

1. Methodische Vorbemerkungen

Das generelle Problem bei der Bildung von Stichproben für telefonische Umfragen in der Bundesrepublik Deutschland besteht darin, dass inzwischen etwas mehr als 50% der Telefonhaushalte nicht in den einschlägigen Verzeichnissen eingetragen sind. Es handelt sich bei diesen Haushalten in erster Linie um solche mit Neuanschlüssen aus den verschiedensten Gründen (Haushaltsgründung, Umzug, Wechsel des Providers, gerade in den Neuen Bundesländern in großem Umfang auch viele Erstanschlüsse nach der Wiedervereinigung, etc.). Diese Haushalte haben einen geringen Anteil "alte" Nummern, die schon seit mehr als 10 Jahren bestehen, und überwiegend "neue" Nummern, die erst in den letzten Jahren freigeschaltet wurden - bei den eingetragenen Privathaushalten ist dieses Verhältnis umgekehrt. Die Eintragsrate der Telefonhaushalte variiert des Weiteren regional stark, sie beträgt beispielsweise in Berlin mittlerweile nur wenig über 30%. Daneben gibt es einen kleinen aber zunehmenden Anteil von Privathaushalten, die überhaupt keinen Festnetzanschluss haben. Es handelt sich dabei überwiegend aber um solche die stattdessen über Mobilfunk erreichbar sind (dazu s. Abschnitt 2.4).

Als im Laufe der 1990er Jahre klar wurde, dass der Anteil nicht-eingetragener Privathaushalte eine nicht zu vernachlässigende Größenordnung angenommen hatten, wurden vermehrt Verfahren mit zufälliger Generierung von Telefonnummern eingesetzt, damit auch diese Privathaushalte erreicht werden konnten.

Siegfried Gabler und Sabine Häder (damals ZUMA Mannheim) haben als Erste die herkömmlichen Verfahren für Telefonstichproben in Deutschland wissenschaftlich analysiert und nachgewiesen, dass bei diesen Stichproben die Auswahlwahrscheinlichkeit der Telefonhaushalte

- nicht gleich, - nicht bekannt und - auch nicht berechenbar sind.

Mit anderen Worten: die herkömmlichen Stichproben - Auswahl nur von

eingetragenen Privathaushalten, Random Digit Dialing (RDD), Random Last Digit (RLD), etc. - sind statistisch verzerrt, und diese Verzerrungen sind weder vernachlässigbar noch korrigierbar. (Näheres hierzu in den Anmerkungen am Ende dieses Textes.)

Sie haben daraufhin ein zu seiner Zeit bahnbrechendes Modell entwickelt, das auf sogenannten Hunderterblocks basiert. Ein valider Hunderterblock ist definiert durch eine eingetragene Telefonnummer eines Privatanschlusses und Variation der letzten beiden Ziffern (etwa der Privatanschluss 08141-94965 erzeugt den Hunderterblock 08141-949xx, d.h. alle 100 Nummern von 08141-94900 bis 08141-94999). Wenn man nun alle Hunderterblocks mit ihrer jeweiligen Eintragsdichte (Anzahl der eingetragenen Anschlüsse pro Hunderterblock) und die Blockdichte (Anzahl valider Nummern pro Hunderterblock) ermittelt oder hinreichend genau (erwartungstreu) schätzt, so kann man daraus Stichproben erstellen, bei denen die Auswahlwahrscheinlichkeiten der Stichprobeneinheiten gleich oder zumindest berechenbar sind.

Leider hat das Gabler-Häder-Verfahren in der ursprünglichen Form eine Reihe von spezifischen Nachteilen:

Die betreffenden Stichproben sind nach den Telefonanschlüssen 'repräsentativ', verteilt - dies ist aber von untergeordnetem Interesse, denn in der Regel bilden **'Privathaushalte'** bei den Bevölkerungsumfragen die **Grund- bzw. Auswahlgesamtheit der sekundären Auswahlstufe.**

- Mehrfacheinträge von Privathaushalten werden dabei nicht berücksichtigt..
- Firmen- und Behördeneinträge werden nur sehr unvollkommen erkannt, etwa über sog. 'Bindestrich-Nummern', also Durchwahlnummern von Telefonanlagen.
- Die naturwüchsigen Aggregate der Einheiten bzw. Hunderterblocks sind Vorwahlnummern, über die keine hinreichenden sekundärstatistischen Daten oder Strukturen verfügbar sind. Die 'normalen' Gebietseinheiten, wie Gemeinden, Kreise, Regierungsbezirke oder Bundesländer sind für dieses System inkompatibel. Die kleinräumige Variation dieser Stichproben ist zudem sehr hoch.
- Der Ausgleich von Non-Response-Fehlern oder differenzierte Redressementgewichtungen sind somit nur eingeschränkt möglich.

Damit ist das "reine" System für Umfragen in der Markt-, Medien- und Sozialforschung in der Regel nicht optimal geeignet.

2. Das Rösch-Telefonstichproben-System (RTS)

2.1 Auswahlgrundlage Für das Rösch-Telefonstichproben-System wurde der grundlegende Ansatz von Gabler und Häder konsequent umgesetzt und um einige wesentliche Aspekte erweitert.

Die Basis des RTS bildet eine Datenbank von allen eingetragenen Telefonanschlüssen, die aus den verschiedensten verfügbaren und kritisch überprüften Quellen **ständig aktualisiert wird**.

Die zugehörige Adressenbasis wird in einem komplexen Verfahren zunächst auf nicht-private Anschlüsse geprüft (Abgleich mit Branchenverzeichnissen spezielle und selbstlernende Erkennungssoftware, etc.) und die betreffenden Anschlüsse werden - genauso wie z.B. erkannte Telefaxanschlüsse - entsprechend gekennzeichnet.

Danach werden Mehrfacheinträge einer Nummer (unter verschiedenen Namen) eliminiert und in mehreren unterschiedlichen Schritten die Zahl der Mehrfacheinträge eines Haushalts ermittelt und festgehalten.

Die enthaltenen Anschlüsse haben ja naturgemäß immer einen Namen und eine (vollständige) Telefonnummer, ansonsten aber durchaus unterschiedliche Informationsniveaus. Diese Informationsniveaus werden optimal ausgeschöpft, um die Anschlüsse mit verschiedenen regionalen Merkmalen wie z.B. Gemeindeschlüssel und Stadtteil (in größeren Städten) anzureichern.

Somit erhalten wir:

- voll qualifizierte Nummern eingetragener Anschlüsse (Gemeindeschlüssel, ggfs. Stadtteil, Postleitzahl, Ortsname, Adresse)

- teilweise qualifizierte Nummern eingetragener Nummern mit unterschiedlichem Niveau:

- Gemeindeschlüssel ohne Stadtteile
- Ortsnetz mit Wahrscheinlichkeiten für mehrere Gemeinden
- Ortsnetz ohne weitere Qualifizierung

- voll qualifizierte Nummern nichteingetragener Anschlüsse (s.o., ohne Adresse)
- teilweise qualifizierte nichteingetragener Anschlüsse (s.o.)

Die nichteingetragenen Nummern in den Hunderterblocks enthalten dabei naturgemäß einen gewissen Anteil nicht existierender Anschlüsse, die bei korrekter Behandlung und unter gewissen Umständen als stichprobenneutrale Ausfälle übergangen werden können.

Das Konzept für die Erstellung dieser Auswahlgrundlage wurde vom Verfasser 1998 entwickelt und ist im Übrigen auch in die Mitte 1999 fertiggestellte Auswahlgrundlage des ADM eingeflossen.

Da in den folgenden Jahren bei der RTS-Auswahlgrundlage ein paar methodische Veränderungen vorgenommen wurden, und diese zudem ständig - und nicht nur in größeren Zeiträumen - aktualisiert und an neuere Sachverhalte, wie z.B. Gebiets- und Bevölkerungsstandsänderungen oder spezielle Anforderungen durch neue Stichprobenverfahren, angepasst wird, sind zwischen den beiden relevanten Auswahlgrundlagen in Deutschland an einigen Punkten im Laufe der Zeit gewisse strukturelle Unterschiede entstanden.

2.2 Nationale Telefonstichproben
Somit ist die RTS-Auswahlgrundlage optimal für die Erstellung von komplexen Stichproben durch spezielle Stichprobenverfahren ausgerichtet, mithin können daraus hinreichend

erwartungstreue und kleinräumig hoch auflösende Stichproben für die Grundgesamtheit "Privathaushalte" (und nicht: "Telefonanschlüsse" oder "Telefonhaushalte") gezogen werden. Bei RTS-Stichproben werden - im Gegensatz zum klassischen Gabler-Häder-Design - zur Erhöhung der Stichprobenqualität grundsätzlich getrennte Teilstichproben für eingetragene und nichteingetragene Privathaushalte erstellt, die zusammengesetzt ein optimales Abbild der Grundgesamtheit ergeben. Nur so lassen sich die für beide Teilgesamtheiten verfügbaren unterschiedlichen Schichtungsmöglichkeiten maximal ausschöpfen. Das konkrete Stichprobenverfahren selbst hängt natürlich von den Zielsetzungen der jeweiligen Forschungsaufgabe ab und ist auf Grund der Gegebenheiten relativ kompliziert (z.B. "erweiterte Mikrostratifikation"). Es stellt aber sicher, dass jede Stichprobe grundsätzlich für mindestens zwei Jahre überschneidungsfrei zu allen vorangegangenen Stichproben des Systems ist.

2.3 Regionale Telefonstichproben

Wenn die Auswahlverfahren für nationale Telefonstichproben auf regionale übertragen werden, ergeben sich in der Regel Komplikationen, die nicht ohne wesentliche Ergänzung der Auswahlgrundlage und einer grundlegenden Modifikation des Stichprobenverfahrens beherrschbar sind. Der Grund liegt in der völligen gegenseitigen Inkompatibilität von Gebieten, die durch Ortsnetze (Vorwahlbereiche), Gebietskörperschaften (Gemeinde, Kreis Regierungsbezirk, Bundesland) und Postleitzahlen oder -bereiche definiert werden. Zur Verdeutlichung der Problematik wollen wir ein einfaches Beispiel betrachten:

Es soll eine Brutto-Stichprobe vom Umfang 100 aus der Gemeinde Stapelfeld (Bz. Hamburg) gezogen werden. Diese Gemeinde liegt komplett im Ortsnetz von Hamburg und sie ist dort mit einem Anteil von ca. einem Promille an den eingetragene Privathaushalten vertreten. Wenn man vereinfachend annimmt, dass der Anteil nichteingetragener Privathaushalte einheitlich bei 50% liegt, dann müsste man 100.000 valide Nummern, also mehr als 200.000 Nummern insgesamt auswählen. Wenn das Stichprobenverfahren nach eingetragenen und nicht-eingetragenen aufgeteilt wird, dann wären immer noch 50.000 (aus 100.000 nichteingetragenen Nummern) Haushaltskontakte erforderlich und davon 49.950 zu verwerfen, um die erforderlichen 50 aus Stapelfeld zu realisieren - man muss kein Fachmann sein, um dies als forschungsökonomisch völlig sinnloses Verfahren zu einzuschätzen. Wenn sich eine Gemeinde auf mehrere Ortsnetze mit unterschiedlichen Anteilen ihrer Privathaushalte und völlig anderen Anteilen innerhalb der jeweiligen Ortsnetzen aufteilt - was relativ häufig vorkommt -, verkompliziert sich dieser Umstand zunehmend.

Ortsnetz anderer	Verteilung der gesamt	der Gemeinde X nach Ortsnetzen	Haushaltsanteil der Gemeinde X am jew. Ortsnetz	Anteil Gemeinden
A :	10%	60%	100%	90%
B :	90%	20%	100%	10%
C :	10%	5%	95%	100%
D :	99%	8%	100%	1%
E :	99,9%	2%	100%	0,1%
		100%		

Im Ortsnetz A (z.B. die Vorwahl 08141) liegen 60% der Privathaushalte der Gemeinde X, der Anteil der Gemeinde A beträgt 90% der Privathaushalte aller Privathaushalte im Ortsnetz A.

Ursprünglich (1998) hatte der Verfasser vorgeschlagen, das Problem mit Hilfe von Auswahlwahrscheinlichkeiten für die Gemeindezugehörigkeiten je Hunderterblock zu lösen. Dies hat sich jedoch in der Praxis als nicht praktikabel und schließlich als Irrweg herausgestellt. Es hat sich nämlich die unterstellte Korrelation in der überwiegenden Mehrzahl der Ortsnetze als nicht hinreichend erwiesen. Daneben wäre die für den Ausgleich der damit verbundenen unterschiedlicher Auswahlwahrscheinlichkeiten erforderliche Designgewichtung durch die große Varianz der Gewichtungsfaktoren völlig untolerierbar gewesen.

Eine teilweise einfache Substitution von nicht-eingetragenen Privathaushalten durch eingetragene im Befragungsgebiet war ein Notbehelf der auf Grund der eingangs aufgeführten Mängel in der Struktur der eingetragenen Haushalte auf Dauer nicht akzeptierbar war. Damit war für regionale Stichproben eine grundsätzliche Neukonzeption des Stichprobensystems erforderlich, um auch im regionalen Rahmen repräsentative Telefonstichproben in der Bevölkerung zu ermöglichen.

Die Überlegungen führten schließlich zu der zusätzlichen Qualifizierung

der Telefonnummern nach dem Alter des Anschlusses und der Anreicherung der RTS-Datenbank mit diesem Merkmal. Damit können ausgewählte nichteingetragene Nummern, für die die Zuordnung zum jeweiligen Gebiet nicht hinreichend eindeutig vorgenommen werden kann, geeignet durch eingetragene Privathaushalte substituiert werden. "Geeignet" heißt in diesem Zusammenhang, dass bei dieser Substitution unter anderem der Anteil "neuer" Nummern mittels disproportionaler Allokation so stark überhöht wird, wie es den ursprünglich in den betreffenden nicht-eingetragenen Telefonnummern enthaltenen Privathaushalten entsprochen hätte.

2.4 Mobilfunk

Der Anteil der Personen, die nur über Mobilfunk erreichbar sind und die im Haushalt keinen Festnetzanschluss besitzen (die sog. "Mobile Onlys"), hat in den letzten Jahren stetig zugenommen, so dass Überlegungen angestellt wurden, auch diesen Personenkreis bei telefonischen Bevölkerungsumfragen zu erfassen. Da die Mobilfunkeinträge ebenso in der RTS-Datenbank enthalten sind, lassen sich daraus auch Mobilfunkstichproben erstellen. Allerdings ist der Aufwand dafür in vielerlei Hinsicht größer und die Stichprobenqualität wegen der spezifischen Gegebenheiten geringer als bei herkömmlichen Telefonstichproben. So lassen sich beispielsweise die nichteingetragenen Mobilfunknummern nicht regional sondern allenfalls nach Provider zuordnen. Der Nummernraum ist verhältnismäßig groß und mithin die Hittrate relativ gering. Die Orte und die Befragungssituation in denen der Angerufene sich gerade befindet sind unkontrollierbar und lassen häufig keine ordnungsgemäße Durchführung eines Interviews zu. Die Abbruchraten und die gesamten Ausschöpfungsraten verringern sich dadurch deutlich. Kurzum, die Kosten für Erhebungen in dieser Teilgruppe sind relativ hoch. Somit stellt sich für allgemeine Bevölkerungsumfragen die Frage nach dem Nutzen einer Aufstockung oder Zusatzerhebung von Mobilfunknutzern hinsichtlich einer Verbesserung der Gesamtergebnisse durch eine Reduktion des systematischen Fehlers ("Bias") gegenüber herkömmlichen Festnetzstichproben.

Bei den "Mobile-Onlys" handelt es sich überwiegend um junge, mobile Personen in Einpersonenhaushalten. Ihr Anteil an den Privathaushalten wird gegenwärtig etwa bei 12%-15% eingeschätzt.

Die entscheidende Frage ist nun, ob sich der Personenkreis der "Mobile-Onlys" hinsichtlich der spezifischen Fragestellung der jeweiligen Forschungsaufgabe und der entsprechenden Zielvariablen von anderen Personen der gleichen Altersgruppe, des gleichen Geschlechts, der gleichen Schulbildung, der gleichen Art der Berufstätigkeit, der gleichen Region, des gleichen Gemeindetyps, etc., kurz: der durch Stichprobenanlage und Gewichtung kontrollierten Stichprobenmerkmale, so

deutlich unterscheiden, dass dadurch wahrnehmbare Abweichungen in den Umfrageergebnissen und den abgeleiteten Marktforschungsanalysen zu erwarten sind.

Ein vereinfachtes Beispiel soll diesen Aspekt verdeutlichen:

Wir nehmen einmal an, dass der wahre Anteil der "Mobile-Onlys" 5% an der Grundgesamtheit beträgt, und dass eine Festnetzstichprobe bei einer Zielvariablen einen 50%-Anteil ergab. Wenn die "Mobile-Onlys" sich diesbezüglich in etwa wie ihr soziodemographisch fassbares Pendant in der Festnetzstichprobe (s.o.) verhält, so wird sich keine Ergebnisverbesserung durch eine zusätzliche Befragung dieser speziellen Teilgruppe ergeben. Ist jedoch das extreme Gegenteil der Fall, sagen wir einmal, eine Abweichung von 30 Prozentpunkten von der entsprechenden Festnetzpopulation, so wirkt sich dies gerade in 1,5 Prozentpunkten Differenz zum Festnetzergebnis aus.

Wenn wir nun diesen Bias in Beziehung zum üblichen zufälligen Fehler einer reinen Festnetzstichprobe setzen, der bei 2000 Befragten im vorliegenden Beispiel immerhin bei etwa 50,0% +/- 3,1% liegt, so würde er völlig von diesem Konfidenzintervall überdeckt. Mehr noch: eine Integration einer "Mobile-Only"-Aufstockung von 100 (=5%) einschlägigen Interviews (möglicherweise aus einer sehr hohen Zahl von kontaktierten Mobilfunk-Nutzern) in die Festnetzstichprobe würde, wegen der geringeren Stichprobenqualität, den gesamten zufälligen Fehler noch zusätzlich erhöhen. Hinzu kommen die spezifischen systematischen Fehler, das teilweise inkompatible Stichprobendesign und die speziellen "Mode-Effekte", auf die wir hier wegen ihrer Komplexität nicht näher eingehen können.

Niemand könnte vor diesem Hintergrund eine Zusatzstichprobe von Mobilfunknutzern, selbst bei etwas höheren Anteilen der "Mobile-Onlys" an den Privathaushalten, rechtfertigen.

Anders verhält es sich allerdings, wenn spezielle Zielgruppen, etwa Jugendliche oder sehr mobile Personen, im Fokus einer Untersuchung stehen. Dann wäre eine Zusatzstichprobe von "Mobile-Onlys" bzw. Mobilfunknutzern sorgfältig in Erwägung zu ziehen und unter Einbeziehung aller im konkreten Einzelfall relevanter statistischer, logistischer und forschungsökonomischer Gesichtspunkte zu entscheiden, ob sich der beträchtliche Mehraufwand als zwingend zur Erreichung des Untersuchungsziels erweist.

2.5 Dual Frame

Für nationale Telefonumfragen wurde um 2005 von Gabler und Häder das Dual-Frame-Verfahren entwickelt. Dabei wird eine herkömmliche Festnetzstichprobe mit einer getrennten Mobilfunkstichprobe über ein komplexes Gewichtungsverfahren so verknüpft, dass eine gemischte repräsentative Gesamtstichprobe entsteht. Im RTS ist das Dual-Frame-Verfahren standardmäßig implementiert und wird für qualitativ anspruchsvolle nationale Umfragen empfohlen.

3. Gewichtung

Telefonstichproben in der Bevölkerung sind häufig zunächst Haushaltsstichproben (meist in der zweiten Auswahlstufe), bei denen jeder Privathaushalt im Prinzip die gleiche Auswahlchance hat. Wenn in den Haushalten dann eine Zielperson zufällig ausgewählt wird, so sind die Auswahlchancen abhängig von der jeweiligen auf die Zielgruppe eingeschränkten Haushaltsgröße. Daher sind die resultierenden Stichproben auf Personenebene zwangsläufig statistisch verzerrt und müssen gewichtet werden, um diese design-gegebene Verzerrung auszugleichen - es sei denn, man beschränkt sich nur auf die Auswertung von haushaltsbezogenen Merkmalen. Den Ausgleich unterschiedlicher Auswahlwahrscheinlichkeiten nennt man "Design-Gewichtung", hier, in diesem speziellen Fall, handelt es sich um eine "Transformation" der Haushalts- in eine Personenstichprobe. Zu dieser Klasse gehört auch die vom Design abhängige Gewichtung der nichteingetragenen Haushalte in der Stichprobe.

Daneben ergeben sich - abhängig vom äußeren Rahmen der Untersuchung (wie z.B. Thema, Interviewlänge, Feldzeit, etc.), durch Nonresponse und andere Ausfälle - Abweichungen von der ursprünglichen modellgetreuen Stichprobe, die erheblich sein können und die in der Regel auch nicht zu vernachlässigen sind (z.B. werden kleinere Haushalte oft schwerer angetroffen, jüngere Personen haben u.U. weniger Zeit, andere weniger Interesse am Thema, usw.)

Deshalb ist (zumindest bei den üblichen Ausschöpfungsraten) eine a-posteriori-Anpassung ("Redressement") der Stichprobe an vorliegende soziodemografische Strukturen (z.B. der amtlichen Statistik oder aus anderen Quellen) dringend erforderlich, um die zwangsläufigen und unvermeidbaren Defizite der realisierten Stichprobe auszugleichen oder wenigstens zu verringern.

Die entsprechenden Merkmale und ihre Ausprägungen hängen insbesondere von der Fallzahl und anderer Gegebenheiten der jeweiligen Umfrage, der Verfügbarkeit aktueller und valider Strukturen sowie der Feinheit der

geplanten Auswertungen ab. Insbesondere können bei RTS-Stichproben die Design-Merkmale "eingetragener / nicht eingetragener Privathaushalt" sowie "Alter der Telefonnummer" zusätzlich zu den üblichen soziodemografischen Redressement-Merkmalen Verwendung finden und zu einer deutlichen Verringerung des Nonresponse-Fehlers beitragen.

Das eingesetzte Gewichtungsverfahren ist nach dem aus der Informationstheorie stammenden Prinzip "des minimalen Informationsverlusts" vom Verfasser weiterentwickelt und stellt sicher, dass die vorgegebenen komplexen Strukturen (die aktiven Merkmale der Gewichtung) exakt und mit minimaler Veränderung der Stichprobe statistisch vollkommen transparent reproduziert werden.

Designgewichtung und Redressement sind in der Regel unverzichtbare Bestandteile qualitativ anspruchsvoller Stichproben in der Umfrageforschung. Sonst ist die Bezeichnung "repräsentativ" jedenfalls nicht gerechtfertigt.

Anmerkungen:

1. RDD

Das reine "Random Digit Dialing"-Verfahren (RDD, alternative, ebenso korrekte und verbreitete Schreibweise: "Random Digit Dialling"), bei dem Telefonnummern mit Zufallszahlengeneratoren erzeugt werden, ist generell weniger optimal und besonders nicht in Ländern angezeigt, in denen die Telefonnummern (wenigstens innerhalb von Ortsnetzen) keine feste Länge haben, was außer in Deutschland z.B. auch in Österreich und Großbritannien der Fall ist. Abgesehen von einer Reihe statistischer Nachteile, ist es wegen der geringen Trefferquote zudem höchst ineffizient. Seine Anwendung ist generell nur dann gerechtfertigt, wenn k einerlei forschungsökonomisch akzeptablen Alternativen zu Verfügung stehen.

2. RLD

Beim sogenannten "Random-Last-Digit"-Verfahren (RLD) werden zunächst eingetragene Nummern ausgewählt (wir wollen hier einmal unterstellen, dass diese Auswahl fachgerecht erfolgt) und es wird beabsichtigt, nichteingetragene Haushalte durch die Variation der letzten Ziffer oder der beiden letzten Ziffern dieser Telefonnummern zu erhalten. Fatalerweise bleibt man dadurch vor allem in den herkömmlichen "alten" Hunderterblocks mit relativ hoher Eintragsdichte jedoch relativ geringer Blockdichte, also mit wenigen validen nichteingetragenen Nummern, wobei man dann bei Treffern valider Nummern von privaten Haushalten überwiegend solche vorfindet, die ihrerseits auch eingetragen sind. Mehrfach eingetragene Haushalte haben dabei natürlich auch eine entsprechend höhere Auswahlchance.

Weil darüber hinaus wesentliche Kenntnisse über den Nummernraum fehlen und diese auch nicht mit statistischen Mitteln sui generis, also aus der Stichprobe selbst heraus geschätzt werden können, sind die Treffer bei so erzeugten Nummern völlig willkürlich. Somit können die Verzerrungen auch nachträglich nicht durch Gewichtung geheilt werden, denn sie entziehen sich jeder formalen statistischen Einschätzung.

Und weil sich die Verzerrungseffekte zudem gegenseitig verstärken, ist RLD nicht wesentlich besser zu bewerten als einfache Stichproben von eingetragenen Telefonhaushalten. Leider findet dieses zwar einfache, aber im Grunde völlig ungeeignete Verfahren, immer noch gelegentlich seine Anwendung.

3. Implizite Gewichtung (durch Aussteuerung im Feld)

Gelegentlich werden Telefonstichproben von Feldinstituten nach soziodemografischen Merkmalen "ausgesteuert", womit häufig der Verzicht auf eine Gewichtung einhergeht. Die realisierten Stichproben werden dann meist als "nach Altersgruppen repräsentativ" oder "nach Bundesländern repräsentativ" bezeichnet, wenn die Verteilungen dieser Merkmale den entsprechenden Strukturen der amtlichen Statistik entsprechen. Dazu ist zu betonen, dass es eine Repräsentativität nach einzelnen Merkmalen nicht geben kann. Die Information wäre auch völlig nutzlos, denn diese Strukturen sind ja bereits bekannt, dazu bedarf es keiner Umfrage. Eine Stichprobe ist entweder auf Grund der Methode und des konkret durchgeführten Erhebungs- und Gewichtungsverfahrens repräsentativ oder nicht. Eine Selektion im Zuge der Feldarbeit nach einzelnen Strukturmerkmalen wie z.B. Alter, Geschlecht, Haushaltsgröße, etc., führt im Allgemeinen zu völlig willkürlichen und lediglich optischen "Verbesserungen".

(So würden beispielsweise junge Personen im Alter von 18-25 Jahren möglicherweise in ihrer Anzahl in der Stichprobe "richtig" repräsentiert, aber die meisten wären solche, die noch bei den Eltern wohnen, weil sie in größeren Haushalten leichter angetroffen werden. Im Gegenzug dazu wären dann die entsprechenden Elternteile von jungen Menschen unterrepräsentiert, da in diesen Haushalten ja nur junge Personen befragt wurden, und schwerer anzutreffende Personen in der Altersgruppe, die bereits in eigenen Haushalten leben, wären kaum vertreten, weil die Quote vorzeitig erfüllt wäre. Diese deutlich verzerrte Stichprobe wäre dennoch im entsprechenden Jargon nach "Altersgruppen und Haushaltsgröße repräsentativ".)

Günther RöschDiplom-Mathematiker

Marktforscher BVM

Maisach
2013

im Dezember

Büro für Erhebungsdesign & Datenanalyse

Tel: 08141-39121

Fax: 08141-39123

guenther.roesch@roesch-research.de

Anhang:

Alters- und Geschlechtsverteilung einer telefonischen Umfrage der deutschen Bevölkerung ab 18 Jahren in Privathaushalten in Bayern im Dezember 2007 (n=1000)

Die Zahlen unterstreichen die soziodemographischen Unterschiede zwischen eingetragenen und nicht-eingetragenen Privathaushalten. Die Vergleichswerte der entsprechenden Wohnbevölkerung dienen nur der Veranschaulichung und dürfen hier schon wegen der speziellen Design-Gegebenheiten (z.B. unterschiedliche Auswahlwahrscheinlichkeiten für Personen in unterschiedlich großen Haushalten) keineswegs als Sollwerte - quasi als Qualitätsmaßstab für die ungewichtete Stichprobe - fehlinterpretiert werden!

	Eingetragene	Nicht eingetr.	Wohnbevölkerung
	ungewichtet	ungewichtet	in Privat-HH
männlich	%	%	%
18-29	5,1	10,1	8,3
30-39	5,7	11,6	8,8
40-49	9,0	9,5	10,4
50-64	9,5	6,5	11,1
65+	8,9	3,0	9,6
weiblich			

18-29	5,7	15,1	8,2
30-39	11,5	12,6	8,5
40-49	16,2	17,1	10,1
50-64	15,7	13,1	11,2
65+	12,6	1,5	13,8
	100,0	100,0	100,0
			0