

Bericht

Repräsentative Bevölkerungsbefragung im Rahmen des BaSiD-Teilvorhabens „Sicherheitsgefährdungen durch Kriminalität“

Methodenbericht

Autoren: Stefan Schiel, Christian Dickmann, Reiner Gilberg, Aneta Malina

infas Institut für angewandte
Sozialwissenschaft GmbH

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

Bericht an

Bundeskriminalamt
Äppelallee 45
65173 Wiesbaden

Projekt

4326
Bonn, April 2013
SI

Vorgelegt von

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

Kontakt

Stefan Schiel
Senior-Projektleiter Sozialforschung

Tel. +49 (0)228/38 22-424
Fax +49 (0)228/310071
E-Mail s.schiel@infas.de

Autoren

Stefan Schiel, Christian Dickmann, Reiner Gilberg, Aneta Malina

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

Vorbemerkung

Im Verbundprojekt „Sicherheiten, Wahrnehmungen, Lagebilder, Bedingungen und Erwartungen – Ein Monitoring zum Thema Sicherheit in Deutschland“ (kurz: „Barometer Sicherheit in Deutschland“ – BaSiD) bearbeiten unterschiedliche Institutionen das Themenfeld Sicherheit. In einem Teilprojekt steht ein Beitrag zur Dunkelfeldforschung im Fokus der Untersuchung. Unter der Federführung des Bundeskriminalamts (BKA) und in enger Zusammenarbeit mit dem Freiburger Max-Planck-Institut für ausländisches und internationales Strafrecht (MPI) wurde eine bundesweite, repräsentative CATI-Befragung über Viktimisierungserfahrungen der Bevölkerung konzipiert. Erstmals wurden rund 35.500 Männer und Frauen in Deutschland zu ihren Opfererfahrungen befragt. Mit der Durchführung der Erhebung wurde das infas Institut für angewandte Sozialwissenschaft beauftragt.

Das BKA und das MPI haben nach der Auftragsvergabe einen sehr umfangreichen Fragebogenentwurf vorgelegt. Dieser wurde zusammen mit infas in einer Vielzahl von Abstimmungsrunden überarbeitet und weiterentwickelt. Nach der intensiven Überarbeitungsphase wurde der Fragebogen in mehreren qualitativen und quantitativen Pretests auf seine Feldfähigkeit hin geprüft. Im Saldo liegt ein ausgiebig getestetes Instrument vor. Insbesondere die Screenerfragen, die für die spätere Hochrechnung der Prävalenzraten zentral sind, wurden in den Pretests intensiv getestet. Alle Testphasen sind in den drei Erhebungssprachen Deutsch, Türkisch und Russisch vollzogen worden.

Der vorliegende Methodenbericht dokumentiert das Erhebungsdesign, die Stichprobenziehung, die Feldarbeit der Hauptstudie sowie die Prüfung und Gewichtung der erhobenen Daten. Die Vorgehensweise in den vorgeschalteten Pretests ist in einem gesonderten Bericht enthalten.

inf Institut für angewandte Sozialwissenschaft GmbH

Inhaltsverzeichnis

1 Erhebungsdesign	7
1.1 Grundgesamtheit	9
1.2 Stichprobenziehung	9
1.2.1 Basisstichprobe	9
1.2.2 Aufstockungsstichprobe	11
1.2.3 Zusatzstichprobe nach onomastischem Verfahren	12
1.3 Das Erhebungsinstrument	14
2 Durchführung der Hauptstudie	19
2.1 Feldzeit	19
2.2 Interviewerschulung	19
2.3 Interviewereinsatz	23
2.4 Supervision zur Qualitätssicherung	25
2.5 Feldsteuerung	26
2.6 Adressnachbearbeitung zur Erhöhung der Ausschöpfung	27
2.7 Stichprobenausschöpfung	28
2.7.1 Basisstichprobe	29
2.7.2 Aufstockung	32
2.7.3 Onomastische Stichprobe	34
2.8 Interviewdauer	37
2.9 Interviewsprache	37
2.10 Kontakthäufigkeit	38
2.11 Fallzahl pro Fragemodul	39
2.12 Einsatz der Vignettensets	40
2.13 Verhältnis Interviews aus der Festnetz- und der Mobilfunkstichprobe in der realisierten Stichprobe	40
2.14 Interviewabbrüche und Abbruchgründe	41
3 Datenprüfung und –gewichtung	45
3.1 Datenprüfung und –aufbereitung	45
3.2 Gewichtung und Hochrechnung	50
3.2.1 Bestimmung der Auswahlwahrscheinlichkeit für die Dual-Frame-Basisstichprobe und die Aufstockerstichproben	51
3.2.2 Integration von Basisstichprobe und Aufstockerstichproben	53
3.2.3 Proportionalisierung des Dual-Use-Anteils	54
3.2.4 Auswahlwahrscheinlichkeit für die Onomastikstichprobe	55
3.2.5 Integration der Onomastikstichprobe	55
3.2.6 Trimmen der Designgewichte	56
3.2.7 Kalibrierung und Hochrechnung	57
3.2.8 Berechnung und ggf. Kalibrierung für die einzelnen Bundesländer	64
4 Literaturverzeichnis	66

Abbildungsverzeichnis

Abbildung 1	Struktur des Fragebogens	16
Abbildung 2	Inhalte des Interviewerhandbuchs	22

Tabellenverzeichnis

Tabelle 1	Synopse	8
Tabelle 2	Aufstockung der Bundesländer	12
Tabelle 3	Regionale Verteilung von Personen mit türkischem Migrationshintergrund im engeren Sinne nach Bundesländern in der Grundgesamtheit und der Bruttostichprobe der onomastischen Stichprobe	13
Tabelle 4	Angezielte Fallzahl pro Modul in der Basisstichprobe	17
Tabelle 5	Fragebogeninhalte	18
Tabelle 6	Telefoninterviewer nach Geschlecht	23
Tabelle 7	Telefoninterviewer nach Alter (gruppiert)	23
Tabelle 8	Realisierte Interviews pro Interviewer	24
Tabelle 9	Anzahl realisierte Interviews pro Interviewer (gruppiert)	24
Tabelle 10	Feldergebnis der Basisstichprobe nach AAPOR-Klassifikation	30
Tabelle 11	Final Outcome nach AAPOR, differenziert nach Teilstichproben Festnetz und Mobilfunk	32
Tabelle 12	Feldergebnis der Aufstockungsstichprobe nach AAPOR-Klassifikation	33
Tabelle 13	Verteilung der auswertbaren Interviews auf die fünf Bundesländer	34
Tabelle 14	Feldergebnis der onomastischen Stichprobe nach AAPOR-Klassifikation	36
Tabelle 15	Interviewdauer	37
Tabelle 16	Basis- und Aufstockungsstichprobe: Interviewsprache (Prozent in Klammern)	38
Tabelle 17	Onomastische Stichprobe: Interviewsprache	38
Tabelle 18	Kontakthäufigkeit der auswertbaren Interviews	39
Tabelle 19	Fallzahl pro Modul in der Basisstichprobe	39
Tabelle 20	Vignettensetverteilung	40
Tabelle 21	Festnetz- und Mobilfunkstichprobe in der Basisstichprobe	40
Tabelle 22	Festnetz- und Mobilfunkstichprobe insgesamt	41
Tabelle 23	Abbruchstellen im Fragebogen	44
Tabelle 24	Vercodung der offenen Ortsangaben	48
Tabelle 25	Prozentualer Verteilungsvergleich ungewichtete, designgewichtete und kalibrierte Stichprobe mit dem Mikrozensus für die Grundgesamtheit (Personen)	59
Tabelle 26	Kennwerte der kalibrierten Gewichte und Hochrechnungsfaktoren	62
Tabelle 27	Prozentualer Verteilungsvergleich ungewichtete, designgewichtete und kalibrierte Stichprobe mit dem Mikrozensus für die Grundgesamtheit (Haushalte)	63
Tabelle 28	Kennwerte der kalibrierten Gewichte und Hochrechnungsfaktoren	64

1 Erhebungsdesign

In einem interdisziplinären Verbundprojekt „Barometer Sicherheit in Deutschland“ (kurz: BaSiD) bearbeiten unterschiedliche Institutionen arbeitsteilig das Themenfeld Sicherheit. Mithilfe von insgesamt neun Forschungsmodulen soll ein Bild der objektiven Sicherheitslage wie auch der subjektiv wahrgenommenen Sicherheit in Deutschland erstellt werden.

Im Fokus des vierten Forschungsmoduls steht ein Beitrag zur Dunkelfeldforschung. Unter der Federführung des Bundeskriminalamts (BKA) und in enger Zusammenarbeit mit dem Freiburger Max-Planck-Institut für ausländisches und internationales Strafrecht (MPI) wurde eine bundesweite, repräsentative CATI-Befragung über Viktimisierungserfahrungen der Bevölkerung konzipiert. Mittels Telefonbefragung wurden erstmals in rund 35.500 Interviews Männer und Frauen nach ihren Opfererfahrungen befragt. Die ermittelten Prävalenzraten lassen sich auf die Bundesrepublik hochrechnen. Mit dieser bundesweiten Dunkelfelderhebung wird eine Forschungslücke geschlossen und an internationale Standards aufgeschlossen. Die Befragung ist als Querschnitterhebung angelegt.

Die Datenerhebung erfolgte mittels computerunterstützter Telefoninterviews (CATI) zentral aus dem infas-Telefonstudio in Bonn. Die Interviews wurden in den Sprachen Deutsch, Türkisch und Russisch durchgeführt. Im Vorfeld der Haupterhebung wurde das Fragenprogramm in einem Zwei-Phasen-Pretest intensiv auf seine Feldfähigkeit getestet. Die Durchführung der Tests ist in einem Pretestbericht¹ dokumentiert.

Schwerpunkte der Befragung sind die Themen Kriminalitätsfurcht, Opfererfahrungen, Anzeigeverhalten, Betrugsdelikte mit EC- und Kreditkarte oder im Internet sowie Gewalterfahrungen. Zudem wurde ein Fragenblock zur Soziodemografie eingesetzt.

Es wurden Personen in Privathaushalten im Alter ab 16 Jahre befragt. Die Gesamtstichprobe für die Haupterhebung setzt sich aus drei Teilstichproben zusammen. Dabei handelt es sich um:

- eine Basisstichprobe (hier sollten n=30.000 Personen in einem Dual-Frame-Ansatz (Mix aus Festnetz- und Mobilfunkstichprobe) befragt werden),
- eine Aufstockungsstichprobe (hier sollten zusätzlich in fünf Bundesländern insgesamt n=4.348 Personen befragt werden) und

¹ Schiel, Stefan, Aneta Malina, Helmut Schröder und Markus Banz (2012): „Repräsentative Bevölkerungsbefragung im Rahmen des BaSiD-Teilvorhabens „Sicherheitsgefährdungen durch Kriminalität“. Pretestbericht“. Bonn Häder, S. und S. Gabler (1998): Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. In Gabler S. , S. Häder und J.H.P. Hoffmeyer-Zlotnik (Hrsg.) Telefonstichproben in Deutschland. Opladen, Westdeutscher Verlag.

- eine Zusatzstichprobe, die nach onomastischem Verfahren (im Weiteren als „onomastische Stichprobe“ bezeichnet) gezogen wurde. Hier sollten möglichst n=1.000 Personen befragt werden, um den Anteil von Personen mit türkischem Migrationshintergrund in der Gesamtstichprobe zu erhöhen.

Während der gesamten Feldphase stand allen Zielpersonen für telefonische Rückfragen eine Mitarbeiterin im MPI und bei infas zur Verfügung. Weitere Informationsmöglichkeiten waren auf der infas-Homepage bereitgestellt.

Die nachfolgende Synopse gibt einen Überblick über das Studiendesign.

Tabelle 1 Synopse

Studientitel	„Lebenssituation und Sicherheit in Deutschland“
Grundgesamtheit	Bevölkerung in Privathaushalten ab 16 Jahre in Deutschland
Erhebungsmethode	computergestützte Telefoninterviews (CATI)
Erhebungsinstrument	CATI-Fragebogen
Erhebungssprachen	Deutsch, Türkisch und Russisch
Pretest	<ul style="list-style-type: none"> – kognitive Pretests in Deutsch, Russisch und Türkisch (n=20 Interviews) – CATI-Pretests in Deutsch, Russisch und Türkisch (n=249)
Stichprobenansatz	<p><u>Basisstichprobe:</u></p> <ul style="list-style-type: none"> – Dual-Frame-Designs, angezieltes Verhältnis zwischen Festnetzstichprobe und Mobilfunkstichprobe 23.000 (76,7%) zu 7.000 (23,3%) Interviews. jeweils zufällig generierte Rufnummern – Ziel n=30.000 Interviews <p><u>Aufstockungsstichprobe für fünf ausgewählte Bundesländer:</u></p> <ul style="list-style-type: none"> – nur Festnetz, jeweils zufällig generierte Rufnummern – Ziel n=4.348 Interviews <p><u>Onomastische Stichprobe:</u></p> <ul style="list-style-type: none"> – Listenauswahl unter Anwendung onomastischer Verfahren – Türkisch, nur Festnetz – Ziel n=1.000 Interviews
Feldzeit Hauptfeld	<p>Basisstichprobe: 11.06.2012 bis 24.10.2012</p> <p>Aufstockungsstichprobe: 5.7. 2012 bis 24.10.2012</p> <p>onomastische Stichprobe: 5.9.2012 bis 3.11.2012</p>
Auswahl der Zielperson	<p>Festnetz: Last-Birthday-Schlüssel</p> <p>Mobilfunk: Hauptnutzer des Telefons</p>
Interviewerschulung	mündliche Schulung und Einsatz eines schriftlichen Interviewerhandbuchs
Auswertbare Interviews	<p>Basisstichprobe: n=30.278</p> <p>Aufstockungsstichprobe: n=4.417</p> <p>onomastische Stichprobe: n=808</p>
Gesamtdauer CATI - Interviews	durchschnittlich 19,6 Minuten
Rückversicherungsmöglichkeit	telefonische Ansprechpartner bei infas und MPI, infas-Homepage
Datenlieferung	<p>gelabelter Gesamtdatensatz: 13.11.2012</p> <p>Gewichte und Hochrechnungsfaktoren: 03.12.2012</p> <p>Tabellenband: 19.12.2012</p>

1.1 Grundgesamtheit

Die Grundgesamtheit der Erhebung setzt sich aus der in Privathaushalten lebenden Wohnbevölkerung in Deutschland zusammen, die zum Zeitpunkt des Interviews mindestens 16 Jahre alt sind und über ausreichende Deutsch-, Türkisch-, oder Russischkenntnisse verfügt.

1.2 Stichprobenziehung

Es kamen drei Teilstichproben zum Einsatz. Eine bundeweite Festnetz- und Mobilfunkstichprobe (Basisstichprobe), eine Aufstockungsstichprobe für ausgewählte Bundesländer sowie eine zusätzliche Stichprobe, um den Anteil Personen mit türkischem Migrationshintergrund zu erhöhen.

1.2.1 Basisstichprobe

Der Studie liegt eine Telefonstichprobe nach dem Häder/Gabler-Verfahren, dem Standardverfahren der ADM²-Institute, zu Grunde³. Bei diesem Verfahren werden alle im Feld eingesetzten Telefonnummern synthetisch generiert. Dieses Vorgehen ist notwendig, da allgemein zugängliche Verzeichnisse wie Telefonbücher, nur einen eingeschränkten Auswahlrahmen bieten. Aufgrund verschiedener Studien muss davon ausgegangen werden, dass nicht alle Personen in allgemein zugänglichen Registern gelistet sind. Beispielsweise waren in Deutschland bereits Ende der 90er Jahre 15-25 Prozent der westdeutschen Haushalte mit einem Festnetzanschluss nicht in Telefonbüchern gelistet. In den ostdeutschen Ländern lag dieser Anteil aufgrund der im Schnitt jüngeren Telefonanschlüsse mit 25-30 Prozent insgesamt sehr viel höher. Am höchsten ist der Anteil nicht gelisteter Festnetzanschlüsse in Großstädten⁴. Darüber hinaus sind jüngere Leute, Single-Haushalte und Personen mit niedrigem Bildungsniveau häufiger nicht im Telefonbuch eingetragen als ältere Personen oder Mehrpersonenhaushalte. Neben gelisteten (in Telefonregistern eingetragenen) Nummern sind daher zufällig generierte Nummern bei der Stichprobenziehung zu berücksichtigen bzw. mit entsprechenden Inklusionswahrscheinlichkeiten zu versehen. In Deutschland wurde in diesem Zusammenhang ein Lösungsvorschlag von Häder/Gabler entwickelt, der in der Praxis in großem Maßstab erprobt ist und mittlerweile zur gängigen Praxis zählt. Die Grundidee wird im Folgenden kurz skizziert:

² ADM = Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.

³ Häder, S. und S. Gabler (1998): Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. In Gabler S., S. Häder und J.H.P. Hoffmeyer-Zlotnik (Hrsg.) Telefonstichproben in Deutschland. Opladen, Westdeutscher Verlag.

⁴ Follmer R., Smid, M.: „Nichteingetragene Telefonnummern: Ergebnisse eines Methodentests“ in: Gabler, S., S. Häder und J.H.P. Hoffmeyer-Zlotnik (Hrsg.) (1998): Telefonstichproben in Deutschland. Opladen, Westdeutscher Verlag, S.43-57.

Ausgangspunkt für die Erzeugung der Auswahlgesamtheit von möglichen Telefonnummern sind allgemein zugängliche Verzeichnisse von Festnetztelefonnummern der Bundesrepublik. Von den in Telefonnummernverzeichnissen gelisteten Telefonnummern werden die beiden letzten Ziffern entfernt und der so entstandene Nummernstamm mit allen zweistelligen Endziffern von 00 bis 99 ergänzt, so dass sich aus jeder eingetragenen Nummer ein sogenannter „Nummernblock“ von 100 aufeinander folgenden potentiellen Telefonanschlüssen ergibt. Rufnummern mit demselben Nummernstamm erzeugen denselben Block, der aber nur einmal gespeichert wird. Zusätzlich werden diese Rufnummernblöcke um die Liste der Rufnummernblöcke der Bundesnetzagentur erweitert. Ist dort ein Nummernblock als vergeben gelistet, dann werden dort auch Rufnummern generiert, wenn es keinen Telefonbucheintrag gibt.

Die Menge aller so erzeugten Blöcke bildet die Auswahlgrundlage, aus der anschließend zufällig einzelne Nummern gezogen werden. Aufgrund dieser Anordnung der Auswahlgesamtheit hat jede darin enthaltene Rufnummer unabhängig von ihrer Länge dieselbe Auswahlchance. Da über die Vorwahl und die ersten Stellen des Rufnummernblocks eine kleinräumige regionale Zuordnung der generierten Nummern möglich ist, kann eine nach regionalen und regionalstrukturellen Merkmalen geschichtete Zufallsauswahl für die Festnetzstichprobe vorgenommen werden, so dass eine gute Streuung über verschieden große Gemeinden und Gemeindetypen sichergestellt werden kann.

Ein neueres Problem stellen Haushalte dar, die nur noch einen Mobilfunkanschluss haben und über keinen Festnetzanschluss mehr verfügen (sogenannte mobile-onlys). Sie haben nach der oben beschriebenen Logik der Stichprobenziehung aus der Menge aller Festnetznummern keine Inklusionswahrscheinlichkeit, wären also aus der Stichprobenziehung ausgeschlossen (Coverage-Problem). Das Problem ist, dass nach neueren Erhebungen etwa 13 Prozent der Bevölkerung (vermutlich 18 Prozent der Haushalte)⁵ nur noch über Mobilfunknummern erreichbar sind. Diese Haushalte unterscheiden sich zudem in ihrer Zusammensetzung stark von denjenigen Haushalten, die nach wie vor über einen Festnetzanschluss erreichbar sind. Es handelt sich demnach um ein systematisches Coverage-Problem.

Vor diesem Hintergrund wurde zusätzlich eine Mobilfunkstichprobe aus synthetisch generierten Mobilfunknummern gezogen. Diese Vorgehensweise, die Ziehung einer Stichprobe aus zwei unterschiedlichen Auswahlrahmen, wird als „Dual-Frame-Ansatz“ bezeichnet. Dies bedeutet, dass zwei unterschiedliche Stichprobenansätze (einer für Nummern mit Festnetzanschluss, ein anderer für die Mobilnummern) miteinander verbunden werden.

⁵ ADM (2012), ADM-Forschungsprojekt „Dual-Frame-Ansätze“ 2011/2012, Forschungsbericht.

Auch bei den Mobilfunknummern handelt es sich um generierte Nummern, da nur ein sehr kleiner Teil der Mobilfunknummern in einem Verzeichnis gelistet ist. Ausgangspunkt für die Erzeugung der Auswahlgesamtheit der Mobilfunknummern bildet die Liste der von der Bundesnetzagentur vergebenen Nummernblöcke. Anders als im Auswahlrahmen für Festnetznummern sind die Rufnummernblöcke im Mobilfunkbereich aber größer gefasst. Es werden nicht 100er-Blöcke generiert (also die letzten beiden Stellen der Rufnummer), sondern 100.000er-Blöcke (also die letzten sechs Stellen der Rufnummer). Dies ist im Mobilfunkbereich insofern unproblematisch, da hier, im Gegensatz zum Festnetzbereich, alle Rufnummern eine einheitliche Länge haben

Aus der Menge aller so erzeugten Mobilfunknummern können anschließend zufällig einzelne Nummern gezogen werden. Im Mobilfunkbereich ist allerdings keinerlei regionale Verortung über die Vorwahl und Rufnummer möglich, so dass keine regionale bzw. regionalstrukturelle Schichtung und auch keine regionale Selektion möglich sind.

Die Auswahlrahmen für Festnetz- und Mobilfunkstichproben wurden von der Arbeitsgemeinschaft ADM-Telefonstichproben⁶ zur Verfügung gestellt. Aus diesen beiden beschriebenen Auswahlrahmen wurde die Bruttostichprobe der Telefonnummern gezogen.

Sowohl in der Festnetzstichprobe als auch in der Mobilfunkstichprobe war aufgrund der beschriebenen Auswahlgrundlage aus generierten Nummern davon auszugehen, dass ein erheblicher Teil der Telefonnummern der Bruttostichprobe nicht geschaltet ist. Für die Berechnung der Bruttostichprobengröße wurde anhand der vorliegenden Erfahrung mit den ADM-Auswahlgrundlagen ein Anteil von 60 Prozent nicht geschalteter Nummern in beiden Stichproben unterstellt.

Die Bruttoeinsatzstichprobe für die Basisstichprobe bestand aus insgesamt 173.850 Mobilfunk- und 576.150 Festnetznummern. Das Mischungsverhältnis von Festnetznummern zu Mobilfunknummern in der Basisstichprobe lag somit bei 77 Prozent Festnetznummern zu 23 Prozent Mobilfunknummern. Die Bruttoeinsatzstichprobe wurde in mehrere Teilstichproben⁷ aufgeteilt und abhängig vom Feldverlauf in einzelnen Tranchen ins CATI-Feld gegeben.

1.2.2 Aufstockungsstichprobe

Um für die spätere Auswertungen auf Bundeslandebene ausreichend viele Fälle vorliegen zu haben, nutzten fünf Bundesländer (Hamburg, Berlin, Brandenburg, Sachsen und Hessen) die Möglichkeit, die Stichprobe für ihr Bundesland aufzustoocken. Für diese Bundesländer sollten insgesamt 4.348 Interviews realisiert

⁶ Die kontinuierliche Aktualisierung des Frames wird von der Arbeitsgemeinschaft „Stichproben“ des Arbeitskreises Deutscher Markt- und Sozialforschungsinstitute (ADM) e.V. organisiert. Zugriff haben nur Institute, die Mitglied dieser AG sind. infas ist Mitglied und hat entsprechend Zugriff auf die aktualisierte Grundgesamtheit.

⁷ Beim Einsatz der ersten Tranche wurde eine Ausschöpfung von rund 40 Prozent unterstellt.

werden. Die Ziehung der Telefonnummern erfolgte bundeslandweise aus dem beschriebenen Auswahlrahmen für ADM-Festnetzstichproben. Jedes der fünf Bundesländer erhielt somit eine eigene Aufstockungsstichprobe, die in Abhängigkeit vom Feldverlauf in einzelnen Tranchen ins Feld gegeben wurde. Die Fallzahlen sehen wie folgt aus:

Tabelle 2 Aufstockung der Bundesländer

Bundesland	Angezielte Interviews	Eingesetzte Telefonnummern
Hamburg	1.345	32.047
Berlin	721	17.171
Brandenburg	1.060	25.313
Sachsen	437	10.434
Hessen	785	18.625

Quelle: Stichprobendatei

Bei den eingesetzten Telefonnummern handelte es sich ausschließlich um Festnetznummern, da durch ihre Vorwahlen Gebietszuordnungen möglich sind. Mobilfunknummern hingegen sind über ihre Vorwahl nicht regionalisiert und können daher nur in bundesweiten Stichproben eingesetzt werden.

1.2.3 Zusatzstichprobe nach onomastischem Verfahren

Zusätzlich zu den beiden bereits beschriebenen Stichproben wurde eine Onomastikstichprobe⁸ für die Aufstockung von Personen mit türkischem Migrationshintergrund gezogen⁹.

Die Stichprobe wurde gezogen von „Dr. Andreas Humpert und Klaus Schneiderheinke GbR“, die über jahrelange Expertise bei der Ziehung von onomastischen Stichproben verfügen.¹⁰ Der Auswahlrahmen wurde dabei so gebildet, dass mithilfe der Onomastik den im aktuellen Telefonverzeichnis eingetragenen Privathaushalten (25,9 Mio. Eintragungen, Januar 2012) eine regionale Herkunft bzw.

⁸ Onomastik = Namenforschung, Namenkunde

⁹ Zur Mitte der Feldzeit lag der Anteil Personen mit türkischem Migrationshintergrund in der realisierten Stichprobe unterhalb des im Mikrozensus ausgewiesenen Anteilswerts. Vor diesem Hintergrund wurde entschieden, eine zusätzliche Stichprobe ins Feld zu geben, um den Anteil Personen mit türkischem Migrationshintergrund entsprechend ihrem Anteil in der Gesamtbevölkerung zu erhöhen. Ein ähnlicher Abgleich zwischen der realisierten Stichprobe und dem Mikrozensus wurde zur Mitte der Feldzeit für Personen aus der ehemaligen Sowjetunion vorgenommen. Die Anteilswerte in der realisierten Stichprobe und dem Mikrozensus waren nahezu identisch, weshalb für diese Migrantengruppe keine separate Stichprobe eingesetzt wurde.

¹⁰ Humpert, A. und K. Schneiderheinke (2002), Stichprobenziehung für telefonische Zuwandererumfragen – Praktische Erfahrungen und Erweiterung der Auswahlgrundlage, in: Gabler, S. und S. Häder (Hg.), Telefonstichproben - Methodische Innovationen und Anwendungen in Deutschland, Münster: Waxmann, S. 187-208. Humpert, A. und K. Schneiderheinke (2000), Stichprobenziehung für telefonische Zuwandererumfragen – Einsatzmöglichkeiten der Namenforschung (Onomastik), in: ZUMA-Nachrichten 47, S36-63.

ein spezifischer Migrationshintergrund zugeordnet wurde. Die Kennzeichnung basiert auf 2,0 Mio. Namen (Vor- und Nachnamen) aus über 536 Quellen der Onomastik. Die zentrale Zuordnungsdatei enthält 25,5 Mio. verschiedene Personennamen (Nachname, Vorname) mit einer Zuordnung der Herkunftsregion. Daraus ergaben sich 193.378 Festnetznummern von eingetragenen Personen, die mit hoher Wahrscheinlichkeit über einen türkischen Migrationshintergrund verfügen. Sie bildeten den Auswahlrahmen für die Stichprobenziehung.

Aus diesem Auswahlrahmen wurde eine regional geschichtete Zufallsstichprobe gezogen. Die regionale Schichtung nach Bundesländern erfolgte anhand der Verteilung von Personen mit türkischem Migrationshintergrund im engeren Sinne laut Angaben aus dem Mikrozensus 2010 (Statistisches Bundesamt, Mikrozensus 2010, Fachserie 1, Reihe 2.2).

Tabelle 3 Regionale Verteilung von Personen mit türkischem Migrationshintergrund im engeren Sinne nach Bundesländern in der Grundgesamtheit und der Bruttostichprobe der onomastischen Stichprobe

Bundesland	Grundgesamtheit (in %)	Bruttostichprobe (in %)
Schleswig-Holstein	2,2	2,2
Hamburg	3,4	3,4
Niedersachsen	7,1	7,1
Bremen	1,6	1,6
Nordrhein-Westfalen	33,1	33,1
Hessen	9,4	9,4
Rheinland-Pfalz	4,3	4,3
Baden-Württemberg	17,2	17,2
Bayern	12,8	12,8
Saarland	0,8	0,8
Berlin	7,2	7,2
Neue Länder zusammen	0,9	0,9

Quelle: Mikrozensus 2010, eigene Berechnung

Es wurden insgesamt 9.000 Telefonnummern für die Bruttoeinsatzstichprobe berücksichtigt. Um Überschneidungen der Telefonnummern zwischen der onomastischen und der Basisstichprobe sowie zwischen der onomastischen und den Aufstockungstichproben zu vermeiden, wurde ein Abgleich der Festnetznummern vorgenommen. Insgesamt gab es 55 Doppelungen, so dass 8.945 statt 9.000 Fälle aus der onomastischen Stichprobe ins CATI-Feld gegeben wurden.

Als Folge der geschilderten Auswahlprinzipien stehen für alle Teilstichproben lediglich Telefonnummern für den Feldeinsatz zur Verfügung, d.h. es liegen weder Namen noch Adressen vor. Die Zielpersonen wurden in den Festnetzstichproben anhand des Last-Birthday-Schlüssels und in der Mobilfunkstichprobe der Hauptnutzer des Telefons ausgewählt. Es bestand somit keine Möglichkeit, mithilfe eines Anschreibens über das Projekt zu informieren und die Betroffenen auf den Anruf vorzubereiten. Der Anruf durch die Interviewer erfolgt also „kalt“.

1.3 Das Erhebungsinstrument

Mit der Auftragsvergabe legte das BKA einen Fragebogenentwurf vor. Dieser Entwurf wurde in den folgenden Monaten in enger Abstimmung zwischen infas, BKA und MPI weiter entwickelt. Die Feldfähigkeit des Fragebogens ist in mehreren Pretests intensiv getestet worden¹¹. Anders als in den meisten sozialwissenschaftlichen Studien üblich, wurden auch die fremdsprachigen Erhebungsinstrumente einem qualitativen und einem quantitativen Pretest unterzogen. Es fanden 20 qualitative Interviews in Deutsch, Russisch und Türkisch statt. Diese kognitiven Pretests¹² dienten dazu, die intersubjektive Verständlichkeit und die Trennschärfe der Screenerfragen zur Ermittlung, ob die Zielperson Opfer einer Straftat geworden ist, zu prüfen. Zudem wurden insgesamt 249 quantitative Pretestinterviews ebenfalls in Deutsch (n=196), Russisch (n=27) und Türkisch (n=26) durchgeführt, um das Erhebungsinstrument unter realen Feldbedingungen zu testen.

Der Fragebogen gliedert sich in acht Bereiche (vgl. Abbildung 1, Bereich A bis H). Im Kontaktgespräch (Bereich A) wurde in den Festnetzstichproben bei Mehrpersonenhaushalten mittels Last-Birthday-Schlüssels die Zielperson zufällig ausgewählt. Bei Einpersonenhaushalten musste keine weitere Auswahl getroffen werden; der Bewohner des Haushalts war gleichzeitig auch Zielperson für das Interview. In der Mobilfunkstichprobe wurde der Hauptnutzer des Handys befragt. Da bei der Festnetzstichprobe nur Privatanschlüsse zur Grundgesamtheit zählen, wurde auch bei der Mobilfunkstichprobe analog vorgegangen, indem Handyanschlüsse, die ausschließlich beruflich genutzt werden, zum Ausschluss von der Befragung des Hauptnutzers führten.

Im Zuge der Zielpersonenauswahl in der Festnetzstichprobe wurde bereits im Kontaktgespräch die reduzierte Haushaltsgröße (Anzahl Personen ab 16 Jahre im Haushalt) erfragt. In der Mobilfunkstichprobe wurde die reduzierte Haushaltsgröße erst im Verlauf des Interviews ermittelt, da für die Auswahl der Zielperson die Haushaltsgröße irrelevant war.

Nachdem im Kontaktgespräch (Bereich A) die Zielperson ausgewählt wurde und sie ihre Bereitschaft erklärt hat, am Interview teilzunehmen, beginnt das inhaltliche Interview (ab Bereich B).

¹¹ Die Vorgehensweise bei den Pretest sind dokumentiert in : Schiel, Stefan, Aneta Malina, Helmut Schröder, Markus Banz (2012): „Repräsentative Bevölkerungsbefragung im Rahmen des BaSiD-Teilvorhabens „Sicherheitsgefährdungen durch Kriminalität“ Pretestbericht“, Bonn.

¹² Zur Funktion und Durchführung von kognitiven Pretest vgl. Prüfer, P. und M. Rexroth (2000): „Zwei-Phasen-Pretesting“ in ZUMA-Arbeitsbericht 2000/08. Mannheim. Prüfer, P. und M. Rexroth (2005): „Kognitive Interviews“ in ZUMA How-to-Reihe, Nr.15, Mannheim.

Die Fragen zu den Themen Lebenszufriedenheit, Gesundheit und Vertrauen wurden allen Zielpersonen gestellt. Im Anschluss daran wurden einem Teil der Zielpersonen Modulfragen gestellt. Es waren sieben Fragenmodule im Fragebogen implementiert (Nummern¹³ 1, 2, 4, 5, 6, 7 und 8, Fragenbereich C, vgl. Abbildung 1). Jedes Fragenmodul beinhaltet mehrere Fragen zu einem spezifischen Thema (zu den Inhalten der einzelnen Module siehe Tabelle 5 am Ende des Kapitels). Diese Fragen wurden jeweils nur einer Teilgruppe gestellt. Die Zuordnung einer Zielperson zu den Teilgruppen erfolgte zufallsgesteuert. Die zufällige Zuordnung, ob ein Modul und welches Modul für diesen Fall vorgelegt wurde, geschah über eine Kennung der Fallnummern in der Bruttostichprobe. Weder das Antwortverhalten der Zielperson noch die Interviewer hatten einen Einfluss auf die Auswahl der Modulfragen.

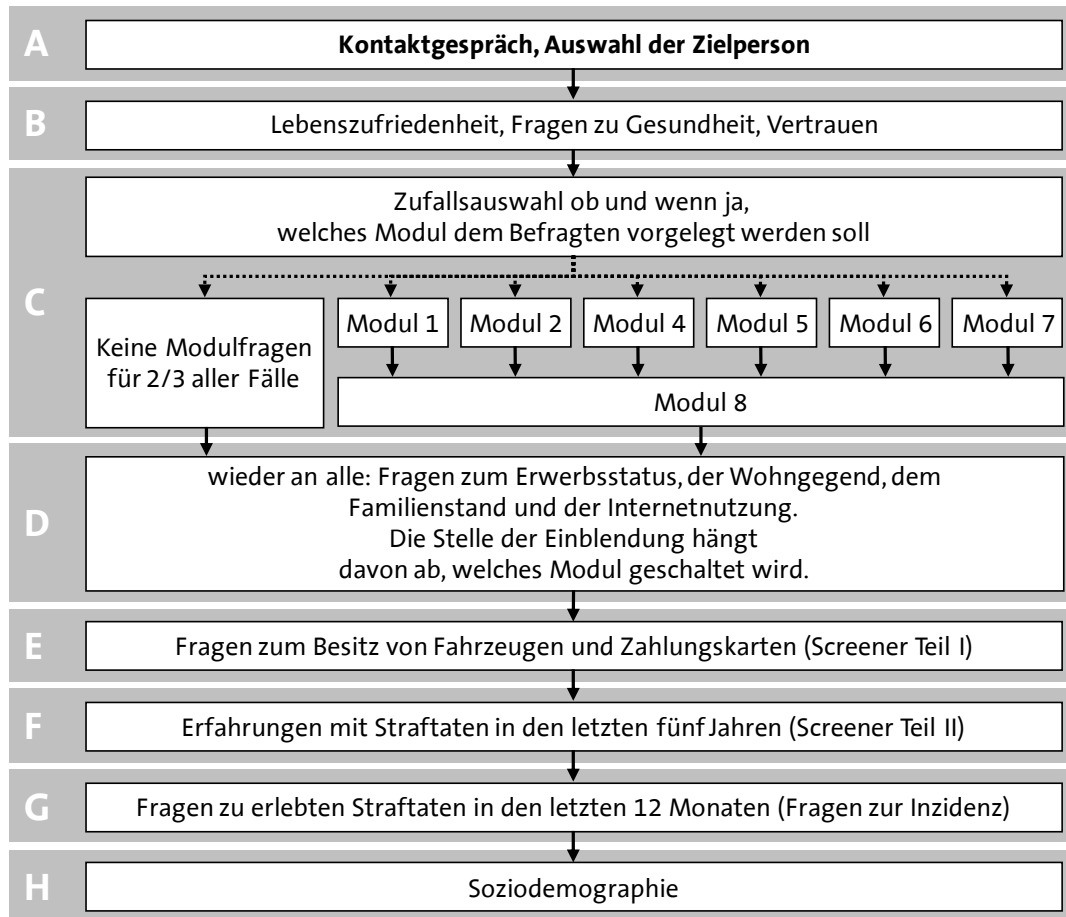
Im Modul 6 sollten u.a. Strafeinstellungen der Zielpersonen gemessen werden. Da Strafeinstellungen komplexe Entscheidungsprozesse zugrunde liegen, werden sie den Zielpersonen in empirischen Erhebungen meist in Form von inhaltlich variierenden Fallbeispielen Zielpersonen vorgelegt. Eine eindimensionale Abfrage zur Messung von Strafeinstellungen greift zu kurz und reicht daher nicht aus. Innerhalb des Moduls Nummer 6 waren unterschiedliche Fallbeispiele vorgesehen (sogenannte Vignetten¹⁴), die die Zielpersonen bewerten sollten. Insgesamt waren sieben verschiedene Vignettensets vorgesehen, die sich aus zwei bis drei verschiedenen Fallbeispielen zusammensetzten. Es erfolgte eine zufällige Zuweisung der Vignettensets zu einzelnen Fallnummern (Case-IDs) über die Bruttostichprobe. Weder das Antwortverhalten der Zielperson noch der Interviewer hatten einen Einfluss darauf, ob welches Vignettenset im Interview vorgelegt wurde.

Die Modulfragen des Moduls Nummer 8 wurden immer in Zusammenhang mit einem anderen Modul im Interview vorgelegt. Die übrigen sechs Module wurden alternativ eingesetzt, so dass entweder nur eines dieser Module in Kombination mit Modul 8 geschaltet wurde oder kein Modul im Interview zum Einsatz kam.

Für ein Drittel der Interviews in der Basisstichprobe sollten Modulfragen vorgelegt werden. Für die Mehrzahl der Interviews (64 Prozent) wurde das Interview ohne Modulfragen durchgeführt. In der Aufstockungsstichprobe wurden ebenfalls keine Module geschaltet. In der Zusatzstichprobe (onomastische Stichprobe) wurde analog zur Basisstichprobe verfahren.

¹³ Das Fragenmodul mit der Nummer 3 ist nach dem ersten CATI-Pretest gestrichen worden. Die ursprüngliche Nummerierung ist für die anderen Module beibehalten worden, um diesbezüglich Eingriffe in den Fragebogen vermeiden zu können.

¹⁴ Zu Möglichkeiten und Grenzen von Vignetten in der sozialwissenschaftlichen Forschung siehe Auspurg, K., Hinz, T. und S. Liebig (2009): „Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey“, in Methoden – Daten – Analysen, Jg. 3, Heft 1, S. 59-96.

Abbildung 1 Struktur des Fragebogens

Das beschriebene Vorgehen garantierte, dass für jedes Modul nach Abschluss der Feldarbeit wie geplant zwischen 1.500 und 10.200 Interviews realisiert wurden und damit ausreichend viele Fälle für spätere Analysen vorliegen. Die angestrebten Fallzahlen pro Modul sind in Tabelle 4 aufgeführt. Andererseits konnte durch die eingeschränkte Vorlage der sieben Fragenmodule bei nur einem Drittel der Interviews die vereinbarte durchschnittliche Interviewdauer von maximal 20 Minuten über alle vorgesehenen Interviews eingehalten werden.

Die Themenfelder, die in der Abbildung 1 unter Punkt D zusammengefasst sind, wurden an unterschiedlichen Stellen im Fragenkomplex Bereich C eingestreut (s. dazu auch Tabelle 5). Es wurde darauf geachtet, dass die Fragen in den Kontext des Interviews passten. Fragen zum Erwerbsstatus oder zum Familienstand sind beispielsweise so platziert, dass die Antworten für eine spätere Steuerung von entsprechenden Einblendungen genutzt werden können. Bei Folgefragen war es in manchen Fällen notwendig zu wissen, ob die Zielperson verheiratet oder ledig, berufstätig oder noch Schüler war.

Im Screener Teil I (Fragenblock E) wurde ermittelt, ob die Zielperson (oder ein anderes Haushaltsmitglied) ausgewählte Fahrzeuge besitzt. Im zweiten Teil der Screenerfragen (Block F) wurden die Zielpersonen zu Beginn gefragt, ob sie persönlich in den vergangenen rund fünf Jahren (d.h. seit Anfang 2007) Erfahrungen mit ausgewählten Straftaten gemacht hatten. Sollten sie Opfer einer der genannten Straftaten geworden sein, dann wurde im Block G gezielter nach Opfererfahrungen in den letzten 12 Monaten gefragt. Das Interview endete mit ausgewählten Fragen zur Soziodemografie.

Den Screenerfragen in Block E und F kam eine zentrale Bedeutung bei der Ermittlung von Prävalenzen zu. Sie sind daher im Vorfeld der Hauptstudie besonders intensiv im Rahmen der qualitativen und quantitativen Pretests getestet und optimiert worden. In Tabelle 5 sind die Themen der einzelnen Frageblöcke bzw. -module aufgeführt. Zudem kann der Tabelle entnommen werden, ob die Fragen an alle Zielpersonen oder nur an einen ausgewählten Kreis gestellt wurden. Der Fragebogen wurde in drei Sprachen (Deutsch, Russisch und Türkisch) eingesetzt.

Tabelle 4 Angezielte Fallzahl pro Modul in der Basisstichprobe

Modulnummer	angestrebte Fallzahl (absolut)	Anteil an Basisstichprobe (in %)
1	1.500	5
2	1.500	5
3	nach Pretest gestrichen	
4	1.500	5
5	1.500	5
6	2.000	6,7
7	2.200	7,3
8	10.200	34

Quelle: Stichprobenplan

Tabelle 5 Fragebogeninhalte

Zielgruppe	Inhalt
alle	Kontaktgespräch und Auswahl der Zielperson (Last-Birthday-Schlüssel, Hauptnutzer des Mobilfunktelefons), Ermittlung der reduzierten Haushaltsgröße bei Festnetz
alle	Gesundheit, Lebenszufriedenheit, Vertrauen
Zufallsstichprobe	Modul 6: Kriminalitätsbezogene Einstellungen
Zufallsstichprobe	Modul 2: Mediennutzung
Zufallsstichprobe	Modul 5: Sozialpsychologische Perspektiven und Einstellungen
alle	Allgemeiner und beruflicher Bildungsabschluss, Erwerbsstatus, Fragen zur Wohngegend
Zufallsstichprobe	Modul 4: Wohngegend, Disorder/Incivilities
alle	Ausgehverhalten, Familienstand, Internetnutzung
Zufallsstichprobe	Modul 7: Sicherheitsgefühl bei Internetnutzung
alle	Standardfrage zur Kriminalitätsfurcht
nur Befragte, denen auch Fragen zu einem anderen Modul gestellt wurden.	Modul 8: Deliktspezifische Kriminalitätsfurcht
Zufallsstichprobe	Modul 1: Kriminalitätsfurcht letzte 7 Tage
Befragte, die Fragen zu Modul 8 erhalten haben	Kriminalitätsfurcht (kognitiv): Fortsetzung Modul 8
alle	<p>Screeenerfragen zu den Themen:</p> <p>Besitz und Diebstahl von Kfz, Krad, Fahrrädern, und Kreditkartenbesitz</p> <p>Diebstahl aus Wohnungen, Wohnungseinbruchsdiebstahl</p> <p>Versuchter Diebstahl aus Wohnungen, Wohnungseinbruchsdiebstahl</p> <p>Fahrzeugdiebstahl</p> <p>Konsumentenbetrug</p> <p>Betrug mittels rechtswidrig erlangter Zahlungskarten</p> <p>Viktimisierung bei Internetnutzung</p> <p>Raub</p> <p>Sonstiger Diebstahl</p> <p>Körperverletzung</p>
Befragte, die positiv auf Screeenerfrage(n) geantwortet haben	Bündel Folgefragen zur Inzidenz, zum einzelnen Ereignis (einfache/mehrfache Viktimisierung)
alle	Soziodemografie
alle	Ausstattung mit Telekommunikationsmedien
alle	Postleitzahl
alle	Verabschiedung

2 Durchführung der Hauptstudie

2.1 Feldzeit

Die Hauptstudie der Basisstichprobe startete am 11. Juni 2012. Die ersten Interviews der Aufstockungsstichprobe für fünf Bundesländer (Hamburg, Berlin, Sachsen, Brandenburg und Hessen) wurden am 5. Juli 2012 durchgeführt. Nach vier Monaten Feldzeit lagen bereits 33.569 Interviews (98 Prozent der angezielten 34.348 Fälle) der Basis- und Aufstockungsstichprobe vor. Für beide Teilstichproben – Basis und Aufstockung – endete die Feldzeit am 24.10.2012. Die Feldzeit der Basisstichprobe umfasst damit 18,5 Wochen bzw. viereinhalb Monate. Die Aufstockungsstichprobe war rund 16 Wochen im CATI-Feld.

Zur Mitte der Feldzeit wurde geprüft, wie hoch der Anteil von Personen mit Migrationshintergrund in den bis dahin realisierten Interviews war. Dabei standen zum einen Männer und Frauen mit türkischem Migrationshintergrund und zum anderen Zielpersonen mit einem Migrationshintergrund aus den ehemaligen Sowjetrepubliken im Vordergrund der Betrachtung. Diese beiden Personengruppen sollten in etwa mit dem Anteil in der realisierten Stichprobe vertreten sein, wie es der Mikrozensus für die Bundesrepublik ausweist. Für die sowjetischen Staaten lagen die Vergleichszahlen sehr nahe beieinander, so dass nach Rücksprache mit dem BKA keine zusätzliche Stichprobe für die Zielgruppe eingesetzt werden sollte. Anders war es jedoch bei den Befragten mit türkischem Migrationshintergrund. Ihr Anteil in der Erhebung lag zu diesem Zeitpunkt unterhalb der Angaben des Mikrozensus. Deshalb wurden auf Wunsch des BKA entsprechende Vorbereitungen für den Einsatz einer Zusatzstichprobe getroffen, die den Anteil der Interviews mit Personen mit türkischem Migrationshintergrund erhöhen sollte. Dazu wurde auf die Erkenntnisse der Namensforschung¹⁵ (Onomastik) zurückgegriffen. Die onomastische Stichprobe kam ab dem 05.09.2012 zum Einsatz im Telefonfeld. Die Interviews wurden ausschließlich von zweisprachigen Interviewern (deutsch und türkisch) realisiert. Das letzte Interview wurde nach acht Wochen, am 03.11.2012, geführt.

2.2 Interviewerschulung

Grundsätzlich haben bei infas alle eingesetzten Interviewer vor ihrem ersten Interview an einer Grundschulung teilgenommen, in der u.a. auch die Faktoren einer erfolgreichen Felddurchführung und Kontaktierung vertieft werden. Insbesondere der Umgang mit nicht teilnahmebereiten Zielpersonen wird dort eingeübt. Dazu besteht ein Schwerpunkt der Grundschulung aus einem speziellen

¹⁵ Humpert, A. und K. Schneiderheinze (2002), Stichprobenziehung für telefonische Zuwandererumfragen – Praktische Erfahrungen und Erweiterung der Auswahlgrundlage, in: Gabler, S. und S. Häder (Hg.), Telefonstichproben - Methodische Innovationen und Anwendungen in Deutschland, Münster: Waxmann, S. 187-208. Humpert, A. und K. Schneiderheinze (2000), Stichprobenziehung für telefonische Zuwandererumfragen – Einsatzmöglichkeiten der Namensforschung (Onomastik), in: ZUMA-Nachrichten 47, S36-63.

Training zur Einwandbehandlung (RAT – Refusal Avoidance Training¹⁶). Die Grundschulungen erfolgen vor dem Ersteinsatz der Interviewer und werden bei Bedarf im Einzelfall zur Vertiefung wiederholt.

Um eine hohe Qualität der Datenerhebung im Feld sicherzustellen und positiv auf die Ausschöpfung der Befragung einzuwirken, haben alle für die Erhebung vorgesehenen Interviewerinnen und Interviewer an einer studienspezifischen Schulung teilgenommen. Die Teilnahme ist für die eingesetzten Interviewer obligatorisch. Besonders inhaltlich komplexe Interviews sowie ein hohes Ausschöpfungsziel erfordern eine sorgfältige Auswahl, Schulung und intensive Betreuung (Supervision) der Interviewer. Um Interviewereffekte möglichst gering zu halten, sollte jeder einzelne Interviewer nur einen geringen Anteil der mehr als 35.000 Interviews durchführen. Deshalb wurde eine Vielzahl an Interviewern projektspezifisch geschult und im Telefonfeld eingesetzt. Insgesamt wurden 600 Interviewer für die Studie in einer mündlichen Schulung auf ihren Einsatz vorbereitet. Davon wurden 570 Interviewer auch eingesetzt und haben mindestens ein Interview realisiert. Es fanden insgesamt 30 Schulungen statt. Die durchschnittliche Gruppenstärke lag bei 20 Interviewern pro mündliche Schulung. Die Schulungsinhalte betrafen folgende Bereiche:

- Studientitel
- Auftraggeber
- Feldzeit
- Befragungshintergrund und Ziel der Studie
- Zielgruppe, Auswahl der Zielperson
- Stichprobenherkunft
- geplante Interviewdauer und Fallzahlen
- Fragebogaufbau und -inhalte
- Besonderheiten des Fragebogens
- studienspezifische Gesprächsvorbereitung
- Ansprache der Kontakt- bzw. Zielpersonen
- studienspezifische Einwandbehandlung bei Schwermotivierbaren

Einen besonderen Stellenwert in der Schulung hatten die Reaktionsmöglichkeiten für die Interviewer, wenn eine Zielperson emotional aufgewühlt auf erlebte Straftaten (z.B. Gewalterfahrung wie Körperverletzung) reagiert. In solchen Situationen galt es, der Situation entsprechend sensibel zu reagieren. Dazu gehörte auch zu klären, ob die Zielperson das Interview noch weiter führen möchte oder kann. Die Option, in solchen Fällen das Interview abubrechen oder zu unterbrechen, bestand zu jeder Zeit. Es wurden weitere mögliche Verhaltensweisen mit den Interviewern besprochen. Zudem lagen den Interviewern die Kontaktdaten der Opferhilfe „Weißer Ring“ vor, die sie bei Bedarf an die Zielperson weitergeben konnten.

¹⁶ Verweigerungsvermeidungstraining: Besondere Aufmerksamkeit wird bei diesem Training den Maßnahmen und Interviewerstrategien zur Vermeidung von Teilnahmeverweigerung zuteil. Die Bemühungen zielen darauf ab, schwer motivierbare Personen zur Teilnahme an der Befragung zu bewegen, um die Stichprobe maximal auszuschöpfen. Es gibt kein allgemeingültiges Argument, mit dem garantiert jede Zielperson überzeugt werden kann, an der Befragung teilzunehmen. Jedes Interview ist individuell und einzeln auszuhandeln.

Das Interview wurde anschließend modulweise erläutert und in einem konkreten Fallbeispiel aktiv durchgespielt. Die Fallbeispiele dienten dazu, die Interviewer mit inhaltlichen Aspekten und mit der technischen Handhabbarkeit des Instruments vertraut zu machen. Im Fallbeispiel wurden alle Modulfragen besprochen. Das bedeutet: Alle Interviewer wurden auf die vollständige Fragebogenfassung mit sämtlichen Modulen geschult. Es wurde explizit auf Besonderheiten des Fragebogens eingegangen und der Hintergrund des Forschungsvorhabens vermittelt. Ein Vertreter des BKA-Projektteams war bei der ersten Schulung anwesend.

Alle Interviewer erhielten zusätzlich ein ausführliches Interviewerhandbuch. Im Interviewerhandbuch waren alle wichtigen Hinweise und Erläuterungen noch einmal schriftlich dargelegt, so dass das Handbuch während der Feldzeit bei Bedarf als Nachschlagewerk diente. Der Inhalt des Handbuchs (vgl. Abbildung 2) wurde mit dem BKA abgestimmt.

Bevor die geschulten Interviewer im Telefonstudio eingesetzt wurden, waren sie gehalten, sich intensiv mit der Testfassung des Fragebogens vertraut zu machen.

Abbildung 2 Inhalte des Interviewerhandbuchs

Inhaltsverzeichnis

1	Überblick über zentrale Kennwerte der Studie	3
1.1	Thema, Bedeutung und Besonderheiten der Studie	3
1.2	Rahmendaten der Studie	4
2	Kontaktierung der Haushalte, der Zielpersonen	6
2.1	Unterschiedlicher Beginn des Kontaktgespräch	6
2.2	Besonderheiten beim Geburtstagschlüssel (Last-Birthday-Verfahren)	7
2.3	Interviewsprachen und Sprachenwechsel	7
3	Aufbau des Fragebogens	8
4	Erläuterungen zu ausgewählten Fragen	10
4.1	Haushaltsdefinition	10
4.2	Interviewerhinweise „Bei Bedarf vorlesen“	10
4.3	Verwendung von Buttons	11
4.4	Wechselnde Bezugszeiträume: Tage, Monate, Jahre	12
4.5	Nur bei einem Teil der Interviews: Beurteilung von Fallbeispielen	12
4.6	Erfahrungen mit erlebten Straftaten	13
4.7	Fragen nach Geburtsland und Staatsangehörigkeit	14
4.8	Frage nach der Einwohnerzahl	16
4.9	Anzahl der Rufnummern im Haushalt und Mobilfunknummern	16
4.10	Angabe der Postleitzahl	16
5	Allgemeine Hinweise	17
5.1	Was mache ich bevor ich anfangen...	17
5.2	Was tue ich, wenn die ZP fragt, was mit den Daten geschieht/ nach Datenschutz fragt?	17
5.3	Was tue ich, wenn die Zielperson fragt, ob sie Ergebnisse der Studie bekommen kann?	18
5.4	Was tue ich, wenn die Zielperson bei der Frage nach Straftaten emotional aufgewühlt reagiert?	18
5.5	Auf wen kann ich bei studienspezifischen Nachfragen verweisen?	19
5.6	Allgemeine Hinweise zur Interviewführung	19

2.3 Interviewereinsatz

570 Interviewer haben mindestens ein Interview vollständig realisiert. Der Anteil Männer und Frauen war unter diesen Interviewern nahezu gleich (Männer 49 Prozent, Frauen 51 Prozent). Es kamen Interviewer aller Altersklassen zwischen 18 und 74 Jahre zum Einsatz (vgl. Tabelle 6 und Tabelle 7). Die Interviewer realisierten – bezogen auf alle drei Teilstichproben (Basis-, Aufstockungs- und onomastische Stichprobe) – zwischen einem und 376 Interviews¹⁷. Im Durchschnitt führten die Interviewer 62 Interviews durch¹⁸.

Tabelle 6 Telefoninterviewer nach Geschlecht

Geschlecht	Anzahl	Prozent
männlich	279	48,9
weiblich	291	51,1
gesamt	570	100,0

Quelle: Einsatzdatenbank

Tabelle 7 Telefoninterviewer nach Alter (gruppiert)

Altersgruppen der Interviewer	Anzahl	Prozent
zwischen 18 und 24 Jahre	279	48,9
zwischen 25 und 34 Jahre	125	21,9
zwischen 35 und 44 Jahre	44	7,7
zwischen 45 und 54 Jahre	54	9,5
zwischen 55 und 64 Jahre	51	8,9
zwischen 65 und 74 Jahre	17	3,0
gesamt	570	100,0

Quelle: Einsatzdatenbank

¹⁷ In der Basisstichprobe hat nur ein Interviewer mit 319 Interviews mehr als 300 Interviews realisiert. Alle anderen Interviewer haben weniger als 300 Interviews (d.h. weniger als ein Prozent der realisierten Interviews) durchgeführt.

¹⁸ Die Standardabweichung beträgt 68,3, d.h. es liegt eine breite Streuung über die Interviewer vor.

Tabelle 8 Realisierte Interviews pro Interviewer

Realisierte Interviews pro Interviewer	Anzahl
durchschnittliche Interviewanzahl pro Interviewer	62,3
maximale Interviewanzahl pro Interviewer	376
minimale Interviewanzahl pro Interviewer	1
Standardabweichung	68,29
Anzahl Interviewer mit mindestens einem realisiertem Interview	570

Quelle: Einsatzdatenbank

Tabelle 9 Anzahl realisierte Interviews pro Interviewer (gruppiert)

Anzahl realisierte Interviews pro Interviewer	Anzahl Interviewer	Prozent
1 bis 50	326	57,2
51 bis 100	134	23,5
101 bis 150	52	9,1
151 bis 200	23	4,0
201 bis 250	21	3,7
251 bis 300	6	1,1
über 300	8	1,4
gesamt	570	100,0

Quelle: Einsatzdatenbank

Um mögliche Interviewereffekte gering zu halten, wurde eine Vielzahl von Interviewern eingesetzt, die möglichst nur einen geringen Anteil an der Gesamtfallzahl von mehr als 35.000 Interviews durchführen sollte. Dies ist gelungen. Nur in wenigen Fällen (achtmal) haben Interviewer mehr als 300 Interviews über alle drei Teilstudien hinweg realisiert. Und lediglich sechs Interviewer haben zwischen 251 und 300 Interviews geführt. Hingegen haben 57 Prozent der Interviewer bis zu 50 Interviews realisiert und weitere 24 Prozent haben 51 bis 100 Interviews durchgeführt (vgl. Tabelle 9).

Die Analysen möglicher Interviewereffekte haben keine Auffälligkeiten erkennen lassen. Zum einen wurde feldbegleitend wöchentlich mögliche Interviewereffekte auf die Prävalenzen der fünf Merkmale „Waren-/Dienstleistungsbetrug“, „Raub“, „persönlicher Diebstahl“, „Körperverletzung“ und „Fahrraddiebstahl“ geprüft. Neben deskriptiven Mittelwertvergleichen wurde in linearen Modellen ein Interviewerkumpungeffekt berechnet. Als Maß wurde der Intracusterkorrelationskoeffizient (ICC) verwendet, wobei die Interviewer als Cluster definiert wurden. Bei keinem der Merkmale konnte ein statistisch signifikanter Effekt nachgewiesen werden.

Zum anderen wurde nach Abschluss der Erhebung zusätzlich ein Gruppenvergleich vorgenommen zwischen den Interviewern, die weniger als 150 Interviews

und denen, die mindestens 150 Interviews durchgeführt haben. Geprüft wurde, ob sich diese beiden Gruppen bezüglich der gemessenen Zwölfmonatsprävalenzen für die fünf Delikte „Waren-/Dienstleistungsbetrug“, „Raub“, „persönlicher Diebstahl“, „Körperverletzung“ und „Fahrraddiebstahl“ signifikant unterscheiden. Entsprechend den im laufenden Feld nicht beobachtbaren linearen Effekten ergeben sich auch bei dieser Kontrastierung keine statistisch signifikanten Unterschiede. Die Prävalenzen sind in beiden Gruppen identisch.

2.4 Supervision zur Qualitätssicherung

Während der gesamten Feldphase fand eine Qualitätssicherung der Interviewer statt. Die Qualitätssicherung der Interviewführung wurde durch geschulte Supervisoren durchgeführt, die mit methodischen Anforderungen der Erhebung und der CATI-Befragungstechnik eng vertraut sind. Diese Supervisoren wurden durch die Projektleiter gezielt inhaltlich geschult, um in enger Rückbindung mit der Projektleitung studienspezifische Rückfragen der Interviewer beantworten zu können. Zielsetzung der Supervision war es, durch kontinuierliche Betreuung und persönliche Ansprechbarkeit eine hohe Qualität der Befragungsdaten zu erhalten und die Ausschöpfung der Studie durch adäquate Kontaktaufnahme zu maximieren. Zu den Aufgaben der Supervision gehörten im Einzelnen:

- Die Kontrolle durch Mithören der Interviews im Telefonstudio und Überprüfung der Eingaben durch Beobachtung auf Supervisionsmaske (Übernahme der Interviewermonitoranzeigen auf den Supervisionsarbeitsplatz).
- Die Durchführung von Einzelgesprächen oder Nachschulungen mit allen Interviewern hinsichtlich studienspezifischer Probleme.
- Die Durchführung von Einzel- und Gruppengesprächen oder auch Nachschulungen direkt im Anschluss an das Interview bei Rückmeldungen von Problemen. Dabei wurden auch schriftliche Rückmeldungen an die Interviewer ausgeteilt.
- Direktes Eingreifen in das Interview bei beobachteten Fehlern bzw. Zuordnungsschwierigkeiten, die zu Fehlsteuerungen und Lücken zu führen drohen. Wobei dies entweder über kurze mündliche oder schriftliche Hinweise während des Interviews geschieht oder bei größeren Problemen und in Überforderungssituationen der Interviewer (meist infolge sehr komplizierter Verläufe oder widersprüchlicher Angaben der ZP) auch Eingriffe der Supervision in das Gespräch zur Folge haben kann.

Die Aufgaben der inhaltlichen Supervision waren sehr vielfältig und dauerten während der gesamten Feldphase an. Es bestand ein enger Informationsfluss zur Projektleitung, so dass inhaltliche Probleme möglichst rasch und auf Basis umfassender Kommunikation aller Beteiligten gelöst werden konnten. Auch eine sofortige Rückmeldung der gemeinsam erarbeiteten Problemlösungen ins Feld hinein war damit gewährleistet.

2.5 Feldsteuerung

Den Verlauf der Erhebung beobachtete die Projektleitung anhand von täglich aktualisierten Feldreportings. Der aktuelle Feldstand wurde regelmäßig in der Arbeitsgruppe bewertet.

Der Feldeinsatz stand unter dem Ziel einer maximal möglichen Ausschöpfung der Stichprobe in der vorgesehenen Feldzeit. Dazu gehört, dass einerseits schwer erreichbare Personen so lange angerufen werden, bis ein Kontakt zustande kommt. Eine zweite wichtige Feldstrategie ist die Motivierung von Befragten und die Vermeidung von Verweigerungen.

Mit dem Ziel, auch Personen einzubinden, die zu Hause selten anzutreffen sind, wurden die Adressen so lange im Telefonfeld bearbeitet, bis ein endgültiger Bearbeitungsstatus feststand. Darüber hinaus ermöglichten Terminabsprachen eine flexible und individuell ausgerichtete Bearbeitung der Stichproben. Die Stichprobe wurde in der gesamten Feldzeit an sechs Tagen in der Woche zwischen 9 Uhr und 21 Uhr (Montag bis Freitag) bzw. 10 bis 18 Uhr (Samstag) telefoniert. Für jede Rufnummer wurde die Tageszeit für den nächsten Kontakt von Tag zu Tag kontinuierlich verändert, um eine möglichst große Variation an unterschiedlichen Wochentagen und Tageszeiten für einen Kontaktversuch zu gewährleisten.

Die Kontaktzeiten zu Beginn des Feldes waren an einigen Tagen durch die Fußball-Europameisterschaft eingeschränkt. Die Interviewer waren gehalten, sorgsam zu reagieren und einen erneuten Anruf anzukündigen, wenn sich Zielpersonen durch die Kontaktaufnahme während eines oder kurz vor einem Fußballspiel gestört fühlten.

Die Feldphase umfasste auch die Sommerferienzeit. Die kontinuierliche Kontaktierung der Stichprobe über die gesamte Feldphase machte es möglich, auch Personen zu erreichen, die in den Sommerferien verreist waren. Es bestand beispielsweise die Möglichkeit, vor Reisebeginn einen Kontakttermin für die Zeit nach dem Urlaub zu vereinbaren.

Sprachenswitch

In der onomastischen Stichprobe wurden ausschließlich zweisprachige Interviewer eingesetzt. Sie hatten so die Möglichkeit, bei Bedarf das Kontaktgespräch direkt auf Türkisch oder Deutsch zu führen. Ihnen lag der Fragebogen in beiden Sprachen vor. Sie konnten auch im späteren Verlauf des Gesprächs auf die jeweils andere Sprache wechseln, wenn sich herausstellte, dass die Zielperson nicht ausreichend Deutsch versteht und/oder spricht.

In der Basis- und Aufstockungstichprobe starteten die Interviewer das Kontaktgespräch auf Deutsch. Wenn sich im Kontaktgespräch herausstellte, dass die Kontaktperson (bzw. Zielperson) nicht ausreichend Deutsch spricht, wurde das Gespräch an dieser Stelle unterbrochen. Die Interviewer legten den Fall ab mit einem Hinweis auf die Fremdsprache, die der Gesprächspartner bzw. die Zielperson spricht. Wenn es sich bei dieser Sprache um Russisch oder Türkisch handelte, dann wurden diese Fälle in eine separate Stichprobendatei überführt. Diese Sub-

stichprobe wurde in den folgenden Tagen und Wochen ausschließlich von zweisprachigen Interviewern bearbeitet. Wenn der Gesprächspartner bzw. die Zielperson keine der drei Interviewsprachen (Deutsch, Russisch oder Türkisch) sprach, dann wurde der Endstatus „keine Verständigung in einer der Zielsprachen möglich“ vergeben.

2.6 Adressnachbearbeitung zur Erhöhung der Ausschöpfung

Die Interviewer wurden in den vorgeschalteten Schulungen mit Zugangsstrategien und der Behandlung von Einwänden vertraut gemacht. Kam es trotzdem zu einer Verweigerung, so erlaubten die differenziert dokumentierten Ausfallgründe eine Analyse, weshalb es nicht zu einem Interview gekommen war. Es wurden „harte“ Verweigerungen – strikte Ablehnung einer Teilnahme – von „weichen“ Verweigerungen, die in der jeweiligen Situation begründet sind – z.B. „keine Zeit“ – unterschieden.

Zielpersonen, die während des ersten Kontakts situativ das Interview verweigerten¹⁹, wurden nach einiger Zeit erneut angerufen und noch einmal um Mitwirkung gebeten (sog. Nonresponse-Studien). Eine solche Nachbearbeitung wurde für die Basisstichprobe am 13.07.2012, für die Aufstockungsstichprobe am 24.07.2012 und für die onomastische Stichprobe am 09.10.2012 gestartet. Für die Nachbearbeitung wurden die entsprechenden Fälle in eine weitere Teilstichprobe separiert. Diese Nachbearbeitungsstichproben wurden von speziell geschulten Interviewern bearbeitet und dauerten bis zum Feldende an. Insgesamt wurden vier solcher Nachbearbeitungsstichproben eingerichtet und wochenweise durch neue Fälle aus den Ausgangsstichproben aufgefüllt:

- für die Basis- und Aufstockungsstichprobe in Deutsch,
- für die Basis und Aufstockungsstichprobe in Türkisch,
- für die Basis und Aufstockungsstichprobe in Russisch,
- und zuletzt für die onomastische Stichprobe.

Diese Nachbearbeitungsstudien erwiesen sich als erfolgreich. Insgesamt konnten durch die intensive feldbegleitende Nachbearbeitung bestimmter Adressen zusätzlich 2.264 Interviews realisiert werden.

Im gesamten Feldverlauf gingen n=50.687 Adressen in die deutschsprachige Nachbearbeitung. Bis zum Ende der Feldzeit konnten n=2.162 Interviews realisiert werden. Dies entspricht einem Anteil von vier Prozent.

In der türkische Nachbearbeitungsstudie wurden insgesamt n=101 Adressen eingesetzt. Daraus konnten n=11 Interviews realisiert werden (elf Prozent). Für die russischsprachigen Interviews lagen am Ende der Feldzeit n=25 Interviews aus der Nachbearbeitung vor. Dies entspricht 18 Prozent der 143 eingesetzten Adressen.

¹⁹ Berücksichtigt wurden die Codes „keine Zeit“ und „kein Interesse“.

Die Nachbearbeitung von schwermotivierbaren Zielpersonen in der onomastischen Stichprobe (n=1.461 Adressen) erbrachte weitere 66 Interviews. Dies entspricht einem Anteil von knapp fünf Prozent.

2.7 Stichprobenausschöpfung

Im Folgenden werden die Ergebnisse der Feldarbeit differenziert für alle drei Teilstichproben (Basis, Aufstockung und Onomastik) diskutiert.

Das Ergebnis der einzelnen Kontaktversuche wurde mit differenzierten Responsecodes erfasst. Die differenzierten Responsecodes orientieren sich an den „Standard Definitions“ der American Association for Public Opinion Research (AAPOR 2011²⁰). Zu Gruppen von Bearbeitungsstatus zusammenfasst, bilden sie die Grundlage für die Berechnung von Kennwerten zur Stichprobenausschöpfung nach dem AAPOR-Standard. Nach AAPOR werden die Bearbeitungsergebnisse folgenden Gruppen zugeordnet:

- nicht Zielgruppe (NE)
- unbekannte Auswahl (Unknown Eligibility, UE)
- Nonresponse – nicht befragbar (NR-NA)
- Nonresponse – nicht erreicht (NR-NC)
- Nonresponse – Verweigerung (NR-R)
- realisierte Interviews (vollständig) (I)
- realisierte Interviews (unvollständig) (IP)

Zur Kategorie „nicht Zielgruppe“ zählen beispielsweise alle Festnetzanschlüsse, die zu keinem Privathaushalt führen, Mobilfunknummern, die ausschließlich beruflich genutzt werden oder der Hauptnutzer jünger als 16 Jahre alt ist, Rufnummern, die nicht geschaltet sind, Fax- bzw. Modemanschlüsse, Haushalte bzw. Zielpersonen, die keine der Interviewsprachen beherrschen sowie Haushalte, die mehrfach in die Stichprobe geraten sind.

Die Gruppe „unbekannte Auswahl“ bilden Zielpersonen, bei denen die Zugehörigkeit zur Zielgruppe unbekannt ist. Keine Aufklärung des Anschlusses gab es bei Kontaktversuchen, bei denen niemand abgehoben hat, lediglich ein Besetzzeichen zu hören, nur ein Anrufbeantworter geschaltet war oder sofort wieder aufgelegt wurde.

Die Gruppe der „Nonresponses“ wird differenziert nach „nicht befragbar“ und „Teilnahmeverweigerungen“. „Nicht befragbar“ war ein Teil der Zielpersonen aus gesundheitlichen Gründen. Daneben gibt es unterschiedliche Gründe für eine Teilnahmeverweigerung. Dazu zählen zum Beispiel kein Interesse an der Befragung, keine Zeit für ein Interview, Bedenken bezüglich des Datenschutzes

²⁰ AAPOR (The American Association for Public Opinion Research) (2011): Standard Definitions. Final Dispositions of Case Codes and Outcome Rate for Survey. Revisited Version 2011.

oder eine grundsätzliche Verweigerung ohne Nennung eines spezifischen Grundes.

Die Kategorie „realisierte Interviews“ enthält alle im Feld durchgeführten Interviews, unabhängig davon, ob sie im Zuge der nachgelagerten Datenprüfung als auswertbar bewertet wurden oder nicht vollständig realisiert werden konnten.

Die im Folgenden ausgewiesenen Bearbeitungs-codes bilden den endgültigen Bearbeitungsstatus einer Telefonnummer ab (Final Outcome). Zur Ermittlung des Final Outcomes wird zwischen allen Feldcodes innerhalb eines Kontaktverlaufs eine Priorisierung vorgenommen. So gilt grundsätzlich der Status aus dem letzten Sprachkontakt zum Haushalt, auch wenn danach noch vergebliche Kontaktversuche unternommen worden sein sollten. Eine Telefonnummer wird z.B. nur dann als „nicht erreicht“ ausgewiesen, wenn im gesamten Kontaktverlauf kein Sprachkontakt zur Zielperson oder ihrem Haushalt hergestellt werden konnte. Telefonische Terminvereinbarungen oder telefonisch erteilte Teilnahmeverweigerungen werden bei der Dokumentation des Feldstatus ausgewiesen, auch wenn danach ein Interviewer noch vergeblich versucht haben sollte, die Zielperson zu erreichen.

2.7.1 Basisstichprobe

Die Bruttoeinsatzstichprobe der Basisstichprobe umfasste 750.000 Telefonnummern. In gut 69 Prozent der Fälle konnte unter der Telefonnummer keine Person erreicht werden, die zur Zielgruppe der Studie gehört (vgl. Tabelle 10, Spalte II). Mit 63 Prozent der eingesetzten Rufnummern führte, wie erwartet (vgl. Abschnitt 1.2.1), der Großteil der generierten Rufnummern zu keinem Anschluss. Bei zwölf Prozent konnte die Zielgruppenzugehörigkeit nicht geklärt werden. Insbesondere bei den acht Prozent der Anschlüsse, bei denen niemand den Hörer abgehoben hat, war abschließend nicht zu klären, ob der Anschluss tatsächlich geschaltet ist oder nicht. In den Fällen, die unter der Kategorie „unbekannte Auswahl“ zusammengefasst sind, kann sich demnach noch ein Teil stichprobenneutrale Ausfälle verbergen.

Zieht man die genannten Fälle von den eingesetzten Rufnummern ab, so verbleiben noch 145.633 Telefonnummern, für die bekannt ist, dass eine Zielperson vorhanden ist. In 383 Fällen (0,3 Prozent der 145.633 Telefonnummern, vgl. Tabelle 10, Spalte IV) konnte mit der Zielperson kein Interview durchgeführt werden, da sie aus gesundheitlichen Gründen nicht befragbar war. 17 Prozent der 145.633 kontaktierten Haushalte bzw. Personen konnten in der Feldzeit nicht erreicht werden. Der weitaus größte Teil (22.620, 15,5 Prozent) ist unter „vager Termin“ verbucht. Vage Termine werden dann vergeben, wenn die Teilnahmebereitschaft der Zielperson noch offen ist, der Interviewer die Zielperson noch nicht persönlich erreichen konnte, beispielsweise lediglich ein Minderjähriger im Haushalt angetroffen wurde und dieser mitteilt, wann ein Erwachsener für ein Gespräch erreichbar ist, u.v.m. Zudem können sich hinter vagen Terminen auch versteckte Teilnahmeverweigerungen verbergen, wenn Zielpersonen z.B. aus Höflichkeit ihre Teilnahme nicht explizit verweigern möchten und den Interviewer auf eine erneute Kontaktaufnahme „vertrösten“. In den genannten Fällen kam es zwar zwischenzeitlich zu mindestens einem Sprachkontakt mit einer Kontakt- oder Zielperson, und es wurde vereinbart, dass sich der Interview-

wer noch einmal melden darf. Es ist jedoch bis zum Feldende weder zu einem definitiven Interviewtermin noch zu einem Interview gekommen.

Im Unterschied zu den „vagen Terminen“ stimmen die Interviewer im Feldverlauf auch konkrete Interviewtermine (in der Regel) mit den Zielpersonen ab. Diese Termine werden als „definitive Termine“ im System abgelegt und zum festgelegten Datum einem freien Interviewer vorgelegt. Definitive Termine führen zwar nicht zu 100 Prozent, aber in vielen Fällen dann auch zu einem Interview. Zu einem sehr geringen Anteil (1,5 Prozent der 145.633 Fälle, für die die Zugehörigkeit zur Zielgruppe bekannt ist) wurde zwar ein definitiver Interviewtermin im Feldverlauf vereinbart. Die Zielperson konnte jedoch bis zum Feldende nicht mehr erneut erreicht werden.

Tabelle 10 Feldergebnis der Basisstichprobe nach AAPOR-Klassifikation

	Anzahl	%	Anzahl	%	Anzahl	%
	I	II	III	IV	V	VI
Bruttostichprobe	750.000	100,00				
nicht Zielgruppe (NE)	516.546	68,9				
kein Anschluss	474.332	63,2				
Anschluss nur Fax/Modem	19.030	2,5				
kein Privathaushalt unter diesem Anschluss	16.761	2,2				
nicht in Zielgruppe lt. Angabe im Kontaktgespräch	4.031	0,5				
bereits befragt	670	0,1				
keine Verständigung in einer der drei Zielsprachen möglich	1.722	0,2				
unbekannte Auswahl (UE)	87.821	11,7				
nicht abgehoben	58.041	7,7				
Anrufbeantworter	3.201	0,4				
besetzt	1.594	0,2				
legt sofort auf	24.985	3,3				
Zugehörigkeit zur Zielgruppe bekannt	145.633	19,4	145.633	100,0		
Nonresponse – nicht befragbar (NR-NA)	383	0,1	383	0,3		
ZP (lt. Auskunft) nicht befragbar/dauerhaft krank/behindert	383	0,1	383	0,3		
Nonresponse – nicht erreicht (NR-NC)	24.896	3,3	24.896	17,1		
definitiver Termin	2.155	0,3	2.155	1,5		
vager Termin	22.620	3,0	22.620	15,5		
Zielperson in Feldzeit nicht zu erreichen	121	0,0	121	0,1		
Zielperson/Haushalt erreicht und befragbar					120.354	100
Nonresponse – Verweigerung (NR-R)	88.401	11,8	88.401	60,7	88.401	73,5
Zielperson verweigert grundsätzlich	24.938	3,3	24.938	17,1	24.938	20,7
Zielperson verweigert: keine Zeit/dauert zu lange	7.643	1,0	7.643	5,2	7.643	6,4
ZP verweigert: will nur persönlich befragt werden	4.688	0,6	4.688	3,2	4.688	3,9
ZP verweigert krank	1.776	0,2	1.776	1,2	1.776	1,5
kein Zugang zu ZP/TN untersagt/lt. Auskunft nicht bereit	1.070	0,1	1.070	0,7	1.070	0,9
KP verweigert jegliche Auskunft	10.032	1,3	10.032	6,9	10.032	8,3
ZP verweigert: kein Interesse, wegen des Themas	26.831	3,6	26.831	18,4	26.831	22,3

	Anzahl	%	Anzahl	%	Anzahl	%
	I	II	III	IV	V	VI
Bruttostichprobe	750.000	100,00				
ZP verweigert: Datenschutzgründe/zu persönlich	1.376	0,2	1.376	0,9	1.376	1,1
ZP verweigert: sonstige Gründe	10.047	1,3	10.047	6,9	10.047	8,3
realisierte Interviews (I)	31.953	4,3	31.953	21,9	31.953	26,5
Interviews vollständig realisiert, ausgeliefert	30.278	4,0	30.278	20,8	30.278	25,2
Interviews vollständig realisiert, nicht auswertbar	29	0,0	29	0,0	29	0,0
Interviews unvollständig realisiert	1.646	0,2	1.646	1,1	1.646	1,4

Quelle: Rücklaufdatenbank, eigene Berechnung

Für 120.354 Rufnummern wurde eine Haushalt bzw. eine Zielperson erreicht, der bzw. die prinzipiell befragbar war. In 31.953 (26,5 Prozent) Fällen ist es zu einem Interview gekommen²¹. 30.307 vollständige Interviews wurden realisiert. Lediglich 29 mussten nach der formalen Prüfung als nicht auswertbar deklariert werden. In 1.646 Fällen ist das Interview vorzeitig abgebrochen worden (vgl. dazu auch die Ausführungen in Abschnitt 2.14). Der Auswertungsdatensatz umfasst somit 30.278 Fälle.

Sofern die Kontakt- bzw. Zielperson einen Grund für die Teilnahmeverweigerung angegeben hat, wurde dieser Verweigerungsgrund vom Interviewer entsprechend verbucht. Die differenzierten Ausfallgründe sind in Tabelle 10 aufgeführt. Als die beiden mit Abstand häufigsten Ausfallgründe notierten die Interviewer eine grundsätzliche Verweigerung (20,7 Prozent) und fehlendes Interesse am Thema (22,3 Prozent)²².

In der folgenden Tabelle 11 ist das Feldergebnis nach der AAPOR-Klassifikation differenziert für die beiden Teilstichproben Festnetz und Mobilfunk der Basisstichprobe dargestellt. In der Mobilfunkstichprobe liegt der Anteil der Anschlüsse, der nicht zur Zielgruppe führte, fünf Prozentpunkte höher als in der Festnetzstichprobe. Der Anteil „unbekannte Auswahl“ liegt dafür in der Festnetzstichprobe höher als bei den Handynummern. Bezogen auf die jeweiligen Einsatz-

²¹ Allgemein liegt die Teilnahmebereitschaft bei Kaltanrufen am Telefon niedriger als bei Studien, bei denen die Zielpersonen vorab schriftlich informiert werden können. Vgl. Aust, F. und H. Schröder: Sinkende Stichprobenausschöpfung in der Umfrageforschung – ein Bericht aus der Praxis. In: Weichbold, Martin, Johann Bacher und Christof Wolf: Umfrageforschung. Grenzen und Herausforderung. 2009. VS Verlag.

²² Auf Basis des endgültigen Bearbeitungsstatus kann alternativ die Responderate gemäß Standarddefinition (AAPOR 2011) berechnet werden. AAPOR unterscheidet mehrere Responderaten. Bei der Wahl der adäquaten Responderate für die vorliegende Studie ist zu berücksichtigen, dass neben dem großer Anteil nicht geschalteter Telefonnummern bei einem erheblichen Anteil der übrigen Fälle die Zielgruppenzugehörigkeit nicht festgestellt werden konnte. Für solche Studien steht die Responderate 4 nach AAPOR zur Verfügung, die hier im Folgenden angewendet wird. In diese Rate fließt eine Schätzung über den Anteil (e) der Telefonnummern mit Status „Zielgruppenzugehörigkeit unbekannt“ ein, der zur Zielgruppe gehört. Die Schätzung unterstellt, dass dieser Anteil genauso hoch ist wie bei den in der Befragung aufgeklärten Fällen. Der Zähler der Rate umfasst vollständig und unvollständig realisierte Interviews. Für den Nenner der Rate wird die Bruttostichprobe bereinigt um Adressen außerhalb der Zielgruppe und um einen geschätzten Anteil der ebenfalls nicht zur Zielgruppe gehörenden Personen aus der Kategorie „Zielgruppenzugehörigkeit unbekannt“. Unter dieser Annahme ergibt sich eine Responderate 4 in Höhe von 19,4 Prozent. Zur Berechnung der Responderaten vgl.: The American Association for Public Opinion Research (AAPOR) (2011): Standard Definitions. Final Dispositions of Case Codes and Outcome Rate for Survey. Revisited Version 2011, S. 44f

bruttostichproben liegen die Ausschöpfungen zwar auf nahezu gleichem Niveau. Das Verhältnis von Interviews zu Verweigerungen unterscheidet sich jedoch spürbar. Unter den Rufnummern, unter denen ein Haushalt bzw. eine Zielperson erreicht wurde und sie auch befragbar waren, haben in der Festnetzstichprobe 25,3 Prozent²³ am Interview teilgenommen, bei der Mobilfunkstichprobe waren es mit 31,2 Prozent²⁴ deutlich mehr.

Tabelle 11 Final Outcome nach AAPOR, differenziert nach Teilstichproben Festnetz und Mobilfunk

Final Outcome	Festnetz		Mobilfunk		insgesamt	
	Anzahl	%	Anzahl	%	Anzahl	%
Bruttostichprobe	576.150	100,0	173.850	100,0	750.000	100,0
nicht Zielgruppe (NE)	389.903	67,7	126.643	72,8	516.546	68,9
unbekannte Auswahl (UE)	73.791	12,8	14.030	8,1	87.821	11,7
Nonresponse – nicht befragbar (NR-NA)	383	0,1	0	0,0	383	0,1
Nonresponse – nicht erreicht (NR-NC)	16.892	2,9	8.004	4,6	24.896	3,3
Nonresponse – Verweigerung (NR-R)	71.091	12,3	17.310	10,0	88.401	11,8
realisierte Interviews (I)	24.090	4,2	7.863	4,5	31.953	4,3

Quelle: Rücklaufdatenbank, eigene Berechnung

2.7.2 Aufstockung

Wie bereits für die Basisstichprobe beschrieben, so ergab sich auch in der Aufstockungsstichprobe für fünf Bundesländer ein hoher Anteil an Rufnummern, die nicht geschaltet sind. Rund 66 Prozent der 103.590 Telefonnummern umfassenden Bruttoeinsatzstichprobe führten zu keinem geschalteten Telefonanschluss (vgl. Tabelle 12). Weitere drei Prozent (3.085 Rufnummern) waren auf ein Fax bzw. Modem geschaltet, und in vier Prozent der Nummern handelte es sich um keinen Privatanschluss. Insgesamt summieren sich die Festnetznummern der Einsatzstichprobe, die zu keinem Zielgruppenhaushalt führten, auf 75.200 (72,6 Prozent der Einsatzstichprobe). Der Anteil liegt nur auf einem leicht höheren Niveau im Vergleich zur Basisstichprobe (vgl. Abschnitt 2.7.1).

Bei acht Prozent (8.081 Telefonnummern) der Einsatzstichprobe konnte nicht festgestellt werden, ob die Rufnummer zu einem Zielgruppenhaushalt führt.

Abzüglich der genannten Fälle verbleiben 20.372 Telefonnummern, für die geklärt werden konnte, dass eine Zugehörigkeit zur Zielgruppe vorliegt. Dies entspricht 20 Prozent der Bruttoeinsatzstichprobe. In 61 Fällen konnte mit der Zielperson aus gesundheitlichen Gründen kein Interview geführt werden. In 16 Prozent der 20.372 Haushalte konnte die Zielperson bis zum Feldende trotz mehrfacher Kontaktversuche nicht wieder erreicht werden. Gleichfalls wie in der Basis-

²³ $(24.090 / (24.090 + 71.091)) * 100 = 25,3\%$

²⁴ $(7.863 / (7.863 + 17.301)) * 100 = 31,2\%$

stichprobe handelt es sich in fast allen Fällen um vage Termine, die bis zum Feldende nicht in einen Interviewtermin bzw. ein Interview mündeten.

In 17.104 Fällen hat der Interviewer mit einer Kontaktperson im Haushalt oder mit der Zielperson persönlich sprechen können. In 27 Prozent ist es zu einem Interview gekommen. 202 Mal wurde das Interview vorzeitig beendet. Zwei der verbleibenden 4.419 vollständig realisierten Interviews wurden nach der formalen Datenprüfung als nicht auswertbar deklariert und aus dem Datensatz entfernt. 4.417 vollständige Interviews wurden an das BKA übergeben.

Die Verweigerungsgründe verteilen sich in gleicher Weise wie bei der Basisstichprobe. Die Hauptgründe für eine Nichtteilnahme liegen bei der grundsätzlichen Teilnahmeverweigerung (22 Prozent) und beim fehlenden Interesse am Thema. Datenschutzgründe wurden nur äußerst selten als Ausfallgrund verbucht²⁵.

Tabelle 12 Feldergebnis der Aufstockungsstichprobe nach AAPOR-Klassifikation

	Anzahl	%	Anzahl	%	Anzahl	%
	I	II	III	IV	V	VI
Bruttostichprobe	103.590	100,00				
nicht Zielgruppe (NE)	75.200	72,6				
kein Anschluss	67.818	65,5				
Anschluss nur Fax/Modem	3.085	3,0				
kein Privathaushalt unter diesem Anschluss	3.967	3,8				
nicht in Zielgruppe lt. Angabe im Kontaktgespräch	14	0,0				
bereits befragt	64	0,1				
keine Verständigung in einer der drei Zielsprachen möglich	252	0,2				
unbekannte Auswahl (UE)	8.018	7,7				
nicht abgehoben	4.530	4,4				
Anrufbeantworter	218	0,2				
besetzt	39	0,0				
legt sofort auf	3.231	3,1				
Zugehörigkeit zur Zielgruppe bekannt	20.372	19,7	20.372	100,0		
Nonresponse – nicht befragbar (NR-NA)	61	0,1	61	0,3		
ZP (lt. Auskunft) nicht befragbar/ dauerhaft krank/behindert	61	0,1	61	0,3		
Nonresponse – nicht erreicht (NR-NC)	3.207	3,1	3.207	15,7		
definitiver Termin	232	0,2	232	1,1		
vager Termin	2.957	2,9	2.957	14,5		
Zielperson in Feldzeit nicht zu erreichen	18	0,0	18	0,1		

²⁵ Analog zum Vorgehen bei der Basisstichprobe wurde auf Basis des endgültigen Bearbeitungsstatus die Responserate 4 gemäß Standarddefinition (AAPOR 2011, S.44f.) berechnet. Für die Aufstockungsstichprobe errechnet sich eine Responserate 4 von 20,9 Prozent.

	Anzahl	%	Anzahl	%	Anzahl	%
	I	II	III	IV	V	VI
Bruttostichprobe	103.590	100,00				
Zielperson/Haushalt erreicht und befragbar					17.104	100
Nonresponse – Verweigerung (NR-R)	12.483	12,1	12.483	61,3	12.483	73,0
Zielperson verweigert grundsätzlich	3.699	3,6	3.699	18,2	3.699	21,6
Zielperson verweigert: keine Zeit/dauert zu lange	928	0,9	928	4,6	928	5,4
ZP verweigert: will nur persönlich befragt werden	715	0,7	715	3,5	715	4,2
ZP verweigert krank	242	0,2	242	1,2	242	1,4
kein Zugang zu ZP/TN untersagt/lt. Auskunft nicht bereit	165	0,2	165	0,8	165	1,0
KP verweigert jegliche Auskunft	1.660	1,6	1.660	8,1	1.660	9,7
ZP verweigert: kein Interesse, wegen des Themas	3.436	3,3	3.436	16,9	3.436	20,1
ZP verweigert: Datenschutzgründe/ zu persönlich	181	0,2	181	0,9	181	1,1
ZP verweigert: sonstige Gründe	1.457	1,4	1.457	7,2	1.457	8,5
realisierte Interviews (I)	4.621	4,5	4.621	22,7	4.621	27,0
Interviews vollständig realisiert, ausgeliefert	4.417	4,3	4.417	21,7	4.417	25,8
Interviews vollständig realisiert, nicht auswertbar	2	0,0	2	0,0	2	0,0
Interviews unvollständig realisiert	202	0,2	202	1,0	202	1,2

Quelle: Rücklaufdatenbank, eigene Berechnung

In Tabelle 13 ist die Anzahl der auswertbaren Interviews für jedes der fünf Bundesländer der Aufstockungsstichprobe aufgeführt. In allen Fällen wurde die angezielte Anzahl Interviews übertroffen. Insgesamt wurden 69 Interviews mehr als geplant realisiert und an das BKA übergeben.

Tabelle 13 Verteilung der auswertbaren Interviews auf die fünf Bundesländer

Bundesland	Interviews	geplante Interviews	Differenz
Hamburg	1.355	1345	+ 10
Hessen	799	785	+ 14
Berlin	730	721	+ 9
Brandenburg	1.092	1060	+ 32
Sachsen	441	437	+ 4
gesamt	4.417	4.348	+ 69

Quelle: Rücklaufdatenbank, Stichprobenplan

2.7.3 Onomastische Stichprobe

Anders als bei den beiden zuvor berichteten Stichproben handelte es sich bei der onomastischen Stichprobe um eine reine Listenauswahl (vgl. Abschnitt 1.2.3). Die Wahrscheinlichkeit, dass sich hinter einer Telefonnummer ein geschalteter Anschluss befindet, ist daher weitaus höher als bei vollständig generierten Rufnummern, wie sie bei der Basis- und Aufstockungsstichprobe eingesetzt wurden. Der Anteil Rufnummern, die zu keinem Anschluss führten, liegt mit 22 Prozent der Bruttoeinsatzstichprobe erwartungsgemäß deutlich unter dem Anteil, wie er

in den beiden vorangegangenen Abschnitten berichtet wird (vgl. Tabelle 14). Der Anteil von Anschlüssen, die zu keinem Privathaushalt führten, liegt bei sechs Prozent. Insgesamt führten 31 Prozent (entspricht $n = 2.762$ Rufnummern) der Bruttoeinsatzstichproben (8.945 Telefonnummern) zu keinem Zielgruppenhaushalt. Der Anteil Rufnummern, bei denen nicht geklärt werden konnte, ob sich dahinter ein Zielhaushalt befindet, liegt mit acht Prozent auf dem gleichen Niveau wie bei der Aufstockungsstichprobe.

Im Vergleich zur Basis- und Aufstockungsstichprobe liegt bei der onomastischen Stichprobe der Anteil derer, die sofort auflegen, mit 7,3 Prozent mehr als doppelt so hoch. Hier zeigt sich bereits, dass die Gewinnung der Zielpersonen bei Haushalten mit türkischem Migrationshintergrund erkennbar schwieriger war als in den beiden anderen Stichproben. In 4.483 Fällen konnte mit einer Ziel- oder Kontaktperson geklärt werden, ob die Zielperson an einem Interview teilnimmt. In 948 Fällen stimmte die Zielperson dem Interviewstart zu. 138 Mal konnte das Interview nicht zu Ende geführt werden.

Mit Hilfe der onomastischen Stichprobe sollte der Anteil Personen mit türkischem Migrationshintergrund in der Gesamtstichprobe aufgestockt werden. Anfang November lagen rund 800 realisierte Interviews vor. Zusammen mit den Interviews aus der Basisstichprobe summierten sich die insgesamt vorliegenden Interviews mit Personen mit türkischem Migrationshintergrund auf rund 1.300. Damit war bereits das angestrebte Ziel erreicht, eine Repräsentanz der Personen mit türkischem Migrationshintergrund in der Stichprobe entsprechend dem Anteil in der Gesamtbevölkerung zu erlangen. Eine weitere Ausschöpfung der Stichprobe wurde nicht weiter verfolgt. Vor dem Hintergrund, dass der Projektabschluss beim BKA keine Verlängerung der Feldzeit erlaubte und weitere Interviews keinen zusätzlichen Nutzen bezgl. einer ausreichenden Repräsentanz in der Stichprobe (s.o.) bedeuteten, entschied das BKA, dass die Feldarbeit für die Onomastikstichprobe am 03.11.2012 abgeschlossen werden sollte, auch wenn die ursprünglich angezielte Fallzahl von 1.000 Interviews noch nicht erreicht war. Nach Abschluss der Feldarbeit wurden dem BKA 808 formalgeprüfte, auswertbare Interviews übergeben²⁶.

²⁶ Auch für die onomastische Stichprobe wurde die Responserate 4 gemäß AAPOR-Standarddefinition berechnet. Der Anteil liegt bei 15,9 Prozent.

Tabelle 14 Feldergebnis der onomastischen Stichprobe nach AAPOR-Klassifikation

	Anzahl	%	Anzahl	%	Anzahl	%
	I	II	III	IV	V	VI
Bruttostichprobe	8.945	100,00				
nicht Zielgruppe (NE)	2.762	30,9				
kein Anschluss	1.989	22,2				
Anschluss nur Fax/Modem	99	1,1				
kein Privathaushalt unter diesem Anschluss	534	6,0				
nicht in Zielgruppe lt. Angabe im Kontaktgespräch	41	0,5				
bereits befragt	39	0,4				
keine Verständigung in einer der drei Zielsprachen möglich	60	0,7				
unbekannte Auswahl (UE)	687	7,7				
nicht abgehoben	38	0,4				
legt sofort auf	649	7,3				
Zugehörigkeit zur Zielgruppe bekannt	5.496	61,4	5.496	100,0		
Nonresponse – nicht befragbar (NR-NA)	7	0,1	7	0,1		
ZP (lt. Auskunft) nicht befragbar/dauerhaft krank/behindert	7	0,1	7	0,1		
Nonresponse – nicht erreicht (NR-NC)	1.006	11,2	1.006	18,3		
definitiver Termin	28	0,3	28	0,5		
vager Termin	962	10,8	962	17,5		
Zielperson in Feldzeit nicht zu erreichen	16	0,2	16	0,3		
Zielperson/Haushalt erreicht und befragbar					4.483	100
Nonresponse – Verweigerung (NR-R)	3.535	39,5	3.535	64,3	3.535	78,9
Zielperson verweigert grundsätzlich	1.399	15,6	1.399	25,5	1.399	31,2
Zielperson verweigert: keine Zeit/dauert zu lange	319	3,6	319	5,8	319	7,1
ZP verweigert: will nur persönlich befragt werden	55	0,6	55	1,0	55	1,2
ZP verweigert krank	57	0,6	57	1,0	57	1,3
kein Zugang zu ZP/TN untersagt/lt. Auskunft nicht bereit	100	1,1	100	1,8	100	2,2
KP verweigert jegliche Auskunft	558	6,2	558	10,2	558	12,4
ZP verweigert: kein Interesse, wegen des Thema	675	7,5	675	12,3	675	15,1
ZP verweigert: Datenschutzgründe/zu persönlich	23	0,3	23	0,4	23	0,5
ZP verweigert: sonstige Gründe	349	3,9	349	6,4	349	7,8
realisierte Interviews(I)	948	10,6	948	17,2	948	21,1
Interviews vollständig realisiert, ausgeliefert	808	9,0	808	14,7	808	18,0
Interviews vollständig realisiert, nicht auswertbar	2	0,0	2	0,0	2	0,0
Interviews unvollständig realisiert	138	1,5	138	2,5	138	3,1

Quelle: Rücklaufdatenbank, eigene Berechnung

2.8 Interviewdauer

Die durchschnittliche Dauer der CATI-Interviews beträgt für die n=35.503 auswertbaren Interviews knapp 20 Minuten²⁷. Für die Basisstichprobe liegt der Wert bei 19,8 Minuten (vgl. Tabelle 15). Die durchschnittliche Dauer der Interviews der Länderaufstockung ist mit 18 Minuten etwas kürzer, da für diese Teilstichprobe keine Fragenmodule zum Einsatz kamen.

Die fremdsprachigen Interviews dauerten meist länger als die deutschsprachigen. Dies hatte sich bereits im Pretest gezeigt. Für die türkischsprachigen Interviews der Basis- und Aufstockungsstichprobe wurden im Durchschnitt 30 Minuten benötigt, bei den russischsprachigen Interviews liegt der Wert bei 28 Minuten.

Die Interviews der onomastischen Stichprobe wurden gut zur Hälfte auf Deutsch und in einem etwas geringeren Umfang auf Türkisch geführt. Die durchschnittliche Interviewdauer lag bei der onomastischen Stichprobe über alle Interviews bei 24 Minuten (vgl. Tabelle 15). Auch hier dauerten die Interviews auf Türkisch (26 Minuten) länger als die Interviews, die in deutscher Sprache geführt wurden (22 Minuten). Nach Auskunft der Interviewer und Supervisoren liegt dies u.a. daran, dass Personen, die kein oder kaum Deutsch sprechen, es nicht gewohnt sind, an Befragungen dieser Art teilzunehmen. Ihnen fehlt offenbar die Übung für das Verhalten in einem sozialwissenschaftlichen Interview.

Tabelle 15 Interviewdauer

	Anzahl Interviews	Durchschnittsdauer in Minuten
Basis	30.278	19,8
– davon auf Türkisch	182	29,8
– davon auf Russisch	369	27,9
Aufstockung	4.417	17,8
Onomastik	808	23,9
gesamt	35.503	19,6

Quelle: Zeitstempel (time stamps) im Fragebogen, eigene Berechnung

2.9 Interviewsprache

Um Personen, die nicht ausreichend Deutsch verstehen und sprechen, auch die Gelegenheit zu geben, am Interview teilzunehmen, wurden die unter Migranten in Deutschland am häufigsten gesprochenen Sprachen Türkisch und Russisch als weitere Interviewsprachen neben Deutsch vorgehalten. Insgesamt wurden 182 Interviews in der Basis- und Aufstockungsstichprobe zusammen auf Türkisch geführt (vgl. Tabelle 16). In doppelt so vielen Fällen (369 Mal) wurde das Inter-

²⁷ Inklusive der Ermittlung der reduzierten Haushaltsgröße im Kontaktgespräch und der Gewinnung der Zielperson.

view in russischer Sprache durchgeführt. Ohne die Vorhaltung der beiden Fremdsprachen wären diese 551 Interviews mit Personen, die der deutschen Sprache nicht (ausreichend) mächtig sind, nicht möglich gewesen. Zusammengekommen wurden 1,6 Prozent der 34.695 Interviews nicht in deutscher Sprache realisiert.

Tabelle 16 Basis- und Aufstockungsstichprobe: Interviewsprache (Prozent in Klammern)

Interviewsprache	Basisstichprobe	Aufstockungsstichprobe	gesamt
Deutsch	29.796 (98,4)	4.348 (98,4)	34.144 (98,4)
Türkisch	165 (0,5)	17 (0,4)	182 (0,5)
Russisch	317 (1,0)	52 (1,2)	369 (1,1)
Interviews gesamt	30.278 (100)	4.417 (100)	34.695 (100)

Quelle: Stichprobeninformation

Weitaus häufiger kam der türkische Fragebogen in der onomastische Stichprobe zum Einsatz. Schließlich bestand diese Zusatzstichprobe mit hoher Wahrscheinlichkeit aus Haushalten, in denen Personen mit türkischem Migrationshintergrund leben. Diese sollten unabhängig von ihren Deutschkenntnissen für ein Interview gewonnen werden. Der Einsatz des fremdsprachigen Fragebogens war für die Realisierung der Interviews in der onomastischen Stichprobe sehr bedeutsam. 42 Prozent der Interviews wurden vollständig und vier Prozent zum Teil auf Türkisch geführt (vgl. Tabelle 17).

Tabelle 17 Onomastische Stichprobe: Interviewsprache

Interviewsprache	Anzahl	Prozent
Deutsch	437	54,1
Türkisch	340	42,1
teils/teils	31	3,8
gesamt	808	100,0

Quelle: Angaben der Interviewer

2.10 Kontakthäufigkeit

Um auch schwer erreichbaren Zielpersonen die Möglichkeit zu geben, an einem Interview teilzunehmen, wurden die Stichproben über die gesamte Feldzeit kontinuierlich kontaktiert. Es wurde darauf geachtet, dass die Rufnummern zu unterschiedlichen Wochentagen und unterschiedlichen Tageszeiten immer wieder angerufen wurden, damit möglichst viele Personen erreicht werden konnten. Die erzielte Ausschöpfung der Stichproben konnte nur durch eine konsequente Kontaktierungsstrategie erreicht werden. Dies belegen die Werte in Tabelle 18. Es war zwar auch möglich, direkt beim Erstkontakt das Interview durchzuführen. Im Durchschnitt über alle auswertbaren Interviews benötigten die Interviewer allerdings sieben Kontakte, um ein Interview zu erhalten. Die maximale

Anzahl Kontakte, die dafür benötigt wurde, liegt bei 50 (vgl. Tabelle 18). Bei den nicht realisierten Interviews liegen die durchschnittliche Kontakthäufigkeit bei knapp neun Versuchen und der Maximalwert bei 52.

Tabelle 18 Kontakthäufigkeit der auswertbaren Interviews

	Anzahl
durchschnittliche Kontakthäufigkeit	6,96
Minimum Kontakthäufigkeit	1
Maximum Kontakthäufigkeit	50
Standardabweichung	5,72
gesamt	35.503

Quelle: Kontaktdatei, eigene Berechnungen

2.11 Fallzahl pro Fragemodul

Im gesamten Fragebogen kamen sieben inhaltlich verschiedene Fragenmodule zum Einsatz. Sie sollten nur einer zufällig ausgewählten Teilstichprobe vorgelegt werden, damit die durchschnittliche Dauer der Befragung im vereinbarten Rahmen von 20 Minuten blieb und gleichzeitig für die spätere Auswertung der Antworten eine Fallzahl zwischen 1.500 und 2.000 als ausreichend erachtet wurde. In der nachfolgenden Tabelle sind die angezielten und die tatsächlich erreichten Fallzahlen pro Modul ausgewiesen.

Tabelle 19 Fallzahl pro Modul in der Basisstichprobe

Modulnr.	realisierte Fallzahl (absolut)	Anteil an Basisstichprobe (in %)	angezielte Fallzahl (absolut)	zusätzliche Fälle (absolut)
1	1.608	5,3	1.500	108
2	1.695	5,6	1.500	195
4	1.649	5,4	1.500	149
5	1.626	5,4	1.500	126
6	2.100	6,9	2.000	100
7*	2.965 2.440	9,8 8,8	--- 2.200	--- 240
8	11.643 11.118**	38,5 37,4	--- 10.200	--- 918

Quelle: auswertbare Interviews, eigene Berechnung

* Modul 7 startet mit einer Filterfrage (Frage 607, Internetnutzung). In der ersten Zeile werden die Fälle ausgewiesen, die die Filterfrage bekommen haben. In der zweiten Zeile sind nur jene Fälle ausgewiesen, die auch die Folgefragen erhalten haben (d.h. nicht überfiltert wurden).

** Für 11.118 Fälle kann Modul 8 zusammen mit einem der Module 1 bis 7 ausgewertet werden. Die Differenz zu den angegebenen 11.643 Fällen ergibt sich auf der Filterung zu Beginn von Modul 7. Vergleiche dazu die Anmerkungen zu Modul 7.

Die angezielten Fallzahlen sind in jedem Modul übertroffen worden. Es konnten zwischen 100 und 240 zusätzliche Interviews pro Modul realisiert werden. Für Modul 8 liegen somit 918 Interviews mehr vor als ursprünglich geplant.

2.12 Einsatz der Vignettensets

Im Modul 6 wurden unterschiedliche Vignettensets eingesetzt (vgl. Abschnitt 1.3). Die Zuordnung zu den Interviews wurde vorab über die Bruttostichprobe organisiert, so dass eine zufällige Einspielung der Vignettensets gewährleistet war. Es liegen zwischen 259 und 359 eingesetzte Sets vor (vgl. Tabelle 20). Insgesamt summieren sich die beantworteten Vignettensets auf 2.100 und damit 100 mehr als ursprünglich geplant. Der Anteil der einzelnen eingesetzten Sets variierte zwischen gut zwölf Prozent (Setnummern 6 und 7) und 17 Prozent für das zweite Vignettenset. Die übrigen Vignettensets haben einen Anteil von 14 bzw. 15 Prozent (rd. 300 Fälle) an der Gesamtanzahl.

Tabelle 20 Vignettensetverteilung

Vignettenset	Anzahl	Prozent
Set 1	326	15,5
Set 2	359	17,1
Set 3	315	15,0
Set 4	291	13,9
Set 5	289	13,8
Set 6	261	12,4
Set 7	259	12,3
Gesamt	2.100	100,0

Quelle: auswertbare Interviews

2.13 Verhältnis Interviews aus der Festnetz- und der Mobilfunkstichprobe in der realisierten Stichprobe

Für die Basisstichprobe war ein Verhältnis der Interviews aus der Festnetzstichprobe und der Mobilfunkstichprobe von 23 zu 7 vorgesehen. Die Einsatzstichproben wurden in diesem Verhältnis ins Feld gegeben (vgl. Abschnitt 1.2.1). Mit 7.385 Interviews konnten 385 Interviews mehr als geplant aus der Mobilfunkstichprobe realisiert werden (vgl. Tabelle 21). Der Anteil liegt bei 24,4 Prozent der realisierten Basisstichprobe. Für die späteren Analysen des BKA und MPI wurden alle drei Teilstichproben zu einer Gesamtstichprobe zusammengeführt. Im integrierten Gesamtdatensatz liegt der Anteil Interviews aus der Mobilfunkstichprobe mit 21 Prozent entsprechend niedriger.

Tabelle 21 Festnetz- und Mobilfunkstichprobe in der Basisstichprobe

Festnetz- und Mobilfunkstichprobe	Anzahl	Prozent
Festnetz	22.893	75,6
Mobilfunk	7.385	24,4
gesamt	30.278	100,0

Quelle: auswertbare Interviews

Tabelle 22 Festnetz- und Mobilfunkstichprobe insgesamt

Festnetz- und Mobilfunkstichprobe	Anzahl	Prozent
Festnetz	28.118	79,2
Mobilfunk	7.385	20,8
gesamt	35.503	100,0

Quelle: auswertbare Interviews

2.14 Interviewabbrüche und Abbruchgründe

Über alle drei Teilstichproben hinweg kam es in 1.986 Fällen zu einem Interviewabbruch. Ein Teil der Abbrüche fand bereits im Kontaktgespräch (Intro) statt, während der Interviewer versuchte, die richtige Zielperson im Haushalt zu ermitteln und sie zu einer Teilnahme an der Befragung zu gewinnen. Die restlichen 1.565 Abbrüche ereigneten sich während des Interviews. 111 Abbrüche erfolgten in der onomastischen Stichprobe und 1.454 in den beiden anderen Teilstichproben (Basis- und Aufstockungstichprobe). Im Folgenden wird dargestellt,

- inwieweit die Interviewabbrüche sich an bestimmten Stellen im Fragebogen häufen oder sie sich eher unspezifisch über den gesamten Bogen verteilen,
- inwieweit es Abbruchgründe gibt, die besonders häufig genannt wurden und/oder möglicherweise mit Besonderheiten des Fragebogens zusammenhängen,
- ob in diesem Zusammenhang Unterschiede zwischen der Basis- und Aufstockungstichprobe auf der einen Seite und der onomastischen Stichprobe auf der anderen Seite zu finden sind?

Die Abbruchquoten (Anteil der abgebrochenen Interviews an der Gesamtzahl vollständig realisierter Interviews plus abgebrochene Interviews) unterscheiden sich zwischen der Basis- und Aufstockungstichprobe und der onomastischen Stichprobe spürbar²⁸. Während die Abbruchquote in der Basis- und Aufstockungstichprobe bei fünf Prozent liegt, bewegt sie sich in der onomastischen Stichprobe mit 17 Prozent deutlich darüber.

Die mit Abstand meisten Interviewabbrüche ereignen sich zu Beginn des Interviews bei den Fragen zu Gesundheit, Lebenszufriedenheit und Vertrauen (Frageblock 100, vgl. Tabelle 23). In der Basis- und Aufstockungstichprobe entfallen 33 Prozent aller Abbrüche auf die ersten Fragen nach dem Interviewstart. Bei

²⁸ Es gibt in der Basisstichprobe keine nennenswerten Unterschiede zwischen den Abbrüchen der Mobilfunk- und Festnetzstichprobe. Die Verteilungen der Abbruchstellen sind von der Struktur her identisch. Die Abbrüche verteilen sich in ähnlichem Umfang über den gesamten Fragebogen.

der onomastischen Stichprobe liegt der Anteil mit 41 Prozent noch etwas höher (vgl. Tabelle 23). Die im Vergleich zu anderen Stellen im Fragebogen hohe Abbruchquote zu Beginn des inhaltlichen Interviews lässt sich auf unterschiedliche Faktoren zurückführen. Dies zeigen die von den Interviewern notierten Erläuterungen zum Abbruchgrund. Sie geben wichtige Hinweise auf die Gesprächssituation und eine Begründung für das vorzeitige Ende der Befragung.

Bevor die einzelnen Abbruchgründe diskutiert werden, sei angemerkt, dass der Start der Befragung eine sehr entscheidende Rolle für die Realisierungschancen sowie den gesamten Verlauf eines Interviews spielt. Zu Beginn muss der Interviewer ein gewisses Maß an Vertrauen aufbauen, damit die Zielperson bereit ist, auf die vielfach persönlichen Fragen zu antworten. Insbesondere bei zweifelnden Personen, die zur Nichtteilnahme tendieren, schlägt der Interviewer oftmals vor, erst einmal mit dem Interview zu beginnen, damit sich die Zielperson ein besseres Bild über die Inhalte der Befragung machen kann. Zum Teil gelingt es dann, das Interview bis zum Ende durchzuführen. In anderen Fällen entscheiden sich die Zielpersonen nach den ersten Fragen doch gegen eine Fortführung des Interviews. Zudem kann in einigen Fällen erst nach den ersten inhaltlichen Fragen festgestellt werden, ob die Zielperson dem Interview geistig, sprachlich oder trotz eingeschränkter Gesundheit überhaupt folgen kann. So stehen am Anfang eines jeden Interviews die Fragen nach der Befragbarkeit der Zielperson und ihrem Interesse, an der Befragung teilzunehmen bzw. die Akzeptanz der Fragen auf Seiten der Zielperson im Raum.

Neben den Abbrüchen zu Beginn der Befragung verteilen sich die übrigen Abbruchstellen über den gesamten Fragebogen. Sie schwanken meistens zwischen 0,1 und fünf Prozent der Abbrüche und erreichen nur einmal einen Anteil von zehn Prozent.

Leicht erhöhte Abbruchquoten finden sich in allen Teilstichproben bei den Fragen zum allgemeinen und beruflichen Bildungsabschluss und zur Erwerbstätigkeit (rund fünf Prozent) und der höchste Wert mit gut zehn Prozent bei Fragen zur Wohngegend. Jeweils rund fünf Prozent der Abbrüche entfallen für die Basis- und Aufstockungsstichproben auf den Screener-Block 900 mit den Fragen zum Wohnungseinbruch und bei der onomastischen Stichprobe auf die Fragen zum Betrug mit Zahlungskarten ebenfalls im 900er Block sowie auf die abschließenden Fragen zur Soziodemografie. Allerdings stehen hinter diesen Prozentangaben in der onomastischen Stichprobe nur sehr kleine Fallzahlen von sechs bzw. sieben Personen.

Die Durchsicht der Interviewerangaben zu den Abbruchgründen ergab eine Vielschichtigkeit von Gründen für ein vorzeitiges Ende des Interviews. Insbesondere wurde zu Beginn des Interviews deutlich, dass die Zielperson aus alters- oder gesundheitlichen Gründen dem Interview nicht folgen konnte und es dadurch zu einem Abbruch kam. Zudem entschieden sich Zielpersonen nach den ersten inhaltlichen Fragen gegen die Fortführung des Interviews, weil ihnen das Thema nicht zusagte, ihnen das Interview zu lange dauerte, sie keine Zeit oder allgemein kein besonderes Interesse hatten.

Im gesamten Interviewverlauf kam es auch mehrfach vor, dass die Zielperson einfach auflegte, ohne dass dem Interviewer ein Grund für das abrupte Ende des Gesprächs genannt wurde. Der Versuch, das Telefonat erneut aufzunehmen, scheiterte in diesen Fällen.

Darüber hinaus berichteten Interviewer, dass es zu Interviewabbrüchen durch Dritte kam, weil diese Person der Zielperson die Fortführung des Interviews untersagte. Diese Abbrüche durch Dritte sind von den Interviewern im Vergleich zur Basis- und Aufstockungstichprobe häufiger in der onomastischen Stichprobe notiert worden. Hier wurde vermehrt jüngeren Zielpersonen durch ihre Eltern oder weiblichen Zielpersonen durch ihre Ehemänner eine Fortführung des Interviews untersagt. Zudem berichten die Interviewer, dass in der onomastischen Stichprobe ein Teil der Zielpersonen offenbar keine Erfahrungen mit Befragungen hatte, dem Wechsel von Fragen und Antworten nicht folgen konnte und aus diesem Grund das Interview nicht fortgeführt werden konnte.

Zusammenfassend lässt sich feststellen, dass gehäuft Abbrüche zu Beginn des Interviews stattgefunden haben. An dieser Stelle des Interviews musste oftmals erst geklärt werden, ob die Zielperson tatsächlich befragbar ist bzw. ob die Befragung auf Akzeptanz bei der Zielperson stößt. Im weiteren Verlauf des Interviews fällt nur eine Stelle auf, die eine erhöhte Abbruchquote aufweist: Es handelt sich um die Fragen zur Wohngegend und zum Zusammenleben in der Wohngegend. In allen Teilstichproben haben mehr Personen an dieser Stelle das Interview abgebrochen als an anderen Stellen im Interview. Die notierten Abbruchgründe unterscheiden sich jedoch nicht von den genannten Gründen an anderen Abbruchstellen.

Zu den am häufigsten genannten Abbruchgründen zählen:

- „einfach aufgelegt“,
- „Zielperson hatte keine Lust/Zeit mehr“,
- „Zielperson verstand Frage nicht“,
- „Zielperson ist zu alt“,
- „Zielperson spricht nicht ausreichend Deutsch“,
- „Zielperson ist gesundheitliche nicht in der Lage, das Interview durchzuführen“
- „Zielperson waren die Fragen zu persönlich“.

Die breite Streuung der Abbrüche über den gesamten Fragebogen und die von den Interviewern notierten Gründe für das vorzeitige Ende der Befragung deuten darauf hin, dass es keine Stelle im Fragebogen gibt, die aufgrund der Fragestellung zu einer deutlich höheren Abbruchquote geführt hat.

Tabelle 23 Abbruchstellen im Fragebogen

Fragebogenabschnitt	Basis- und Aufstockungsstichprobe		Onomastische Stichprobe	
	Anzahl	Prozent	Anzahl	Prozent
Frage 1-1a (Alter)	10	0,7	-	-
Block 100	480	33,0	46	41,4
Modul 6	22	1,5	3	2,7
Modul 6 - Vignetten	20	1,4	-	-
Modul 2	13	0,9	1	0,9
Modul 5	28	1,9	-	-
Allgemeiner u. beruflicher Bildungsabschluss/Erwerbsstatus	76	5,2	5	4,5
Wohngegend Soziodemografie	148	10,2	12	10,8
Modul 4	12	0,8	1	0,9
Ausgehverhalten	19	1,3	1	0,9
Familienstand	45	3,1	4	3,6
Internetnutzung	14	1,0	1	0,9
Modul 7	4	0,3	1	0,9
Kriminalitätsfurcht	16	1,1	-	-
Modul 8	67	4,6	5	4,5
Modul 1	1	0,1	-	-
Fortsetzung Modul 8	46	3,2	3	2,7
Block 800 KFZ	20	1,4	1	0,9
Block 800 KB	51	3,5	2	1,8
Screenener – Block 900 W	72	5,0	2	1,8
Screenener – Block 900 VW	20	1,4	1	0,9
Screenener – Block 900 KFZ oder B	29	2,0	1	0,9
Screenener – Block 900 KB, IT oder R	36	2,5	7	6,3
Screenener – Block 900 D	35	2,4	1	0,9
Screenener – Block 900 KV	67	4,6	5	4,5
Viktimsierungsblöcke	47	3,2	2	1,8
Soziodemografie	56	3,9	6	5,4
gesamt	1.454	100,0	111	100,0

Quelle: abgebrochene Interviews, eigene Berechnungen

3 Datenprüfung und –gewichtung

3.1 Datenprüfung und –aufbereitung

Die ersten Schritte der Datenprüfung begannen bereits vor der eigentlichen Feldphase. Nachdem der abgestimmte Fragebogen programmiert war, wurde die korrekte Ablage der Daten geprüft. Dabei wurde der gesamte Filterverlauf im programmierten Instrument nachvollzogen. Für alle Fragen wurde zudem auf die Einhaltung des vorgesehenen Wertebereichs geachtet, damit bei geschlossenen Skalen lediglich die in der Fragebogenvorlage vorgesehen Codes auftreten können. Bei offenen Zahlenangaben durften die erhobenen Werte nur im vorgegebenen Range auftreten.

Auch die Richtigkeit der Steuerung einzelner Fragenmodule, die kombinierte Schaltung der Module (sofern vorgesehen) und die Zuweisung der Vignetten im Modul 6 wurden in der Datenprüfung nachvollzogen und auf ihre Regelkonformität hin geprüft. Diese intensiven Prüfungen der korrekten Datenablage gewährleisten eine fehlerfreie technische Erhebung der Daten im Feld. Diese Prüfungen erfolgten vor Feldstart anhand von speziell dafür eingegebenen Testfällen und wurden anhand der realisierten Fälle sowohl nach dem Pretest als auch zu Beginn des Hauptfelds fortgesetzt.

In einem weiteren Schritt wurden die Häufigkeitsauszählungen durchgesehen. Bei den Fragen nach offenen Zahlenangaben (z.B. Anzahl Haushaltsmitglieder, Häufigkeit der einzelnen Opfererfahrungen, Anzahl Telefonnummern im Haushalt) kam es in sehr wenigen Fällen zu vergleichsweise hohen Angaben. Nach Rücksprache mit dem BKA werden diese Angaben im Rahmen der Auswertung im BKA auf ihre Plausibilität hin geprüft und bei Bedarf bereinigt.

An den im Folgenden beschriebenen Variablen sind in enger Abstimmung mit dem BKA Recodierung vorgenommen worden, die jeweils in einer neuen Variable abgelegt wurden. Die Ursprungsangabe der Zielperson ist in allen Fällen in der Originalvariable enthalten geblieben, so dass es zu keinem Informationsverlust gekommen ist.

Angabe zum Geburtsland

In den Fragen zum Geburtsland der Zielperson (SD12/SD13), der Mutter (SD15/SD16) und des Vaters der Zielperson (SD18/SD19) haben die Zielpersonen in einigen wenigen Fällen erst bei der offenen Nachfrage zum Geburtsland angegeben, dass ein Elternteil in Deutschland geboren wurde. Für diese Fälle wurden die Korrekturvariablen SD12k/SD13k, SD15k/SD16k sowie SD18k/SD19k angelegt. In der Eingangsfrage SD12k (bzw. SD15k oder SD18k) wurden diese Fälle auf Code 1 „Bundesrepublik Deutschland oder DDR“ umgesetzt. Die Angabe aus der Nachfrage SD13k (bzw. SD16k oder SD19k) wurde gelöscht. Die Originalvariablen SD12, SD13, SD15, SD16, SD18 und SD19 blieben hingegen unverändert. Die betroffenen Fälle können also durch den Abgleich der Original- und der Korrekturvariablen nachvollzogen werden. Die kategorisierten Variablen zum Geburtsland SD12t, SD15t und SD18t basieren auf den Korrekturvariablen.

Recodierung beim Ort, wo sich Vorfall ereignete

Für die Fälle, in denen die Zielperson nicht angeben konnte, in welchem Bundesland sie Opfer einer Straftat geworden war, wurde mit einer offenen Frage die Ortschaft oder die Stadt ermittelt. Diese Ortsangaben konnten im Rahmen der Datenaufbereitung eindeutig einzelnen Bundesländern zugeordnet werden (vgl. Tabelle 24). Nach Rücksprache mit dem BKA wurden die Recodierungen im Datensatz vorgenommen. Auch diese Umsetzungen können in den Daten nachvollzogen werden. Das nachfolgende Beispiel illustriert die Vorgehensweise, die bei allen entsprechenden Fragen analog angewendet wurde:

Wenn einer Zielperson lediglich einmal das Fahrrad gestohlen wurde und sie in Frage FD6 das Bundesland nicht angegeben hat (Code 97 „verweigert“/98 „weiß nicht“), dann wurde ihr die Nachfrage zur Stadt bzw. Ortschaft des Vorfalls (hier FD7) gestellt. Sofern die Zielperson hier einen Ort nennen konnte, ist in FD7 der Code 1 „offene Nennung erfassen“ abgelegt. Es gilt also allgemein, dass die Nachfrage zur Stadt bzw. Ortschaft nur gestellt wird, wenn in der vorherigen Abfrage des Bundeslands einer der Codes 97 „verweigert“ oder 98 „weiß nicht“ vorlag. Bei den nachcodierten Fällen wird hiervon abgewichen. Bei der Nachcodierung wurde in der Frage nach dem Bundesland von den Codes 97 oder 98 auf einen der Codes 1 bis 16 recodiert. Die Nachfrage nach der Stadt bzw. Ortschaft wurde in diesen Fällen nicht gelöscht. Nachcodierte Angaben zum Bundesland, in dem sich ein Vorfall ereignet hat, lassen sich also allgemein daran erkennen, dass in der Frage zum Bundesland einer der Antwortcodes 1 bis 16 vorliegt und in der Nachfrage zur Stadt bzw. Ortschaft der Code 1 „offene Nennung erfassen“ vorliegt.

Bei der Nachfrage zur Stadt bzw. Ortschaft beim Delikt „Waren- oder Leistungsbetrug“ (nur ein berichteter Vorfall, Variable B9) wurde in mehreren Fällen angegeben, dass diese Frage so nicht beantwortet werden kann, weil sich der Vorfall über das Internet ereignet hat. Diese Fälle wurden in Variable B9 mit dem Sondercode 2 „offene Nennung: über das Internet“ gekennzeichnet.

Tabelle 24 Vercodung der offenen Ortsangaben

Ortsangabe aus dem Interview	Vercodetes Bundesland
Aachen	NRW
Bayreuth	Bayern
Berlin	Berlin
Düsseldorf	NRW
Emden	Niedersachsen
Estorf	Niedersachsen
Frankfurt	Hessen
Gelsenkirchen	NRW
Hamburg	Hamburg
Hamm	NRW
Köln	NRW
Konstanz	Baden-Württemberg
Köpenick	Berlin
Langeoog (Insel)	Niedersachsen
Lüneburg	Niedersachsen
Mannheim	Baden-Württemberg
Neckar-Odenwald-Kreis	Baden-Württemberg
Schopfheim	Baden-Württemberg
Schwerin	Mecklenburg-Vorpommern
Stade	Niedersachsen
Werne	NRW
Westoverledingen	Niedersachsen
Wismar	Mecklenburg-Vorpommern
Wolfsburg	Niedersachsen
Wurmlingen	Baden-Württemberg
Zirndorf	Bayern
Gemeinde liegt im Landkreis Barnim*	Brandenburg
Gemeinde liegt im Kreis Nordfriesland*	Schleswig-Holstein

* Da es sich bei der Angabe um einen sehr kleinen Ort handelt, wird an dieser Stelle der Kreis aufgeführt, in dem sich die Gemeinde befindet. Somit ist eine Re-Identifikation der Zielperson nicht möglich.

Negative Missingcodes

In allen Fragen, bei denen Länderlisten hinterlegt waren, um eine schnelle Erfassung der Nennungen zu gewährleisten, wurden die Missingcodes für „weiß nicht“ (98) und „verweigert“ (97) mit einem negativen Vorzeichen versehen. Dies war notwendig, da die Codes 97 und 98 zu den gültigen Codes der 321 Staaten und Territorien in der Länderliste zählen. Die Codierung mit „-97“ bzw. „-98“ betrifft die Variablen SD11, SD11b, SD13, SD16 und SD19.

Darüber hinaus werden bei den Fragen zur Staatsangehörigkeit SD11 und SD11b zwei weitere Sondercodes verwendet. Es handelt sich dabei um die Codes „-96“ für „ohne Staatsangehörigkeit“ und der im Fragebogen nicht aufgeführte Code „-95“ „nicht in Liste genannte Staatsangehörigkeit“.

Abgleich Zuzugsjahr mit Geburtsjahr

Am Ende des Fragebogens wurden die Zielpersonen, die nicht in Deutschland geboren wurden, gefragt, in welchem Jahr sie nach Deutschland (heutige Gebiet der Bundesrepublik) zugezogen sind (Frage SD14). Es wurde eine Plausibilitätsprüfung vollzogen, indem geprüft wurde, ob das Zuzugsjahr vor oder nach dem Geburtsjahr der Zielperson liegt. Da für die Zielpersonen nur eine Angabe zum Alter und nicht zum Geburtsjahr vorliegt, wurde eine entsprechende Umrechnung aufgrund der Angabe zum Alter vorgenommen. Der Einfachheit halber wurde bei allen Zielpersonen unterstellt, dass diese im Befragungsjahr noch nicht Geburtstag hatten. Für eine Zielperson, die ihr Alter mit 20 Jahren angegeben hat, wurde beispielsweise das Geburtsjahr 1991 angenommen. Sofern nur eine Angabe der Altersgruppe aus Frage 1a vorliegt, wurde ein Alter am oberen Ende der Skala für die Prüfung berücksichtigt²⁹. Für Personen mit der Altersgruppenangabe „75 Jahre und älter“ wurde auf einen Abgleich der Altersangabe mit der Antwort aus Frage SD14 verzichtet. Unter den genannten Rahmenbedingungen ergab sich lediglich für zehn Fälle eine Situation, bei der das Zuzugsjahr vor dem errechneten Geburtsjahr lag. Die Entscheidung über den Umgang mit diesen wenigen Fällen obliegt dem BKA. Daher wurden keine Änderungen im Datensatz vorgenommen.

Zuspielung der Kreiskennziffer (KKZ)

Um regionale Einflüsse bei der Datenanalyse berücksichtigen zu können, wurde die Kreiskennziffer (KKZ) den Befragungsdaten – sofern möglich - zugespielt. Das im Folgenden beschriebene Verfahren war sehr erfolgreich. Es konnte für 98,2 Prozent der Interviews eine KKZ und zugleich auch eine Bundeslandkennung zugespielt werden. Die KKZ wurde für alle Fälle aus den Festnetzstichproben über die Stichprobe anhand der Vorwahl zugespielt. Für die Fälle der Mobilfunkstichprobe hingegen liegt in der Stichprobe keine Information zur KKZ vor. Für die realisierten Fälle der Mobilfunkstichprobe wurde die KKZ aus der am Ende

²⁹ Für eine Zielperson, die ihr Alter beispielsweise mit „zwischen 25 und 34 Jahren“ angegeben hat, wurde ein Alter von 34 Jahren und somit das Geburtsjahr 1977 angenommen.

der Befragung von der Zielperson genannten Postleitzahl ihres Wohnorts generiert. Voraussetzung hierfür war, dass die Zielperson die vollständige fünfstellige Postleitzahl genannt hat. Erfahrungsgemäß sind nicht alle genannten Postleitzahlen korrekt. Gründe für die Nennung von Postleitzahlen, die sich als nicht existent herausstellen, können Erinnerungsfehler der Zielperson, genauso wie verdeckte Verweigerungen der Angabe, und damit verbundene fehlerhafte Angaben der Postleitzahl sein. Im letzten Fall werden dann fiktive Postleitzahlen genannt. Fälle, bei denen sich eine von einer Zielperson der Mobilfunkstichprobe genannte fünfstellige Postleitzahl bei der Generierung der KKZ als fehlerhaft herausgestellt hat, wurden entsprechend gekennzeichnet. Die Kennung kann der Variable „plzang“ entnommen werden.

3.2 Gewichtung und Hochrechnung

Die Berechnung der Gewichte und Hochrechnungsfaktoren erfolgte in einem mehrstufigen Prozess. Dabei war zu berücksichtigen, dass die gesamte realisierte Endstichprobe aus unterschiedlichen Auswahlrahmen (Frames) stammte, die durch einen gemeinsamen Gewichtungsfaktor in eine Stichprobe integriert werden mussten. Grundlegend war die Entscheidung, die Basisstichprobe, bestehend aus einer Festnetzstichprobe und einer Mobilfunkstichprobe, die Aufstockungsstichproben aus fünf Bundesländern und die onomastische Stichprobe von Personen mit türkischem Migrationshintergrund in eine gesamte Stichprobe zu integrieren.

Die Auswahlrahmen waren:

- ADM-Telefonnummerndatei Festnetz gesamt
- ADM-Telefonnummerndatei Festnetz für die fünf Aufstockerbundesländer
- ADM-Telefonnummerndatei Mobilfunk
- Telefonbucheinträge Festnetz nach onomastischen Gesichtspunkten für Personen mit türkischem Migrationshintergrund³⁰.

Zudem erfolgte eine Umgewichtung der Stichproben für Personen und für Haushalte. Während es sich bei einer Festnetzstichprobe um eine Haushaltsstichprobe handelt, lässt sich eine Mobilfunkstichprobe als Personenstichprobe auffassen. Alle Gewichtungsschritte und die Berechnung der Hochrechnungsfaktoren erfolgten also getrennt für Personen und Haushalte. Die personenbezogenen Fragen können mit dem Personengewicht ausgewertet und mit dem Personenhochrechnungsfaktor hochgerechnet werden auf alle Personen in der Grundgesamtheit. Die (wenigen) haushaltsbezogenen Fragen (z.B. Fahrraddiebstahl oder Kfz-Diebstahl) können mit dem Haushaltsgewicht ausgewertet und dem Haushaltshochrechnungsfaktor auf die Grundgesamtheit aller Haushalte (mit mindestens einem 16-jährigen oder älteren Haushaltsmitglied) hochgerechnet werden.

³⁰ Die Stichprobe wurde gezogen von Humpert & Schneiderheinze GbR.

Die einzelnen Gewichtungsschritte erfolgten jeweils getrennt für Haushalte und Personen und lassen sich wie folgt kurz beschreiben:

1. Bestimmung der Auswahlwahrscheinlichkeit für die Dual-Frame Basisstichprobe und die Festnetz-Aufstockerstichproben.
2. Integration von Basisstichprobe und Aufstockerstichproben über die Bestimmung der gemeinsamen Auswahlwahrscheinlichkeit.
3. Proportionalisierung des Dual-Use Anteils (sowohl über Festnetz als auch über Mobilfunk erreichbar) in der Festnetz- und der Mobilfunkstichprobe für die integrierte Stichprobe (Adjustierung der Dual-Frame-Gewichtung).
4. Bestimmung der Auswahlwahrscheinlichkeit für die Onomastikstichprobe.
5. Integration der Onomastikstichprobe über eine Konvexverbindung mit türkischen Migranten in der integrierten Basis- plus Aufstockerstichprobe.
6. Trimmen der integrierten adjustierten Designgewichte.
7. Kalibrierung der integrierten Stichprobe mittels GREG (generalized regression estimator).
8. Fallzahlnormierung der Gewichte und Berechnung der Hochrechnungsfaktoren für die Grundgesamtheit.
9. Fallzahlnormierung und, falls notwendig, Kalibrierung für die einzelnen Bundesländer mit dem kalibrierten Endgewicht aus Schritt 7 als Ausgangsgewicht.

Diese Vielzahl an notwendigen Gewichtungsschritten und die Mehrzahl an Auswahlrahmen sowie die Integration der disproportionalen Stichproben in eine gemeinsame Stichprobe führt im Ergebnis zu einer großen Varianz der Gewichtungsfaktoren und mithin einer geringen Effizienz der Gewichte. Aufgrund der sehr hohen Fallzahl der realisierten Stichprobe bewegen sich allerdings die korrigierten Stichprobenfehler immer noch in engen Grenzen.

3.2.1 Bestimmung der Auswahlwahrscheinlichkeit für die Dual-Frame-Basisstichprobe und die Aufstockerstichproben

Für die Basisstichprobe wurden Telefonnummern aus zwei Auswahlrahmen gezogen, nämlich aus der „ADM-Auswahlgrundlage für Telefonstichproben“ (Festnetz) und der „ADM-Auswahlgrundlage für Mobilfunkstichproben“ (Mobilfunk). Die beiden Stichproben lassen sich in Anlehnung an das von Gabler/Häder³¹ vorgeschlagene Verfahren grundsätzlich so zusammenführen, dass

³¹ Hoffmann, H. (2007): Kombinierte Stichproben für Telefonumfragen - Ansätze in Europa. In: S. Gabler und S. Häder (Hg.): Mobilfunktelefonie - Eine Herausforderung für die Umfrageforschung. ZUMA-Nachrichten Spezial Band 13. Mannheim, S. 47-57; Häder, S., Häder, M. (Hrsg.): Telefonbefragungen über das Mobilfunknetz: Konzept, Design und Umsetzung einer Strategie zur Datenerhebung. Wiesbaden: VS Verlag für Sozialwissenschaften 2009.

die beiden Auswahlwahrscheinlichkeiten (Auswahlwahrscheinlichkeit in Festnetzstichprobe und Auswahlwahrscheinlichkeit in Mobilfunkstichprobe) addiert werden. Die gemeinsame Auswahlwahrscheinlichkeit (in beiden Stichproben zu sein) kann an dieser Stelle vernachlässigt werden, da diese bei bundesweiten Stichproben extrem gering ist.

Bestimmende Faktoren für die Auswahlwahrscheinlichkeit sind die Anzahl der Zielpersonen im Haushalt (reduzierte Haushaltsgröße), da bei mehreren Zielpersonen eine Zufallsauswahl (über das Last-Birthday-Verfahren) erfolgte und entsprechend die Auswahlwahrscheinlichkeit bestimmt werden muss. Zudem definiert auch die Anzahl der Festnetznummern sowie der Mobilfunknummern die Auswahlwahrscheinlichkeit eines Haushalts bzw. einer Person. Über je mehr Nummern ein Haushalt bzw. eine Person erreichbar ist, desto größer ist auch die Chance, in die Stichprobe zu gelangen.

Eine Auswahlwahrscheinlichkeit für die Basisstichprobe hatten dabei selbstverständlich nicht nur die Personen bzw. Haushalte der Basisstichprobe, sondern auch die Personen bzw. Haushalte der Aufstockungsstichproben, diese allerdings nur über die Festnetzstichprobe. Umgekehrt hatten auch die Personen bzw. Haushalte der Basisstichprobe über die Festnetzstichprobe eine Auswahlwahrscheinlichkeit > 0 , in einer der Aufstockerstichproben zu sein, sofern sie in einem der Bundesländer mit einer Aufstockungsstichprobe lebten.

Personen

Die Auswahlwahrscheinlichkeit für Personen, in der Basisstichprobe enthalten zu sein, lässt sich für die Basisstichprobe und die Aufstockerstichproben bestimmen über:

$P(\text{Personen, Basis}) = ((k_F * (m_F / M_F) * (1 / z)) + (k_C * (m_C / M_C)))$, wobei:

k_F = Anzahl Festnetznummern, über die der Haushalt erreichbar ist

m_F = realisierte Stichprobe im Festnetz

M_F = Anzahl Festnetznummern im Auswahlrahmen (bereinigt um ungültige Nummern)

z = Anzahl Zielpersonen im Haushalt

k_C = Anzahl Mobilfunknummern, über die der Befragte persönlich erreichbar ist

m_C = realisierte Stichprobe im Mobilnetz

M_C = Anzahl Mobilfunknummern im Auswahlrahmen (bereinigt um ungültige Nummern)

Im Unterschied zur ursprünglichen Gabler/Häder-Formel wurden dabei Modifikationen vorgenommen, die mittlerweile dem ADM-Standard entsprechen, der von infas maßgeblich mitentwickelt wurde³². So wurde der Auswahlrahmen für Festnetz- und Mobilfunknummern um die jeweils nicht existierenden sowie geschäftlichen resp. Faxnummern bereinigt³³.

Da die Aufstockungsstichproben reine Festnetzstichproben sind, lässt sich die Auswahlwahrscheinlichkeit für Personen der Basisstichprobe und der Aufstockungsstichproben, in einer der Aufstockerstichproben zu sein, bestimmen über:

$$P_j(\text{Personen, Aufstockung}) = (k_F * (m_{F,j} / M_{F,j}) * (1 / z))$$

für j = Hamburg, Hessen, Berlin, Brandenburg und Sachsen

Nur die Fälle der Basisstichprobe in den genannten Bundesländern mit einer Aufstockungsstichprobe haben also eine Wahrscheinlichkeit > 0, in einer der Aufstockerstichproben enthalten zu sein.

Haushalte

Analog ergibt sich für die Haushalte die Auswahlwahrscheinlichkeit über:

$$P(\text{Haushalte, Basis}) = (k_F * (m_F / M_F)) + (k_{ch} * (m_c / M_c)), \text{ wobei:}$$

k_{ch} = Anzahl Mobilfunknummern, über die der Haushalt erreichbar ist

Die Wahrscheinlichkeit für Haushalte der Basis- und Aufstockerstichproben, in einer der Aufstockerstichproben ausgewählt zu werden, ergibt sich über:

$$P_j(\text{Haushalte, Aufstockung}) = k_F * (m_{F,j} / M_{F,j})$$

3.2.2 Integration von Basisstichprobe und Aufstockerstichproben

Die Integration der Basisstichprobe und der Aufstockerstichproben in eine gemeinsame Stichprobe beruht auf der Idee, für jede Einheit der Vereinigungsstichprobe die Inklusionswahrscheinlichkeit zu bestimmen. Voraussetzung dafür ist, dass jede Einheit der Grundgesamtheit die Chance hatte, in mindestens eine der beiden Stichproben zu gelangen.

Nach der oben beschriebenen Berechnung der Auswahlwahrscheinlichkeiten für Personen und Haushalte erfolgt dann die Integration der Stichproben über die Berechnung der Inklusionswahrscheinlichkeit in der Vereinigungsstichprobe für Personen und Haushalte. Diese ist gegeben durch die Auswahlwahrscheinlichkeit für die Basisstichprobe plus die Auswahlwahrscheinlichkeit für die Aufsto-

³² ADM (2012), ADM-Forschungsprojekt „Dual-Frame-Ansätze“ 2011/2012, Forschungsbericht.

³³ Der Anteil dieser Nummern liegt entsprechend dem Ergebnis des ADM-Forschungsprojekts im Festnetzauswahlrahmen bei 60 Prozent und im Mobilfunkauswahlrahmen bei 72 Prozent.

ckerstichproben minus die Auswahlwahrscheinlichkeit für Basis- und Aufstockerstichproben:

$$P(\text{BasisAufstockung}) = P(\text{Basis}) + P(\text{Aufstockung}) - P(\text{Basis}) * P(\text{Aufstockung}).$$

Die inverse Inklusionswahrscheinlichkeit ist dann das reine Designgewicht für die integrierte Stichprobe (Basis- und Aufstockerstichproben), womit erwartungstreue Schätzungen für die aktuelle Grundgesamtheit möglich sind. Die Summe aus den Stichprobenwerten, die jeweils mit der inversen Inklusionswahrscheinlichkeit gewichtet werden, entspricht dem sogenannten Horwitz-Thompson-Schätzer.

3.2.3 Proportionalisierung des Dual-Use-Anteils

Allerdings enthält das Verfahren für die Gewichtung von Dual-Frame-Stichproben (Festnetz- und Mobilfunkstichprobe) einen weiteren Schritt, der die Daten der Gesamtstichprobe sehr viel robuster in Bezug auf das Stichprobenverhältnis Mobilfunk – Festnetz gestaltet. Wie eine ausführliche Simulationsstudie, die infas im Rahmen des ADM-Projekts „Dual-Frame“ durchgeführt hat, zeigt, ist dieser zusätzliche Schritt insbesondere bei ungleichen Mobilfunk- und Festnetzanteilen in der Stichprobe von erheblichem Vorteil für die Robustheit gewichteter Schätzungen.

In diesem Schritt erfolgt eine Proportionalisierung der Dual-Use-Gruppen (sowohl über Festnetz als auch über Mobilfunk erreichbar) in den beiden Stichproben. Insofern wurde eine weitere Transformation durchgeführt. Unter der Annahme, dass sowohl die Festnetzstichprobe als auch die Mobilfunkstichprobe für sich in ihrer jeweiligen Zielgruppe repräsentativ sind und damit auch die Potenziale der Dual-Use-Gruppe gleich sein müssen, wurden dabei gleiche Potenziale für die Teilgruppe „Festnetz und Mobil“ hergestellt.

Der zusätzliche Transformationsfaktor, mit dem die reinen Designgewichte der integrierten Stichprobe zum Ausgleich der Potenziale multipliziert wurden, berechnet sich über:

$$\text{Festnetzstichprobe: } F_{\text{Festnetz}} = \sqrt{\frac{\beta_M}{\beta_F}}$$

$$\text{Mobilfunkstichprobe: } F_{\text{Mobilfunk}} = \frac{1}{F_{\text{Festnetz}}}, \text{ wobei:}$$

β_M = Teilgruppe „Dual-Use“ aus Mobilfunkstichprobe (designgewichtet)

β_F = Teilgruppe „Dual-Use“ aus Festnetzstichprobe (designgewichtet)

Diese Proportionalisierung erfolgte sowohl für die Personen als auch für die Haushalte. Abschließend wurde die Summe der so adjustierten Designgewichte auf die ungewichtete Fallzahl normiert. Das Ergebnis der Berechnungen sind adjustierte Designgewichte für die integrierte Basis- und Aufstockungsstichprobe für Personen und für Haushalte, in der auch ausschließlich über Mobilfunk erreichbare Personen bzw. Haushalte enthalten sind.

3.2.4 Auswahlwahrscheinlichkeit für die Onomastikstichprobe

Bei der Onomastikstichprobe handelt es sich um eine reine Festnetzstichprobe eingetragener Telefonnummern mit Angaben zu den Namen. Die Stichprobe war nach Bundesländern geschichtet. Die Auswahlwahrscheinlichkeit für die realisierte Onomastikstichprobe (m_{FO}) wurde entsprechend bundeslandweise über den Auswahlrahmen (M_{FO}) je Bundesland für Personen berechnet über:

$$P_j (\text{Personen, Onomastik}) = (k_F * (m_{FO,j} / M_{FO,j}) * (1 / z))$$

für $j = 1, 2, \dots, 16$ (Bundesländer).

Für Haushalte erfolgte die Berechnung über:

$$P_j (\text{Haushalte, Onomastik}) = k_F * (m_{FO,j} / M_{FO,j})$$

3.2.5 Integration der Onomastikstichprobe

Die Auswahl der im Telefonbuch eingetragenen Festnetznummern wurde für die Onomastikstichprobe über die eingetragenen Namen vorgenommen. Für die generierten Telefonnummern der Basis- und Aufstockerstichprobe liegen keine Namen vor. Daher lässt sich die Auswahlwahrscheinlichkeit der Basis- und Aufstockerstichprobeneinheiten für die Onomastikstichprobe nicht bestimmen. Die Berechnung gemeinsamer Auswahlwahrscheinlichkeiten ist hierbei nicht möglich.

Ist die Voraussetzung erfüllt, dass zwei Stichproben vorliegen, die erwartungstreue Schätzungen für den wahren Parameter erlauben, dann ist jede Konvexkombination, d.h. die Berechnung von Konvexgewichten, die sich zur Gesamtheit summieren, der erwartungstreuen Schätzer ebenfalls erwartungstreu³⁴. Eine Konvexverbindung führt demnach zu erwartungstreuen Schätzern, kann aber ggf. die Varianz der Schätzer deutlich erhöhen.

Für die Konvexkombination wurde nur der Teil der integrierten Basis- und Aufstockerstichprobe verwendet, für den eine mehrfache Auswahlwahrscheinlichkeit unterstellt werden kann. Deshalb wurden also nur Personen mit türkischem Migrationshintergrund mit der Onomastikstichprobe konvex verbunden, d.h. nur deren Designgewichte wurden angepasst. Alle anderen behalten ihre adjustierten Designgewichte für Personen und für Haushalte. Da rd. 92 Prozent der Onomastikstichprobe tatsächlich auch einen türkischen Migrationshintergrund hatten, ist der damit verbundene Fehler relativ gering.

³⁴ Rendtel, U. (1999). The Application of the Convex Weighting Estimator to Household Panel Surveys. Mimeo, Frankfurt; Spiess, M. & U. Rendtel (2000). Combining an ongoing panel with a new cross-sectional sample, DIW Discussion Paper No. 198, Berlin.

Bei der Konvexgewichtung werden die ursprünglichen Ausgangsgewichte jeder Stichprobe mit einem beliebigen Faktor multipliziert, wobei die Gewichte der Stichprobe 1 (türkische Migranten der integrierten Basis- und Aufstockerstichprobe) mit einem beliebigen Faktor α , die Gewichte der Stichprobe 2 (Onomastikstichprobe) mit $1 - \alpha$ multipliziert wird. Im vorliegenden Fall wurde die Stichprobe 1 multipliziert mit: $\alpha = n_1 / (n_1 + n_2)$ und Stichprobe 2 wurde multipliziert mit: $1 - \alpha = n_2 / (n_1 + n_2)$.

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

Nach diesem letzten Integrationsschritt konnten alle Stichproben in einer gemeinsamen Stichprobe integriert werden. Allen Fällen der gemeinsamen Stichprobe ist ein integriertes adjustiertes Designgewicht auf Personenebene und ein integriertes adjustiertes Designgewicht auf Haushaltsebene zugeordnet.

3.2.6 Trimmen der Designgewichte

Das Hauptziel der Gewichtung ist der Ausgleich von Verzerrungen, die durch das Stichprobendesign (mit ungleichen Auswahlwahrscheinlichkeiten) und durch selektive Teilnahme bzw. Nichtteilnahme entstehen. Unter Verwendung der Gewichte lassen sich Populationswerte aus der Stichprobe erwartungsgetreu schätzen. Weisen die Gewichte allerdings eine starke Streuung auf, so kann dies zu einer großen Varianz der Schätzfunktionen führen. Hier tritt der in der Statistik immer wieder anzutreffende Trade-off zwischen Bias und Varianz auf. Die Gewichtung reduziert den Bias, eine allzu große Steigerung der Varianz durch die Gewichtung sollte allerdings auch vermieden werden. Aus diesem Grunde wird versucht, besonders große Gewichtungsfaktoren (und nachrangig auch besonders kleine Faktoren) möglichst zu umgehen und ggf. geeignete Korrekturen an den Gewichten vorzunehmen.

Ausschließlich für die Kalibrierung wurden deshalb die Eingangsgewichte (die beschriebenen finalen Designgewichte) vor der Kalibrierung getrimmt, d.h. sie wurden durch neue Eingangsgewichte ersetzt. Unter- bzw. Obergrenze der getrimmten Designgewichte wurden durch bestimmte Perzentile der Verteilung, in Abhängigkeit von der Verteilung der Designgewichte, ermittelt. Die Perzentile wurden dabei so gewählt, dass die Trimmung keine erheblichen Veränderungen der gewichteten Verteilungen bewirkte, gleichzeitig aber die Gewichtungsfaktoren mit den höchsten und niedrigsten Werten gekappt wurden. Die Personengewichte und die Haushaltsgewichte wurden jeweils am 2%- und am 98%-Perzentil ihrer Verteilung getrimmt und anschließend so reskaliert, dass ihre Summe wieder die Summe der ungetrimmten Designgewichte ergab.

3.2.7 Kalibrierung und Hochrechnung

Anschließend wurden die getrimmten Designgewichte als Ausgangsgewichte für die Randanpassung verwendet. Die Kalibrierung der Designgewichte erfolgte, getrennt auf Personen- und auf Haushaltsebene, mittels GREG (generalized regression estimator)³⁵.

Kalibriert wurde an die Eckwerte des Mikrozensus 2010 des Statistischen Bundesamts für die Personen am Ort der Hauptwohnung in Privathaushalten im Alter von 16 Jahre oder älter sowie an die Eckwerte des Mikrozensus 2010 für Haushalte am Ort der Hauptwohnung (hier ist auch die verschwindend geringe Anzahl der Haushalte enthalten, in denen ausschließlich Personen unter 16 Jahre leben). Für die Eckwerte zum Migrationsanteil wurde ebenfalls auf Eckwerte aus dem Mikrozensus 2010 des Statistischen Bundesamts zurückgegriffen, die 2011 in der Fachserie 1, Reihe 2.2 veröffentlicht wurden. Hier ist allerdings nur eine Abgrenzung ab dem Alter 15 Jahre möglich.

Um die Varianz der Gewichte zu begrenzen, wurde neben der Trimmung das Intervall der Gewichte begrenzt, d.h. es wurde eine obere und eine untere Grenze für die Gewichte festgelegt. Festgelegt wurde dabei die Spannbreite der reinen Kalibrierungsgewichte (sog. g-weights), die sich aus dem Verhältnis von ursprünglichen Gewichten (getrimmte Eingangsgewichte der Kalibrierung) zum kalibrierten Endgewicht ergibt. Bei dieser Begrenzung der Gewichte ist allerdings zu beachten, dass engere Grenzen für die Gewichte in geringerer Streuung und damit geringerer Varianz der Schätzungen resultieren, bei zu eng gesetzten Grenzen kann allerdings die Kalibrierung aller Eckwerte³⁶ schlicht unmöglich werden.

Um die Kalibrierung durchzuführen, muss jede Eckwertvariable für jede Record im Datensatz einen gültigen Wert besitzen. Deshalb wurde der geringe Item-Nonresponse vor der Kalibrierung imputiert. Die Imputation erfolgte über den Mittelwert bzw. den Modalwert der jeweiligen Variable. Da die Imputation einzig der Durchführbarkeit der Kalibrierung dient, wurden die imputierten Werte nach der Kalibrierung wieder auf fehlende Werte zurückgesetzt.

Beim Bundesland und bei der BIK-Gemeindegrößenklasse erfolgte keine Imputation, hier wurde die Kategorie „fehlende Werte“ als zusätzliche Eckwertkategorie definiert und der Anteil der gewichteten fehlenden Angaben als Zielwert festgelegt. Die Kalibrierung verändert also den (designgewichteten) Anteil der fehlenden Angaben nicht. Die Anteile der Bundesländer und der BIK-Gemeindegrößenklassen wurden proportional um den Anteil der fehlenden Angaben reduziert, so dass sich in der Summe wieder 100 Prozent ergaben.

³⁵ J.C. Deville, C.E. Särndal (1992): Calibration Estimators in Survey Sampling. Journal of the American Statistical Association, 82, 376-381; Särndal, C.-E., Swensson, B., Wretman, J. (1992): Model Assisted Survey Sampling, Springer.

³⁶ Z.B. Soll-Verteilung aus dem Mikrozensus.

Personen

Bei den Personengewichten wurde das Verhältnis der Gesamthochrechnungsfaktoren nach Kalibrierung zu getrimmten Designgewichten nach unten auf minimal 0,3 und nach oben auf maximal 3,5 begrenzt.

Für folgende Merkmale wurde kalibriert:

- Geschlecht x Altersgruppen mit den Altersgruppen 16-24 Jahre, 25-34 Jahre, 35-44 Jahre, 45-54 Jahre, 55-64 Jahre, 65-74 Jahre und 75 Jahre und älter.
- Höchster allgemeinbildender Schulabschluss x Altersgruppen mit den Kategorien: niedrig = Förderschul- oder Volksschul- oder Hauptschulabschluss oder Abschluss der POS8 oder kein Schulabschluss, mittel = Mittlere Reife oder Abschluss POS10 oder noch in schulischer Ausbildung oder sonstiger Schulabschluss, hoch = Fachhochschul- oder Hochschulreife.
- Erwerbstätigkeit dichotomisiert in erwerbstätig vs. nicht erwerbstätig.
- Arbeitslosigkeit dichotomisiert in arbeitslos vs. nicht arbeitslos.
- Migrationshintergrund mit den Kategorien: kein Migrationshintergrund, türkischer Migrationshintergrund, Migrationshintergrund ehemalige Sowjetunion und sonstiger Migrationshintergrund. Dabei wurde Migrationshintergrund definiert über eigene Migrationserfahrung und/oder ausschließlich nicht deutsche Staatsangehörigkeit.
- Haushaltsgröße mit den Kategorien Personen in Einpersonenhaushalten, Personen in Zweipersonenhaushalten, Personen in Dreipersonenhaushalten und Personen in Vier- und Mehrpersonenhaushalten
- Bundesland
- BIK-Gemeindegrößenklasse

Das Ergebnis der Kalibrierung sind zum einen fallzahlnormierte kalibrierte Endgewichte für die Personengewichtung und zum anderen Hochrechnungsfaktoren auf die Grundgesamtheit der 16-jährigen und älteren Bevölkerung. Sowohl das kalibrierte Endgewicht als auch der kalibrierte Hochrechnungsfaktor wurden dem Datensatz zugespielt.

In der folgenden Tabelle ist das Ergebnis der Gewichtung zusammengefasst. Dargestellt sind die Prozentverteilungen der aktiven Gewichtungsmerkmale

- i) ungewichtet,
- ii) designgewichtet (ungetrimmt vor Kalibrierung) und
- iii) kalibriert.

Tabelle 25 Prozentualer Verteilungsvergleich ungewichtete, designgewichtete und kalibrierte Stichprobe mit dem Mikrozensus für die Grundgesamtheit (Personen)

Merkmale	ungewichtet (%)	designgewichtet (%)	kalibriert (%)	MZ 2010 (Soll) (%)
Geschlecht x Altersgruppe				
Mann 16-24 Jahre	5,3	7,3	6,2	6,2
Mann 25-34 Jahre	5,5	8,0	7,1	7,1
Mann 35-44 Jahre	7,1	7,8	8,7	8,7
Mann 45-54 Jahre	10,1	11,5	9,3	9,3
Mann 55-64 Jahre	8,4	8,3	7,1	7,1
Mann 65-74 Jahre	6,9	5,7	6,5	6,5
Mann 75 Jahre u. älter	3,4	2,5	3,9	3,9
Frau 16-24 Jahre	4,7	6,1	5,9	5,9
Frau 25-34 Jahre	5,9	7,0	6,9	6,9
Frau 35-44 Jahre	8,9	8,6	8,5	8,5
Frau 45-54 Jahre	12,0	11,4	9,3	9,3
Frau 55-64 Jahre	9,8	8,0	7,3	7,3
Frau 65-74 Jahre	7,8	5,3	7,4	7,4
Frau 75 Jahre u. älter	4,4	2,6	6,1	6,1
Bildung x Altersgruppe				
niedrig 16-24 Jahre	1,3	2,1	3,8	3,8
niedrig 25-34 Jahre	1,5	2,4	3,4	3,4
niedrig 35-44 Jahre	2,3	2,7	5,0	5,0
niedrig 45-54 Jahre	4,2	5,3	6,6	6,6
niedrig 55-64 Jahre	4,9	5,1	6,9	6,9
niedrig 65-74 Jahre	5,2	4,2	9,3	9,3
niedrig 75 Jahre u. älter	3,6	2,5	7,7	7,7
mittel 16-24 Jahre	4,6	5,9	4,2	4,2
mittel 25-34 Jahre	3,6	4,9	4,5	4,5
mittel 35-44 Jahre	6,1	6,6	6,6	6,6
mittel 45-54 Jahre	8,8	9,1	7,0	7,0
mittel 55-64 Jahre	6,7	5,7	4,3	4,3
mittel 65-74 Jahre	4,8	3,5	2,5	2,5
mittel 75 Jahre u. älter	2,0	1,3	1,2	1,2
hoch 16-24 Jahre	4,0	5,4	4,0	4,0
hoch 25-34 Jahre	6,3	7,7	6,1	6,1
hoch 35-44 Jahre	7,6	7,2	5,7	5,7
hoch 45-54 Jahre	9,0	8,5	5,0	5,0
hoch 55-64 Jahre	6,7	5,5	3,2	3,2
hoch 65-74 Jahre	4,6	3,3	2,1	2,1
hoch 75 Jahre u. älter	2,1	1,4	1,1	1,1

Merkmale	ungewichtet (%)	designgewichtet (%)	kalibriert (%)	MZ 2010 (Soll) (%)
Erwerbstätigkeit				
erwerbstätig	52,2	55,8	53,5	53,5
nicht erwerbstätig	47,8	44,2	46,5	46,5
Arbeitslosigkeit				
arbeitslos	2,8	3,5	4,2	4,2
nicht arbeitslos	97,2	96,5	95,8	95,8
Migrationshintergrund				
kein Migrationshintergrund	88,6	87,8	83,4	83,4
Migrationshintergrund türkisch	2,8	2,9	3,0	3,0
Migrationshintergrund ehem. SU	3,1	3,2	3,1	3,1
Migrationshintergrund sonstiger	5,5	6,1	10,5	10,5
Haushaltsgröße				
1 Person	29,5	21,9	22,6	22,6
2 Personen	48,8	47,5	38,4	38,4
3 Personen	13,1	16,6	17,8	17,8
4 und mehr Personen	8,7	14,0	21,2	21,2
Bundesland				
Schleswig-Holstein	3,1	3,2	3,3	3,4
Hamburg	5,8	2,0	2,1	2,2
Niedersachsen	7,5	8,2	9,3	9,6
Bremen	0,7	0,7	0,8	0,8
Nordrhein-Westfalen	18,6	20,6	21,0	21,7
Hessen	8,6	6,7	7,2	7,4
Rheinland-Pfalz	4,2	4,6	4,7	4,9
Baden-Württemberg	10,4	11,1	12,5	13,0
Bayern	14,1	16,7	14,6	15,1
Saarland	0,9	1,0	1,2	1,3
Berlin	6,6	4,5	4,2	4,3
Brandenburg	5,9	3,1	3,0	3,2
Mecklenburg-Vorpommern	1,7	2,3	2,0	2,1
Sachsen	5,7	5,2	5,1	5,3
Sachsen-Anhalt	2,6	3,3	2,9	3,0
Thüringen	2,3	2,9	2,7	2,8
Fehlend	1,4	3,9	3,5	----

Merkmale	ungewichtet (%)	designgewichtet (%)	kalibriert (%)	MZ 2010 (Soll) (%)
BIK-Gemeindegrößenklasse				
unter 2.000 EW	1,5	1,6	1,9	2,0
2.000 - unter 5.000 EW	2,3	2,6	2,8	2,9
5.000 - unter 20.000 EW	7,5	7,7	7,7	7,9
20.000 - unter 50.000 EW	10,0	11,3	11,2	11,6
50.000 - unter 100.000 EW Rand	7,2	7,5	7,7	8,0
50.000 - unter 100.000 EW Kern	2,1	2,1	2,3	2,4
100.000 - unter 500.000 EW Rand	13,1	13,5	14,7	15,2
100.000 - unter 500.000 EW Kern	13,5	14,6	15,9	16,5
500.000 und mehr EW Rand	9,8	9,1	8,3	8,6
500.000 und mehr EW Kern	31,7	26,1	24,2	25,0
Fehlend	1,5	4,0	3,6	----

Quelle: Mikrozensus, eigene Berechnungen

Insgesamt betrachtet, sind die meisten untersuchten Merkmale in der realisierten (ungewichteten) Stichprobe im Vergleich zur Grundgesamtheit (Soll-Verteilungen nach Mikrozensus, vgl. Tabelle 25) sehr gut repräsentiert. Lediglich bei der formalen Schulbildung zeigt sich, dass Personen mit niedriger formaler Bildung erkennbar unter- und Personen mit hoher formaler Bildung überrepräsentiert sind. Dieses Phänomen ist in sozialwissenschaftlichen Befragungen oft zu beobachten. Zudem haben sich allgemein Personen mit Migrationshintergrund seltener an der Befragung beteiligt. Auch dieses Ergebnis entspricht den Erfahrungen einer Reihe anderer Befragungen. Des Weiteren sind Ein- und Zweipersonenhaushalte über- und Mehrpersonenhaushalten (ab drei oder mehr Personen) unterrepräsentiert. Die Berücksichtigung einer Mobilfunkstichprobe dürfte dies mit verursacht haben, da die Wahrscheinlichkeit, eine Person aus einem Einpersonenhaushalt zu erreichen, in der Mobilfunkstichprobe größer ist als in der Festnetzstichprobe. Die Differenzen zur Sollverteilung laut Mikrozensus sind durch die beschriebenen Gewichtungsschritte ausgeglichen worden.

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

$$E = (n' / n) * 100, \text{ wobei: } n' = ((\sum_i g_i)^2 / \sum_i g_i^2)$$

Tabelle 26 Kennwerte der kalibrierten Gewichte und Hochrechnungsfaktoren

Kennwert (Personen)	kalibriertes Gewicht	Hochrechnungsfaktor
Mittelwert	1	1955,2
Standardabweichung	1,45	2826,2
Minimum	0,44	85,18
Maximum	25,58	50021,0
Fallzahl	35.503	
Effektivitätsmaß	32,4%	
effektive Fallzahl	11.492	

eigene Berechnungen

Die Effektivität ist relativ gering bzw. der Designfaktor ist vergleichsweise hoch. Dies bedeutet eine deutliche Reduktion der effektiven Fallzahl und mithin eine Vergrößerung des Standardfehlers bei gewichteten Populationsschätzern. Dies ist insgesamt ein Ergebnis des Stichprobendesigns, da in der Gesamtstichprobe unterschiedliche Frames und disproportionale Stichproben integriert wurden. Der Ausgleich der daraus resultierenden unterschiedlichen Auswahlwahrscheinlichkeiten und Disproportionalitäten erhöht zwangsläufig die Varianz der Gewichte.

Es bleibt allerdings festzuhalten, dass aufgrund der sehr hohen Fallzahl die effektive Fallzahl immer noch sehr groß, der Stichprobenfehler entsprechend, trotz des Designeffekts, gering ist.

Haushalte

Bei den Haushaltsgewichten wurde das Verhältnis der Gesamthochrechnungsfaktoren nach Kalibrierung zu getrimmten Designgewichten nach unten auf minimal 0,3 und nach oben auf maximal 4,0 begrenzt. Da keine Daten über den Haupteinkommensbezieher eines Haushalts vorliegen, konnten die Haushaltsgewichte an nur vergleichsweise wenigen Verteilungen aus dem Mikrozensus kalibriert werden.

Für folgende Merkmale wurde kalibriert:

- Haushaltsgröße mit den Kategorien Einpersonenhaushalte, Zweipersonenhaushalte, Dreipersonenhaushalte und Vier- und Mehrpersonenhaushalte.
- Bundesland
- BIK-Gemeindegroßenklasse

Das Ergebnis der Kalibrierung sind zum einen fallzahlnormierte kalibrierte Endgewichte für die Haushaltsgewichtung und zum anderen Hochrechnungsfaktoren auf die Grundgesamtheit aller Haushalte. Sowohl das kalibrierte Endgewicht als auch der kalibrierte Hochrechnungsfaktor für die Haushalte wurden dem Datensatz zugespielt. Sinnvoll verwendbar sind sie aber ausschließlich für Merkmale, die für den gesamten Haushalt erhoben wurden, also haushaltsbezo-

gene Fragestellungen. In der folgenden Tabelle ist das Ergebnis der Gewichtung zusammengefasst.

Tabelle 27 Prozentualer Verteilungsvergleich ungewichtete, designgewichtete und kalibrierte Stichprobe mit dem Mikrozensus für die Grundgesamtheit (Haushalte)

Merkmale	ungewichtet (%)	designgewichtete (%)	kalibriert (%)	MZ 2010 (Soll) (%)
Haushaltsgröße				
1 Person	29,5	39,1	39,6	39,6
2 Personen	48,8	43,4	34,5	34,5
3 Personen	13,1	10,6	12,8	12,8
4 und mehr Personen	8,7	7,0	13,1	13,1
Bundesland				
Schleswig-Holstein	3,1	3,3	3,3	3,4
Hamburg	5,8	2,1	2,4	2,4
Niedersachsen	7,5	8,1	9,3	9,6
Bremen	0,7	0,8	0,9	0,9
Nordrhein-Westfalen	18,6	20,3	20,8	21,6
Hessen	8,6	6,5	7,0	7,3
Rheinland-Pfalz	4,2	4,4	4,5	4,7
Baden-Württemberg	10,4	10,5	12,0	12,5
Bayern	14,1	16,1	14,3	14,8
Saarland	0,9	0,9	1,2	1,2
Berlin	6,6	5,0	4,8	5,0
Brandenburg	5,9	3,2	3,0	3,1
Mecklenburg-Vorpommern	1,7	2,5	2,1	2,1
Sachsen	5,7	5,6	5,3	5,5
Sachsen-Anhalt	2,6	3,6	2,9	3,0
Thüringen	2,3	2,8	2,7	2,8
Fehlend	1,4	4,4	3,6	----
BIK-Gemeindegrößenklasse				
unter 2.000 EW	1,5	1,4	1,7	1,8
2.000 - unter 5.000 EW	2,3	2,5	2,6	2,7
5.000 - unter 20.000 EW	7,5	7,1	7,2	7,4
20.000 - unter 50.000 EW	10,0	10,9	10,7	11,1
50.000 - unter 100.000 EW Rand	7,2	7,2	7,3	7,6
50.000 - unter 100.000 EW Kern	2,1	2,2	2,4	2,4
100.000 - unter 500.000 EW Rand	13,1	12,5	13,8	14,3
100.000 - unter 500.000 EW Kern	13,5	15,2	16,8	17,4
500.000 und mehr EW Rand	9,8	8,8	7,9	8,2
500.000 und mehr EW Kern	31,7	27,8	26,1	27,1
Fehlend	1,5	4,5	3,7	----

Quelle: Mikrozensus, eigene Berechnungen

Die Kennwerte für die Haushaltsgewichte sind in der folgenden Tabelle dargestellt.

Tabelle 28 Kennwerte der kalibrierten Gewichte und Hochrechnungsfaktoren

Kennwert (Haushalte)	kalibriertes Gewicht	Hochrechnungsfaktor
Mittelwert	1	1117,5
Standardabweichung	1,29	1446,0
Minimum	0,87	96,88
Maximum	17,43	19483,0
Fallzahl	35.503	
Effektivitätsmaß	37,4%	
effektive Fallzahl	13.276	

eigene Berechnungen

Die Effektivität ist bei den Haushaltsgewichten, die nur für wenige haushaltsbezogene Fragen relevant ist, etwas höher als bei den Personengewichten. Dies ist eine Folge der geringeren Zahl der kalibrierbaren Merkmale auf der Haushaltsebene. Hier sind demnach zwar nicht so viele Merkmale angepasst, die Varianz der Gewichte aber entsprechend auch geringer.

3.2.8 Berechnung und ggf. Kalibrierung für die einzelnen Bundesländer

Für die fünf Bundesländer mit einer Aufstockungstichprobe wurden getrennte Datensätze und entsprechend getrennte Gewichtung- und Hochrechnungsfaktoren erstellt. Die jeweiligen Datensätze enthalten die (identifizierbaren) Fälle des jeweiligen Bundeslands aus der Basisstichprobe und der Onomastikstichprobe sowie die Aufstockerstichprobe des Bundeslands. In den Stichproben sind demzufolge auch Fälle aus der Mobilfunkstichprobe, der Anteil der Mobilfunkstichprobe ist aber in den Aufstockerbundesländern geringer als in der Basisstichprobe, da die Aufstockerstichproben reine Festnetzstichproben waren. Entsprechend vergrößert sich tendenziell auch die Varianz der Gewichtungsfaktoren.

Da die Aufstockerstichproben bereits in einer Gesamtstichprobe integriert und die Auswahlwahrscheinlichkeiten bestimmt wurden, konnten als bundeslandspezifische Ausgangsgewichte die kalibrierten Endgewichte und die kalibrierten Hochrechnungsfaktoren verwendet werden.

Diese wurden zunächst fallzahlnormiert auf die realisierten Fallzahlen der einzelnen Bundesländer bzw. die Hochrechnungsfaktoren auf die Grundgesamtheiten der jeweiligen Bundesländer umgerechnet (fehlende Werte bei der Angabe des Bundeslands waren hier per Definition ausgeschlossen).

Anschließend erfolgte eine Prüfung der bundeslandspezifischen Verteilungen gegen die Verteilungen des Mikrozensus. Nur falls die gewichteten Verteilungen deutlich von den Verteilungen des Mikrozensus abwichen, erfolgte eine weitere zusätzliche bundeslandspezifische Kalibrierung.

Bei den Bundesländern Hamburg, Hessen und Berlin erwiesen sich die Ergebnisse der Basiskalibrierung als sehr genau, eine weitere Anpassung war hier nicht notwendig. Bei den Bundesländern Brandenburg und Sachsen war jedoch eine zusätzliche Anpassung einiger Verteilungen notwendig. Diese zusätzliche Kalibrierung erfolgte erneut mittels GREG-Schätzer, so dass die bundeslandspezifisch kalibrierten Gewichte so gering wie möglich von den kalibrierten Basisendgewichten abweichen.

In Brandenburg ergaben sich bei Personen Abweichungen in der Geschlechtsverteilung, der Verteilung der Haushaltsgröße und der BIK-Gemeindegrößenklasse. Diese drei Merkmale wurden zusätzlich kalibriert. Für die Haushaltsgewichte in Brandenburg wurden zusätzlich die Haushaltsgröße und die BIK-Gemeindegrößenklasse kalibriert.

In Sachsen ergaben sich Abweichungen bei der Haushaltsgröße und bei der Geschlechtsverteilung, die für die Personengewichte in Sachsen zusätzlich angepasst wurden. Die Haushaltsgewichte wurden zusätzlich nur an die Verteilung der Haushaltsgröße kalibriert.

Bei den Bundesländern Hamburg, Hessen und Berlin sind also die gewichteten prozentualen Verteilungen mit den bundeslandspezifischen Gewichten identisch mit den gewichteten Verteilungen der nach dem Bundesland selektierten Gesamtstichprobe. In den Bundesländern Brandenburg und Sachsen kann es hier zu geringen Abweichungen bei den Prozentverteilungen kommen.

4 Literaturverzeichnis

AAPOR (The American Association for Public Opinion Research) (2011): Standard Definitions. Final Dispositions of Case Codes and Outcome Rate for Survey. Revisited Version 2011.

Arbeitskreis Deutscher Marktforschungsinstitute (ADM) (2012), ADM-Forschungsprojekt „Dual-Frame-Ansätze“ 2011/2012, Forschungsbericht.

Aust, F, Schröder, H. (2009): Sinkende Stichprobenausschöpfung in der Umfrageforschung – ein Bericht aus der Praxis. In: Weichbold, Martin, Johann Bacher und Christof Wolf: Umfrageforschung. Grenzen und Herausforderung. VS Verlag.

Auspurg, K., Hinz, T. und S. Liebig (2009): „Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey“, in Methoden – Daten – Analysen, Jg. 3, Heft 1, S. 59-96.

Deville, J.C., Särndal, C.E.(1992): Calibration Estimators in Survey Sampling. Journal of the American Statistical Association, 82, 376-381; Särndal, C.-E., Swensson, B., Wretman, J. (1992): Model Assisted Survey Sampling, Springer.

Follmer R., Smid, M. (1993): „Nichteingetragene Telefonnummern: Ergebnisse eines Methodentests“ in: Gabler, S., S. Häder und J.H.P. Hoffmeyer-Zlotnik (Hrsg.) (1998): Telefonstichproben in Deutschland. Opladen, Westdeutscher Verlag, S.43-57.

Hoffmann, H. (2007): Kombinierte Stichproben für Telefonumfragen - Ansätze in Europa. In: S. Gabler und S. Häder (Hg.): Mobilfunktelefonie - Eine Herausforderung für die Umfrageforschung. ZUMA-Nachrichten Spezial Band 13. Mannheim, S. 47-57;

Häder, S., Häder, M. (Hrsg.): Telefonbefragungen über das Mobilfunknetz: Konzept, Design und Umsetzung einer Strategie zur Datenerhebung. Wiesbaden: VS Verlag für Sozialwissenschaften 2009

Häder, S., Gabler, S. (1998): Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. In Gabler S. , S. Häder und J.H.P. Hoffmeyer-Zlotnik (Hrsg.) Telefonstichproben in Deutschland. Opladen, Westdeutscher Verlag.

Humpert, A., Schneiderheinze, K. (2002), Stichprobenziehung für telefonische Zuwandererumfragen – Praktische Erfahrungen und Erweiterung der Auswahl-grundlage, in: Gabler, S. und S. Häder (Hg.), Telefonstichproben - Methodische Innovationen und Anwendungen in Deutschland, Münster: Waxmann, S. 187-208.

Humpert, A., Schneiderheinze, K (2000), Stichprobenziehung für telefonische Zuwandererumfragen – Einsatzmöglichkeiten der Namenforschung (Onomastik), in: ZUMA-Nachrichten 47, S36-63.

Prüfer, P., Rexroth, M. (2005): „Kognitive Interviews“ in ZUMA How-to-Reihe, Nr.15, Mannheim.

Prüfer, P., Rexroth, M. (2000): „Zwei-Phasen-Pretesting“ in ZUMA-Arbeitsbericht 2000/08. Mannheim.

Rendtel, U. (1999). The Application of the Convex Weighting Estimator to House-hold Panel Surveys. Mimeo, Frankfurt; Spiess, M. & U. Rendtel (2000). Combining an ongoing panel with a new cross-sectional sample, DIW Discussion Paper No. 198, Berlin.

Schiel, S., Malina, A., Schröder, H., Banz, M. (2012): „Repräsentative Bevölkerungsbefragung im Rahmen des BaSiD-Teilvorhabens „Sicherheitsgefährdungen durch Kriminalität“. Pretestbericht“. Bonn