nach Designgewichtung	64,4%	64,4%
Effektivität insgesamt	49,5%	49,4%
Standardabweichung	1,009	1,013
Strukturinkongruenz gesamt	16,6	7,6
Strukturinkongruenz im Durchschnitt	0,3	0,1

Die Designgewichtung wird vor der Kalibration durchgeführt und dient – wie beschrieben – als Eingangsgewicht für die weiteren Schritte. Demnach ist die Effektivität der Designgewichtung für beide Varianten gleich. Mit rund 64 % ist sie etwas niedriger als in der Erhebung 2012, dies liegt in erster Linie am neuen Dual-Frame-Ansatz. Bei diesem ist das Designgewichtungsmodell wesentlich komplexer als bei einer reinen Festnetzstichprobe, u. a. wird neben der Anzahl der Festnetznummern auch die Anzahl der Mobilfunknummern für den zugrundeliegenden Algorithmus verwendet. Dies führt dazu, dass die Effektivität etwas niedriger ist. Allerdings ist die Stichprobenqualität bei einem Dual-Frame-Ansatz besser, da durch die Mobilfunkinterviews mehr Personen der Grundgesamtheit erreicht werden können.

Die anschließende Kalibrationsgewichtung verringert in der Regel das Konfidenzintervall der Messvariablen und verbessert die Güte einer Stichprobe. Diese hat jedoch nur Einfluss auf die anschließende Varianz der Gewichtungsfaktoren, jedoch nicht auf den effektiven Stichprobenumfang. Eine „feiner“ angelegte Kalibrationsgewichtung erhöht in der Regel die Güte der Stichprobe, ohne jedoch den effektiven Stichprobenumfang weiter zu verringern. Aus diesem Grund sind bei der vorliegenden Untersuchung mit einer entsprechend hohen Fallzahl nicht nur ausgewählte Gewichtungsränder, sondern auch -zellen angepasst worden.

Variante 2 wurde 2012 als die finale Gewichtungsvariante verwendet und ist auch 2018 unsere Empfehlung. Die Effektivität sowie die Standardabweichung unterscheiden sich bei beiden Gewichtungsmodellen nur geringfügig. Dagegen sind die Strukturen bei dem Merkmal „Stellung im Beruf“ deutlich verbessert und die Strukturinkongruenz dementsprechend niedriger.