
Zufall oder Quote?

Das Sample als Herausforderung in der Online-Forschung

Martin Welker, MHMK München

21. Mai 2010, Methodenforum Universität Wien

Überblick

- Allgemeines zu und Bedeutung von Stichproben
- A. Personenstichproben (Online-Befragung):
 - 7 wichtige Typen
 - a. zufallsbasiert
 - b. nicht-zufallsbasiert
 - c. Probleme & Lösungsansätze
- B. Inhaltsstichproben (Online-Inhaltsanalyse)
 - a. zufallsbasiert
 - b. nicht-zufallsbasiert
 - c. Probleme & Lösungsansätze
- C. Offene Fragen

Zwei Alternativen

- Der Forscher untersucht tatsächlich alle Elemente in Bezug auf die interessierenden Merkmale (Vollerhebung) oder

- Der Forscher untersucht einen Teil der Elemente und greift anschließend auf eine Regel zurück, mit deren Hilfe man von den Ergebnissen der Teiluntersuchung auf alle Elemente schließen kann.



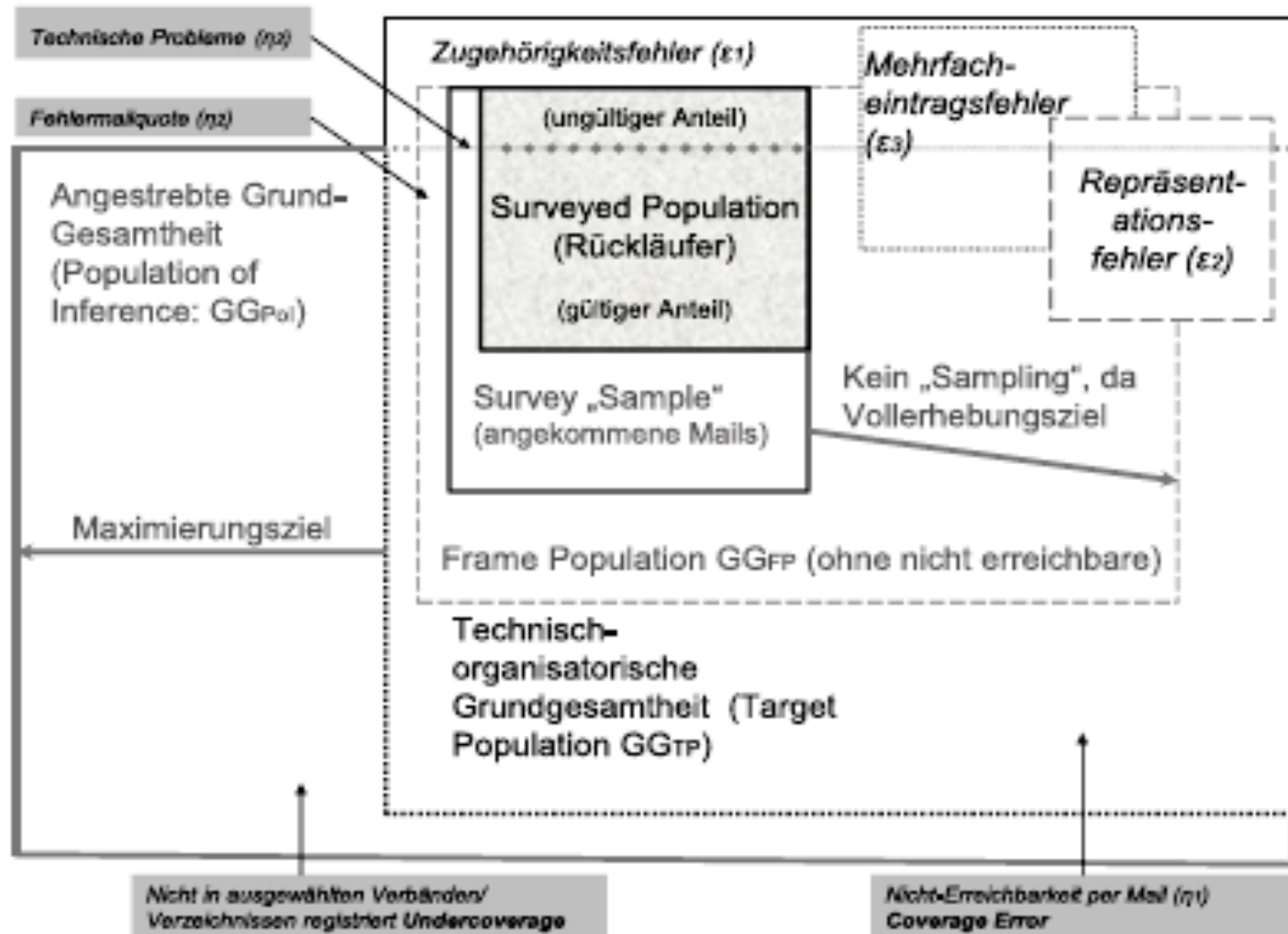
Repräsentationsschluss

Um Schlüsse auf eine (größere) Grundgesamtheit zu ziehen, muss die Stichprobe best. Kriterien genügen (vgl. auch Maurer & Jandura 2009: 62ff.; Baur & Florian 2009: 109 ff.):

- Die Zielgesamtheit (Target Pop.) muss bekannt und definierbar sein
- Die Auswahlgesamtheit (Frame Pop.) muss identifiziert werden
- Under- & Overcoverage sind zu vermeiden
- Die Stichprobe muss ausreichend ausgeschöpft werden
- Inferenzstatistik: Alle Elemente der Auswahlgesamtheit müssen die gleiche, von Null verschiedene Chance haben, in die Stichprobe (Sample) aufgenommen zu werden (Anwendung Multivariater Verfahren)

ABBILDUNG 1

Grundgesamtheit – Auswahlgesamtheit – Stichprobe



Quelle: Studie ZdJ

Nicht-zufallsbasierte Verfahren

- Die Gruppe der nichtzufälligen Verfahren bedient sich sehr verschiedener Argumentationen und Strategien, um eine Grundgesamtheit möglichst genau abzubilden.
- Die meisten greifen dabei auf a priori vorhandene Informationen über die Grundgesamtheit wie bspw. deren Struktur bezüglich relevanter Merkmale zurück. Es werden dann systematisch solche Elemente in die Stichprobe aufgenommen, welche für die bereits bekannten Eigenschaften der Grundgesamtheit >typisch< sind.
 - Diese Strategie wird entweder auf jedes Element einzeln angewandt (z. B. bei der Auswahl typischer Fälle) oder auf die Aggregatmerkmale der Stichprobe (bspw. das Quotenverfahren).

Zufallsbasierte Verfahren

- Zufallsbasierte Verfahren stützen sich auf die Regeln der Wahrscheinlichkeitsrechnung.
- Hat jedes Element x aus n der Grundgesamtheit von der Größe N die gleiche (von Null verschiedene) Wahrscheinlichkeit, in der Untersuchung berücksichtigt zu werden, genügt eine hinreichend große Anzahl an untersuchten Elementen, um mit einer hohen Wahrscheinlichkeit auf die >wahren< Ergebnisse in der Grundgesamtheit zu schließen.



Bildquelle: WPGS

A. Online-Personenstichproben

In Anlehnung an Couper & Coutts (2006) werden sieben in der Forschungspraxis verbreitete Verfahren unterschieden, Stichproben bei Online-Erhebungen zu rekrutieren.

- a. Nicht-zufallsbasierte Verfahren
- b. Zufallsbasierte Verfahren
- c. Einige Probleme und Lösungen

a. nicht-zufallsbasiert

1. „Offene WWW-Umfragen“:

□ Hierbei handelt es sich um die einfachste Möglichkeit, Stichproben online zu generieren. Im WWW wird auf einer Seite zur Teilnahme an einer Umfrage aufgefordert. Jeder Besucher dieser Seite entscheidet dann selbst, ob und wie oft er an dieser Umfrage teilnimmt.

2. „Befragungen aus Freiwilligen-Panels“:

□ Ähnlich wie beim Typ 1 werden hier über einen Aufruf (on- oder offline) Teilnehmer für ein Access Panel gesucht. Die Aufnahme in das Panel unterliegt der Selbstselektion.

□ In einem zweiten Schritt werden dann für einzelne Befragungen innerhalb dieses Panels einzelne Personen entweder über ein Zufallsverfahren oder anhand bestimmter Merkmale (Quoten) ausgewählt.

Wird in der Praxis häufig eingesetzt

„Journalismus 2009“

- Online-Befragung, Feldzeit März 2009
- Stichprobe n = 983
- Online-Panel der yougov/Psychonomics AG, N = 80.000
Aus dieser Grundgesamtheit wurden die Teilnehmer zunächst per Zufallsauswahl eingeladen und dann nach den Strukturen der deutschen Bevölkerung in Anlehnung an den Mikrozensus für die Altersgruppen 16 bis 65 Jahre quotiert (Teilnehmer, die nicht in die Quoten passen, wurden ausgescreent). Zusätzlich zu den demographischen Quoten wurden in Anlehnung an die MA-Reichweiten der AG.MA (Arbeitsgemeinschaft Leseranalyse) Quoten für bestimmte Lesergruppen definiert. Journalisten waren ausgeschlossen.
- Diese Befragung basiert somit nicht auf einer Zufalls-, sondern auf einer Quotenstichprobe.

Ergebnisse und Vergleich

- Journalist ist ein durchaus gesellschaftlich geschätzter Beruf (**entgegengesetzt** zu Allensbach).
- Journalisten werden als mächtig wahrgenommen (ganz ähnlich: Donsbach et al. 2009).
- Journalisten werden für beeinflussbar gehalten und als nur bedingt an einer wahrhaften und richtigen Berichterstattung interessiert: Das Vertrauen in den Journalismus ist insgesamt somit nicht sonderlich hoch (siehe auch Donsbach et al. 2009: 74).
- Entscheidende erklärende Variable ist Bildung. Je höher der formale Bildungsgrad, desto höher die **Skepsis** (**entgegengesetzt** zu Donsbach et al. 2009: 133).
- Das Wissen über Journalisten und Journalismus ist insgesamt eher gering (ähnlich auch Donsbach et al. 2009: 132).

b. zufallsbasiert

3. „Intercept-Verfahren“:

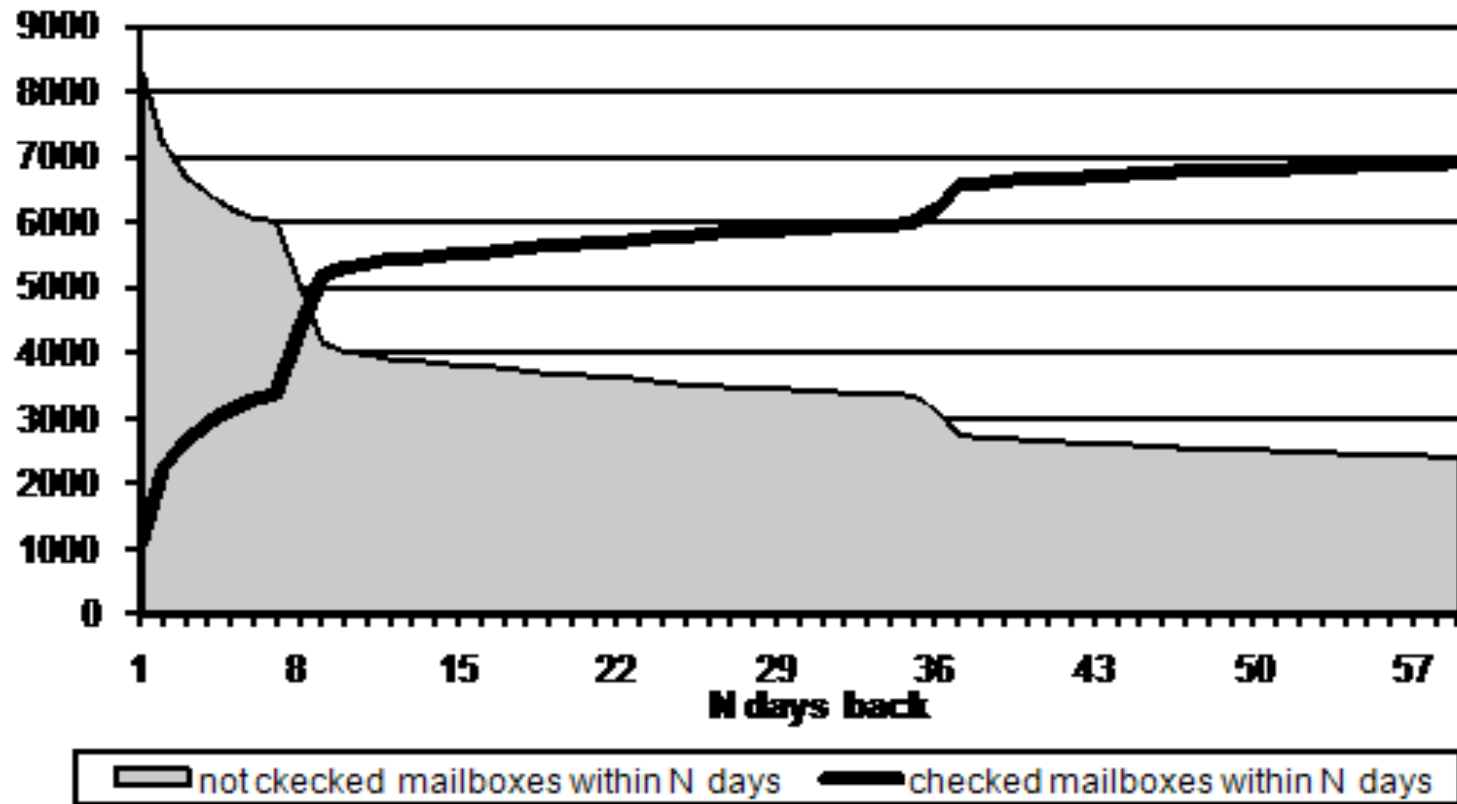
□ Verbreitetes Rekrutierungsverfahren. Hier wird nach einem Zufallsprinzip (in der Regel jeder n-te Besucher) in die Stichprobe aufgenommen. Die Online-Nutzung wird durch eine Aufforderung zur Teilnahme unterbrochen. Hierbei handelt es sich um eine echte Zufallsauswahl. Problematisch sind bei diesen Auswahlen die recht geringen Rücklaufquoten zwischen 2 und 15 Prozent (Theobald 2003).

4. „Befragungen mit listenbasierten Stichproben aus Populationen mit einem hohen Abdeckungsgrad“:

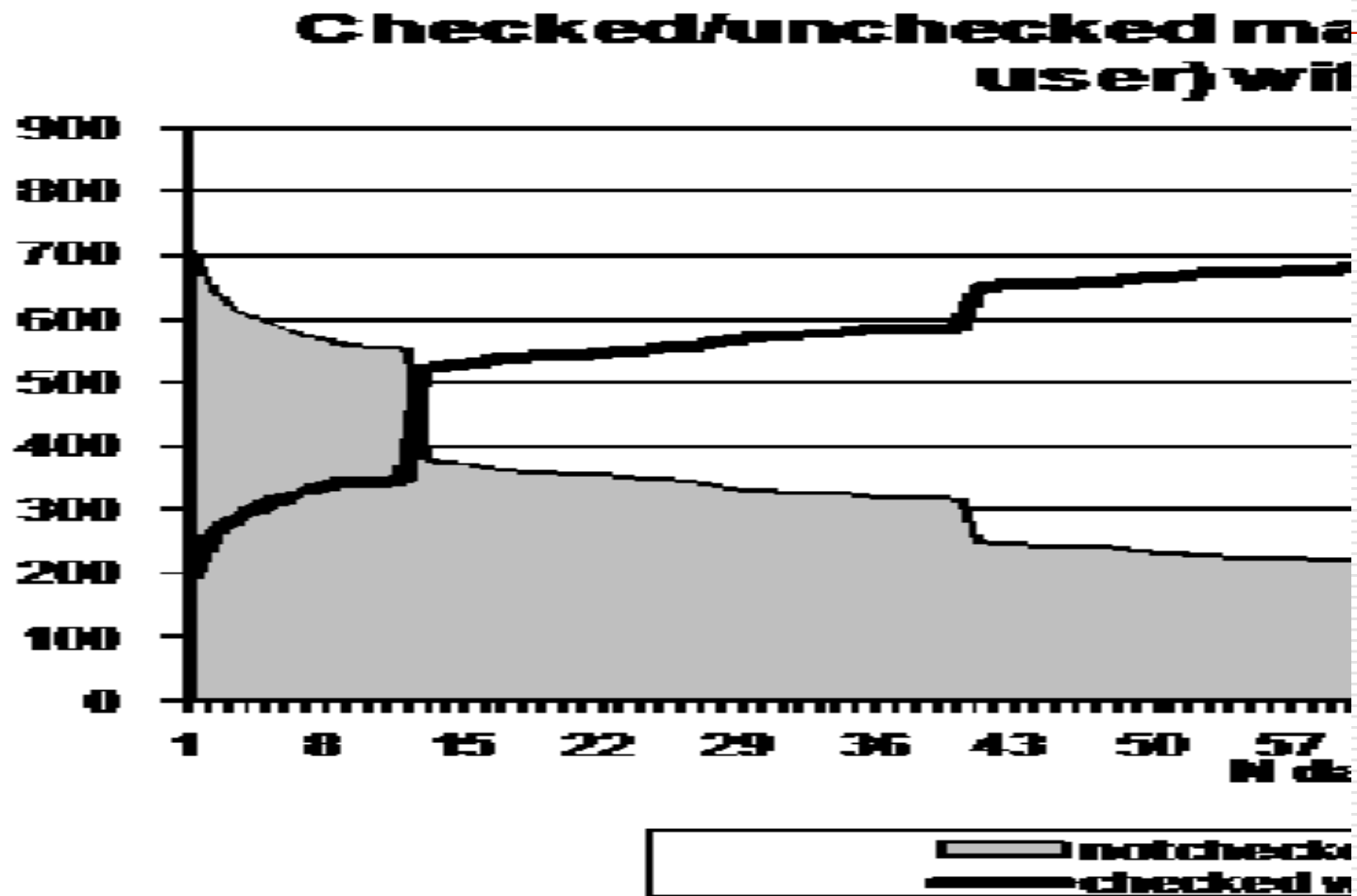
□ Die Besonderheit dieses Verfahren liegt darin, dass (nahezu) alle Elemente der avisierten Grundgesamtheit bekannt und online erreichbar sind. Dies ist bspw. in einem Unternehmen oder einer Universität der Fall, wenn eine Liste mit allen E-Mail-Adressen vorliegt. Aus dieser vollständigen Liste kann dann nach einem Zufallsverfahren eine Stichprobe ausgewählt und online kontaktiert werden.

E-Mail Survey 1998, Universität Mannheim

Number of checked/unchecked mailboxes of 9328 users within 60 days



Sample: 900



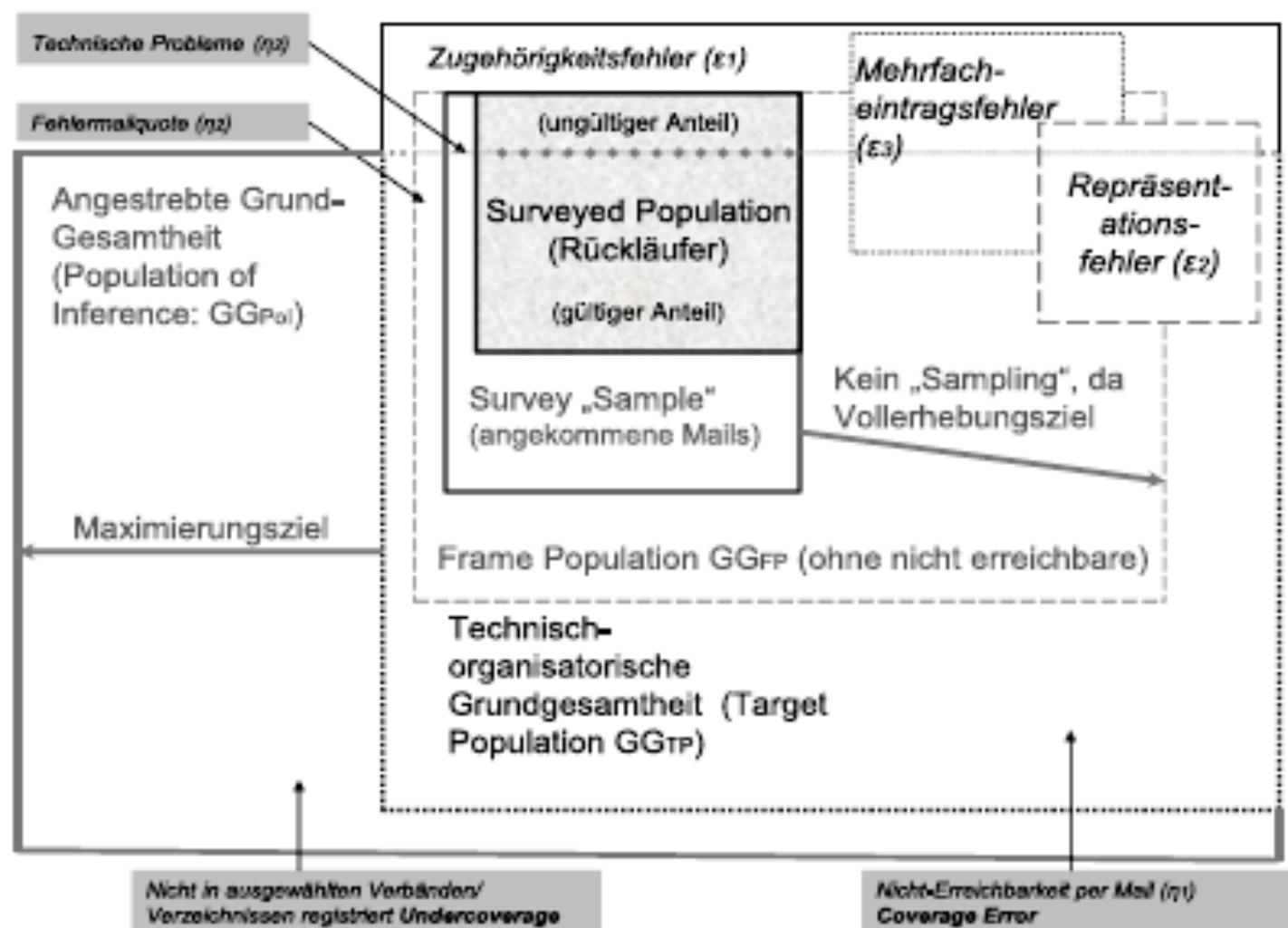
„Zukunft des Journalismus“, Universität Leipzig

- Datenerhebung Sommer 2005
- Online-Befragung von Journalisten
- Einladung per E-Mail; einschlägige Verbandslisten (DJV, Landesverbände, Pressekonferenz Berlin ...)

- Def. Journalist: mind. eine journalistische Stellung und mind. 2 Stunden pro Tag journalistisch tätig (somit auch max. Anzahl freier Journalisten erfasst)
- N = 3.766 Journalisten per Def.

ABBILDUNG 1

Grundgesamtheit – Auswahlgesamtheit – Stichprobe



Quelle: Studie ZdJ

Fehlerschätzungen

- (2) $ggFP = 51.297$, bei $\varepsilon_1, \varepsilon_2, \varepsilon_3 = 0$
- (3) $ggFP(O) = 51.297 \pm (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ bei ε_1 und $\varepsilon_3 = 0$
- (4) $ggFP(U) = 51.297 \pm (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ bei $\varepsilon_2 = 0$
- (5) $\varepsilon_{max} = 47.552$, bei ε_2 und $\varepsilon_3 = 0$
- (6) $\varepsilon_{max} = 89.703$, bei ε_1 und $\varepsilon_3 = 0$
- (7) $\varepsilon_{max} = 34.218$, bei ε_1 und $\varepsilon_2 = 0$

ε_1 Overcoverage- bzw. Zugehörigkeitsfehler; ε_2 Repräsentationsfehler; ε_3 Mehrfacheintragsfehler

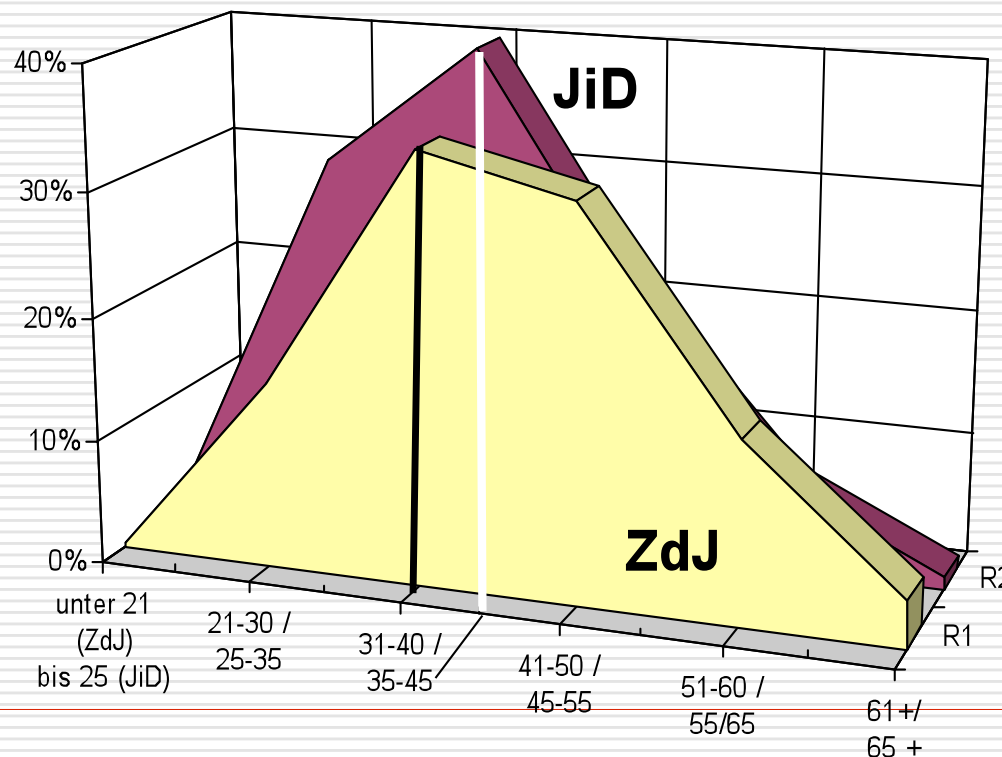
- Wie groß können die drei ε -Fehler maximal sein?
-

Mögliche Fehlergrößen und plausible Gründe

Fehlerart	angenommene Größe	Grund; plausible Annahme	Auswirkungen auf unbereinigte Frame Population (GG_{FP})
Undercoverage	30.000	Differenz von angenommener Target Population zur Frame Population; falls die mittlere Größe der Target Population mit 80.000 angenommen wird	$51.297 + 30.000$
ϵ_1 Overcoverage bzw. Zuehörigkeitsfehler	7.576	15 Prozent der Frame Population; DJV weist in seinen Mitgliederbeständen rund 15 Prozent PR-/Öffentlichkeitsarbeiter aus (ohne Freie, da Medienbereich unbekannt)	$51.297 - 7.576$
ϵ_2 Repräsentationsfehler	89.703	Obergrenze der Population of Inference: 141.000 Personen (maximaler, aber nicht plausibler Fehler)	$51.297 + 89.703$
ϵ_3 Mehrfachversandfehler	5.000	Mehrfacheinträge, Mehrfachmitgliedschaften	$51.297 - 5.000$
γ unbekannter neutraler Ausfall: Nicht Abfragen des Postfachs	15.000	Etwa ein Drittel der Frame Population; nach früheren Untersuchungen (vgl. Welker 2001)	$51.297 - 15.000$
η Fehlermails	3.749	8,04 Prozent Ermittlung der durchschnittlichen Fehlermailquote der Versandwege ohne Angabe über vorhandene Fehlermailquote	$51.297 - 3.749$

Soziodemographie: Vergleich von Zukunft d. Journ./Journ. in Dtl.

- Geschlecht: 33% weiblich (JiD: 37%)
- Bildung: 67% Hochschulabschluss (JiD: 66%)
- Freie Mitarbeiter: 33% (JiD: 25% → Diff.!)
- Alter:



Alterskategorien jeweils um 5 Jahre verschoben; Peak jeweils um die 40 Jahre

[zufallsbasiert Fortsetzung]

5. „Befragungen mit einem vorrekrutierten Panel von Internet-Nutzern“:

□ mit einer vorgeschalteten zufallsbasierten Auswahl der Panelteilnehmer bspw. durch das Intercept-Verfahren oder telefonisch durch das RDD-Verfahren. Die Schwierigkeit: geringe Teilnahmebereitschaft: 2 bis 15 Prozent einer offline rekrutierten Ausgangsstichprobe (Couper & Coutts 2006; Bandilla et al. 2009).

6. „Befragungen mit einem vorrekrutierten Panel aus der Gesamtbevölkerung“:

□ Hoher finanzieller und technischer Aufwand. Die Stichprobe wird zunächst offline rekrutiert (z.B. Face-to-Face oder RDD). Personen mit Teilnahmebereitschaft aber Internet-Zugang bekommen anschließend die notwendige technische Infrastruktur (Computer oder TV-Set-Top-Box mit Internet-Anbindung) gratis zur Verfügung gestellt. Damit wird in dieser Stichprobe auch die Gruppe der Offliner berücksichtigt, welche bei allen anderen vorgestellten Verfahren per se systematisch ausgeschlossen ist. In Deutschland wurde ein solches Panel beispielsweise von forsa in Form des forsa-omninet-Panels aufgebaut (Forsa 2009).

[zufallsbasiert Fortsetzung]

7. „Web-Option bei Mixed-Mode-Befragungen“:

- Bei diesem Typ stellt die Online-Befragung lediglich einen von verschiedenen Modi dar, welcher den Befragten für die Teilnahme angeboten wird.
- Die Stichprobenziehung erfolgt dabei in der Regel nicht online, sondern beispielsweise telefonisch mit Hilfe von Random Digital Dialing (RDD) oder anhand des ADM-Stichprobensystems (ADM 2009). Damit hängt auch die Art und Qualität der Stichprobe vom letztlich eingesetzten Auswahlverfahren ab. Prinzipiell kommen hier alle bekannten Auswahlverfahren in Frage.
- Ein solches Mixed-Mode-Verfahren kann jedoch nur bei solchen Untersuchungen eingesetzt werden, die keinen Einfluss des Befragungsmodus erwarten lassen. Sonst wären die Ergebnisse von zwei Teilnehmern, die sich für verschiedene Modi entschieden haben, nicht vergleichbar.

Mixed Mode

Steffen Kolb: Journalisten - eine schwer zu befragende Spezies,
Universität Hamburg 2002

Teilnahmebereitschaft

■ telefonisch

- gute Teilnahmebereitschaft
- kein Interviewabbruch

■ schriftlich

- durchschnittliche bis schlechte Teilnahmebereitschaft

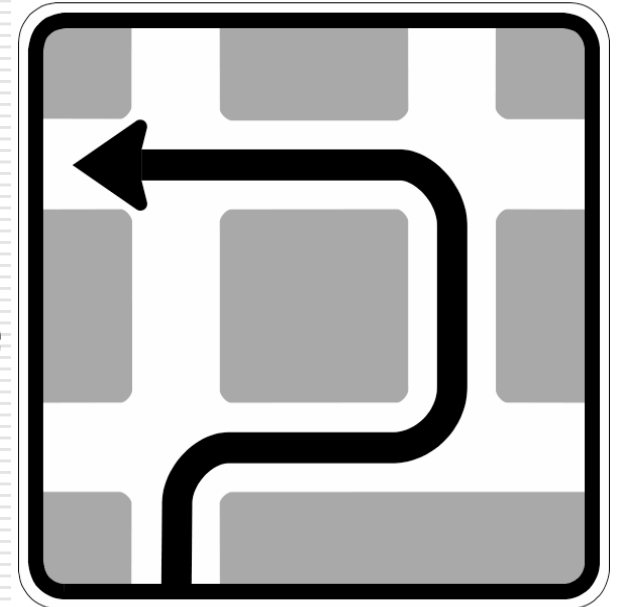
■ E-Mail

- schlechte Teilnahmebereitschaft
- technische Probleme

2002 -> 2010?

c. Schwierigkeiten

- ❑ Allg. Sozialforschung: Bei Grundgesamtheiten, die „Offliner“ beinhalten, steht eine Online-Stichprobe vor Problemen. Onliner und Offliner unterscheiden sich noch z.B. in den Merkmalen Alter, Bildung und Einkommen (vgl. Initiative D21 & TNS 2009).
- ❑ Schwierig sind demnach Grundgesamtheiten, die Offliner enthalten.
- ❑ Insofern ist auch die Fragestellung der Untersuchung wichtig

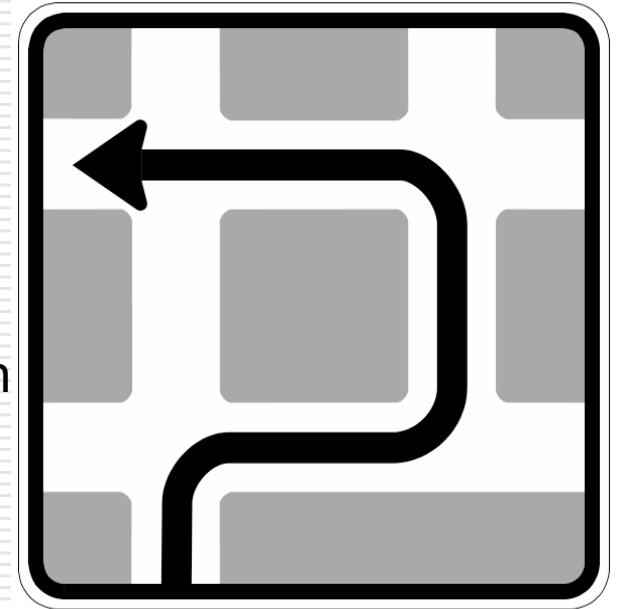


c. Weitere Schwierigkeiten

„Die Stichprobenziehung ist (nicht nur) bei Online-Befragungen ein sehr komplexer Vorgang.

Es ist praktisch unmöglich, eine echte Zufallsstichprobe zu erzielen, da zu Coverage-Problemen Ausfälle, fehlende Werte und Abbrüche hinzukommen, so dass die Stichprobenqualität bei vielen Online-Befragungen zweifelhaft ist und zumindest auf die Inferenzstatistik verzichtet werden sollte.

Dies gilt jedoch nicht generell, sondern muss jeweils auf Basis einer konkreten Stichprobe für eine spezifische Forschungsfrage und Zielpopulation entschieden werden.“ (Baur & Florian 2009: 126)



Einige Lösungs-/ Verbesserungsmöglichkeiten;

ex-ante und ex-post

- ❑ Auswahlgesamtheit und Fragestellung prüfen:
Benötige ich Aussagen über die Gesamtbevölkerung oder reichen bspw. die 14-29 Jährigen?
- ❑ Müssen Offliner in meiner Grundgesamtheit sein, oder beziehe ich mich nur auf Onliner?
- ❑ Propensity-Gewichtung: Feststellung der Neigung, im Online-Panel zu sein durch parallele telefonische/mündl. Befragung und entsprechende Gewichtung; umstritten
- ❑ Vergleich und Abgleich der Ergebnisse der Online-Untersuchung mit vergleichbaren Untersuchungen

Lösungs-/ Verbesserungsmöglichkeiten

- Kontrolle eines Mehrfachversandfehlers bei unterschiedlichen Datenquellen durch Abfrage der Verbandszugehörigkeit bzw. des Eintrags im Branchenbuch,
- Kontrolle der neutralen Ausfälle (insb. die Frage, zu welchem Anteil Postfächer abgefragt werden und ob sich daraus systematische Fehler ergeben können) durch Stichproben in Zusammenarbeit mit Verbänden und Datenbankbetreibern,
- Kontrolle und Abschätzung des Overcoverage durch Stichproben bei den Verbänden,
- Falls zeitlich und finanziell möglich: Flankierende Telefonbefragung mit kleinerer Zufallsstichprobe für Vergleichsdaten,
- Falls Datenschutz eine untergeordnete Rolle spielt: Individualisierung einer jeden Einladungsmail mittels Linkcode (DASCHMANN/HARTMANN 2005: 259; VIETH 2003: 107),

Lösungs-/ Verbesserungsmöglichkeiten

- Kontrolle aller methodischen Parameter „so gut es geht“
 - Over- und Undercoverage durch begleitende Stichproben bspw. bei Verbänden
 - Mehrfachversandfehler bei E-Mail Befragungen
 - Neutrale Ausfälle (Ausmaß der Abfrage von E-Mailpostfächern)
 - Non-response durch Linkcodes u.ä.
 - Nicht-Stichprobenbezogen
 - Meta- und Paradata
-

B. Online-Inhaltsanalyse

- Auch für die Analyse von Online-Inhalten kann in der Regel nur eine Stichprobe (einer größeren Grundgesamtheit) untersucht werden.
- Auf Grund der Flüchtigkeit und Dynamik der Inhalte verändert sich die Grundgesamtheit einer Stichprobe von Online-Inhalten in der Regel kontinuierlich. Bei der Definition der Grundgesamtheit ist daher entweder ein statischer ‚Schnappschuss‘ zu definieren oder die Zeitdimension wird explizit als eine Dimension der Grundgesamtheit betrachtet und stellt beispielsweise ein Schichtungskriterium der Stichprobe dar. In beiden Fällen ist darauf zu achten, eine trennscharfe Abgrenzung zwischen zwei Versionen eines Angebots zu definieren

SPIEGEL ONLINE

Frankfurter Allgemeine
FAZ.NET

 heise online

WELT  **ONLINE**

Stefan Niggemeier

Bild.de

ZEIT  **ONLINE**

a. Nicht-zufallsbasiert

1. mehrstufige und nicht zufällige Auswahl

- Wie stellen sich die Nutzer des deutschen studiVZ in ihren Profilen dar?
 - Grundgesamtheit ist klar definiert, allerdings nicht komplett erreichbar. Die einzige Möglichkeit, eine Liste mit allen Profilen (die Grundgesamtheit) zu erzeugen, ist die Nutzung der Suchfunktion von studivz. Dennoch reicht die Trefferliste – die nur bis zu 300 Profile auflistet – nicht aus, um alle Profile aufzuführen.
 - Durch Suchfunktion nicht-zufällige Ziehung aus den Profilen.
-

[Fortsetzung Nicht-zufällige Auswahl]

2. systematische Auswahl; Konzentrationsprinzip

- ❑ Welche Themen werden in der deutschen Blogosphäre behandelt?
- ❑ Die Grundgesamtheit dieser Fragestellung – die >deutsche Blogosphäre< – ist sowohl definitorisch als auch physisch schwer zu fassen.
- ❑ Rückgriff auf Rankings bspw. auf die Deutschen Blogcharts.
- ❑ Nicht-repräsentative Stichprobe

3. Auswahl bei unbekannter Grundgesamtheit

- ❑ Wie wird im Internet über die Universität Wien berichtet und geschrieben?
 - ❑ Unbekannte, nicht erreichbare Grundgesamtheit
 - ❑ Beschaffung möglichst umfassender Informationen: Kombination verschiedener Such- und Metasuchmaschinen
-

b. zufallsbasiert

4. echte Zufallsauswahl

- Wie wurde die Bundeskanzlerin Angela Merkel 2007 im Forum auf *Spiegel-Online* beurteilt?
- Hier bietet sich eine echte, einfache Zufallsauswahl aus allen Beiträgen an. Dies kann z. B. durch die Auswahl jedes n-ten Beitrages in einer chronologisch der zufällig sortierten Liste aller Beiträge erfolgen.

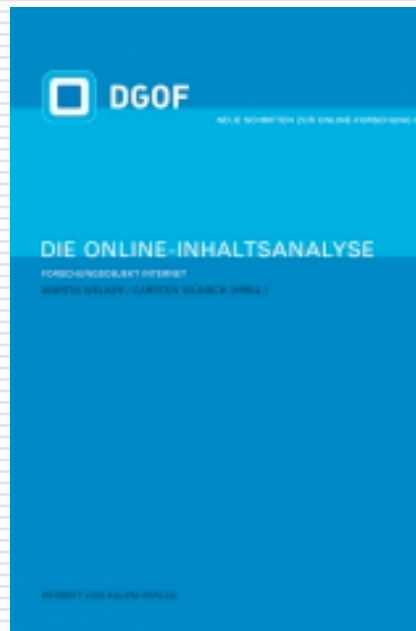
5. mehrstufige, zufällige Auswahl

- Wie berichteten *faz.Net*, *Spiegel.de* und *Handelssblatt.de* 2008 über die Finanz- und Wirtschaftskrise?
 - Auch in diesem Beispiel ist die Grundgesamtheit prinzipiell bekannt und erreichbar: alle Beiträge über die Finanz- und Wirtschaftskrise der drei Online-Medien im Jahr 2008.
 - Mit einer Liste einschlägiger Begriffe und der Suchfunktion der Archive dieser Online-Angebote kann ein Zugriff auf relevante Beiträge erfolgen.
-

C. Kontroversen/Fragen

- Multivariate Verfahren nur bei „echten“ Zufallsstichproben?
- Bald 100 Prozent Online-Abdeckung und damit Bevölkerungsstichproben möglich?
- Verfall der etablierten Stichprobenverfahren?

Literaturauswahl



Danke für Ihre Aufmerksamkeit!

Prof. Dr. Martin Welker

Macromedia Hochschule für
Medien und Kommunikation (MHMK)
München
m.welker@macromedia.de

DGOF e.V.
Köln/Hürth
m.welker@dgof.de

www.martin-welker.de

