

Dr. habil. Rüdiger Jacob
Methoden und Techniken der empirischen Sozialforschung
Vorlesung mit Diskussion

7. Auswahlverfahren

Grundgesamtheit und Stichprobe, Repräsentativität, einfache und komplexe Zufallsauswahlen, willkürliche und bewusste Auswahlen

Grundgesamtheit: (Population) Zielgruppe einer Untersuchung. Wichtig ist die definitorische Abgrenzung der Grundgesamtheit, um eine nachvollziehbare Auswahl treffen und exakt angeben zu können, für wen die Untersuchungsergebnisse Gültigkeit beanspruchen können.

Abgrenzungskriterien:

- **Untersuchungsziel**
- **Zugänglichkeit**
- **Kosten**

Grundsätzlich gilt:

Ergebnisse von Stichprobenerhebungen können unter Verwendung mathematisch-statistischer Verfahren nur dann auf eine Grundgesamtheit verallgemeinert werden, wenn sie das Ergebnis einer Zufallsauswahl sind. Dieser induktive Schluss von einer numerisch kleineren auf eine numerisch größere Gesamtheit wird als Repräsentationsschluss bezeichnet.

Stichprobenergebnisse sind keine Punktschätzer für Werte der Grundgesamtheit (Bsp.: Stimmanteile bei Wahlen), sondern stehen für einen

Wertebereich (Intervallschätzer), dessen Ober- und Untergrenzen bestimmte, berechenbare Wahrscheinlichkeiten haben (Vertrauensintervalle).

Mit zunehmender Größe der Stichprobe (dem N) verringert sich die Breite der Vertrauensintervalle

Repräsentativität:

Eigenschaft einer Zufallsstichprobe. Innerhalb angegebener Fehlergrenzen ist es dabei möglich, von der Verteilung bestimmter Merkmale und Zusammenhänge in der Stichprobe auf die Verteilung dieser Merkmale und Zusammenhänge in der eigentlich interessierenden Grundgesamtheit zu schließen.

Eine Stichprobe kann nur dann korrekt als "repräsentativ" bezeichnet werden, wenn sie auf einer Zufallsauswahl basiert. Eine repräsentative Auswahl ist kein exaktes verkleinertes Abbild einer Grundgesamtheit, sondern nur ein näherungsweise genaues Modell.

Die Größe der Stichprobe ist außerdem für Subgruppen-Analysen wichtig.

Richtwerte für Stichprobengrößen (Netto-N)

2.000 bis 3.000 Personen:

Repräsentative allgemeine Bevölkerungsumfragen, bei denen auch differenziertere Subgruppen-Analysen möglich sein sollen.

1.000 Personen:

Repräsentative allgemeine Bevölkerungsumfragen mit vorwiegend deskriptiver Zielsetzung.

500 Personen:

Repräsentative Umfragen in spezifischen homogenen Gruppen (z. B. Studierende einer Hochschule oder Angehörige bestimmter Berufe).

100 bis 200 Personen:

Repräsentative Umfragen in spezifischen homogenen Gruppen mit vorwiegend deskriptiver Zielsetzung.

Zufallsauswahlen (Random-Samples)

Der statistische Zufall entscheidet darüber, welche Elemente der Grundgesamtheit in die Stichprobe gelangen.

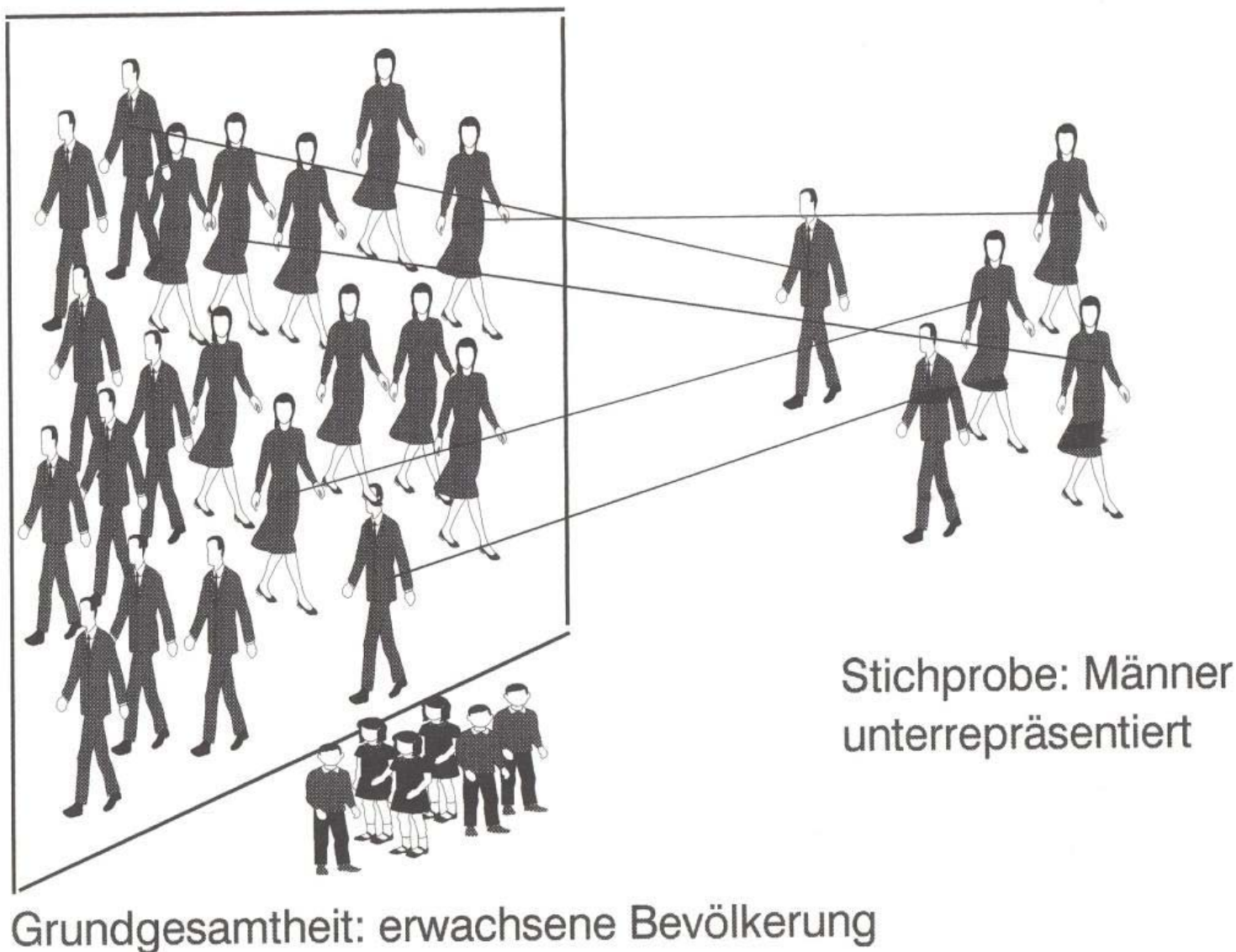
Kriterien:

- Die Grundgesamtheit muss bekannt und exakt definiert sein
- Die Grundgesamtheit muss physisch oder symbolisch präsent und manipulierbar sein (sie muss sich durchmischen lassen, jedes Element muss entnehmbar sein)
- Jedes Element darf nur einmal in der Grundgesamtheit vertreten sein
- Die Auswahl muss so erfolgen, dass jedes Element die gleiche oder eine berechenbare Auswahlchance hat, in die Stichprobe zu gelangen.

Einfache Zufallsauswahlen

Jedes Element der Grundgesamtheit hat die gleiche Auswahlchance

- Urnenauswahl
 - Karteiauswahl (Listenauswahl)
Stichprobe: 1.000 Personen
Startzahl: Nummer 326
Schrittweite: 45
Beginnend bei Eintragung 326 wird solange jede 45. Eintragung gezogen, bis man insgesamt 1.000 Namen zusammen hat.
 - Buchstabenauswahl
 - Geburtstagsauswahl
- ➔ Die Definition der Grundgesamtheit hängt nicht nur von theoretischen Überlegungen ab, sondern auch von der Art ihrer symbolischen Präsenz.



Telefonstichproben

- eingeschränkte Grundgesamtheit: Personen mit Telefon
- Anschlussdichte und Eintragungsdichte
- Doppeleintragungen
- Geschäftsanschlüsse
- Aktualisierungsintervalle
- ISDN, Anrufbeantworter
- Haushaltsstichprobe!!

Ziehungsverfahren: Listenauswahl, RDD, RLD, Gabler-Häder-Design

Problem: Ziffernlängen von 8 bis 12

Hitraten bei generierten Nummern:

Zahl der Treffer: Verhältnis der existierenden und der insgesamt

gewählten Nummern

RDD	unter 1%
RLD	72%
Gabler-Häder	66%

Gabler-Häder-Design

Einteilung von Nummern in 100er Blocks:

Stamm einer Telefonnummer, der nach dem Abschneiden der beiden letzten Ziffern verbleibt

0611 6036|14 —————> Block von: 0611 6036 00

bis: 0611 6036 99

Aus Blöcken mit wenigstens einer eingetragenen Nummer werden zufällig Nummern ausgewählt.

Verhältnis generierter Nummern zu Netto-N: 5 zu 1

Komplexe Zufallsauswahlen

Bei komplexen Zufallsauswahlen hat jedes Element der Grundgesamtheit nicht mehr die gleiche, sondern nur noch eine berechenbare Auswahlchance

Geschichtete Auswahl

Die Grundgesamtheit wird in verschiedene Gruppen oder Schichten unterteilt. Danach werden per einfacher Zufallsauswahl Elemente aus jeder Schicht gezogen.

Proportionale und disproportionale Schichtung

In der Praxis dominieren disproportional geschichtete Auswahlen → kostengünstige Methode, statistisch gut auswertbare Teilstichproben von Gruppen zu realisieren, die in der Grundgesamtheit nur schwach vertreten sind

Durch Schichtung lässt sich der Gesamtstichprobenumfang minimieren, was wiederum die Erhebungskosten senkt →

Erwünscht: Schichtungseffekt

GRHGRHGR H GRHSGR
HGRHGRHGRHSGRHR R
RGRHHRR SRG S GGG
GRRRRRRRHGHHSRR
RRRHGG HRRRRSGG
GHHRRRSGHRS

Schichtung



Gymnasium: 200
Realschule: 250
Hauptschule: 170
Sonderschule: 50



Gymnasium: 20
Realschule: 20
Hauptschule: 20
Sonderschule: 20

Grundgesamtheit:
Lehrer der Sekundarstufe 2

Disproportional
geschichtete Stichprobe

Bei disproportionaler Schichtung:
unterschiedliche Auswahlchancen für die Mitglieder der
Grundgesamtheit →
getrennte Datenanalyse für jede Schicht
Gewichtung

Klumpenauswahl
Einfache Zufallsauswahlen, die aber nicht auf einzelne Elemente
sondern auf bereits zusammengefasste Elemente, z. B. komplette
Wohnanlagen oder Stadtteile, zugreift

Problem: Klumpeneffekt
Klumpen sollen intern möglichst heterogen sein, um die Bandbreite
der Merkmalsausprägungen der Grundgesamtheit auch adäquat
erfassen zu können →

In der Realität sind Klumpen sozialstrukturell häufig sehr homogen.
Beispiel Stadtteile/Wohngebiete/Wohneinheiten: Segregation
Beispiel Vorlesungen: Fachspezifische Selektion

Mehrstufige Auswahlverfahren

Kombination mehrerer Zufallsauswahlen

Ermittlung der Zielpersonen auf der letzten Stufe

**Dabei ist die jeweils entstehende
Zufallsstichprobe Auswahlgrundlage für die
nächste Auswahlstufe**

Anwendungsfall: unbekannte Grundgesamtheiten

**ADM-Master-Sample oder ADM-Design für
bundesweite Bevölkerungsumfragen**

GRUNDGESAMTHEIT	Erwachsene deutsche Wohnbevölkerung in Privathaushalten
1. AUSWAHLSTUFE: Primäreinheiten	<i>Primäreinheiten:</i> Stimmbezirke für die Wahl zum Deutschen Bundestag. Auswahl einzelner Stimmbezirke durch einfache Zufallsauswahl. Auswahl von Startadressen innerhalb der ausgewählten Stimmbezirke durch einfache Zufallsauswahl.
2. AUSWAHLSTUFE: Haushalte	<i>Haushalte:</i> Auflistung von 23 Haushalten ab Startadresse durch Random-Route. Auswahl von 8 Haushalten aus Adressprotokoll durch einfache Zufallsauswahl.
3. AUSWAHLSTUFE: Personen	<i>Zielpersonen:</i> Auswahl von Zielpersonen in Haushalten durch Schwedenschlüssel.

Listen Sie zunächst fortlaufend 23 Privathaushalte auf. Sie beginnen in der vorgegebenen Straße (Startstr.) bei der vorgegebenen Hausnummer. Achten Sie darauf, die Hausnummern wie angegeben zu begehen (gerade oder ungerade). Sollte die Anzahl der Haushalte nicht ausreichen, listen Sie auf der gegenüberliegenden Straßenseite weiter auf. Fehlen Ihnen dann immer noch Haushalte, listen Sie die fehlenden Adressen in der ersten Querstraße rechts auf.

Bei kleineren Orten ist es möglich, daß wir Ihnen keine Straße vorgegeben haben, sondern einen Buchstaben. Suchen Sie sich in diesem Fall eine Startstraße, die mit dem angegebenen Buchstaben beginnt. Ist dies nicht möglich, wählen Sie eine Straße mit dem im Alphabet folgenden Buchstaben.

Es gibt auch Ortschaften, in denen es keine Straßennamen gibt, beginnen Sie Ihre Auflistung dann bei einem öffentlichen Gebäude, z. B. bei der Post oder bei der Kirche.

Übertragen Sie die Adressen, die mit einem Kreuz versehen sind auf die beigelegte große weiße Adressenliste; in diesen Haushalten sollen Sie die Interviews durchführen. Leben zwei oder mehr Personen der Zielgruppe in einem Haushalt, muß die Befragungsperson nach dem auf der Adressenliste angegebenen Auswahlsschlüssel bestimmt werden.

WICHTIG !!! ⇒ DIE ADRESSENLISTE IST IN GUT LESERLICHER DRUCKSCHRIFT AUSZUFÜLLEN !

LFD-NR	FAMILIENNAME	STRASSE	H-NR	ETAGE	TÜR
01 X	> <				
02					
03					
04 X	> <				
05					
06					
07 X	> <				
08					
09					
10 X	> <				
11					
12					
13 X	> <				
14					
15					
16 X	> <				
17					
18					
19 X	> <				
20					
21					
22 X	> <				
23					

Schwedenschlüssel

Ermitteln Sie die Befragungsperson, indem Sie alle Personen der Zielgruppe im Haushalt dem Alter nach geordnet - älteste Person zuerst - auflisten.

Befragungsperson lt. Auswahlziffern bestimmen.
(siehe Anleitung)

ZIELPERSON(EN) IM HAUSHALT

ZAHL D. ZIELPERSONEN

Alter m/w

Alter m/w

Auswahlziffern

1. <u>75</u>	<u>W</u>	6. <u> </u>	<u> </u>
2. <u>42</u>	<u>W</u>	7. <u> </u>	<u> </u>
3. <u>39</u>	<u>M</u>	8. <u> </u>	<u> </u>
4. <u>19</u>	<u>W</u>	9. <u> </u>	<u> </u>
5. <u>18</u>	<u>M</u>		

Haushaltsgröße (ZP)

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

1	2	1	3	5	4	2	8	6
---	---	---	---	---	---	---	---	---

Auswahlziffer

Probleme

- Die Zahl der Privathaushalte pro Stimmbezirk ist nicht genau bekannt, sondern muss geschätzt werden
- Trotz umfangreicher Vorgaben zur Adressenaufstellung haben Interviewer einen schwer kontrollierbaren und nicht zu beseitigenden Entscheidungsspielraum
- Dies gilt auch für die Realisierung von Interviews, wenn z. B. statt der zur Befragung eigentlich vorgesehenen Person eine eher erreichbare oder auskunftswilligere Person befragt wird

AIIBUS 1994: zweistufige Zufallsauswahl

- 1. Stufe: Auswahl von Gemeinden**
- 2. Stufe: Personenstichprobe aus Karteien der
 Meldeämter**

Vorteile:

- Interviewer haben keinen Einfluss auf die Stichprobenbildung und die Auswahl der Zielpersonen**
- Ausschöpfungsquote kann durch gezielte Anschreiben optimiert werden**
- Ausfälle sind besser kontrollierbar, da über alle Personen der Bruttostichprobe zumindest einige Informationen aus den Einwohnermelderegistern zur Verfügung stehen**

Nachteile:

- hoher Zeitaufwand, aufwendige Feldarbeit**
- hohe Kosten durch Zeitaufwand und Gebühren für die Ziehung von Adressen in den Einwohnermeldeämtern**

Grundgesamtheiten bei Bevölkerungsumfragen:

Nie das gesamte „Volk“ eines Landes oder die innerhalb bestimmter administrativer oder nationaler Grenzen lebende Bevölkerung.

Aus verschiedenen Gründen werden bei der Definition der Grundgesamtheit bestimmte Gruppen ausgeschlossen, z. B.:

- **Ausländer**
- **Minderjährige**
- **Anstaltsbevölkerung**
- **Deutsche im Ausland**
- **Zivildienstleistende und Wehrpflichtige**

Nicht-zufällige Auswahlen

Willkürliche Auswahl

Auswahl aufs Geradewohl: Für wissenschaftliche Untersuchungen völlig ungeeignet:

- **Keine Definition der Grundgesamtheit**
- **Kein Auswahlplan**
- **Keine Interviewerkontrolle**

„Vorteil“: Kostengünstig und schnell zu realisieren

Klassisches Beispiel: Passantenbefragungen

Bewusste Auswahlen

Auswahlverfahren, bei denen subjektive Kalküle die Auswahl der Zielpersonen steuern. Generalisierungen auf der Basis mathematisch-statistischer Modelle sind bei bewussten Auswahlen nicht möglich, weil kein entsprechend fundiertes Auswahlmodell existiert.

Auswahl typischer Fälle

Fälle, die man hinsichtlich bestimmter Merkmale als besonders charakteristisch für die Grundgesamtheit ansieht.

Anwendungsfall: Expertenbefragungen

Schneeballauswahl

unbekannte und/oder unzugängliche Grundgesamtheit

Quotenauswahl

Standardverfahren der kommerziellen Markt- und Meinungsforschung

Konstruktion von Personenstichproben, so dass die Verteilung bestimmter Merkmale in der Stichprobe der Verteilung dieser Merkmale in der Grundgesamtheit exakt entspricht.

Voraussetzung:

Kenntnisse über die Verteilung bestimmter Merkmale in der Grundgesamtheit, die man der amtlichen Statistik oder einer Verwaltungsstatistik (etwa der Statistik des Studierendensekretariates) entnehmen kann.

Auswahlgrundlage:

Quotenplan mit Quotenmerkmalen

Bei allgemeinen Bevölkerungsumfragen in der Regel das Alter (als klassiertes Merkmal), das Geschlecht, die Berufsgruppenzugehörigkeit und der formale Bildungsabschluss

Einfache Quotenauswahl

unabhängige Quoten → Randverteilungen der Quotenmerkmale müssen den Verteilungen in der Grundgesamtheit entsprechen

→ unterschiedliche Stichproben möglich

Kombinierte Quotenauswahl

Sowohl die Randverteilungen der Quotenmerkmale als auch alle Kombinationen dieser Merkmale müssen den Werten der Grundgesamtheit entsprechen

→ Nur eine Stichprobe erfüllt diese Bedingung

	BWL	VWL	SOZ.	
GRUND-STUDIUM	931 40%	304 13%	121 5%	1356 58%
HAUPT-STUDIUM	718 30%	168 7%	116 5%	1002 42%
	1649 70%	472 20%	237 10%	<u>2358</u> 100%

unabhängige Quote

	BWL	VWL	SOZ.	
GRUND-STUDIUM	58	-	-	58
HAUPT-STUDIUM	12	20	10	42
	70	20	10	100

kombinierte Quote

	BWL	VWL	SOZ.	
GRUND-STUDIUM	40	13	5	58
HAUPT-STUDIUM	30	7	5	42
	70	20	10	100

Quotenvorgaben:

Zahl der zu befragenden Personen mit bestimmten Merkmalkombinationen

Vorteil:

Zeit- und Kostenaufwand niedriger als bei Random-Stichproben

Nachteile:

- **Interviewer wählen Zielpersonen aus**
- **unzureichende Interviewerkontrollen**
- **Ausfälle und Verweigerungen werden nicht dokumentiert**

→ Repräsentationsschlüsse sind nicht möglich

Werden Ungleichheiten als ungerecht empfunden?

**Universität Trier sucht
Personen aus Ost- und
Westdeutschland zur
Durchführung einer Studie.**

Trier. Mit Unterstützung der Deutschen Forschungsgemeinschaft wird an der Universität Trier eine sozialwissenschaftliche Studie zur Lebensqualität in Ost- und Westdeutschland durchgeführt. Es soll untersucht werden, in welchen Lebensbereichen von den Bürgern fünf Jahre nach der Wiedervereinigung noch Unterschiede zwischen den alten und den neuen Bundesländern wahrgenommen werden, und wie diese erklärt werden. Die einzelnen Fragen beziehen

sich auf die fünf Themenbereiche: Arbeits- und Berufsleben, finanzielle Situation und die Wohlbildung, zwischenmenschliche Beziehungen, Qualität der Wohnungen und Städte.

Für die Untersuchung, die mittels umfangreicher, aber einfach zu beantwortender Fragebögen erfolgt, werden freiwillige Teilnehmerinnen und Teilnehmer gesucht. Die Befragung wird völlig anonym auf dem Postwege durchgeführt.

Wer sich beteiligen möchte, sollte die den Forschern auf einer Postkarte kurz mitteilen. Anschrift: Forschungsprojekt GiP, Dr. Manfred Schmitt, Fachbereich I, Universität Trier, 54286 Trier. 2056

Befragung in Trier zur Attraktivität des ÖPNV

Untersuchungsfragen:

- **Nutzung des ÖPNV**
- **Alternativen**
- **andere Taktzeiten, Versorgungsdefizite**
- **Preisobergrenze für Einzelfahrscheine**

Zielgruppe: Trierer Bevölkerung

Design 1: Passantenbefragung auf dem Hauptmarkt und dem Viehmarkt

Grundgesamtheit: **Personen, die man dort antrifft**

Auswahlchance: ?

Struktur der Stichprobe: Abhängig vom Erhebungszeitpunkt
Bei Vormittagsbefragungen dürften Frauen zwischen ca. 30 und 65 Jahren dominieren

Bei Befragungen ab 20.00 dürfte das Geschlechterverhältnis ausgeglichener sein, dafür sind vermutlich Personen ab 60 unterrepräsentiert

Im Sommer sind mehr Touristen in der Stichprobe als im Winter

Design 2: Einfache Zufallsauswahl

Grundgesamtheit: In Trier gemeldete Personen ab 16 Jahren (60.000 Personen)

Stichprobe: Karteiauswahl aus der Datei des Einwohnermeldeamtes.

Auswahlchance: 1 zu 60.000

Struktur der Stichprobe entspricht bei korrekter Feldarbeit weitgehend der der Grundgesamtheit

Design 3: Mehrstufige Zufallsauswahl

**Grundgesamtheit: In Trier lebende Personen ab 16 Jahren
(ca. 100.000 Personen, GG nicht
gelistet)**

Stichprobe: ADM-Design

Auswahlchance: nach Design zu berechnen

**Struktur der Stichprobe entspricht bei korrekter Feldarbeit
weitgehend der der Grundgesamtheit**

Design 4: Telefonstichprobe

**Grundgesamtheit: In Trier lebende Personen ab 16 Jahren
(ca. 100.000 Personen, GG teilweise nicht gelistet)**

Stichprobe: **a: Listenauswahl aus Telefonbuch/Next
Birthday**
b: Gabler-Häder/Next Birthday

Auswahlchance: nach Design zu berechnen

**Struktur der Stichprobe entspricht bei korrekter Feldarbeit
weitgehend der der Grundgesamtheit**

versitäten.

Zusammen mit dem Bielefelder Emnid-Institut entwickelten Daniel und Hornbostel im Laufe des vergangenen Jahres drei Fragebögen für die Studenten in West- und Ostdeutschland sowie für die Professoren. Am Montag, dem 4. Januar 1993, schwärmten 438 Emnid-Mitarbeiter in 57 westdeutsche und 12 ostdeutsche Universitäten aus. Sie befragten genau 10 637 West- und 1191 Ost-Studenten. Parallel lief in der Bielefelder Emnid-Zentrale die telefonische Befragung von insgesamt 1185 Professoren an. Die letzten Interviews führte Emnid am 26. Februar.

Befragt wurden ausschließlich erfahrene Hauptfachstudenten im 5. bis 10. Semester, die an der derzeitig besuchten Universität mindestens ein Semester studiert hatten und somit kompetent urteilen konnten. In Westdeutschland konzentrierte sich die Umfrage –

ten wurde der SPIEGEL-Fragenkatalog vorgelegt. Die Interviewer wählten die zu Befragenden nach dem Zufallsprinzip aus. Festgelegt war nur, daß in jeder Fakultät mindestens an drei verschiedenen Standorten gefragt werden mußte – zum Beispiel vor einem Hörsaal, einem Labor und einer Bibliothek.

Für die SPIEGEL-Rangliste berücksichtigt wurden die Fragen 1, 4 bis 17 und 19 bis 21. Jede dieser 18 Fragen mußte mit Hilfe einer Punkteskala beantwortet werden, die dem klassischen Schulnotensystem entspricht – also die 1 für den Idealfall und die 6 für das negative Extrem.

Die Beurteilung eines einzelnen Fachbereichs entspricht dem Durchschnittswert aller von den 18 Studenten vergebenen Noten. Miteinander verglichen und in eine gemeinsame Rangliste aufgenommen wurden nur die Ergebnisse aller Fachbereiche desselben Fachs. So entstanden zunächst 19 Rang-

listen für die Fächer Maschinenbau, Humanmedizin, Wirtschaftswissenschaften, Betriebswirtschaftslehre (BWL), Volkswirtschaftslehre (VWL), Jura, Erziehungswissenschaften, Chemie, Physik, Informatik, Geschichte, Mathematik, Anglistik, Sozialwissenschaften, Soziologie, Politologie, Elektrotechnik, Germanistik und Biologie.

Nach den Regeln der Statistik konnten in jeder Fachrangliste drei Gruppen gebildet werden: Spitzengruppe, Mittelfeld und Schlußlichter. An den Gruppengrenzen kommt es zuweilen zu Überschneidungen. Fakultäten, die von den Studenten extrem unterschiedlich bewertet wurden, stehen im Mittelfeld, obwohl sie beispielsweise bessere Durchschnittsnoten aufweisen als einige Fachbereiche der Spitzengruppe. Genauso entschieden wurde im Grenzbe-
reich zu den Schlußlichtern: Nur Fakultäten mit einheitlich negativen Noten erscheinen dann in der Schlußgruppe.

„... bei einer Random-Auswahl scheinen alle subjektiven Einflüsse systematisch soweit wie möglich ausgeschaltet zu sein, so daß sich das Wahrscheinlichkeitsgesetz ungehindert auswirken kann. Trotz zahlreicher empirischer Bewährungsproben ist ein theoretischer Beweis bisher noch nicht erbracht - und läßt sich vielleicht auch nie erbringen -, daß bei richtig gehandhabter Quotenauswahl ... die Ergebnisse gleichfalls innerhalb entsprechender Fehlerspannen verallgemeinert werden dürfen. Lediglich die Wahrscheinlichkeit, daß eine solche Verallgemeinerung zulässig ist, kann durch systematische Experimente zunehmend untermauert werden. Das Institut für Demoskopie Allensbach hat während der letzten Jahrzehnte mehrmals Vergleichsexperimente durchgeführt, bei denen mit dem gleichen Fragebogen je ein Random- und ein Quotenquerschnitt zum selben Zeitpunkt befragt wurde. Jedesmal stimmen die Ergebnisse fast vollständig überein.“ (Noelle-Neumann und Petersen 1996, S. 264 f).

Was ist zu dieser Begründung für „repräsentative Outenauswahlen“ des IFD zu sagen?