

Dual-Frame-Telefonstichproben Entwicklung, Handhabung und Gewichtung

Matthias Sand

GESIS-Technical Reports 2014|02

Dual-Frame-Telefonstichproben

Entwicklung, Handhabung und Gewichtung

Matthias Sand

GESIS-Technical Reports

GESIS – Leibniz-Institut für Sozialwissenschaften

Survey Design and Methodology

Postfach 12 21 55

68072 Mannheim

Telefon: (0621) 1246 - 201

Telefax: (0221) 1246 - 100

E-Mail: matthias.sand@gesis.org

ISSN: 1868-9043 (Print)

ISSN: 1868-9051 (Online)

Herausgeber,

Druck und Vertrieb: GESIS – Leibniz-Institut für Sozialwissenschaften
Unter Sachsenhausen 6-8, 50667 Köln

Zusammenfassung

Die stetig steigende Verbreitung von Mobilfunkanschlüssen und deren Verwendungsmöglichkeiten sowie die Weiterentwicklungen im Bereich der Festnetztelefonie, die klassische Festnetzanschlüsse verdrängen, bewirken, dass klassische Festnetzstichproben die Möglichkeit, zuverlässige Aussagen über die Grundgesamtheit der Haushalte in Deutschland zu treffen, verfehlen können. Ein weiterer wichtiger Einflussfaktor ist der wachsende Anteil derjenigen Haushalte, die nur noch über das Mobilfunknetz erreichbar sind. Dies ist dadurch zu begründen, dass die ausschließliche Beschränkung auf den Auswahlrahmen Festnetz zunehmend zu Undercoverage führt. Dadurch besteht die Gefahr, dass Haushalte, die sich in wesentlichen Charakteristika von denen innerhalb des Auswahlrahmens Festnetz unterscheiden, bei einer solchen Erhebung ausgelassen werden. Die Konsequenz hieraus sind dann verzerrte Schätzergebnisse.

Eine Möglichkeit dieses Undercoverage-Problem zu beheben, besteht in dem Übergang zu Multiple-Frame-Ansätzen, bei der sich eine Erhebung aus mehreren Stichproben, gezogen aus unterschiedlichen Auswahlrahmen, zusammensetzt. Am Beispiel Telefonumfragen sollen hierfür Dual-Frame-Ansätze, die sowohl den Festnetz- als auch den Mobilfunkbereich berücksichtigen, untersucht werden. Dazu wird in Kapitel 1 kurz die derzeitige Situation im Bereich der Telekommunikation innerhalb Deutschlands dargestellt, um anschließend in Kapitel 2 einen Überblick über unterschiedliche Dual-Frame-Ansätze sowie deren Auswahlrahmengestaltung zu liefern. Kapitel 3 beinhaltet einen Exkurs, der den Aufbau von Mobilfunknetzwerken sowie die Identifikation und Bereinigung von Mobilfunkrufnummerlisten via HLR Lookup erörtert. Das vierte Kapitel behandelt schließlich die Designgewichtung solcher Dual-Frame-Ansätze. Ein Ausblick der möglichen Neuerungen sowie den weiterführenden Forschungsbedarf aufzeigt, wird anhand des fünften Kapitels gegeben.

Inhaltsverzeichnis

1	Entwicklung der Festnetz- und Mobilfunkanschlüsse in Deutschland	7
2	Dual-Frame-Ansätze	11
3	Exkurs: HLR Lookup	14
4	Gewichtung von Dual-Frame-Stichproben	18
4.1	Single Frame Ansätze	20
4.2	Composite Weighting	22
4.3	Pseudo-Maximum Likelihood Schätzer	26
4.4	Multiplicity Schätzer	27
4.5	Beurteilung der Ansätze	28
5	Ausblick	32
	Literaturverzeichnis	36
A	Anhang	39

1 Entwicklung der Festnetz- und Mobilfunkanschlüsse in Deutschland

Anhand des Tätigkeitsberichtes der Bundesnetzagentur (BNA) von 2010/2011 bezüglich der Verbreitung der Festnetzanschlüsse innerhalb Deutschlands zeigt sich, dass diese seit 2005 kumuliert auf einem Wert von ca. 38 Mio. Zugängen stagniert. Bei genauerer Betrachtung dieser kumulierten Anschlussmengen, die sowohl Analog- und ISDN-Basisanschlüsse als auch Anschlüsse im Bereich der Kabel-TV-Netze sowie das vom DSL entbundene Voice over IP (VoIP)¹ enthalten, lässt sich allerdings eine Verlagerung der *klassischen* Festnetzanschlüsse (analog und ISDN) zugunsten neuerer Anschlussmöglichkeiten feststellen. So besteht ein Rückgang der *klassischen* Festnetzanschlüsse von ca. 1,7 Mio. im Vergleich zum Vorjahr, bzw. 9,7 Mio. bezogen auf 2005, während neuere Anschlussmöglichkeiten seit 2006 deutlich zunehmen (vgl. BUNDESNETZAGENTUR [2011], S. 30f.).²

Im Gegensatz dazu steht die Entwicklung auf dem Markt für Mobilfunktelefonie. Hier wuchs die Marktpenetration von 96,1% in 2005 auf 137,0% im dritten Quartal 2011, bzw. von 79,27 Mio. auf 112,03 Mio., im vierten Quartal 2012 sogar auf insgesamt 113,16 Mio. Anschlüsse in allen vier Mobilfunknetzen an.³ Die BNA gibt diesbezüglich zu bedenken, dass die zunehmende Verbreitung von Surf-Sticks sowie Tablet-PCs als Zweitgeräte den Anstieg dieser Anschlüsse positiv beeinflusst (vgl. BUNDESNETZAGENTUR [2011], S. 50; BUNDESNETZAGENTUR [2013]). Der Anteil der Nutzer, die ihren Mobilfunkanschluss dabei ausschließlich für Funktionen wie mobiles Internet nutzen, lag BITKOM zufolge bei ca. 3%. 2011 verfügten 83% der deutschen Bevölkerung über einen Mobilfunkanschluss (vgl. BITKOM [2011], S.1). Im Bereich der Festnetztelefonie zeigt sich seit der Aufhebung der Eintragspflicht ein stetiger Rückgang der gelisteten Telefonnummern im Telefonbuch, weswegen der Auswahlrahmen Telefonbuch, bspw. der einer Telefonbuch-CD, zunehmend schlechter wird. So waren 2009 nur noch 57,9% aller deutschen Privathaushalte eingetragen, bzw. lediglich 4,2% aller nach Blöcken möglichen Nummern. Hierbei werden VoIP-Nummern nicht mit in Betracht gezogen, da bei diesen in vielen Fällen keine regionale Zuordnung möglich ist (vgl. HECKEL und WIESE [2012], S. 105f.). Daher ist ein Ansatz wie das Gabler-Häder-Design, das sowohl eingetragene als auch nicht eingetragene Telefonnummern zu einem handhabbaren Aufwand in den Auswahlrahmen einschließt, notwendig (vgl. HÄDER [2009b], S. 21f.). Dabei werden für jeden der 5200 deutschen Ortsnetzbereiche genau diejenigen 100er-Blöcke ermittelt, die mindestens einen Eintrag enthalten. Die 100er-Blöcke sind hierbei als Rufnummernstamm zu verstehen, der um die letzten beiden Ziffern gekürzt wurde. Im Anschluss an die Ermittlung der besetzten Rufnummernblöcke werden alle möglichen (100) Zahlenfolgen generiert. Aus dem hieraus entstandenen Auswahlrahmen wird dann uneingeschränkt oder geschichtet gezogen. Die Entwicklung der Eintragsdichte, die zur Gewinnung eines angemessenen Auswahlrahmens notwendig ist, wird in Abbildung 1 von April 1999 bis Januar 2013 dargestellt.

1 VoIP bezeichnet dabei die Telekommunikation über das Internet. Üblicherweise fangen solche Rufnummern mit der Vorwahl 032 an.

2 Der deutlichste Zuwachs lässt sich hierbei von 2008 auf 2009 verzeichnen. Dieser lag nach Bundesnetzagentur in beiden Bereichen bei ca. 50%. Von 2010 auf 2011 zeigt sich ein Anstieg um ca. 26%, 2011 besteht ein Zuwachs von ca. 14% bei VoIP-Anschlüssen und 24% bei Kabel-TV-Netz-Anschlüssen. Somit stieg die Anzahl der VoIP-Anschlüsse von 0,3 Mio. 2006 auf 5,7 Mio. 2011 und die Anzahl der Kabel-TV-Netz-Anschlüsse von 0,1 Mio. 2006 auf 3,6 Mio. 2011 an.

3 2001 bestand eine Marktpenetration von lediglich 68,1%, bzw. 56,13 Mio. Anschlüssen. Im Vorjahr (2011) lag dieser Wert bei 133,1%, bzw. 108,85 Mio. Anschlüssen. Anhand einer in diesem Zeitraum, unter anderem von der Telekom Deutschland neu eingeführten konsequenten Zählweise, wurden zur Bestimmung dieser Zahlen inaktive Prepaidteilnehmer bereits ausgebucht.

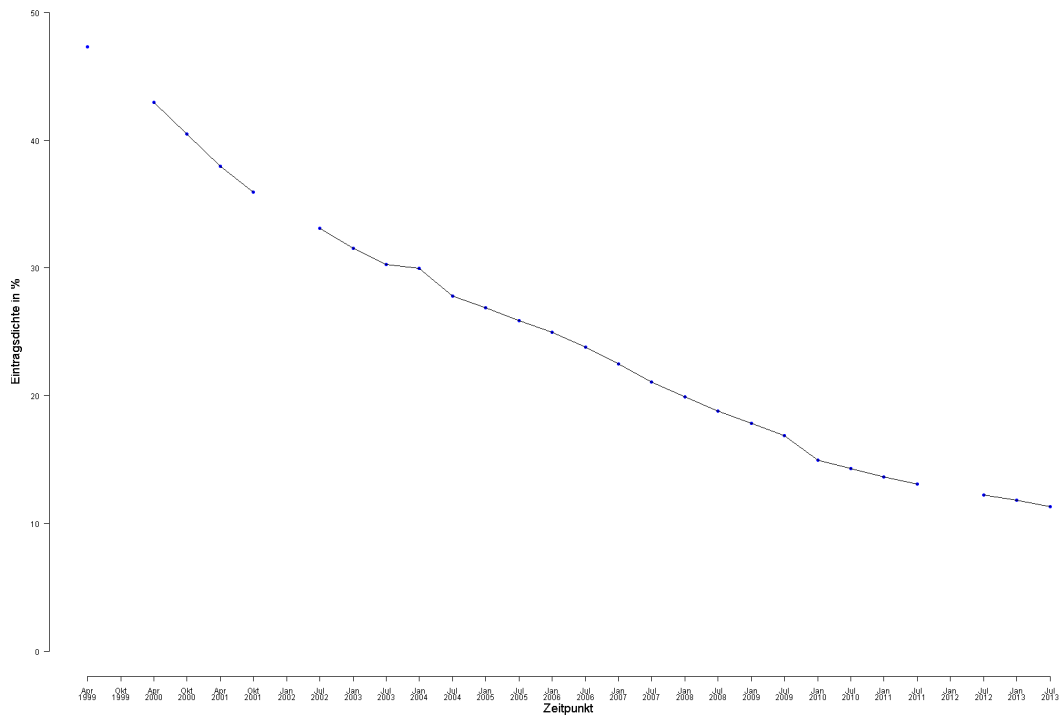


Abbildung 1: Verlauf der Eintragsdichte in Deutschland im GESIS Auswahlrahmen

Bei deren Betrachtung lässt sich eine deutliche Abnahme der Eintragsdichte von 47,33% (April 1999) auf 11,81% (Januar 2013) bestimmen. Neben der Aufhebung der Eintragspflicht kann dabei ebenfalls die Liberalisierung des deutschen Telekommunikationsmarktes ursächlich für den Rückgang der Eintragsdichte sein. Heckel und Wiese (2012) begründen dies unter anderem damit, dass ca. 97% der gelisteten Telefonnummern aus dem Bereich des ehemaligen Monopolisten DBP Telekom stammen, was die Vermutung zulässt, dass beim Wechsel zu einem neuen Anbieter die Telefonnummer nicht mehr eingetragen wird (vgl. HECKEL und WIESE [2012], S. 106).

Unter Berücksichtigung von Tabelle 1 lässt sich überdies feststellen, dass der starke Zuwachs an vergebenen Blöcken, der eine stetig sinkende Anzahl eingetragener Rufnummern gegenüber steht, als eine der Hauptursachen der mittlerweile geringen Eintragsdichte angeführt werden kann. Ein sich daraus ergebendes Problem kann eine sinkende Hitrate, also weniger tatsächlich vergebene und demnach kontaktierbare Rufnummern bei Festnetzstichproben unter Verwendung von Verfahren, wie bspw. der Gabler-Häder-Methode, sein. Als Konsequenz kann sich somit eine sinkende Effizienz gerade für kleine Umfrageinstitute einstellen, da mehr Auswahlversuche pro Anschluss notwendig werden (vgl. HÄDER [2009b], S. 33).⁴

⁴ Für größere Umfrageinstitute sollte aufgrund der Verwendung von Pre-Dialern kein gravierender Unterschied bestehen.

Tabelle 1: Verlauf der Eintragsdichte der 100er Blöcke in Deutschland

Datum	AZ Blöcke	Eingetr. Nummern	Eintragsdichte
April 1999	647.636	30.654.324	47,33
April 2000	721.146	30.970.423	42,95
Oktober 2000	757.092	30.635.916	40,47
April 2001	805.447	30.560.982	37,94
Oktober 2001	844.745	30.341.964	35,92
Juli 2002	904.748	29.937.636	33,09
Januar 2003	946.616	29.833.453	31,52
Juli 2003	975.656	29.542.697	30,28
Januar 2004	1.015.461	30.452.423	29,99
Juli 2004	1.037.791	28.863.646	27,81
Januar 2005	1.063.323	28.608.286	26,90
Juli 2005	1.091.142	28.258.557	25,90
Januar 2006	1.118.360	27.918.774	24,96
Juli 2006	1.152.673	27.446.726	23,81
Januar 2007	1.199.744	26.974.431	22,48
Juli 2007	1.253.148	26.396.587	21,06
Januar 2008	1.297.258	25.839.619	19,92
Juli 2008	1.339.481	25.182.055	18,80
Januar 2009	1.371.845	24.442.031	17,82
Juli 2009	1.403.314	23.657.613	16,86
Januar 2010	1.393.663	20.850.949	14,96
Juli 2010	1.426.246	20.389.266	14,30
Januar 2011	1.468.007	20.021.124	13,64
Juli 2011	1.500.180	19.634.173	13,09
Juli 2012	1.561.668	19.092.396	12,23
Januar 2013	1.588.791	18.762.736	11,81

Als problematisch bezüglich der Generierung von Rufnummern für telefonische Befragungen im Festnetz kann darüber hinaus die Tatsache erachtet werden, dass seit dem 03.05.2010 neu zugeteilte Rufnummernbereiche grundsätzlich mindestens elf Stellen lang sind und diese in einer solchen Form auch nur noch von der Bundesnetzagentur angegeben werden, ohne dass diese eine eindeutige Unterscheidung zu kürzeren Rufnummernbereichen zulassen. Die einzige Ausnahme bilden die vier dreistelligen Ortsnetzbezeichnungen für Berlin, Frankfurt, Hamburg und München (vgl. BUNDESNETZAGENTUR [2011], S. 30). Wird nun bei einer Generierung von Rufnummern, die nicht auf der Basis von tatsächlichen Einträgen (bspw. bei einer Telefonbuch-CD) basiert, sondern die lediglich die Veröffentlichungen der Bundesnetzagentur über die bereits vergebenen 1000er Blöcke berücksichtigt, so kann dies zu gravierenden Fehlern in der Stichprobenauswahl führen, da aufgrund der so generierten „zu kurzen“ Blöcken Nummern generiert werden, bei denen ein kein Kontakt zustande kommen kann und bei „zu langen“ Blöcken Rufnummernbereiche eine höhere Auswahlwahrscheinlichkeit erhalten als auf der Basis von tatsächlichen Einträgen.

Aufgrund der hohen Marktpenetration der Mobilfunkanschlüsse liegt überdies die Vermutung nahe, dass ein mittlerweile nicht unerheblicher Teil der bundesdeutschen Bevölkerung ausschließlich über das Medium Mobiltelefon zu erreichen ist. Eine Studie des ADM-Forschungsprojektes „Dual-Frame-Ansätze“ ermittelt diesbezüglich einen Anteil von 12,4% der deutschen Bevölkerung, der nur noch über Mobilfunk zu erreichen ist (*Mobile-Onlys*), während 18,6% nur über einen Festnetzanschluss verfügen (*Landline-*

Onlys) und 69% über beide Medien zu erreichen sind (*Dual-Users*) (vgl. ADM [2012a], S. 10; ADM [2012b], S. 5).⁵

Gemäß des ZRW-Rahmenschemas für Medien- und Reichweitenforschung besteht, folgt man den Angaben des ADMs, demnach bei der Verwendung einer reinen Festnetzstichprobe noch keine Verletzung des Qualitätskriteriums, das festlegt, dass mindestens 85% der Grundgesamtheit abgedeckt sein müssen (vgl. SCHEFFLER [2013]). Da allerdings nicht auszuschließen ist, dass der Anteil der Mobile-Onlys in den kommenden Jahren weiter steigen wird und man darüber hinaus annehmen kann, dass sich Mobile-Onlys in wesentlichen Charakteristika von Dual-Usern und Landline-Onlys unterscheiden, sollten diese, zur Gewährleistung der Repräsentativität, bei der Ziehung von Telefonstichproben mit einbezogen werden (vgl. GRAESKE [2009], S. 57ff.; SCHEFFLER [2013]).

Eine weitere wichtige Neuerung im Bereich des deutschen Mobilfunknetzes besteht seit September 2011 in der Möglichkeit, bei allen vier Netzbetreibern auch über eine virtuelle Festnetznummer erreichbar zu sein, wodurch auch Haushalte, die lediglich über einen Mobilfunkanschluss verfügen, in den Auswahlrahmen einer Festnetzstichprobe gelangen können (vgl. BUNDESNETZAGENTUR [2011], S. 53). Als mögliche Konsequenz daraus könnten sich, in Abhängigkeit von der Inanspruchnahme dieser mobilen Festnetznummern, Probleme in der Stichprobenziehung, sowohl im Bereich Festnetz als auch Mobilfunk, ergeben. Der Grund hierfür ist, dass diese Nummern ebenso im Auswahlrahmen für Festnetzstichproben auftauchen, während sich die dabei befragte Zielperson selbst, wenn lediglich das Mobiltelefon (inklusive virtueller Festnetznummer) als Telekommunikationsmedium zur Verfügung steht, zum Bereich der Mobile-Onlys zählen würde. Darüber hinaus ist es denkbar, dass sich Zielpersonen nicht bewusst sind, dass sie über eine solche virtuelle Festnetznummer erreichbar sind oder gerade kontaktiert wurden.

Auch seitens des Interviewers entstehen hierbei Probleme, da die Zielpersonenauswahl, wie bspw. durch einen Schwedenschlüssel, nicht, bzw. erschwert, anwendbar sind.⁶ Darüber hinaus können durch technische (Akkulaufzeit, Empfang, etc.) und kommunikative Beeinträchtigungen wegen des Aufenthaltsortes der Zielperson Komplikationen auftreten, die eine Befragung erschweren (vgl. SCHNEID [2007], S. 85ff.). Aufgrund dieser neuen Anforderungen werden Telefonstichproben zunehmend über Dual-Frame-Befragungen durchgeführt, die im folgenden Kapitel erläutert werden.

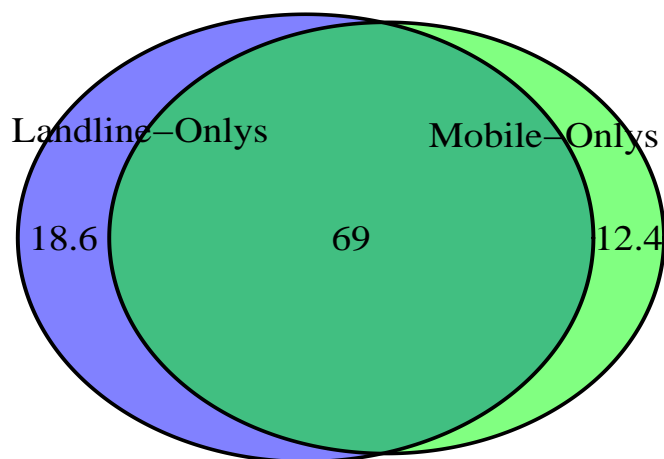
5 Der ADM merkt allerdings bezüglich der Größenordnung der Mobile-Onlys an, dass keine gesicherten Erkenntnisse aus unabhängigen Studien vorliegen.

6 In der Praxis werden Kontakte mit virtuellen Telefonnummern wie Einpersonenhaushalte behandelt.

2 Dual-Frame-Ansätze

Buskirk und Best (2012) zufolge dienen Dual-Frame-Stichproben der Erhöhung der *Coverage*, da diese Modelle darstellen, die aus zwei unterschiedlichen Auswahlrahmen ziehen. Allgemein können die beiden Frames disjunkt sein oder einen Überlappungsbereich aufweisen (vgl. BUSKIRK und BEST [2012], S. 3697). Lohr merkt dazu an, dass Multiple-Frame-Surveys eine „bessere“ Coverage zu geringeren Kosten erzielen können. Dies trifft nach Ansicht der Autorin vor allem dann zu, wenn innerhalb von Befragungen niedrige Response-Raten herrschen, weswegen aufgrund von Undercoverage und Messfehlern Stichprobenschätzwerte stark von dem verwendeten Modell abhängen und durch Nonresponse die Kosten der Datenerhebung steigen (vgl. LOHR [2011], S. 197).

Abbildung 2 stellt nochmals die Grundgesamtheit innerhalb einer Telefonstichprobe unter Verwendung eines solchen Dual-Frame-Ansatzes, aufgeteilt in die beiden Auswahlrahmen Festnetz und Mobilfunk, dar. Die innerhalb dieser Grafik verwendeten Anteile der drei Bereiche (Onlys und Dual-User) entsprechen den vom ADM ermittelten Werten.



Angaben in Prozent

Abbildung 2: Darstellung der Dual-Frame Telefonstichproben

Der Überlappungsbereich der Dual-User stammt dabei aus dem Anteil an der Grundgesamtheit, der sowohl über den Auswahlrahmen Festnetz, als auch durch den des Mobilfunks in eine Stichprobe gelangen kann, während Mobile- und Landline-Onlys lediglich über das jeweilige Sample erreicht werden können. Die Ziehung aus den jeweiligen Auswahlrahmen kann wahlweise ein vorheriges *Screening* bemühen oder, wie bspw. in den CELLA-Studien angewandt, über einen *Nonscreening Ansatz* erfolgen. Bei der Verwendung eines Screening Ansatzes werden im Falle einer Telefonstichprobe entsprechende Mobile-Onlys identifiziert und anschließend via Mobilfunkstichprobe kontaktiert, während der restliche Anteil (Dual-User und Landline-Onlys) anhand einer Festnetzstichprobe gezogen wird. Eine Erhebung auf Basis eines Nonscreening Ansatzes zieht die Stichproben zu einem zuvor festgelegten Anteil aus den jeweiligen Auswahlrahmen, wobei das Interview mit dem Medium durchgeführt wird, mit dem der Erstkontakt zustande gekommen ist. Auf den Umgang bezüglich der Gewichtung der beiden Vorgehensweisen wird in Kapitel 4 näher eingegangen.

Probleme werden in beiden Ansätzen eingeräumt. Ein Screening-Ansatz erschwert die Kontaktaufnah-

me von Personen, die zwar generell über einen Festnetzanschluss verfügen, ihr Mobiltelefon allerdings hauptsächlich als Kommunikationsmedium verwenden, während der Nonscreening-Ansatz die Erreichbarkeit von Dual-Usern, die aber hauptsächlich über Festnetz erreichbar sind und im Rahmen der Mobilfunkstichprobe kontaktiert wurden, überschätzt (vgl. HÄDER [2009b], S. 25f.).

Eine weitere Möglichkeit wäre, die beiden Data-Frames nicht in Onlys und Dual-User zu unterteilen, sondern in Landline- und Mobile-Mainlys. Hierbei wären Onlys lediglich eine extreme Form der Mainlys, während der Überlappungsbereich sich nur noch auf diejenigen innerhalb der Stichproben beschränken würde, denen kein dominantes Kommunikationsmedium zugeordnet werden kann. Die in Abbildung 3 dargestellte Einteilung in Mainlys könnte dabei dazu beitragen, die Probleme zu lösen, die aufgrund der Verwendung eines Nonscreening-Ansatzes sowie durch virtuelle Festnetznummern erwachsen.

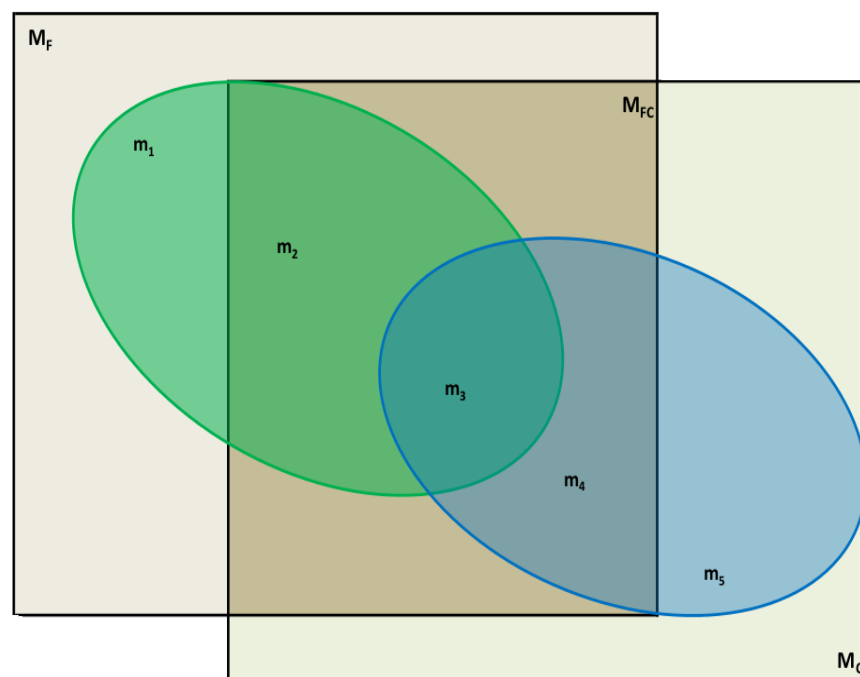


Abbildung 3: Neuzuordnung des Dual-Frames

Innerhalb dieser Zuordnung stellt M_F den Auswahlrahmen der Festnetzstichprobe, M_C den der Mobilfunkstichprobe und M_{FC} den Überlappungsbereich dar. m_1 ist somit der Teil einer Stichprobe, der in der vorherigen Vorgehensweise durch Landline-Onlys abgedeckt wurde. Respektive enthält m_5 nur Mobile-Onlys. m_2 bis m_4 ist der Anteil, der aus dem Überlappungsbereich gezogen wurde und bei dem daher eine Kontaktaufnahme sowohl anhand der Mobilfunk- als auch durch die Festnetzstichprobe erfolgen kann. m_2 sind dabei diejenigen, bei denen das Kommunikationsmedium Festnetz vornehmlich verwendet wird und m_4 die, die hauptsächlich über Handy zu erreichen sind. m_3 kann als derjenige Anteil verstanden werden, bei dem sich kein dominantes Kommunikationsmedium festlegen lässt und demnach den „neuen“, aber wahrscheinlich kleineren Überlappungsbereich darstellt und den man daher als „wahre“ Dual-User bezeichnen kann. Die Einteilung in Mainlys gestaltet sich demnach wie folgt:⁷

7 Eine ähnliche Aufteilung wurde in dem Artikel von Winglee et al. (2007) angewendet.

- Landline-Mainlys: $m_1 + m_2$
- Mobile-Mainlys: $m_4 + m_5$
- Dual-User: m_3

Ein Problem bei dieser Art der Einteilung könnte die genaue Zuordnung der Mainlys in den Bereichen m_2 und m_4 sein, da diese sowohl aus dem Auswahlrahmen Festnetz oder auch Mobilfunk stammen können, das Medium, mit dem diese kontaktiert wurden, aber nicht das dominante sein muss. Eine genaue Zuordnung nach der Stichprobenziehung ist somit erforderlich. Um dies zu erreichen, muss allerdings bereits während des Interviews anhand geeigneter Fragen eine solche Zuordnung ermöglicht werden, was wiederum höhere Anforderungen an das Fragebogendesign stellt.

Betrachtet man nun die Auswahlrahmen, aus denen die beiden Stichproben gezogen werden, so fällt auf, dass deren Größe sich in Abhängigkeit des für die Stichprobenziehung beauftragten Institutes unterscheidet. Als Grund hierfür kann die Grundlage der jeweiligen Auswahlrahmen angeführt werden, die in Abhängigkeit des jeweiligen Institutes entweder nur auf Einträgen innerhalb eines Telefonverzeichnisses beruht, oder noch zusätzliche Quellen bemüht. 2011 bestand dieses *Nummernuniversum* beim ADM aus insgesamt 118,1 Mio. Rufnummern, wobei diese sich aufteilen in 76,8 Mio., basierend auf 100er-Blöcken mit mindestens einem Eintrag und 41,27 Mio., basierend auf Blöcken ohne Eintrag (vgl. ADM [2012a], S. 11f.). Zum Vergleich dazu bestand der Auswahlrahmen der GESIS 2011, die diesen anhand des Gabler-Häder-Verfahrens erstellt, aus ca. 150 Mio. und zum Januar 2013 aus ca. 158,88 Mio. Einträgen.

Für Mobilfunkstichproben stellt sich die Gewinnung dieses Nummernuniversums schwieriger dar, da die Anzahl der Einträge mit ca. 2,1 Mio. deutlich geringer ist, als im Bereich der Festnetzstichprobe. Bei insgesamt 16 vergebenen Mobilfunkvorwahlen⁸ (Stand der Gajek-Liste vom 15.02.2012) und der sieben- bis achtstelligen Rufnummern ergibt sich, unter Berücksichtigung der Gajek-Liste, ein potentiell Nummernuniversum von mehr als 250 Mio. Rufnummern, bzw. dem ca. 2,2-fachen aller vergebenen Rufnummern (vgl. GAJEK [2012]).⁹ Die Bildung des Auswahlrahmens erfolgt sowohl bei GESIS als auch beim ADM auf ähnliche Weise, wobei GESIS eine Einteilung in 10.000-Blöcken, der ADM eine von 100.000-Blöcken verwendet. Zur Ermittlung der vergebenen Blöcke werden die Einträge der Telefonbuch-CD verwendet sowie durch weitergehende Recherche ermittelte Rufnummern unter Ausschluss technischer Sperren und Sonderrufnummern herangezogen (vgl. HÄDER [2009b], S. 43; ADM [2012a], S. 12f.). Der Auswahlrahmen bei GESIS besteht dabei aus 306,61 Mio.¹⁰, der des ADMs aus 247,9 Mio. Rufnummern.

Bei der Zusammensetzung der Stichproben aus den beiden Auswahlrahmen wurde im Verlauf beider CELLA-Studien, die GESIS in Kooperation mit der TU Dresden erstmals im Oktober 2007 durchgeführt hat, um Mixed-Mode-Studien im Fest- und Mobilfunknetz genauer zu untersuchen, eine Aufteilung von ca. 50% aus den jeweiligen Rahmen gewählt (vgl. HÄDER [2009b], S. 27; SCHNEIDERAT und SCHLIZIG [2012], S. 127). Dieser Anteil wurde angesichts fehlender Forschungsergebnisse in diesem Bereich gewählt, um erste Erkenntnisse über Dual-Frame Ansätze für Telefonbefragungen in Deutschland zu erlangen. Der ADM untersucht in seinem Forschungsbericht zu Dual-Frame-Stichproben anhand einer Proportionalisierung der einzelnen Designeffekte sowie unter Berücksichtigung der Kosten einer Erhebung das optimale Mischverhältnis zwischen Mobilfunk- und Festnetzstichprobe und stellt diesbezüglich fest, dass sich bei einem Mobilfunkanteil zwischen 20% und 60% „die Effektivität der adjustierten Designgewichte in einem vertretbaren Rahmen“ (ADM [2012a], S. 31) bewegen, die optimale Aufteilung aus 40% Mobilfunk und 60% Festnetz bestünde, aus Kostengründen allerdings ein Anteil der Mobilfunknummern von lediglich 30% zu empfehlen sei, da die gesteigerten Kosten eines größeren Anteils der Mobilfunkstichprobe nicht den geringen Gewinn an Effektivität rechtfertigen würden (vgl. ADM [2012a], S. 29 ff.).

⁸ zuzüglich drei reservierten Vorwahlen.

⁹ Der ADM geht 2011 in der Berechnung dieser potentiell vergebenen Nummern von 340,1 Mio. aus, (vgl. ADM [2012a], S. 13) bzw. dem 3,1-fachen aller tatsächlich vergebenen Rufnummern (vgl. HECKEL und WIESE [2012], S. 113).

¹⁰ Dieser enthält auch die von der Gajek-Liste als reserviert, allerdings noch nicht in Verwendung gekennzeichnete Vorwahl 0157.

3 Exkurs: HLR Lookup

Als neues Instrument zur Bereinigung der Liste der ausgewählten Rufnummern einer Mobilfunkstichprobe sowie zur Verbesserung der Berechnungsgrundlage der Response Rate hat sich in den vergangenen Jahren das HLR Lookup bewährt. Dieses Instrument kann aber (zumindest theoretisch) weit mehr leisten, als lediglich die Auswahl der Mobilfunkrufnummern zu bereinigen. Um die Möglichkeiten dieses Instrumentes zu erfassen, ist es allerdings von Bedeutung, den eigentlichen Aufbau eines Mobilfunknetzes zu verstehen. Dieser soll anhand dieses Exkurses kurz dargestellt werden, um im Anschluss auf die Potentiale des Instrumentes HLR Lookup einzugehen.

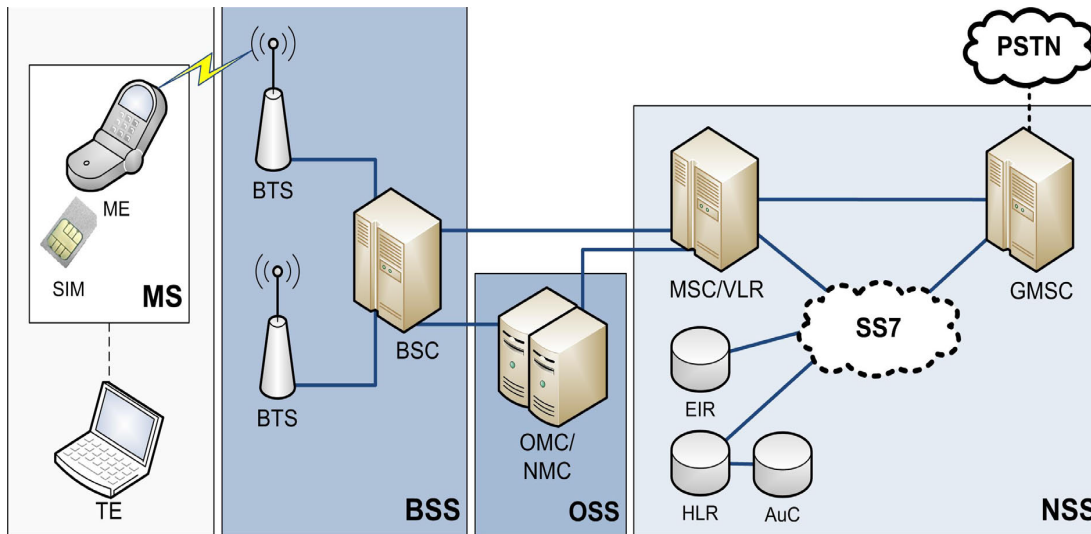
Bei den insgesamt vier dezentralen Mobilfunknetzen in Deutschlands werden innerhalb bestimmter Knotenpunkte, der MSCs (*Mobile Switching Center*), Gesprächsanfragen im gesamten Ablauf gesteuert. Das HLR (*Home Location Register*) ist in diesem Zusammenhang, ebenso wie das VLR (*Visitor Location Register*), als Teil dieser Netzwerke zu verstehen, das eng mit dem MSC in Verbindung steht. Da für die Erfassung von Rechnungsdaten relevant ist, ob ein vom MSC eingeleiteter Ruf von einem innerhalb dieses Netzes eingebuchten Teilnehmer ausgeht oder nicht, bedarf es zur Überprüfung schnelle Datenbanken.

HLR und VLR sind zur Ausfallsicherung redundant ausgelegt, wobei HLR Kundendaten ähnlich der Form eines Stammdatenblattes speichert und ebenso vermerkt, ob ein Kunde derzeit im (Heim-) Netz eingebucht ist, bzw. wo er zuletzt eingebucht war. Dabei geben in der Regel die ersten beiden Ziffern der Mobilfunknummer das für das Endgerät zuständige HLR an (vgl. Gajek). Es verwaltet die IMSI (*International Mobile Subscriber Identity*), eine eindeutige Kennung des Mobilfunkteilnehmers, das MSISDN (*Mobile Station ISDN Number*), also die Telefonnummer des Teilnehmers, gebuchte Zusatzdienste sowie die VLR Nummer zur Identifikation des aktuell für den Teilnehmer zuständigen VLRs.

Das VLR hält unter anderem fest, ob ein „Netzfremder“ sich derzeit in einem Netz eingebucht hat, um so abgehende Gespräche ohne Rückfrage des HLRs zu ermöglichen. Die wichtigste Aufgabe dieses Registers besteht allerdings darin, das MSC beim Aufbau der Verbindung zu unterstützen. Dazu speichert das VLR die Daten aller Mobilstationen, die sich derzeit im Geltungsbereich eines MSCs befinden und stellt zur Authentifizierung die vom HLR erhaltenen Nutzerdaten bereit. Es speichert demnach IMSI, MSISDN und die entsprechenden Sicherheitsparameter, die netzintern festgelegt sind. Darüber hinaus verwaltet es noch die TMSI (*Temporary Mobile Subscriber Identity*), die sich bei jedem Ortswechsel¹¹ ändert, die HLR-Nummer sowie die eindeutige Identifikation des Ortes (LAI für *Location Area Identity*), in der sich der Teilnehmer gerade befindet (vgl. ERICSSON GMBH [2001], S. 8f.; WINNHOVEN [2005], S. 15f.).

Bei der mittlerweile veralteten Netzinfrastruktur GSM (*Global System for Mobile Communication*) bestand pro Netz, das durch das NSS (*Network Subsystem*) definiert wurde, ein HLR und pro MSC ein VLR. Abbildung 4 zeigt eine vereinfachte Darstellung der GSM-Netzwerkarchitektur.

¹¹ Wechsel der Local Area, in der sich der Teilnehmer gerade befindet.



Quelle: BUNDESAMT FÜR SICHERHEIT IN DER INFORMATIONSTECHNIK [2008], S. 9.

Abbildung 4: Vereinfachte Darstellung der GSM-Netzwerkarchitektur

Beim Rufaufbau eines Mobilfunkteilnehmers über ein Endgerät, bestehend aus einem Mobiltelefon und einer SIM-Karte,¹² das sich durch Einschalten und Authentifizierung in das Netz des Anbieters einwählt, erfolgt eine Integration des Endgerätes über die Einwahl ins BSS (*Base Station Subsystem*). Dieses BSS besteht dabei aus mehreren BTS (*Base Transceiver Station*), die zum Senden und Empfangen von Signalen zuständig sind¹³ und einem *Base Station Controller* (BSC), der die entsprechenden Signale an das OSS (*Operations and Support System*) weiterleitet.¹⁴ Das OSS stellt anschließend Kontakt mit dem MSC her, das sich wiederum im NSS (*Network Subsystem*) befindet. Das NSS ist hierbei für die gesamte Kommunikation innerhalb eines Netzes verantwortlich. Diese erfolgt über das SS7 (*Signaling System Number 7*) und enthält die netzintern festgelegten Protokolle zur Telekommunikation. Das MSC erfragt nun mittels VLR über das SS7 die zum Verbindungsaufbau relevanten Daten des Nutzers, die im HLR gespeichert sind. Wird die Verbindung zu einem anderen Nutzer des gleichen Netzes aufgebaut, wird im Anschluss lediglich das dafür relevante MSC kontaktiert und das Signal weitergeleitet. Bei einer Verbindung in ein anderes Mobilfunknetz oder das Festnetz wird das Signal an das Gateway-MSC weitergeleitet (vgl. BUNDESAMT FÜR SICHERHEIT IN DER INFORMATIONSTECHNIK [2008], S. 9ff.).¹⁵

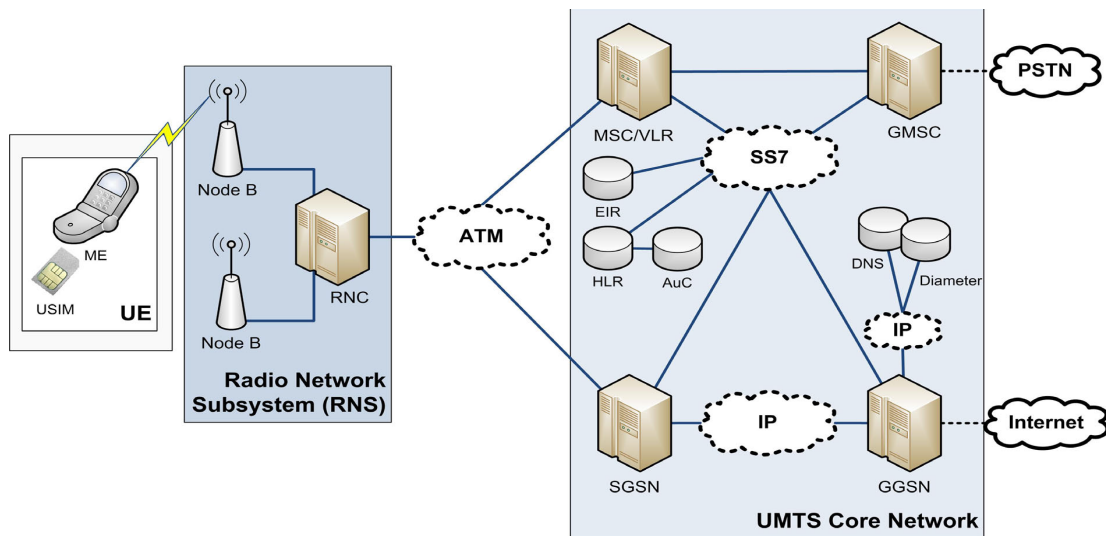
Mittlerweile wird das UMTS-Netzwerk (*Universal Mobile Telecommunication System*) verwendet, das eine ähnliche Architektur wie das GSM-Netzwerk hat. Abbildung 5 stellt eine vereinfachte Darstellung dieser Architektur dar.

¹² SIM steht diesbezüglich für *Subscriber Identity Module* und speichert die Identifikationsinformationen des Inhabers sowie das zwischen Teilnehmer und Netzbetreibers bestehende „Shared Secret“. Dieses ist ein nicht auslesbarer Algorithmus zur Authentifizierung des Teilnehmers.

¹³ Also einer Sendeantenne.

¹⁴ Die Summe aller BSS bildet den Sendebereich des GSM-Netzwerkes.

¹⁵ Da innerhalb dieser Arbeit lediglich der Kommunikationsweg im Mobilfunknetz relevant ist, wird auf eine spezifische Beschreibung der gesamten Architektur an dieser Stelle verzichtet. Diese ist bspw. nachzulesen bei BUNDESAMT FÜR SICHERHEIT IN DER INFORMATIONSTECHNIK [2008], S. 9ff.



Quelle: BUNDESAMT FÜR SICHERHEIT IN DER INFORMATIONSTECHNIK [2008], S. 35.

Abbildung 5: Vereinfachte Darstellung der UMTS-Netzwerkarchitektur

Der Unterschied besteht darin, dass dieses Netzwerk in ein UMTS Core Network und ein RNS (*Radio Network Subsystem*) aufgeteilt ist und das Endgerät initiativ bei einem Verbindungsaufbau mit dem RNS Kontakt aufnimmt. Das RNS teilt sich in einzelne Sendestationen (sog. Nodes) auf, die im Gegenzug zu denen des GSM-Netzwerkes über keine feste Sendereichweite verfügen, sondern in Abhängigkeit der Anzahl der Verbindungsanfragen ihre Reichweite und das Datenvolumen anpassen. Das im Anschluss über den RNC kontaktierte Core Network enthält die wesentlichen Elemente des GSM- sowie des GPRS- (*General Packet Radio Service*), bzw. HSDPA- (*High Speed Downlink Packet Access*) Netzwerkes.¹⁶ Zur netzinternen Kommunikation wird demnach ebenso das SS7 für den Bereich der Telefonie nach einem zum GSM vergleichbaren Schema verwendet. Ein Unterschied findet sich somit hauptsächlich in den Protokollschichten sowie im Bereich der RNS (vgl. BUNDESAMT FÜR SICHERHEIT IN DER INFORMATIONSTECHNIK [2008], S. 35f.; ERICSSON GMBH [2001], S. 16).

Das derzeit geplante und in einigen Städten bereits getestete LTE-Netzwerk (*Long Term Evolution*), das eine wesentlich höhere Datenrate erzielen soll und in seiner Architektur einfacher als das heutige UMTS-Netzwerk konzipiert ist, setzt im Gegenzug zu dem GSM-System nicht mehr auf die Speicherung der Kundendaten innerhalb eines HLRs, sondern hinterlegt diese in einem HSS (*Home Subscriber Server*), der dem Endgerät auch sein entsprechendes Gateway zuweist (vgl. GUTT [2010], S. 23ff.).

Ein HLR Lookup erlaubt innerhalb einer solchen Netzwerkarchitektur eine Abfrage multipler HLRs für einzelne oder gebündelte Rufnummern und somit das Erfragen, ob eine bestimmte Rufnummer vergeben ist. Darüber hinaus wird auch der Status einer Rufnummer bei Abfrage, der MNC/ MCC (*Mobile Network/ Country Code*) sowie das MSC¹⁷, das zumindest theoretisch eine approximative Bestimmung der geografischen Position ermöglicht, angegeben. So wäre es bspw. denkbar, ein HLR Lookup nachts, wenn man davon ausgehen kann, dass ein Großteil der Endgerätebesitzer sich an ihrem tatsächlichen Wohnort befinden, durchzuführen, um so die Position des entsprechenden VLRs in Erfahrung zu bringen. Der MNC und MCC gewinnen besonders im Zuge der Rufnummernportabilität an Relevanz, da der MNC Netz- und der MCC Landeszugehörigkeit beschreibt. Bei einer HLR Lookup-Anfrage wird allerdings standardmäßig nur die Nummer, deren Status sowie deren MNC und MCC weitergeleitet (vgl. ROUTO MESSAGING [2009], S. 4ff.). Der Anbieter dieser Dienstleistung wendet sich dabei direkt mit den entsprechenden Nummern der Endgeräte an das SS7 und greift darüber auf die Datenbank des jeweiligen HLRs

¹⁶ Auf eine Erläuterung des GPRS-, bzw. HSDPA-Netzwerkes, wird an dieser Stelle, da es sich um ein Netzwerk zur Datenübertragung über ein TCP/IP-Protokoll handelt, verzichtet. Dies ist bspw. unter BUNDESAMT FÜR SICHERHEIT IN DER INFORMATIONSTECHNIK [2008], S. 29ff. nachzulesen.

¹⁷ Die MSC werden allerdings bei Anfragen innerhalb Deutschlands aus rechtlichen Gründen nicht mit angegeben.

zu (vgl. INFOBIB [2011], S. 1).¹⁸

Eine Möglichkeit zur Kostenreduktion innerhalb der Mobilfunkbefragung besteht durch die Inanspruchnahme eines HLR Lookups dahingehend, dass bereits im Vorfeld der Befragung, nachdem eine Stichprobe aus dem Mobilfunkrahmen gezogen wurde, nicht vergebene Rufnummern identifiziert und die für die Befragung relevanten Listen an Mobilfunknummern bereits um nicht vergebene Einträge bereinigt werden können. Die Kosten einer solchen Anfrage beginnen bei 0,7 Cent pro Lookup (vgl. MOBIMEX [2013]).¹⁹ Maximal können sich diese auf bis zu ca. 7,5 Cent pro Lookup belaufen (vgl. NUMBERPORTABILITY LOOKUP [2013]).

Bezüglich der Genauigkeit der Identifikation dieser Anfragen stellen Struminskaya et al. innerhalb einer Untersuchung des Mobile-Samples der CELLA2 Studie fest, dass Fehlklassifikationen des Status von Mobilfunknummern sehr gering ausfallen und diese evtl. auch durch die Zeitverzögerung, die zwischen dieser Untersuchung und der eigentlichen Erhebung besteht, begründet werden können. Ein eindeutiger Vorteil des HLR Lookups ist daher die Kostenersparnis innerhalb der Stichprobe als Resultat der reduzierten Interviewerarbeitszeit. Ein weiterer, nicht unwesentlicher Vorteil besteht in der Verbesserung der Berechnungsgrundlage der Response-Rate, da unzulässige Rufnummern bereits im Vorfeld eliminiert werden können. Ein möglicher Nachteil bei einer solchen Vorgehensweise können die bereits im Vorfeld anfallenden Kosten sein. Diese sollten sich allerdings durch die verringerte Interviewerarbeitszeit aufgrund der bereinigten Rufnummernliste zumindest amortisieren.

Eine geografische Einordnung anhand eines HLR Lookups kann wegen der Verzichtserklärungen der Serviceanbieter über die angebotenen Informationen der Anschlüsse problematisch sein, da bei einem solchen Lookup lediglich Auskünfte über die Telefonnummer, nicht aber über persönliche Daten wie bspw. den Wohnort, bereitgestellt werden (vgl. STRUMINSKAYA [2011], S. 7ff.).²⁰

18 Neuere Untersuchungen zeigen, dass anhand der Mobilfunknummer zumindest eine grobe Einordnung des Wohnortes bei Rufnummernvergabe in den Netzen von D1 und D2 möglich ist. Siehe hierzu: ENGEL [2013], S. 11ff.

19 Die Kosten, den gesamten Auswahlrahmen für Mobilfunknummern bspw. der GESIS überprüfen zu lassen, würde einmalig ca. 2.146.270 Euro kosten, wobei die Gültigkeit dieser Kontrolle aufgrund der ständigen Neuvergabe von Rufnummernbereichen nur von beschränkter Dauer wäre. Daher empfiehlt sich eine fallweise Überprüfung der einzelnen Stichprobenziehungen.

20 So ist eine approximative Bestimmung des Wohnortes anhand dieser Daten nicht möglich, da das MSC zwar von den Serviceanbietern abgefragt werden kann, diese Angaben aber in der Regel nicht übermittelt werden.

4 Gewichtung von Dual-Frame-Stichproben

Bezüglich der Fragestellung, ob innerhalb einer Dual-Frame-Stichprobe eine Gewichtung notwendig ist, stellt die Cell Phone Taskforce der AAPOR 2010 fest, dass bei einer Grundgesamtheit, bei der über mehrere Wege die Möglichkeit besteht in eine Stichprobe zu gelangen, nur in wenigen Ausnahmefällen keine Gewichtung von Nöten ist. Als Beispiel für Stichproben, die nach Ansicht dieser Taskforce auch ohne eine Gewichtung auskommen²¹, wird der Fall angeführt, bei dem lediglich ein Auswahlrahmen verwendet wird und kein Versuch unternommen wird, generalisierte Aussagen auch über den anderen Auswahlrahmen zu treffen.²² Darüber hinaus muss dieser Taskforce zufolge bei der Einführung eines neuen Modes der Surveyverwaltung, der auf Fortschritten im Bereich der Telekommunikationstechnologien beruht, keine Gewichtung vorgenommen werden (vgl. AAPOR CELL PHONE TASK FORCE [2010], S. 62ff.).

Eine Erschwernis innerhalb der Gewichtung einer Dual-Frame-Telefonstichprobe kann die zugrunde liegende Annahme der beiden einzelnen Auswahlrahmen darstellen. So wird eine Festnetzstichprobe als eine Haushaltsstichprobe, bei der eine Transformationsgewichtung auf Personenebene erfolgt, aufgefasst, während es sich bei der Befragung im Mobilfunknetz um eine Personenstichprobe handelt (vgl. HÄDER [2009a], S. 18).²³

Der Frage, ob es sich bei Mobiltelefonen tatsächlich um personalisierte Gegenstände handelt und man mit einer solchen Stichprobe nur die Zielperson erreicht, nehmen sich Busse und Fuchs 2013 an. In ihrer Untersuchung stellen die Autoren dabei fest, dass es keine allgemeingültige Definition der gemeinschaftlichen Nutzung von Mobiltelefonen gibt, wodurch sich die Diskrepanz verschiedener Untersuchungen begründen lässt. Darüber hinaus schließen einige Studien in ihrer definitorischen Grundlage auch Haushaltsmitglieder ein, die nicht erhebungsrelevant sind (vgl. BUSSE [2013], S.3.). Aufgrund dessen teilen sie in ihrer Studie Personen, die auch über Mobilfunkgeräte anderer erreichbar sind und Personen, über deren Mobilfunkgerät auch andere erreicht werden können, in aktive und passive „Teiler“ ein. Als Ergebnis halten die beiden Autoren fest, dass bei einigen Fällen zwar kein anteiliges, gleichberechtigtes Teilen von Mobilfunkgeräten gibt, potentiell allerdings bei ca. der Hälfte der Befragten der Kontakt bei einer Mobilfunkstichprobe mit einer Person aufgebaut werden kann, die nicht die eigentliche Zielperson darstellt (vgl. BUSSE [2013], S. 10ff.).

Als problematisch kann zusätzlich die Inklusion virtueller Festnetznummern betrachtet werden, da diese im Vorfeld nicht als solche identifiziert werden können und somit die Möglichkeit besteht, dass der Anteil der 1-Personen-Haushalte überschätzt wird. Grund hierfür ist, dass eigentlich ein Mobilfunkgerät erhoben wird, bei dem die Umstände der Befragung sich von denen einer „regulären“ Festnetzstichprobe unterscheiden.

Anhand der beiden Auswahlrahmen sind vier mögliche Sampling-Designs, die eine entsprechende Gewichtung benötigen, möglich. So ist eine Ziehung aus beiden Frames, entweder mit oder ohne Überlappungsbereich, als auch die Beschränkung auf den Auswahlrahmen Festnetz, bzw. Mobilfunk denkbar.²⁴ Im Falle des Designs ohne Überlappungsbereich wird dabei davon ausgegangen, dass zuvor ein Screening vorgenommen wurde, um so zu bestimmen, welcher Teil der Grundgesamtheit ausschließlich über

21 Als Beispiel hierfür wäre der Fall anzuführen, bei dem alle Elemente der Stichprobe die gleiche Inklusionswahrscheinlichkeit haben ($\pi_i = c$)

22 Wobei auch in diesem Fall von AAPOR angemerkt wird, dass eine Gewichtung vorgenommen werden sollte, wenn nicht-ignorierbare Unterschiede in den Auswahlwahrscheinlichkeiten vorliegen oder eine hohe Nonresponse-Rate besteht.

23 Innerhalb des Forschungsberichtes Dual-Frame-Ansätze des ADMs wird auch bei der Mobilfunkstichprobe eine Gewichtung auf der Basis der Haushaltsgröße vorgenommen (vgl. ADM [2012a], S. 18). Aufgrund mehrerer Untersuchungen sowie anhand weitläufiger Diskussionen innerhalb der einschlägigen Literatur herrscht allerdings der allgemeine Konsens, dass es sich bei einem Mobiltelefon um einen personalisierten Gegenstand handelt (siehe bspw. GABLER [2012], S. 150ff.; KUUSELA und SIMPANEN [2012], S. 41; AAPOR CELL PHONE TASK FORCE [2010], S. 65ff), weswegen auf die Anpassung anhand der Haushaltsgröße bei Mobilfunkstichproben an dieser Stelle verzichtet wird.

24 Aufgrund der Ausrichtung dieser Arbeit wird an dieser Stelle auf eine nähere Untersuchung der Beschränkung auf lediglich einen Auswahlrahmen verzichtet.

einen Auswahlrahmen erreicht werden kann. Im Anschluss wird dann nur dieser Anteil anhand des entsprechenden Auswahlrahmens gezogen, während der restliche Anteil der Grundgesamtheit, bestehend aus Dual-Usern sowie der Onlys des anderen Auswahlrahmens, anhand des anderen (evtl. günstigeren) Auswahlrahmens untersucht wird.

Innerhalb dieses Screening-Ansatzes kann der AAPOR-Taskforce zufolge das Design wie eine geschichtete Stichprobe aufgefasst werden, bei der man die drei Bereiche wie einzelne Schichten behandelt. Die Gewichte werden dafür entsprechend der drei Bereiche gewählt und anschließend kombiniert. Das sich daraus ergebende Problem besteht darin, dass die Gewichte innerhalb des hauptsächlich verwendeten Auswahlrahmens so zu wählen sind, dass diese auch gruppenspezifische Nonresponse aufgrund der Wahl eines dominanten Telekommunikationsmediums berücksichtigen, da Dual-User nur in die Stichprobe gelangen, wenn diese auch durch den entsprechenden Auswahlrahmen erreicht wurden. Daher besteht die Gefahr, dass Mainlys des vorher durch Screening gesondert erhobenen Auswahlrahmens systematisch unterschätzt werden (vgl. AAPOR CELL PHONE TASK FORCE [2010], S. 62ff.).

Bei einem Non-Screening-Ansatz muss während der Durchführung der Befragung ermittelt werden, ob bei der betreffenden Befragung die Möglichkeit besteht, dass der Befragte auch über den anderen Auswahlrahmen in die Stichprobe gelangen kann. Die Gewichtung muss dementsprechend unter der Berücksichtigung beider Inklusionswahrscheinlichkeiten erfolgen. Als problematisch innerhalb dieser Vorgehensweise ist besonders die Gewichtung der Beobachtungen zu erachten, die durch beide Auswahlrahmen in die Erhebung gelangen können, da bei diesen die Möglichkeit besteht, dass ein solcher Haushalt über beide Stichproben erhoben wird. Die Wahrscheinlichkeit, dass dies eintritt, wird allerdings als sehr gering eingeschätzt und kann im Falle von Telefonstichproben als vernachlässigbar betrachtet werden (vgl. AAPOR CELL PHONE TASK FORCE [2010], S. 69f.; GABLER [2012], S. 51ff.).

Wird nun bei einer Erhebung, bei der der Bereich M^{FC} nicht leer ist (siehe Abb. 3), eine Zufallsstichprobe mit dem Umfang n^F aus der Grundgesamtheit vom Umfang N^F mit dem Auswahlrahmen M_F sowie eine davon unabhängige Zufallsstichprobe n^C aus N^C des Auswahlrahmens M_C gezogen, ergibt sich ein Gesamtstichprobenumfang von $n = n^F + n^C$, bei der die i -te Einheit der Stichprobe $S(F)$ (mit $i = 1, \dots, n$) eine Inklusionswahrscheinlichkeit von π_i^F aufweist und die Inklusionswahrscheinlichkeit der Stichprobe $S(C)$ sich analog durch π_i^C berechnen lässt. Die Horvitz-Thompson Gewichte w_i^F und w_i^C werden dann durch

$$w_i^F = 1/\pi_i^F$$

und

$$w_i^C = 1/\pi_i^C$$

ermittelt. Daher entspricht

$$E\left(\sum_{i \in S(F)} w_i^F * y_i\right) = Y^F + Y^{FC}$$

und

$$E\left(\sum_{i \in S(C)} w_i^C * y_i\right) = Y^C + Y^{FC}.$$

y_i entspricht dabei der Ausprägung der interessierenden Variable der Einheit i , Y^F und Y^C stellen die Totalwerte dieser Variable aus den jeweiligen Only-Bereichen dar und Y^{FC} den des Überlappungsbereiches. Aufgrund dessen führt die Kombination der beiden Stichproben mit den ursprünglichen Gewichten durch

$$\hat{Y} = \sum_{i \in S(F)} w_i^F * y_i + \sum_{i \in S(C)} w_i^C * y_i$$

zu einer verzerrten Schätzung des Totalwertes Y (vgl. LOHR [2011], S. 200). Um dies zu vermeiden, werden im Folgenden unterschiedliche Verfahren zum Umgang mit diesem Überlappungsbereich vorgestellt.

4.1 Single Frame Ansätze

Eine Möglichkeit zur Gewichtung innerhalb eines Dual-Frame-Szenarios besteht in dem erstmals von Bankier (1986) und Kalton und Anderson (1986) hervorgebrachten Ansatz, die Beobachtungen aus beiden Stichproben zu einer Art „Super-Stratum“ zu kombinieren und so zu behandeln, als stammen diese aus nur einer Stichprobe. Dabei müssen die Gewichte des Überlappungsbereiches so angepasst werden, dass die Inklusionswahrscheinlichkeit für Erhebungseinheiten in beiden Stichproben berücksichtigt wird. Lohr (2006) zufolge bedarf es, um einen Schätzer nach Bankier zu erlangen, der eindeutigen Identifikation der Schichtzugehörigkeit der einzelnen Erhebungseinheiten. Dazu sind allerdings die entsprechenden Daten zur Einordnung von Bedeutung (vgl. LOHR und RAO [2006], S. 1023).²⁵ Ein Vorzug dieser Vorgehensweise besteht nach Lohr und Rao (2000) darin, dass zur Bestimmung dieser Schätzwerte keine zusätzlichen Informationen bezüglich der Größen der Grundgesamtheit N^F und N^C benötigt werden. Die Anpassung kann hierfür durch einen Raking-Verhältnisschätzer oder einen Regressionsschätzer erfolgen. Ein solches Gewicht kann demnach durch

$$w_{i,1} = \begin{cases} w_i^F = (\pi_i^F)^{-1} & \text{wenn } i \in m_1 \\ w_i^{FC} = (\pi_i^F + \pi_i^C)^{-1} & \text{wenn } i \in m_2 \cup m_3 \cup m_4 \\ w_i^C = (\pi_i^C)^{-1} & \text{wenn } i \in m_5 \end{cases} \quad (4.1)$$

bestimmt werden (vgl. LOHR und RAO [2000], S. 272).

Ein ähnliches Gewichtungsverfahren wurde bspw. in den CELLA 1 und 2 Studien verwendet. Die auf die Fallzahl normierten Gewichte stellen sich nach

$$w_i = n \frac{\frac{1}{\pi_i}}{\sum_{j \in S} \frac{1}{\pi_j}},$$

mit $i \in S$, dar. Unter Berücksichtigung der jeweiligen Stichproben- und Auswahlrahmenumfänge ergab sich im Zuge der CELLA1-Studie folgendes Verhältnis

$$\frac{M_F}{n^F} * \frac{n^C}{M_C} = 1,04$$

wobei n^F und n^C die jeweiligen Stichprobenumfänge bezeichnen. Das entsprechende Gewicht w_i lässt sich demnach ungefähr durch

$$w_i \approx \frac{n}{(1,04 * \frac{k_i^F}{z_i} + k_i^C) * \sum_{j \in S} \frac{1}{1,04 * \frac{k_j^F}{z_j} + k_j^C}}$$

beschreiben. Dabei meinen k_i^F und k_i^C die Anzahl der Telefonnummern aus den beiden Auswahlrahmen über die das Stichprobenelement i erreicht werden kann²⁶ und z_i die Haushaltsgröße des Haushaltes der i -ten Einheit (vgl. HÄDER [2009b], S. 29ff.).

Der ADM verwendet zur Gewichtung ihrer Dual-Frame-Telefonstichproben ein dem der CELLA-Studien nachempfundenen Verfahren. Dafür wird die Inklusionswahrscheinlichkeit auf eine ähnliche Weise bestimmt, mit dem Unterschied, dass auch die Mobilfunkstichprobe die Haushaltsgröße berücksichtigt und somit die Annahme gestellt wird, dass ein Mobilfunkanschluss von mehreren Personen verwendet wird.²⁷

²⁵ Aus den Erfahrungen der CELLA1 und 2 Studien heraus, sollte eine solche Zuordnung als unproblematisch erachtet werden.

²⁶ Im Verlauf der CELLA1-Studie wurde einem Analog-Anschluss im Festnetzbereich der Wert 1 zugeordnet, für eine ISDN-Anlage wurde der Mittelwert 2,5 verwendet. Aufgrund der Neuberechnung der Mittelwerte in CELLA2 zeigt sich allerdings, dass die zuvor verwendeten Werte als unpräzise und sich häufig ändernd zu betrachten sind. So wurde für die Anzahl der Analoganschlüsse ein Mittelwert von 1,36, für ISDN-Anschlüsse ein Mittelwert von 1,69 errechnet. Die Autoren merken dazu an, dass die Angaben diesbezüglich häufig falsch sind, da Befragte die exakte Antwort oft nicht wissen (vgl. Gabler, S. 153f.).

²⁷ Siehe Fußnote 22.

Darüber hinaus wird in einem anschließenden Schritt erst eine Proportionalisierung unter der Annahme, dass die in beiden Stichproben erhobenen Dual-User richtig abgebildet werden, vorgenommen und die Teilgruppe der Festnetzstichprobe auf die Größenordnung der Mobilfunkstichprobe umgerechnet.²⁸ Darauf folgend wird eine Gewichtung auf der Basis der Demographie vorgenommen. Anhand der Proportionalisierung wird der gesamte Stichprobenumfang n_{ges} auf²⁹

$$n_{ges} = n^C + 2n_{FC}^C + \frac{n^F * n_{FC}^C}{n_{FC}^F}$$

umgerechnet. n^C beschreibt den Stichprobenumfang der Teilgruppe Mobile-Only, n^F den der Landline-Onlys, n_{FC}^C den Umfang der Dual-User aus der Mobilfunkstichprobe und n_{FC}^F entsprechend den der Dual-User aus der Festnetzstichprobe. Bei einer getrennten Betrachtung der Designgewichte der beiden Stichproben beinhaltet dieses Verfahren das ursprüngliche Designgewicht basierend auf der Inklusionswahrscheinlichkeit nach $1/\pi_i$, die Normierung der Gesamtgewichte auf die ungewichtete Fallzahl, sowie die Anpassung auf die Größe der Mobilfunkstichprobe.³⁰

Im Anschluss an die Proportionalisierung wird anhand der Vorgaben des Mikrozensus nach soziodemographischen Strukturen durch iterative Randsummenverfahren gewichtet. In seinem Forschungsbericht 2012 stellt der ADM fest, dass sich anhand der Proportionalisierung der beiden Dual-Use-Gruppen die Daten der Gesamtstichprobe „robuster im Bezug auf das Stichprobenverhältnis Mobil - Festnetz gestalten(n)“ (ADM [2012a], S. 29). Das zweite Gewichtungsverfahren soll die unterschiedliche Teilnahmebereitschaft und Erreichbarkeit der beiden Gruppen reflektieren. Hierbei gilt anzumerken, dass unter den drei verschiedenen Schätzvorgängen unterschiedliche Werte bezüglich des Anteils der Onlys berechnet werden. Der ADM verweist in diesem Kontext auf den Zusammenhang des Erhebungsinstrumentariums und des geschätzten Wertes (vgl. ADM [2012a], S. 23ff.).

Ein weiterer Vorschlag in der Klasse der Single-Frame-Gewichte wurde von Bankier (1986) hervorgebracht. Bei diesem werden die Beobachtungen im Überlappungsbereich, die aus beiden Stichproben gezogen wurden, in der Designgewichtung berücksichtigt. Die Gewichte bestimmen sich dann nach:

$$w_{i,2} = \begin{cases} w_i^F = (\pi_i^F)^{-1} & \text{wenn } i \in m_1 \\ w_i^{FC} = (\pi_i^F + \pi_i^C - \pi_i^F * \pi_i^C)^{-1} & \text{wenn } i \in m_2 \cup m_3 \cup m_4 \\ w_i^C = (\pi_i^C)^{-1} & \text{wenn } i \in m_5 \end{cases} \quad (4.2)$$

Nach Winglee et al. (2007) sind dabei die Ergebnisse aus 4.1 und 4.2 vergleichbar, wenn $\pi_i^F * \pi_i^C$ klein ist und sich relativ wenige Beobachtungen der Stichprobe innerhalb des Überlappungsbereiches befinden (vgl. WINGLEE [2007], S. 3197). Aufgrund der Annahme, dass es als unwahrscheinlich zu erachten ist, dass der gleiche Haushalt über beide Stichproben erhoben wird, wählte man auch innerhalb der CELLA-Studien ein Vorgehen nach 4.1 anstelle von 4.2 (vgl. HÄDER [2009b], S. 28f).

Winglee et al. beschreiben darüber hinaus noch ein weiteres Verfahren, das diese als *Pseudo Single-Frame Schätzung* bezeichnen. Mit den Schwierigkeiten in der Berücksichtigung des Überlappungsbereiches wird wie folgt umgegangen:

$$w_{i,3} = \begin{cases} w_i^F & \text{wenn } i \in m_1 \cup m_2 \cup m_3 \\ 0 & \text{wenn } i \in m_4 \\ w_i^C & \text{wenn } i \in m_5 \end{cases} \quad (4.3)$$

28 Beim Dual-Frame-Design des ADMs wird eine Zusammensetzung von 70% Festnetz und 30% Mobilfunk angewendet.

29 Mit der Umrechnung der Festnetzstichprobe auf die Mobilfunkstichprobe durch $n^C + \frac{n^F * n_{FC}^C}{n_{FC}^F}$.

30 Auf eine Ausführung der getrennten Stichproben wird an dieser Stelle verzichtet. Siehe hierzu: ADM [2012a], S. 24f.

Ein Totalwert Y nach diesem Verfahren würde sich demnach durch

$$\hat{Y}_3 = \sum_{i \in m_1 \cup m_2 \cup m_3} w_i^F * y_i^F + \sum_{i \in m_5} w_i^C * y_i^C$$

schätzen lassen. Eine solche Vorgehensweise entspricht der, bei der durch ein Screening Überlappungsbereiche im Vorfeld beseitigt wurden.³¹ Als Nachteil der Gewichtung nach 4.3 bei einem Non-Screening Ansatz kann allerdings der Verlust an Daten durch die Auslassung des Bereiches m_4 betrachtet werden. Als letztes wird von Winglee et al. eine Gewichtung nach einem *Modified Single-Frame Ansatz*, bei dem die Gewichtung nach

$$w_{i,4} = \begin{cases} w_i^F & \text{wenn } i \in m_1 \cup m_2 \\ 1 & \text{wenn } i \in m_3 \cup m_4 \\ w_i^C & \text{wenn } i \in m_5 \end{cases} \quad (4.4)$$

vorgeschlagen. Der Totalwert Y ließe sich demnach durch

$$\hat{Y}_4 = \sum_{i \in m_1 \cup m_2} w_i^F y_i^F + \sum_{i \in m_3 \cup m_4} y_i + \sum_{i \in m_5} w_i^C y_i^C$$

schätzen. Im Gegenzug zur Totalwertberechnung des Auswahlrahmens Festnetz im Pseudo Single-Frame Ansatz berechnet diese durch

$$\hat{Y}_4^F = \sum_{i \in m_1 \cup m_2 \cup m_3 \cup m_4} w_{i,4} y_i.$$

Die Autoren kommen dabei zu dem Schluss, dass, um den Verlust von Informationen in m_4 zu vermeiden, die Gewichte der beiden Gruppen m_3 und m_4 dahingehend zu beschränken sind, dass diese keinen Wert kleiner 1 annehmen, was zu $w_{i,m_3} = w_{i,m_4} = 1$ führt. Akzeptiert man den Verlust von Informationen für m_4 , so gilt $w_{i,m_3} = w_i^F$ und demnach eine Gewichtung nach 4.3 (vgl. WINGLEE [2007], S. 3196 ff.).

4.2 Composite Weighting

Eine weitere Möglichkeit zur Anpassung der Gewichte innerhalb einer Dual-Frame Erhebung besteht in der Verwendung von Composite Weights, bei denen anhand einer Konvexkombination die Gewichtung des Überlappungsbereiches entsprechend angepasst werden. Um dies zu erlangen, werden die Gewichte der einzelnen Stichproben so modifiziert, dass die Schätzer approximativ unverzerrt sind. Diese neuen Gewichte lassen sich anhand von $\tilde{w}_i^F = c_{i,\lambda}^F * w_i^F$ und $\tilde{w}_i^C = c_{i,\lambda}^C * w_i^C$ beschreiben.

Die diesbezüglichen Gewichte entsprechen dann der Form:

$$c_{i,\lambda}^F = \begin{cases} 1 & \text{wenn } i \in m_1 \\ \lambda & \text{wenn } i \in m_2 \cup m_3 \end{cases} \quad (4.5)$$

$$c_{i,\lambda}^C = \begin{cases} 1 & \text{wenn } i \in m_5 \\ 1 - \lambda & \text{wenn } i \in m_3 \cup m_4 \end{cases}$$

mit $\lambda \in [0,1]$, da die Schätzer approximativ unverzerrt sind, wenn $c_{i,\lambda}^F \approx 1$ für $i \in m_1$, $c_{i,\lambda}^C \approx 1$ für $i \in m_5$ und $c_{i,\lambda}^F + c_{i,\lambda}^C \approx 1$ für $i \in m_2 \cup m_3 \cup m_4$, gilt.

31 Siehe S. 18.

Der Schätzer für den Totalwert Y berechnet sich anhand dieses Verfahrens durch

$$\begin{aligned}\hat{Y}(\lambda) &= \sum_{i \in S(F)} c_{i,\lambda}^F * w_i^F * y_i + \sum_{i \in S(C)} c_{i,\lambda}^C * w_i^C * y_i \\ &= \hat{Y}_{m_1}^F + \lambda * \hat{Y}_{m_2 \cup m_3} + (1 - \lambda) * \hat{Y}_{m_3 \cup m_4} + \hat{Y}_{m_5}^C.\end{aligned}\quad (4.6)$$

Zur Bestimmung von λ führt Hartley (1962) das vergleichsweise einfache Verfahren des *Fixed Weight Adjustments* an, bei dem ein zuvor festgelegter Wert verwendet wird. Als vorteilhaft bei dieser Verfahrensweise kann die dadurch erlangte interne Konsistenz, durch die für alle Variablen einer Erhebung die gleiche Menge an angepassten Gewichten verwendet wird, erachtet werden. Darüber hinaus ist dieser Ansatz einfach in der Handhabung und der Implementierung. Lohr (2011) stellt fest, dass die Effizienz dieser Methode stark von den zuvor gewählten Gewichten abhängt. Dabei empfehlen Brick et al. (2006) die Verwendung von $\lambda = 0,5$, da bei ihren Untersuchungen des Current Population Surveys (CPS) und des National Health Interview Surveys (NHIS) ein optimaler Wert für λ durch die Annäherung anhand der Stichprobengröße in Höhe von 0,42 errechnet wurde und es der Auffassung der Autoren entspricht, dass die Einbußen der Effizienz durch die lediglich geringe Abweichung vom Optimalwert verhältnismäßig klein ausfallen. Die Wahl des Wertes 0 oder 1 führt bei diesem Verfahren zum Ausschluss von einem der beiden Auswahlrahmen im Überlappungsbereich (vgl. LOHR [2011], S. 200, BRICK [2006], S. 786). Nach Xia et al. (2010) führt dieser Ansatz allerdings zu ineffizienten Schätzungen, wenn die Auswahlwahrscheinlichkeiten in den beiden Auswahlrahmen sich stark unterscheiden.

Aufgrund dessen werden von den Autoren weitere Möglichkeiten zur Bestimmung von λ angeführt. Anhand der Überlegung, dass beide Stichproben dieselbe Grundgesamtheit darzustellen versuchen, besteht ein mögliches Vorgehen in der proportionalen Anpassung λ 's an die Stichprobengröße der beiden Zufallsstichproben. λ berechnet sich demnach durch

$$\lambda = \frac{n^F}{n^C + n^F}.\quad (4.7)$$

Der Vorzug dieses Vorgehens besteht nach Auffassung der Autoren darin, dass dies größere Gewichte der kleineren Stichproben stärker reduziert als kleine Gewichte der größeren Stichprobe. Sind die beiden Stichproben gleich groß, so entspricht dieses Vorgehen der Empfehlung von Brick et al.

Eine weitere Alternative bei stark variierenden Gewichten besteht darin, λ mittels der *effektiven Stichprobengröße* anzupassen, so dass λ die Informationen, die aus den einzelnen Stichproben stammen, berücksichtigt und nicht nur die reine Stichprobengröße. λ berechnet sich dabei proportional zur effektiven Stichprobengröße (der Festnetzstichprobe) durch das Verhältnis der Stichprobengrößen n^F und n^C zu den jeweiligen Designeffekten d^F und d^C nach

$$\lambda = \frac{\frac{n^F}{d^F}}{\frac{n^C}{d^C} + \frac{n^F}{d^F}},\quad (4.8)$$

wobei d^F und d^C durch

$$\begin{aligned}d^F &= 1 + (CV(w_i^F \in S(F)))^2 \\ d^C &= 1 + (CV(w_i^C \in S(C)))^2\end{aligned}$$

ermittelt werden können. CV steht hierbei für den Variationskoeffizienten der jeweiligen Gewichte. Da sich die Coverage innerhalb der beiden Stichproben unterscheidet und dieser anhand der summierten Gewichte geschätzt werden kann, wird überdies angeführt, λ durch die *gewichtete Stichprobengröße* anhand von

$$\lambda = \frac{\sum_{S(F)} w_i^F}{\sum_{S(C)} w_i^C + \sum_{S(F)} w_i^F}\quad (4.9)$$

zu berechnen. Aufgrund der Überlegung, dass es wahrscheinlicher ist, Mainlys in der Stichprobe zu erreichen, die deren dominantes Telekommunikationsmedium verwendet, wird innerhalb einer weiteren Option zur Bestimmung von λ durch Xia et al. (2010) eine verzerrungskorrigierende Berechnung vorgeschlagen. Dazu werden die Stichprobenumfänge dahingehend angepasst, dass sie dem in der Grundgesamtheit angenommenen Anteil der Mainlys entspricht. u bezeichnet dabei die Verwendungskategorie, die sich durch die Nutzung der jeweiligen Telekommunikationsmediem bestimmen lässt, (P_u^F, P_u^C) die für die Grundgesamtheit angenommenen Anteile der Festnetz- und Mobile-Mainlys und (p_u^F, p_u^C) die innerhalb der Erhebung beobachteten Anteile der jeweiligen Mainlys. λ berechnet sich demnach durch

$$\lambda_u = \frac{\left(\frac{P_u^F}{p_u^F}\right)}{\left(\frac{P_u^F}{p_u^F}\right) * n^F + \left(\frac{P_u^C}{p_u^C}\right) * n^C} \quad (4.10)$$

In der von Xia et al durchgeführten Untersuchung der fünf verschiedenen Verfahren zur Bestimmung von λ unter der Verwendung der Daten des Minnesota Health Access Surveys (MNHA) wird festgestellt, dass die Varianz der Schätzwerte des Totalwertes direkt mit der Varianz der Gewichte zusammenhängt. Darüber hinaus wird angenommen, dass die Verwendung des Ansatzes, der lediglich die Stichprobenumfänge berücksichtigt (4.7), unverzerrte Schätzwerte liefert. Nach Untersuchungen bezüglich der Zusammenstellung der Stichproben und deren Kombination mit den unterschiedlichen Verfahren zur Berechnung von λ anhand von Varianz und MSE, kommen die Autoren zu dem Schluss, dass die effektive Stichprobengröße (4.8) innerhalb dieser Verfahren das beste darstellt, da diese den höchsten Wert für λ erzeugen und somit der Mobilfunkstichprobe geringere Gewichte zuordnet. Als Konsequenz weisen die Schätzwerte dieser Untersuchung zufolge einen geringeren Standardfehler auf. Das am schlechtesten abschneidende Verfahren ist hier das, das λ anhand der gewichteten Stichprobengröße (4.9) ermittelt. Bei der verzerrungskorrigierenden Rechengrundlage (4.10) stellen die Autoren fest, dass diese, vergleicht man sie mit derjenigen, die die effektive Stichprobengröße verwendet, eine höhere Variabilität der Schätzwerte aufweist (vgl. XIA [2010], S. 3914ff.).

Da beide Auswahlrahmen unabhängig von einander sind, berechnet sich die Varianz von $\hat{Y}(\lambda)$ nach

$$V(\hat{Y}(\lambda)) = V(\hat{Y}_{m_1} + \lambda * \hat{Y}_{m_2 \cup m_3}) + V((1 - \lambda) * \hat{Y}_{m_3 \cup m_4} + \hat{Y}_{m_5}).$$

Hartley (1962, 1974) führt diesbezüglich an, λ so zu wählen, dass die Varianz von $\hat{Y}(\lambda)$ minimiert wird. Zur Berechnung dieses *optimalen Schätzers* wird λ durch

$$\lambda_{opt} = \frac{V(\hat{Y}_{m_3 \cup m_4}) + Cov(\hat{Y}_{m_5}, \hat{Y}_{m_3 \cup m_4}) - Cov(\hat{Y}_{m_1}, \hat{Y}_{m_2 \cup m_3})}{V(\hat{Y}_{m_2 \cup m_3}) + V(\hat{Y}_{m_3 \cup m_4})} \quad (4.11)$$

als varianzminimierender Wert verwendet (vgl. LOHR [2007], S. 3259). Da Varianzen und Kovarianzen normalerweise unbekannt sind, müssen diese anhand der verfügbaren Daten geschätzt werden, wodurch sich ein $\hat{\lambda}_{opt}$ von

$$\hat{\lambda}_{opt} = \frac{\hat{V}(\hat{Y}_{m_3 \cup m_4}) + \widehat{Cov}(\hat{Y}_{m_5}, \hat{Y}_{m_3 \cup m_4}) - \widehat{Cov}(\hat{Y}_{m_1}, \hat{Y}_{m_2 \cup m_3})}{\hat{V}(\hat{Y}_{m_2 \cup m_3}) + \hat{V}(\hat{Y}_{m_3 \cup m_4})}$$

ergibt. Die nach diesem Vorgehen modifizierten Gewichte $\tilde{w}_{i,opt}^F$ und $\tilde{w}_{i,opt}^C$ sind dabei als Funktionen der Varianzen und Kovarianzen der geschätzten Bereichstotalwerte zu bewerten und müssen somit als Zufallsvariablen betrachtet werden, weswegen deren Variabilität in der Berechnung von Standardfehlern der Schätzwerte Rechnung getragen werden muss. Darüber hinaus unterscheiden sich aufgrund dessen die optimalen Gewichtsmodifikationen für unterschiedliche Variablen, was demnach zu interner Inkonsistenz führt, da sich die optimalen Gewichte für verschiedene Variablen unterscheiden (vgl. LOHR [2011], S. 200).

Eine weitere, nicht unbedeutende, Überlegung ist die Aufteilung und Gewichtung unter finanziellen Re-

striktionen. Callegaro et al. (2011) untersuchen dies am Beispiel der CELLA1-Studie. Die Kostenrestriktion wurde dabei durch $n^L e^L + n^C e^C \leq g$ angegeben. Hierbei werden die durchschnittlichen Befragungskosten pro Erhebungseinheit der Festnetzstichprobe durch e^L , die der Mobilfunkstichprobe durch e^C beschrieben. g legt hier das gesamt, verfügbare Budget fest. Die Autoren merken diesbezüglich an, dass sich die Bestimmung der Durchschnittskosten in der Praxis schwerer gestaltet, da diese von der Response-Rate abhängig sind. So verursachen bspw. Non-Contacts weniger Kosten als tatsächlich durchgeführte Interviews. Um unter Kostenrestriktion dennoch einen varianzminimierenden Schätzer für den Totalwert \hat{Y} unter Budgetrestriktion zu erhalten, kommen die Autoren zu folgendem Verhältnis zur Bestimmung der Stichprobenzusammensetzung:

$$\frac{n^L}{n^C} = \sqrt{\gamma \frac{e^C}{e^L}}$$

mit

$$\gamma = \frac{V_{\psi^L}^L}{V_{\psi^C}^C} = \frac{\sum_{i \in U^L} \psi_i^L \left(\frac{\theta_i^L}{\psi_i^L} - t_{\theta}^L \right)^2}{\sum_{i \in U^L} \psi_i^C \left(\frac{\theta_i^C}{\psi_i^C} - t_{\theta}^C \right)^2},$$

$$\theta_i^L = c_i^L * y_i^L; \theta_i^C = c_i^C * y_i^C$$

und

$$\psi_i^L = \frac{\pi_i^L}{n^L} \text{ für } i \in U^L; \psi_i^C = \frac{\pi_i^C}{n^C} \text{ für } i \in U^C.$$

U^L und U^C beschreiben dabei den tatsächlich vorhandenen Umfang an Haushalten mit Anschlüssen aus den jeweiligen Telekommunikationsbereichen. Die Autoren kommen bei Ihrer Untersuchung der CELLA1-Studie zum Ergebnis, dass die Zusammensetzung dieser Stichprobe (die aus ca. 50% Festnetz- und 50% Mobilfunkstichprobe bestand) unter der Hinzunahme von Budgetrestriktionen evtl. anders aufzuteilen wäre (vgl. CALLEGARO [2011], S. 18ff.).

Fuller und Burmeister (1972) schlagen eine Modifikation des von Hartley angebrachten Schätzers durch die Inklusion der Stichprobengröße $\hat{N}_{m_2 \cup m_3}$ und $\hat{N}_{m_3 \cup m_4}$ als zusätzliche Informationen vor. Eine Schätzung des Totalwertes ergäbe sich dann nach

$$\hat{Y}_{FB}(\lambda) = \hat{Y}_{m_1} + \lambda_1 * \hat{Y}_{m_2 \cup m_3} + (1 - \lambda_1) * \hat{Y}_{m_3 \cup m_4} + \hat{Y}_{m_5} + \lambda_2 (\hat{N}_{m_2 \cup m_3} - \hat{N}_{m_3 \cup m_4}). \quad (4.12)$$

Die Varianz von λ wird dabei durch

$$V_{FB}(\lambda) = V(\hat{Y}_{m_1} + \lambda_1(\hat{Y}_{m_2 \cup m_3}) + \lambda_2 \hat{N}_{m_2 \cup m_3}) + V((1 - \lambda_1)\hat{Y}_{m_3 \cup m_4} + \hat{Y}_{m_5} - \lambda_2 \hat{N}_{m_3 \cup m_4})$$

berechnet.

Für den Fall, dass Simple Random Sampling in beiden Auswahlrahmen angewendet wird, zeigen Rao (1983) und Skinner (1991), dass \hat{Y}_{FB} nach dem Maximum-Likelihood Prinzip abgeleitet werden kann. Hierbei werden wie bei Hartley λ_1 und λ_2 varianzminimierend gewählt, so dass

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = - \begin{pmatrix} V(\hat{Y}_{m_2 \cup m_3} - \hat{Y}_{m_3 \cup m_4}) & Cov(\hat{Y}_{m_2 \cup m_3} - \hat{Y}_{m_3 \cup m_4}, \hat{N}_{m_2 \cup m_3} - \hat{N}_{m_3 \cup m_4}) \\ Cov(\hat{Y}_{m_2 \cup m_3} - \hat{Y}_{m_3 \cup m_4}, \hat{N}_{m_2 \cup m_3} - \hat{N}_{m_3 \cup m_4}) & V(\hat{N}_{m_2 \cup m_3} - \hat{N}_{m_3 \cup m_4}) \end{pmatrix}^{-1} \begin{pmatrix} Cov(\hat{Y}_{m_1} + \hat{Y}_{m_5} + \hat{Y}_{m_3 \cup m_4}, \hat{Y}_{m_2 \cup m_3} - \hat{Y}_{m_3 \cup m_4}) \\ Cov(\hat{Y}_{m_1} + \hat{Y}_{m_5} + \hat{Y}_{m_3 \cup m_4}, \hat{N}_{m_2 \cup m_3} - \hat{N}_{m_3 \cup m_4}) \end{pmatrix}$$

gilt, wobei auch hier die gleichen Probleme wie bei der Schätzung nach Hartley herrschen und die optimalen Werte der Parameter aus den vorhandenen Daten geschätzt werden müssen (vgl. LOHR [2007], S. 3259f).

4.3 Pseudo-Maximum Likelihood Schätzer

Skinner und Rao (1996) entwickelten ausgehend vom Ansatz von Fuller und Burmeister einen alternativen Schätzer, der eine einheitliche Menge an Gewichten für alle Variablen y verwendet. Daher besteht nicht mehr das vorherige Problem interner Inkonsistenz. Ein weiterer Vorzug im Gegensatz zu den vorherigen optimalen Verfahren ist, dass kein Simple Random Sampling vorausgesetzt wird, sondern dieser Schätzer auch auf komplexe Ziehungsmodelle angewendet werden kann. Der Schätzer für den Totalwert nach einer solchen Pseudo-Maximum Likelihood Methode wird durch

$$\hat{Y}_{PML} = (N^F - \hat{N}_{m_2 \cup m_3 \cup m_4, PML}) \hat{\mu}_{m_1} + (N^C - \hat{N}_{m_2 \cup m_3 \cup m_4, PML}) \hat{\mu}_{m_5} + \hat{N}_{m_2 \cup m_3 \cup m_4, PML} \hat{\mu}_{m_2 \cup m_3 \cup m_4} \quad (4.13)$$

mit

$$\hat{\mu}_{m_2 \cup m_3 \cup m_4} = \frac{\frac{n^F}{N^F} \hat{N}_{m_2 \cup m_3} \hat{\mu}_{m_2 \cup m_3} + \frac{n^C}{N^C} \hat{N}_{m_3 \cup m_4} \hat{\mu}_{m_3 \cup m_4}}{\frac{n^F}{N^F} \hat{N}_{m_2 \cup m_3} + \frac{n^C}{N^C} \hat{N}_{m_3 \cup m_4}}$$

bestimmt. Ferner wird zur Bestimmung von $\hat{N}_{m_2 \cup m_3 \cup m_4, PML}$ die kleinste Wurzel der Gleichung

$$px^2 - qx + r = 0 \quad (4.14)$$

mit

$$\begin{aligned} p &= n^F + n^C \\ q &= n^F N^F + n^C N^C + n^F \hat{N}_{m_2 \cup m_3} + n^C \hat{N}_{m_3 \cup m_4} \\ r &= n^F \hat{N}_{m_2 \cup m_3} N^C + n^C \hat{N}_{m_3 \cup m_4} N^F \end{aligned}$$

herangezogen. Die Autoren merken diesbezüglich an, dass die Wurzeln aus 4.14 immer im Bereich der realen Zahlen liegen, da $\hat{N}_{m_2 \cup m_3} < N^F$ und $\hat{N}_{m_3 \cup m_4} < N^C$ und somit $\hat{N}_{m_2 \cup m_3 \cup m_4, PML}$ wohldefiniert ist (vgl. SKINNER und RAO [1996], S. 350 ff.).

Lohr und Rao (2000) und Lohr (2011) entwickelten diesen Schätzer weiter, so dass

$$\hat{Y}_{PML}(\lambda) = \frac{N^F - \hat{N}_{m_2 \cup m_3 \cup m_4, PML}(\theta)}{\hat{N}_{m_1}} \hat{Y}_{m_1} + \frac{N^C - \hat{N}_{m_2 \cup m_3 \cup m_4, PML}(\theta)}{\hat{N}_{m_5}} \hat{Y}_{m_5} + \frac{\hat{N}_{m_2 \cup m_3 \cup m_4, PML}(\theta)}{\hat{N}_{m_2 \cup m_3 \cup m_4}(\theta)} \hat{Y}_{m_2 \cup m_3 \cup m_4}(\theta). \quad (4.16)$$

gilt. Die Umfänge der Grundgesamtheit sowie die folgenden Varianzen müssen dabei nach Annahmen der Autoren anhand der Daten geschätzt werden, da sie in der Regel unbekannt sind. So dient

$$\hat{\lambda}_{PML} = \frac{\hat{N}_{m_1} \hat{N}^C \hat{V}(\hat{N}_{m_3 \cup m_4})}{\hat{N}_{m_1} \hat{N}^C \hat{V}(\hat{N}_{m_3 \cup m_4}) + \hat{N}_{m_5} \hat{N}^F \hat{V}(\hat{N}_{m_2 \cup m_3})} \quad (4.17)$$

als Schätzer für λ_{PML} . Anhand dieses Schätzers ergibt sich nun eine Anpassung der Gewichte, nach Vorgabe des in Kapitel 4.2 zu Beginn erläuterten Verfahrens, von:

$$c_{i, PML}^F = \begin{cases} \frac{N^F - \hat{N}_{m_2 \cup m_3 \cup m_4, PML}(\hat{\theta}_{PML})}{\hat{N}_{m_1}} & \text{wenn } i \in m_1 \\ \frac{\hat{N}_{m_2 \cup m_3 \cup m_4, PML}(\hat{\theta}_{PML})}{\hat{\theta}_{PML} \hat{N}_{m_2 \cup m_3} + (1 - \hat{\theta}_{PML}) \hat{N}_{m_3 \cup m_4}} \hat{\theta}_{PML} & \text{wenn } i \in m_2 \cup m_3 \end{cases} \quad (4.18)$$

$$c_{i, PML}^C = \begin{cases} \frac{N^C - \hat{N}_{m_2 \cup m_3 \cup m_4, PML}(\hat{\theta}_{PML})}{\hat{N}_{m_5}} & \text{wenn } i \in m_5 \\ \frac{\hat{N}_{m_2 \cup m_3 \cup m_4, PML}(\hat{\theta}_{PML})}{\hat{\theta}_{PML} \hat{N}_{m_2 \cup m_3} + (1 - \hat{\theta}_{PML}) \hat{N}_{m_3 \cup m_4}} (1 - \hat{\theta}_{PML}) & \text{wenn } i \in m_3 \cup m_4 \end{cases}$$

Durch Linearisierung des Ergebnisses aus 4.17 erhält man für $\hat{N}_{m_2 \cup m_3 \cup m_4}(\theta_{PML})$
 $\approx \hat{N}_{m_2 \cup m_3 \cup m_4, PML}(\phi)$, mit

$$\phi = \frac{V(\hat{N}_{m_3 \cup m_4})}{V(\hat{N}_{m_2 \cup m_3}) + V(\hat{N}_{m_3 \cup m_4})}.$$

(Vgl. LOHR und RAO [2000]S. 272; LOHR [2011], S. 201)

4.18 zufolge bestimmt ein PML-Schätzer demnach zwei unterschiedliche Gewichte für den Überlappungsbereich in Abhängigkeit davon, aus welchem Rahmen eine Erhebungseinheit gezogen wurde (vgl. WINGLEE [2007], S. 3196).

Nach Einschätzung von Rao und Skinner (1996) besteht der Hauptvorteil des Pseudo-Maximum Likelihood Schätzers in seiner typischerweise einfachen Form und der Designkonsistenz, was gerade für Erhebungen mit vielen Variablen an Relevanz gewinnt. Als Nachteil wird vermutet, dass er asymptotisch nicht effizient sein könnte, wobei die Autoren annehmen, dass der Effizienzverlust in der Praxis vergleichsweise gering ausfällt (vgl. SKINNER und RAO [1996], S. 350; S. 355).

4.4 Multiplicity Schätzer

Im Gegensatz zu Optimum-, PML- und SF- Schätzer, führt Mecatti (2007) mit seinem *Single Frame Multiplicity Schätzer* ein Verfahren ein, das mit wenig zusätzlichen Informationen und ohne eindeutige Zuordnung der Stichprobenzugehörigkeit auskommt. Dieser macht sich die Eigenschaft, dass die einzelnen Bereiche gegenseitig ausschließend sind, zunutze und berücksichtigt lediglich die Anzahl der Auswahlrahmen, in denen sich eine Erhebungseinheit befindet. Ein sich daraus ergebender Vorteil besteht in der Möglichkeit, diesen Ansatz in der praktischen Anwendung vergleichsweise einfach zu implementieren, da lediglich die Anzahl an Auswahlrahmen, in denen sich eine Erhebungseinheit befindet, erfragt werden muss.³²

Betrachtet man nun die $q = 1, \dots, Q$ an einer Erhebung beteiligten Auswahlrahmen M^q , aus denen jeweils eine unabhängige Stichprobe S^q gezogen wird und beschreibt h_i ferner die Multiplizität der Stichprobeneinheit i über diese Auswahlrahmen, so folgt, da

$$\sum_q \sum_{i \in M^q} y_i = \sum_{i \in \bigcup_q M^q} h_i * y_i,$$

dass sich der Totalwert nach

$$Y = \sum_{q=1}^Q \sum_{i \in M^q} y_i * h_i^{-1}$$

bestimmen lässt. Die innerhalb dieser Erhebung gewonnenen Daten lassen sich hierbei in $2^Q - 1$ nicht-überlappende Bereiche einteilen. Da diese aufgezeigten Bereiche nur eine faktische, aber evtl. unbekannte Aufteilung der Grundgesamtheit darstellen, die eigentlichen Stichproben allerdings aus den jeweiligen Auswahlrahmen gezogen werden, ergibt sich ein Multiplicity Schätzer nach

$$\hat{Y}_M = \sum_{q=1}^Q \sum_{i \in S(q)} w_i^q * y_i * h_i^{-1}. \quad (4.19)$$

(Vgl. MECATTI [2007], S. 152f.)

Die grundsätzliche Idee dieses Ansatzes besteht demnach darin, alle Q Auswahlrahmen in einem zusammenzufassen. Dabei werden alle doppelten Einheiten beibehalten, wobei y_i durch $y_i * h_i^{-1}$ ersetzt

³² Im Bereich der Telefonstichproben kann es aufgrund von virtuellen Festnetznummern, etc. zu Schwierigkeiten kommen.

wird. Im Falle eines Dual-Frame-Ansatzes, wie es bei einer Telefonstichprobe der Fall sein kann, gilt $h_i = 1$, für $i \in m_1 \cup m_5$ und $h_i = 2$ für $i \in m_2 \cup m_3 \cup m_4$ (vgl. RAO und WU [2010], S. 1497f.). Dieses Vorgehen ist mit dem von BRICK et al. (2006), dass λ gleich 0,5 setzt, vergleichbar.

Bei Simple Random Sampling in den einzelnen Auswahlrahmen zeigt sich, dass sich die gezogenen Werte nach einem spezifischen Rahmenkoeffizienten $(w_i^q h_i)^{-1}$ gewichtet lassen. Als Konsequenz kann dieser Schätzer, so Mecatti, als genauer erachtet werden als der einfache SF-Schätzer. Aufgrund der Horvitz-Thompson-Struktur kann die Varianz hier in geschlossener Form abgeleitet werden. Im Falle einer uneingeschränkten Zufallsstichprobe lässt sich diese durch

$$V(\hat{Y}_M) = \sum_{q=1}^Q \frac{N_q - n_q}{n_q(N_q - 1)} (N_q \sum_{i \in M^q} y_i^2 h_i^{-2} - (\sum_{i \in M^q} y_i h_i^{-1})^2)$$

bestimmen. Ein unverzerrter Schätzer der Varianz für alle Auswahlrahmen ist dann

$$\hat{V}(\hat{Y}_M) = \sum_{q=1}^Q \frac{N_q - n_q}{n_q(N_q - 1)} (N_q \sum_{i \in S(q)} y_i^2 h_i^{-2} - f_q^{-1} (\sum_{i \in S(q)} y_i h_i^{-1})^2). \quad (4.20)$$

(Vgl. MECATTI [2007], S. 152f.)

Nach Rao und Wu (2010) lässt sich anhand dieses Ansatzes ein unverzerrter Schätzer für den Umfang der Grundgesamtheit nach

$$\hat{N}_M = \sum_{q=1}^Q \sum_{i \in S(q)} \frac{w_i^q}{h_i^q}$$

bestimmen. Darüber hinaus wird angeführt, dass dieser Schätzer durch Umstellen auch unter einem anderen Blickwinkel betrachtet werden kann. So ergibt sich unter

$$\hat{Y}_M = \sum_{q=1}^Q \sum_{i \in S(q)} \frac{w_i^q}{h_i^q} y_i$$

ein *Weight Sharing - Schätzer*³³, bei dem das „Basisgewicht“ w_i^q der Einheit i im Rahmen M^q bei der gleichen Einheit für alle h_i^q Auswahlrahmen, in denen sie vorkommt, geteilt wird.

In ihrer Untersuchung von Erhebungen mit multiplen Auswahlrahmen zeigen Rao und Wu (2010), dass unter der Verwendung eines Dual-Frame Ansatzes der Hájek-Schätzer des Mittelwertes der Grundgesamtheit, der durch $\hat{\bar{Y}} = \hat{Y}_M / \hat{N}_M$ gegeben ist, asymptotisch dem poststratifizierten *Pseudo Empirical Likelihood Schätzer*, den die Autoren zur Schätzung für multiple Auswahlrahmen anführen, entspricht, wenn N^F , N^C und N^{FC} unbekannt sind und die beiden, für diesen PEL-Schätzer erforderlichen Konstanten $\eta = \nu = 0,5$ gesetzt werden. Daher gehen die beiden Autoren davon aus, dass der Multiplicity-Schätzer nicht zwangsweise optimal sein muss. Der Vorzug aufgrund der Einfachheit in der Anwendung für mehrere Auswahlrahmen wird aber nicht abgestritten (vgl. RAO und WU [2010], S. 1498).

4.5 Beurteilung der Ansätze

Im Bereich der SF-Schätzer stellen Winglee et al. (2007) bei ihrer Untersuchung der Erhebungsdaten der National Incidence Study von 1997 fest, dass der modifizierte SF-Schätzer nach 4.4 am besten abschneidet. Da allerdings von den Autoren angemerkt wird, dass bei dieser Untersuchung keine Bemühungen zur Bestimmung des Effektes von Fehlklassifikationen der Bereichszugehörigkeit, einer Veränderung der

33 Anstelle des vorherigen *Variable Sharing - Schätzers*.

Größe des Überlappungsbereiches sowie des Einflusses des Designeffektes auf die Schätzergebnisse vorgenommen wurde, sollte im Hinblick auf die Menge der Einflussfaktoren, die im Bereich einer Dual-Frame Stichprobe bei Telefonstichproben zu Fehlklassifikationen führen können³⁴, ein Verfahren, wie es in 4.1, bzw. 4.2 dargestellt wird³⁵, nicht ausgeschlossen werden. Als weiteres Ergebnis dieser Untersuchung zeigen die Autoren, dass der klassische SF-Schätzer in ihrer erstmaligen Simulation dem PML-Schätzer vorzuziehen ist (vgl. WINGLEE [2007], S. 3201f.).³⁶ Clark et al. (2007) zeigen in diesem Kontext bei der Untersuchung derselben Studie, dass der relative Bias unter Verwendung der PML-Methode geringer als bei der SF-Methode ausfällt. Als mögliche Begründung hierfür wird eingeräumt, dass innerhalb dieser Schätzung der Auswahlrahmen mit der höheren Varianz einen geringeren Anteil an der Gesamtschätzung hat. Als Ergebnis halten die Autoren dahingehend fest, dass unter Berücksichtigung des relativen Bias und des EMSEs die PML-Methode am besten abschneidet. So weisen die Schätzungen nach 4.3 und 4.4 zwar eine geringere Varianz, aber eine höhere Verzerrung auf (vgl. CLARK [2007], S. 3237f.).

Skinner (1991) zeigt Überdies, dass innerhalb der Verwendung eines SF-Schätzers ein *Raking Ratio Adjustment* bezüglich N^q in der Regel eine höhere Effizienz der Schätzergebnisse liefert, als die einfache Anwendung eines SF-Schätzers. Allerdings wird zu bedenken gegeben, dass die Ergebnisse der Varianzschätzungen für die Überlappungsbereiche sehr nahe beieinander liegen können (vgl. SKINNER [1991], S. 782). Daher sollte berücksichtigt werden, dass im Bereich der Telefonstichproben die tatsächliche Größe des Überlappungsbereiches in der Grundgesamtheit lediglich geschätzt werden kann. Darüber hinaus diskutieren Lohr und Rao (2000), dass Raking nicht immer zu einer Verringerung der Varianz beiträgt (vgl. LOHR und RAO [2000], S. 274).

Innerhalb der Klasse der „einfachen“ Composite-Gewichte stellen Xia et al. (2010) fest, dass, wie bereits in Kapitel 4.2 erwähnt, die Schätzung anhand der effektiven Stichprobengröße die besten Ergebnisse liefert (vgl. XIA [2010], S. 3914 ff.). Bezüglich der optimalen Gewichte von Hartley bzw. Fuller und Burmeister zeigen Lohr und Rao (2000), dass der Schätzer nach Fuller und Burmeister unter den untersuchten Verfahren derjenige ist, der die geringste asymptotische Varianz aufweist und somit, zumindest unter diesen beiden, den effizienten Schätzer darstellt (vgl. LOHR und RAO [2000], S. 274). Zu bedenken ist dabei, dass dieser die gleichen Probleme wie Hartleys Schätzer aufweist, weswegen die modifizierten Gewichte demnach als Zufallsvariablen zu betrachten sind. Dies führt dazu, dass einerseits die Anwendung gerade bei Erhebungen mit mehreren Variablen aufgrund der Rechenintensität erschwert wird und dieser andererseits bei komplexen Ziehungsverfahren zu interner Inkonsistenz führt (vgl. LOHR [2011], S. 201).

Darüber hinaus stellt Skinner (1991) in seiner Untersuchung des Einflusses von Raking fest, dass dieses Verfahren sich weniger einfach auf einen FB-Schätzer anwenden lässt, da bei steigender Anzahl von Schichten und Rahmen die Stichprobengrößen in einigen Überlappungsbereichen klein oder Null werden können (vgl. SKINNER [1991], S. 782).

Bezüglich der Höhe der Varianz der Schätzergebnisse einzelner Schätzverfahren geben Lohr und Rao (2000) zumindest für einige der Verfahren eine eindeutige Anordnung an. Demnach besteht folgendes Verhältnis:

- $V(\hat{Y}_{FB}) < V(\hat{Y}_H) < V(\hat{Y}_{SF})$
- $V(\hat{Y}_{FB}) < V(\hat{Y}_{SFrake})$ ³⁷
- $V(\hat{Y}_{FB}) < V(\hat{Y}_{PML})$

34 So stellt bspw. die Existenz einer virtuellen Festnetznummer eine potentielle Gefahrenquelle zur Fehlklassifikation im Bereich der Mobilfunkstichprobe dar.

35 In Abhängigkeit der Größe von $\pi^F * \pi^C$.

36 Diesbezüglich sei anzumerken, dass innerhalb dieser Untersuchung von einer unbekanntem Größe des zweiten Auswahlrahmens ausgegangen wurde und die effektive Stichprobengröße deshalb anhand eines geschätzten Designeffektes berechnet wurde. Daher können sich Ergebnisse anderer Studien wohl durchaus von Ergebnissen dieser Studie unterscheiden.

37 $V(\hat{Y}_{SFrake})$ steht dabei für die Varianz eines durch Raking geschätzten SF-Schätzers.

Für andere Verhältnisse kann den Autoren zufolge keine eindeutige Anordnung getroffen werden (vgl. LOHR und RAO [2000], S. 274).

In einer weiteren Simulationsstudie, innerhalb derer diese unterschiedlichen Verfahren nochmals untersucht werden, stellen die beiden Autoren 2006 fest, dass ein Hartley-Schätzer bei geringen Stichprobengrößen und lediglich zwei Auswahlrahmen vergleichsweise gut abschneidet, allerdings in Szenarien mit mehr als zwei Auswahlrahmen nicht nur vergleichsweise schlechtere Ergebnisse liefert, sondern auch zur Instabilität neigt, weswegen dieser erst bei höheren Stichprobengrößen wieder stabiler wird und erst dann einen geringeren empirischen Mean-Squared Error (EMSE) aufweist. Die Autoren merken dazu an, dass innerhalb ihrer Studie der Hartley-Schätzer immer einen höheren EMSE aufweist als der PML-Schätzer.

Der SF-Schätzer ohne Raking lieferte innerhalb der meisten Simulationen dieser Untersuchung dürrtige und teilweise widersprüchliche Ergebnisse, da sich bspw. bei einigen dieser Simulationen bei steigender Stichprobengröße auch der EMSE erhöht. Unter der Verwendung eines Raking-SF-Schätzers sind die Ergebnisse zwar deutlich besser, schneiden aber ebenfalls schlechter als PML-Schätzer ab (vgl. LOHR und RAO [2006], S. 1025 ff.).

Lohr (2011) beschreibt, dass die Verwendung des Multiplicity-Schätzers bei Fehlklassifikation einer Erhebungseinheit in einem der Überlappungsbereiche keinen Einfluss auf die Güte der Schätzung hat, solange diese die gleiche Anzahl an beteiligten Rahmen aufweist. Befindet sich die Fehlklassifikation allerdings in einem Bereich, an dem weniger oder mehr Auswahlrahmen als die der tatsächlichen Zugehörigkeit beteiligt sind, so kann dies zu einer Verzerrung der Schätzergebnisse beitragen (vgl. LOHR [2011], S. 209).³⁸

Mecatti (2007) stellt bei der Untersuchung von Fehlklassifikation, bei der ein bestimmter Anteil derjenigen Stichprobe, die den größten Anteil aufweist, ausgewählt und zufällig auf die anderen Bereiche aufgeteilt wurde, fest, dass der Raking-SF-Schätzer ähnlich effizient wie der PML-Schätzer ist. Diese beiden Methoden liefern dabei bei Abwesenheit von Fehlklassifikationen genauere Ergebnisse als der Multiplicity Schätzer. Bei geringer Fehlklassifikation im Bereich von 1% zeigen sich allerdings, den Resultaten dieser Untersuchung zufolge, verzerrte Ergebnisse der beiden anderen Methoden, wodurch deren Effizienz im Vergleich zum Multiplicity Schätzer abnimmt. Bei einem steigendem Anteil an Fehlklassifikation, dessen Effekt hier anhand des relativen Bias sowie des Effizienzverhältnisses der einzelnen Schätzer zum Multiplicity Schätzer untersucht wird, lässt sich feststellen, dass die Auswirkungen bei beiden Schätzern recht stark sind, der PML-Schätzer jedoch noch am besten abschneidet. Darüber hinaus zeigt sich, dass die Verwendung des PML-Schätzers bei einem höheren Anteil von Fehlklassifikation besser abschneidet als bei einem niedrigen Anteil (vgl. MECATTI [2007], S. 155f.).

Lohr und Rao (2006) stellen dahingehend bei ihrer Untersuchung fest, dass die meisten Veränderungen des EMSE im Rahmen von Fehlklassifikation bei den Varianzen, nicht bei der Verzerrung entstehen. Die Varianzen unter Verwendung der PML-Methode steigen dabei stärker an als bei den Schätzern, die durch die Methoden von Hartley oder Fuller und Burmeister gewonnen wurden. Darüber hinaus ist der Effekt abhängig von dem verwendeten Ziehungsmodell.

Als Fazit stellen die Autoren dennoch fest, dass lediglich die Raking-SF-Methode und die PML-Methode in den meisten der beobachteten Fällen gute Ergebnisse liefern. Daher empfehlen sie, bei Multiple-Frame-Ansätzen mit mehr als zwei Auswahlrahmen, andere Methoden zu meiden, da Schätzer nach der Methode von Fuller und Burmeister oder Hartley bei mehr als zwei Auswahlrahmen dazu neigen, instabil zu werden und SF-Schätzer ohne Raking in ihrer Untersuchung in den meisten Fällen wesentlich schlechter abschneiden als PML-Schätzer. Darüber hinaus führt bei der Verwendung von Klumpenstichproben in einem oder mehreren Auswahlrahmen eine Erhöhung der Stichprobengröße entgegen der Erwartung zu einer Erhöhung der Varianz bei SF-Schätzern. Dieser Zusammenhang ist nach Auffassung der Autoren auch unter Verwendung von Raking möglich, wird allerdings als selten und weniger stark erachtet (vgl. LOHR und RAO [2006], S. 1027 ff.).

Bezüglich der Verwendung eines Screening-Ansatzes, bei dem zuvor Einheiten aus M^C in den Berei-

38 Solche Fehlklassifikationen sind aufgrund der zuvor beschriebenen Schwierigkeiten allerdings zu erwarten.

chen m_2 bis m_4 ausgeschlossen und über den Auswahlrahmen M^F erhoben werden, stellt Lohr (2011) fest, dass keine Verzerrung aufgrund von Fehlklassifikation entsteht, solange diese nicht im Auswahlrahmen M^C besteht. Trifft dies aber nicht zu, so zeigt sich, dass dadurch die Verzerrung der Schätzwerte erheblich steigt. Die Autorin begründet dies damit, dass unter Verwendung dieses Ansatzes zusätzliche Informationen anhand des zweiten Auswahlrahmens vergeudet werden (vgl. LOHR [2011], S. 208).

Mit Ausnahme des Multiplicity Schätzers gehen alle der hier vorgestellten Methoden davon aus, dass die Zugehörigkeit der Erhebungseinheit zu einem bestimmten Bereich genau determiniert werden kann.³⁹ Aufgrund dessen stellen Lohr und Rao (2006) fest, dass all diese Gewichtungsverfahren anfällig auf Fehlklassifikationen reagieren und Schätzwerte verzerrt sein können, wenn sich die Mittelwerte bei einer solchen Fehlklassifikation innerhalb der Bereiche stark unterscheiden. Daher empfehlen sie, bei der Bestimmung von Schätzwerten mögliche Fehlklassifikation ebenso wie Nonresponse-Fehler und die Varianzen der Schätzer zu berücksichtigen (vgl. LOHR und RAO [2006], S. 1029).

³⁹ Für den Multiplicity Schätzer ist nicht die Bestimmung der Zugehörigkeit einer Erhebungseinheit zu einem bestimmten Bereich relevant, sondern lediglich die Anzahl der Auswahlrahmen, denen eine bestimmte Erhebungseinheit angehört.

5 Ausblick

Unter Berücksichtigung des stetig abnehmenden Anteils an *klassischen* Festnetzanschlüssen innerhalb Deutschlands sowie der immer größer werdenden Anzahl an Haushalten, die nur noch über einen Mobilfunkanschluss⁴⁰ verfügen, zeigt sich, dass Erhebungen, die nur auf Festnetzstichproben basieren, zu Undercoverage führen können. Darüber hinaus ist davon auszugehen, dass Mobile-Only-Haushalte sich in wesentlichen Charakteristika von Dual-Usern unterscheiden. Daher ist der Übergang zu Dual-Frame-Befragungen, die sowohl Mobilfunk-, als auch Festnetzstichproben verwenden, zumindest für bundesweite Erhebungen unumgänglich, um „repräsentative“ Erhebungen auf der Basis von Telefonstichproben durchzuführen.

Problematisch erweisen sich allerdings Erhebungen auf regionaler Ebene, da Mobilfunkvorwahlen keine geografische Einteilung zulassen. Eine Möglichkeit zumindest nach der Ziehung von Mobilfunknummern eine approximative geografische Einordnung vorzunehmen, bietet das Abfragen der Identifikationsnummern des MSCs, das gerade für ein Endgerät zuständig ist. Diese Nummern könnten dabei im Rahmen eines HLR Lookups in Erfahrung gebracht werden. Diese werden bei einem solchen Lookup aber nicht standardmäßig bereitgestellt.

Bezüglich der geografischen Einteilung von Mobilfunknummern wurde in einer Präsentation von Engel (2013) aufgezeigt, dass zumindest für die beiden Netze Telekom D1 und Vodafone D2 im Vorwahlbereich 0171 und 0172 eine grobe Einordnung anhand der vergebenen Rufnummer möglich ist, die sich aber auf die Städte Berlin, Hamburg, Frankfurt, Stuttgart und München beschränkt. Hierbei soll eine Rufnummer bei der Vergabe lediglich über die nächstgrößere Stadt, in deren Nähe die Rufnummer zugeteilt wurde, lokalisierbar sein. Es kann demnach nicht mit absoluter Sicherheit davon ausgegangen werden, dass sich der betreffende Haushalt noch in dieser Region befindet, da ein Wegzug aus dieser Region zu einem späteren Zeitpunkt nicht ausgeschlossen werden kann. Darüber hinaus lassen sich keine genauen Aussagen über die Vergabeschemata der beiden E-Netze von Telefonica O2 und E-plus treffen (vgl. ENGEL [2013], S. 13ff.). Bei einer Untersuchung verschiedener Rufnummern aus den beiden Vorwahlbereichen 0171 und 0172 in Kombination des originären Ortes des Vertragsabschlusses fällt überdies auf, dass die Erkenntnisse von Engel oftmals nicht zutreffen. Allerdings zeigt sich in vielen Fällen, dass im Vorwahlbereich 0172 noch das „alte“ Vergabeschema von Mannesmann (D2) wiederfinden lässt, bei dem die ersten beiden Ziffern der Mobilfunkrufnummer denen der Ortsnetzbezeichnungen der Gemeinde entsprechen, in der die Rufnummer zugeteilt wurde.

Die unzulängliche Möglichkeit zur geografischen Einordnung von Mobilfunknummern stellt daher im Gegenzug zu anderen Ländern wie bspw. den USA, deren Mobilfunknummern in das dort bestehende Vorwahlsystem mit eingegliedert sind, eine Herausforderung für die Umfrageforschung innerhalb Deutschlands dar.

Forschungsbedarf innerhalb von Dual-Frame Stichproben besteht darüber hinaus in der Vorgehensweise der Stichprobenziehung und Gewichtung. Bei der Verwendung von Screenings zur Identifikation der entsprechenden Only-Haushalte liegt bspw. ein Trade-Off zwischen den Screeningkosten und der Kostenersparnis, die sich durch den vermehrten Einsatz des vermeintlich günstigeren Erhebungsverfahrens ergibt, vor. Darüber hinaus kann das Wegfallen des Überlappungsbereiches durch die Beschränkung Dual-User lediglich über ein Medium zu kontaktieren, möglicherweise die Response-Rate verringern, da die Kontaktaufnahme derjenigen Erhebungseinheiten, die zwar generell über beide Kommunikationsmedien verfügen, jedoch nur eines dominant nutzen, bei einer Stichprobe über das entsprechend andere Medium, erschwert wird.

Bei einem Nonscreening-Ansatz besteht im Gegenzug die Problematik, wie mit dem Überlappungsbereich verfahren wird, der sich aufgrund der Möglichkeit der Inklusion einer Erhebungseinheit durch beide

40 Mit oder ohne virtuelle Festnetznummer.

Auswahlrahmen ergibt. Dabei wird die Wahrscheinlichkeit, dass der gleiche Haushalt über beide Stichproben in die Erhebung gelangt, als vernachlässigbar erachtet. Der Überlappungsbereich kann aber im Zuge der Gewichtung ein Problem darstellen, da unter anderem die unterschiedlichen Erhebungsverfahren Einfluss auf die Inklusionswahrscheinlichkeit haben. Die in diesem Kontext dargestellte Neueinteilung der Erhebung in „wahre“ Dual-User, bei denen kein dominantes Kommunikationsmedium zu beobachten ist, und Mainlys, wobei letztere Onlys als extreme Ausprägung dieser beinhalten, kann in diesem Kontext dazu dienen, den Überlappungsbereich zu verkleinern, wodurch sich ein einfacher Umgang mit den gewonnenen Daten vermuten lässt. Auch hier muss aber berücksichtigt werden, dass diese Einteilung nach der tatsächlichen Erhebung erfolgt, da sich erst im Zuge der Befragung ein dominantes Telekommunikationsmedium ausmachen lässt. Somit bleibt die Gefahr, dass eine Erhebungseinheit über das von ihr weniger genutzte Telekommunikationsmedium in die Stichprobe gelangt, bestehen. Darüber hinaus kann man die eindeutige Zuordnung in Mainlys und Onlys auf der Basis der Angabe von Erhebungseinheiten als problematisch erachten, da dafür ein zuverlässiges Instrumentarium gewählt werden muss, die diese Zuordnung erlaubt.

Die im vorherigen Kapitel ausführlich erläuterten Verfahren zur Gewichtung einer solchen Dual-Frame-Erhebung lassen sich durch Kombinationen der Multiplicity-Gewichtung mit anderen Gewichtungsverfahren ergänzen. Singh und Mecatti (2011) untersuchen dazu die Kombination aus der Anpassung aufgrund der Multiplizität mit den Designgewichten, die sich durch proportionale Zusammensetzung, Konvexkombination oder „einfacher“ Multiplizität ergeben, anhand der verfügbaren Informationen über die Bereichszugehörigkeit und der Inklusionswahrscheinlichkeit. Der Informationsgehalt, der zur Beurteilung einer Multiple-Frame-Erhebung zur Verfügung steht, wird dazu in folgende Klassen aufgeteilt:

- *Basic Frame Level Information*, bei der die Multiplizität der Erhebungseinheiten sowie deren Inklusionswahrscheinlichkeiten der Bereiche, aus denen diese erhoben wurden, bekannt sind. Diese Informationen sind allerdings nicht für diejenigen Bereiche verfügbar, durch die die Erhebungseinheiten nicht in die Stichprobe gelangt sind.
- *Partial Frame Level Information* beinhaltet die vorherigen Informationen sowie Angaben über die Zugehörigkeit von Erhebungseinheiten aller Auswahlrahmen, allerdings ohne Auskünfte über die Inklusionswahrscheinlichkeit von Einheiten aus Auswahlrahmen, durch die diese nicht in die Erhebung gelangt sind.
- *Full Frame Level Information* setzt sich aus den vorherigen beiden Informationsgrundlagen und den Angaben über die Inklusionswahrscheinlichkeiten, durch die Erhebungseinheiten in die Erhebung gelangen könnten, zusammen.

Unter Berücksichtigung der unterschiedlichen Informationsgehalte, werden von den Autoren noch zwei Modelle, die auf der hybriden Zusammensetzung des Informationsgehaltes basieren, untersucht. Bei ihrer Untersuchung der unterschiedlichen Gewichtungsverfahren, die sich aus den fünf soeben beschriebenen Kombinationen sowie einem Single-Frame-Schätzer, bei dem die Multiplizität von Einheiten in der kombinierten Stichprobe nicht zugelassen wird, kommen sie daher auf sechs unterschiedliche Schätzer. Die Simulationsstudie untersucht dabei einerseits den Fall von Simple Random Sampling, andererseits das Vorliegen von Unequal Probability Samples.

Als Ergebnis der Untersuchung der Varianzen dieser Schätzer stellen die Autoren fest, dass geringe Inklusionswahrscheinlichkeiten innerhalb einer Stichprobe den relativen Bias sowie den relativen Root Mean Squared Errors (RRMSE) des einfachen Multiplizitätsschätzers nach Mecatti am stärksten erhöhen. Wohl auch aufgrund dessen empfehlen die Autoren beim Vorliegen von unterschiedlichen Informationsgehalten innerhalb der Auswahlrahmen die Verwendung der hybriden Schätzer, da diese nicht so sehr zur Instabilität neigen. Der von den Autoren vorgestellte Composite-Multiplicity-Schätzer (CM) schließt innerhalb ihrer Untersuchung ähnlich gut ab wie der Proportional-Multiplicity-Schätzer. Der CM berechnet

sich dabei nach:

$$\hat{Y}_{CM} = \sum_{q=1}^Q \sum_{i \in M_q} y_i \alpha_{q(i)}^{CM} \frac{\mathbf{1}_{i \in S(q)}}{\pi_{i(q)}} \quad (5.1)$$

mit

$$\alpha_{q(i)}^{CM} = \lambda_i^{CM} \frac{1}{h_i} + (1 - \lambda_i^{CM}) \frac{\pi_{i(q)}}{h_i \bar{\pi}_{i(q)}}$$

und

$$\bar{\pi}_i = m_i^{-1} \sum_{q' \in K(i)} \pi_{i(q')}$$

sowie

$$\lambda_i^{CM} = \frac{\sum_{q \in K(i)} (1 - \pi_{i(q)}^{-1} \bar{\pi}_i) \pi_{i(q)} (1 - \pi_{i(q)})}{\sum_{q \in K(i)} (1 + \pi_{i(q)}^{-2} \bar{\pi}_i^2 - 2\pi_{i(q)}^{-1} \pi_{i(q)} (1 - \pi_{i(q)})}$$

$K(i)$ beschreibt hierbei den Rahmen, in der sich die i -te Einheit befindet (vgl. SINGH und MECATTI [2011], S. 635ff.).⁴¹ Im Falle von Telefonstichproben, bei denen die Wahrscheinlichkeit, dass ein Haushalt über mehrere Stichproben in die Erhebung gelangt, als vernachlässigbar angesehen wird, bleibt die Notwendigkeit eines solchen Ansatzes allerdings fragwürdig.

Basierend auf dem Gedanken, unterschiedliche Ansätze zu kombinieren, ist es aber vorstellbar, im Bereich von Dual-Frame-Telefonstichproben das in den CELLA Studien verwendete Gewichtungsverfahren, das auf einem Single-Frame-Ansatz beruht, dahingehend zu modifizieren, dass der Neueinteilung nach Abbildung 3 Rechnung getragen wird. Eine Möglichkeit dies zu erreichen besteht darin, Dual-Usern, wenn diese Mainlys sind, ein stärkeres Gewicht für ihr dominantes Telekommunikationsmedium zuzuteilen, als für ihr rezessives. Dabei wird die Erhebung in die in diesem Fall fünf disjunkten Bereiche eingeteilt und die Designgewichte entsprechend der „neuen“ Inklusionswahrscheinlichkeit für die jeweiligen Bereiche bestimmt. Ein solches Gewicht könnte sich dabei wie folgt darstellen:

$$w_{i,q} = n_q \frac{\frac{1}{\pi_{i,q}}}{\sum_{j \in S(q)} \frac{1}{\pi_{j,q}}} \quad (5.2)$$

Die bereichsabhängigen Gewichte weisen dabei einen Horvitz-Thompson-Charakter auf und sind demnach über die Inverse der Inklusionswahrscheinlichkeit zu bestimmen. Zu beachten ist dabei, dass aufgrund der Neueinteilung in die, ausgehend von einer Dual-Frame-Telefonstichprobe, fünf Bereiche, n_q ebenfalls als Zufallsvariable zu betrachten ist. Die bereichsabhängige Inklusionswahrscheinlichkeit bestimmt sich dabei durch:

$$\pi_{i,q} = \lambda_q \pi_{i,q}^F + (1 - \lambda_q) \pi_{i,q}^C \quad (5.3)$$

mit:

$$\lambda_q = \begin{cases} 1 & \text{für } q \in m_1 \\ \delta & \text{für } q \in m_2 \\ 0,5 & \text{für } q \in m_3 \\ 1 - \delta & \text{für } q \in m_4 \\ 0 & \text{für } q \in m_5 \end{cases} \quad (5.4)$$

mit $\delta \leq 1$. λ wird hierbei für den Bereich m_3 , der die „wahren“ Dual-User enthält, der Wert 0,5 zugeteilt unter der Annahme, dass aufgrund des Fehlens eines dominanten Telekommunikationsmediums die Erreichbarkeit⁴² eines Haushaltes für beide Auswahlrahmen gleich hoch ist. Für die Bereiche m_1 und m_5 nimmt λ den Wert 1, bzw. 0 an, da Haushalte innerhalb dieses Bereiches lediglich über einen Auswahlrahmen in die Erhebung gelangen können. Der für die Bereiche m_2 und m_4 relevante Wert von δ ist für

41 Eine Darstellung der restlichen Schätzer befindet sich im Anhang.

42 Diese sollte in diesem Kontext nicht mit der Auswahlwahrscheinlichkeit verwechselt werden

diese Art der Gewichtung so zu wählen, dass dieser der vorherrschenden Auffassung der Erreichbarkeit einer Einheit über das dominante Telekommunikationsmedium entspricht. Eine stark vereinfachte Möglichkeit bestünde bspw. darin, δ den Wert 0,75 zuzuteilen. Wie δ in Abhängigkeit der Erreichbarkeit der Mainlys allerdings genau zu bestimmen ist, sollte weiterführend untersucht werden.

Der hier vorgeschlagene Ansatz lässt sich dabei mit der Berechnung der Inklusionswahrscheinlichkeit, wie sie für die CELLA-Studien verwendet wurde, kombinieren und berechnet sich dann nach:

$$\pi_{i,q} = \lambda_q \left(k_i^F \frac{m^F}{M_F} * \frac{1}{z_i} \right) + (1 - \lambda_q) \left(k_i^C \frac{m^C}{M_C} \right) \quad (5.5)$$

(Vergleiche hierzu auch 4.19 und folgende).

Der Schätzer für den Totalwert für eine Dual-Frame Erhebung nach einem solchen Gewichtungsverfahren lässt sich dann nach

$$\hat{Y} = \sum_{q=1}^Q \sum_{i=1}^{n_q} w_{i,q} y_i \quad (5.6)$$

berechnen.

Es lässt sich abschließend feststellen, dass die Beschäftigung mit der Gewichtung und dem Umgang mit Multiple-Frame-Ansätzen in näherer Zukunft aufgrund der Veränderung der Verwendung (neuer) Telekommunikationsmedien unumgänglich sein wird. Hierzu besteht auch Handlungsbedarf in der Bestimmung und Zusammensetzung des Auswahlrahmens. Darüber hinaus sollten auch Problemstellungen im Umgang mit Nonresponse nicht unbeachtet bleiben.

Literaturverzeichnis

- AAPOR CELL PHONE TASK FORCE (2010). *New Considerations for Survey Researchers when Planning and Conducting RDD Telephone Surveys in the U.S. with Respondents Reached via Cell Phone*. Technischer Bericht, AAPOR, American Association for Public Opinion Research.
- ADM, ARBEITSKREIS DEUTSCHER MARKT- UND SOZIALFORSCHUNGSINSTITUTE (2012a). *Forschungsprojekt Dual-Frame-Ansätze - Forschungsbericht*. Technischer Bericht, ADM Forschungsprojekt Dual-Frame-Ansätze.
- ADM, ARBEITSKREIS DEUTSCHER MARKT UND SOZIALFORSCHUNGSINSTITUTE (2012b). *Jahresbericht 2011*.
- BITKOM (2011). *Presseinformation 02.08.2011 - Allzweckgerät Handy: Wecker, Kamera, Terminplaner*.
- BRICK, J. M.; DIPKO, S.; PRESSER S. TUCKER C.; YUAN Y. (2006). *Nonresponse Bias in a Dual Frame Sample of Cell and Landline Numbers*. *Public Opinion Quarterly*, 70:780 – 793.
- BUNDESAMT FÜR SICHERHEIT IN DER INFORMATIONSTECHNIK (2008). *Öffentliche Mobilfunknetze und ihre Sicherheitsaspekte*. Bundesamt für Sicherheit in der Informationstechnik.
- BUNDESNETZAGENTUR (2011). *Nummerierungskonzept 2011*.
- BUNDESNETZAGENTUR (2011). *Tätigkeitsbericht 2010/2011 Telekommunikation*.
- BUNDESNETZAGENTUR (2013). *Teilnehmerentwicklung im Mobilfunk nach Netzen pro Quartal*.
- BUSKIRK, T.D. und J. BEST (2012). *Venn Diagrams, Probability 101 and Sampling Weights Computed for Dual Frame Telephone RDD Designs*. American Statistical Association - Proceedings of the Survey Research Methods Section, S. 3696 – 3710.
- BUSSE, B. & FUCHS, M. (2013). *Prevalence of Cell Phone Sharing*. *Survey Methods: Insights from the Field*, S. 1 – 15. Retrieved from <http://surveyinsights.org/?p=1019>.
- CALLEGARO, M; AYSHAN, O.; GABLER S. HÄDER S. VILLAR A. (2011). *Combining Landline and Mobile Phone SSample - A Dual Frame Approach*. GESIS Working Paper, 13:5 – 30.
- CLARK, J.; WINGLEE, M.; LIU B. (2007). *Handling Imperfect Overlap Determination in a Dual-Frame Survey*. American Statistical Association - Proceedings of the Survey Research Methods Section, S. 3233 – 3238.
- ENGEL, T. (2013). *Locating Mobile Phones using Signalling System No. 7*.
- ERICSSON GMBH (2001). *Die Infrastruktur von Mobilfunk - Funktion und Elemente*.
- GABLER, S.; HÄDER, S.; LEHNHOFF I.; MARDIAN E. (2012). *Weighting for Unequal Inclusion Probabilities and Nonresponse in Dual Frame Telephone Surveys*. In: *Telephone Surveys in Europe*, S. 147 – 167. Häder, S.; Häder, M.; Kühne, M.
- GAJEK, H. (2012). *Gajek-Liste, Vorwahl.faq*.
- GRAESKE, J.; KUNZ, T. (2009). *Stichprobenqualität der CELLA-Studie unter besonderer Berücksichtigung der Mobile-onlys*. In: *Telefonbefragungen über das Mobilfunknetz*, S. 57–70. Häder, M.; Häder, S.
- GUTT, H. (2010). *LTE - Neue Dimension mobiler Breitbandnutzung - Eine technische Einführung*.
- HÄDER, M. (2009a). *Modell einer Mixed-Mode-Studie mit Mobilfunk- und Festnetzbefragung. Genereller*

- Problemaufriss*. In: *Telefonbefragungen über das Mobilfunknetz - Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*, S. 17 – 20. Häder, M.; Häder, S.
- HÄDER, S.; GABLER, S.; HECKEL C. (2009b). *Stichprobenziehung für die CELLA-Studie*. In: *Telefonbefragungen über das Mobilfunknetz - Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*, S. 21 – 49. Häder, M.; Häder, S.
- HECKEL, C. und K. WIESE (2012). *Sampling Frames for Telephone Surveys in Europe*. In: *Telephone Surveys in Europe*, S. 103 – 119. Häder, S.; Häder, M.; Kühne, M.
- INFOBIB (2011). *HLR Lookup - Smart Probe for Intelligent Decisions*.
- KALTON, G und D. ANDERSON (1986). *Sampling Rare Populations*. Journal of the Royal Statistical Society, Series A, 149:65 – 82.
- KUUSELA, V. und M. SIMPANEN (2012). *Spotlights on selected European Countries: Finland*. In: *Telephone Surveys in Europe*, S. 37 – 45. Häder, S.; Häder, M.; Kühne, M.
- LOHR, S. und J. K. RAO (2000). *Inference from Dual Frame Surveys*. Journal of the American Statistical Association, 95:271 – 280.
- LOHR, S. und J. RAO (2006). *Estimation in Multiple-Frame Surveys*. Journal of the American Statistical Association, 101:475:1019–1030.
- LOHR, S. L. (2007). *Recent developments in multiple frame surveys*. American Statistical Association - Proceedings of the Survey Research Methods Section, S. 3257 – 3264.
- LOHR, SHARON L. (2011). *Alternative survey sample designs: Sampling with multiple overlapping frames*. Statistics Canada - Survey Methodology, 37:197–213.
- MECATTI, F. (2007). *A single frame multiplicity estimator for multiple frame surveys*. Statistics Canada - Survey Methodology, 33:151 –157.
- MOBIMEX (2013). <http://mobimex.de/HLR%20%28IMSI%29%20Lookup.html>.
- NUMBERPORTABILITY LOOKUP (2013). <http://numberportabilitylookup.com/home>.
- RAO, J. N. K. und C. WU (2010). *Pseudo-Empirical Likelihood Inference for Multiple Frame Surveys*. Journal of the American Statistical Association, 105:1494 – 1502.
- ROUTO MESSAGING (2009). *HLR LOOKUP User Guide*.
- SCHEFFLER, H. (2013). *Mobilfunkanteil bei Telefonbefragungen: Hartmut Scheffler im Interview*.
- SCHNEID, M.; STIEGLER, A. (2007). *Virtuelle Festnetznummern: Stolpersteine der Umfragesforschung?*. In: *Mobilfunktelefonie - Eine Herausforderung für die Umfragesforschung*, S. 81 – 90. Gabler, S.; Häder, S.
- SCHNEIDERAT, G. und T. SCHLINZIG (2012). *Mobile- and Landline-Onlys in Dual-Frame-Approaches: Effects on Sample Quality*. In: *Telephone Surveys in Europe*, S. 121 – 143. Häder, S.; Häder, M.; Kühne, M.
- SINGH, A. C. und F. MECATTI (2011). *Generalized Multiplicity-Adjusted Horvitz-Thompson Estimation as a Unified Approach to Multiple Frame Surveys*. Journal of Official Statistics, 27:633 – 650.
- SKINNER, C. J. (1991). *On the Efficiency of Raking Ratio Estimation for Multiple Frame Surveys*. Journal of the American Statistical Association, 86:415:779 – 784.
- SKINNER, C.J. und J. RAO (1996). *Estimation in Dual Frame Surveys with Complex Designs*. Journal of the American Statistical Association, 91:433:349 – 356.
- STRUMINSKAYA, B.; SCHAURER, I.; BANDILLA W.; GABLER S.; HÄDER S. & KACZMIREK L., Hrsg. (2011).

Improved cost-effectiveness in mobile surveys using HLR-Lookup. Paper presented at the GOR 11 Conference. GESIS.

WINGLEE, M; PARK, I.; RUST K.; SHAPIRO G. (2007). *A Case Study in Dual-Frame Estimation Methods.* American Statistical Association - Proceedings of the Section on Survey Research Methods, S. 3195 – 3202.

WINNHOVEN, K. (2005). *Mobilfunk - Funktionen und Elemente am Beispiel von GSM.* In: *Vortrag im Rahmen des DV-Seminars.*

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

A Anhang

Einfacher Multiplicity-Schätzer:

$$\hat{Y}_{SM} = \sum_{q=1}^Q \sum_{i \in M_q} \frac{y_i \mathbf{1}_{i \in S(q)}}{h_i \pi_{i(q)}} \quad (\text{A.1})$$

Der hier angeführte Schätzer ist Äquivalent zu 4.19 zu betrachten.

Proportionaler Multiplicity-Schätzer:

$$\hat{Y}_{SM} = \sum_{q=1}^Q \sum_{i \in M_q} \frac{y_i \mathbf{1}_{i \in S(q)}}{h_i \bar{\pi}_{i(q)}} = \sum_{q=1}^Q \sum_{i \in M_q} y_i \alpha_{q(i)}^{PM} \frac{\mathbf{1}_{i \in S(q)}}{\pi_{i(q)}} \quad (\text{A.2})$$

mit

$$\alpha_{q(i)}^{PM} = \frac{\pi_{i(q)}}{\sum_{q' \in K(i)} \pi_{i(q')}}$$

Dieser Schätzer ist dem Verfahren von Kalton und Anderson (1986) nachempfunden und bedarf vollen Informationsgehalt (vgl. KALTON und ANDERSON [1986], S. 65ff.).

Hybrider Multiplicity-Schätzer:

$$\hat{Y}_{HM} = \sum_{q=1}^Q \sum_{i \in M_q} y_i \alpha_{q(i)}^{HM} \frac{\mathbf{1}_{i \in S(q)}}{\pi_{i(q)}} \quad (\text{A.3})$$

mit

$$\alpha_{q(i)}^{HM} = \alpha_{q(i)}^{PM} \mathbf{1}_{i \in M(q)_{full}} + \frac{1}{h_i} \mathbf{1}_{i \in M(q)_{basic}}$$

Dieser Schätzer setzt sich aus der Verwendung des proportionalen Multiplicity-Schätzers für den Bereich der Full Frame Level Information und dem einfachen Multiplicity-Schätzers für Basic Frame Level Information zusammen. Der zweite innerhalb dieser Untersuchung verwendete hybride Schätzer setzt sich aus dem einfachen und dem Composite-Multiplicity-Schätzer zusammen. Diese Vorgehensweise wird verwendet, um die Instabilität des einfachen Multiplicity-Schätzers bei geringen Inklusionswahrscheinlichkeiten auszugleichen.

Non-Multiplicity-Schätzer:

Der von Bankier (1986) entwickelte Schätzer erlaubt keine Multiplizität, weswegen die Stichproben aus den unterschiedlichen Auswahlrahmen $S(1), \dots, S(Q)$ in eine einzige Stichprobe S^* mit paarweise verschiedenen Erhebungseinheiten überführt werden. Dies wird dadurch erreicht, dass die Erhebungseinheiten zuerst in die disjunkten Bereiche M_q eingeteilt und Duplikate im Anschluss entfernt werden. Zu bedenken ist dabei, dass der sich ergebende Stichprobenumfang n^* nun ebenfalls als Zufallsvariable aufgefasst werden kann. Ein Schätzer für den Totalwert berechnet sich dann nach:

$$\hat{Y}_{NM} = \sum_{i \in S^*} \frac{y_i}{1 - \prod_{q \in K(i)} (1 - \pi_{i(q)})} \quad (\text{A.4})$$

Bei einer geringen Wahrscheinlichkeit, dass Haushalte durch unterschiedliche Stichproben in die Erhebung gelangen, entspricht der Non-Multiplicity-Schätzer approximativ dem proportionalen Schätzer (vgl. SINGH und MECATTI [2011], S.640ff.).