

an objektivierbare Größen wie etwa das Wachstum der gesamtwirtschaftlichen Wertschöpfung und der Intensität des Warenaustausches zwischen den Mitgliedsländern der EG vorzusehen. Alle Vereinbarungen bei einer jetzt zu verwirklichenden Anschluß- oder begrenzten Regelkabotage, die einer solchen Flexibilität entgegenstehen, können nicht als zielkonformer Einstieg in die vom Europäischen Gerichtshof im Hinblick auf die Dienstleistungsfreiheit geforderte Kabotageregulation angesehen werden.

### Summary

In this report the Scientific Advisory Board of the Federal Ministry of Transport - Transportation Section - makes a statement on the future regulation of the cabotage within the EC. The abolition of the cabotage reservations is justified with the national economic disadvantages of such a barrier to market entry. The permission of the cabotage is a necessary consequence of a liberalization of the common European transport market.

In principle an unlimited regular cabotage is required. Since this cannot be accomplished in one step, transitional regulations are suggested. Two options are considered: the "Anschlußkabotage", which means that one transport within another country is permitted, subsequent to a transport across the border, and the limited regular cabotage. Both options are investigated in their possible forms and effects. The choice between the two options should be determined by the political suitability. The regulation, however, should be oriented from the beginning towards the goal of the unlimited liberalized service sector which is to be achieved later.

## Haushaltsbefragungen vom KONTIV-Typ: Hochrechnung - Gewichtung - Genauigkeitsbeurteilung\*) - Teil II -

VON HEINZ HAUTZINGER

### Inhaltsübersicht

#### Vorwort

1. Haushaltsbefragungen vom KONTIV-Typ
  - 1.1 Grundlagen
  - 1.2 Charakterisierung der Haushaltsbefragungen vom KONTIV-Typ
  - 1.3 Alternative Stichprobendesigns
  - 1.4 Grundprinzipien der Hochrechnung und Genauigkeitsbeurteilung
    - 1.4.1 Design „Unabhängige Teilstichproben“
    - 1.4.2 Design „Partielle Rotation“
2. Erhebungsfehler und ihre Berücksichtigung bei der Hochrechnung
  - 2.1 Fehlertypologie
    - 2.1.1 Stichprobenbedingte Fehler
    - 2.1.2 Verfahrensbedingte Fehler
  - 2.2 Verfahrensbedingte Fehler bei Haushaltsbefragungen vom KONTIV-Typ
    - 2.2.1 Nichtübereinstimmung von Zielkollektiv und Auswahlgesamtheit
    - 2.2.2 Eingeschränkte Objektivität, Reliabilität und Validität des Meßverfahrens
    - 2.2.3 Nichterreichbarkeit eines Teils der Stichprobenhaushalte
    - 2.2.4 Nichtteilnahme (Antwortverweigerung) eines Teils der Stichprobenhaushalte
    - 2.2.5 Angabefehler der Befragten

#### Anschrift des Verfassers:

Prof. Dr. Heinz Hautzinger  
Institut für Angewandte Verkehrs-  
und Tourismusforschung e. V.  
an der Fachhochschule Heilbronn  
Max-Planck-Str. 39  
7100 Heilbronn

\*) Die vorliegende Arbeit wurde wesentlich gefördert durch die intensiven und fruchtbaren Diskussionen, die mit Unterstützung durch das Bundesverkehrsministerium im Rahmen von zwei speziellen Fachgesprächen geführt worden sind. Teilnehmer an diesen Fachgesprächen waren die Herren Dr. J. Grevsmühl (Bundesverkehrsministerium), Dr. R. Herz (Universität Karlsruhe), Dr. M. Herry (Innovaplan, München), U. Holz (Technische Universität Berlin), H. Kuhfeld (DIW, Berlin), Prof. Dr.-Ing. E. Kutter (Technische Universität Berlin) sowie Frau B. Tassaux (IVT Heilbronn) und der Verfasser.

Der Gedankenaustausch in dieser Diskussionsrunde hat entscheidend zur Vertiefung des Problemverständnisses und damit zur Ausgewogenheit der Darstellung beigetragen. Für den Inhalt der Arbeit zeichnet allerdings allein der Autor verantwortlich.

- 2.3 Gewichtung als Mittel zur Fehlerreduzierung bzw. Fehlerkorrektur
3. Maßnahmen zur Reduzierung des Stichprobenfehlers
  - 3.1 Gebundene Hochrechnung: Regressions- und Verhältnisschätzung
  - 3.2 Nachträgliche Schichtung (Gewichtung im engeren Sinne)
4. Korrektur von ausfallbedingten Verzerrungen
  - 4.1 Verzerrungskorrektur durch nachträgliche Schichtung
  - 4.2 Verzerrungskorrektur auf der Basis von Informationen über das Verhalten von Nichtantwortern
  - 4.3 Extrapolationsmethoden zur Verzerrungskorrektur: Darstellung und Kritik
5. Korrektur von Verzerrungen durch Angabefehler
  - 5.1 Antwortvariabilität ohne Antwortverzerrung
  - 5.2 Antwortverzerrung: Korrektur auf der Basis von Nacherhebungen
  - 5.3 Antwortverzerrung: Korrektur auf der Basis exogener Daten
  - 5.4 Verzerrungskorrektur bei Angabefehlern bezüglich der Zahl der Wege
  - 5.5 Verzerrungskorrektur bei Angabefehlern bezüglich Wegelänge und Wegedauer
  - 5.6 Verzerrungskorrektur bei der Schätzung von verkehrsmittelspezifischen Verkehrsleistungswerten
6. Genauigkeitsbeurteilung bei ungewichteten Daten
  - 6.1 Varianzschätzung unter dem Design „Unabhängige Teilstichproben“
  - 6.2 Praktische Beispiele für die Berechnung von Konfidenzintervallen
  - 6.3 Quantifizierung des Design-Effekts
7. Genauigkeitsbeurteilung bei gewichteten Daten
  - 7.1 Einfluß der Gewichtung auf die Unverzerrtheit der Schätzung
  - 7.2 Einfluß der Gewichtung auf den Standardfehler der Schätzung
  - 7.3 Vereinfachte Fehlerrechnung bei komplexen Stichprobendesigns
  - 7.4 Ein Anwendungsbeispiel
8. Resümee und Ausblick

[Fortsetzung aus Heft 4/1989]

#### 4. Korrektur von ausfallbedingten Verzerrungen

##### 4.1 Verzerrungskorrektur durch nachträgliche Schichtung

Ausfallbedingte Verzerrungen entstehen dann, wenn die Haushalte in der Erhebungsgesamtheit kein gleichartiges Antwortverhalten aufweisen, sondern wenn statt dessen die Antwortbereitschaft in der Gesamtheit der Haushalte variiert. In dieser Situation wird nämlich die Stichprobenstruktur sich in charakteristischer Weise von der Struktur der Erhebungsgesamtheit unterscheiden. Für die Behandlung des Problems ist es wesentlich, ob die Antwortbereitschaft lediglich mit gewissen soziodemographischen Haushaltsmerkmalen im Zusammenhang steht (Fall I) oder ob zusätzlich - bei gegebenen soziodemographischen Merkmalen des Haushalts - die Antwortbereitschaft auch noch vom Verkehrsverhalten des Haushalts bzw. seiner Mitglieder abhängt (Fall II).

Im Fall I weichen die Stichprobenverteilungen gewisser soziodemographischer Haushaltsmerkmale systematisch von den entsprechenden Grundgesamtheitsverteilungen ab. Wenn man der Einfachheit halber einmal das Merkmal „Gemeindegröße“ betrachtet und nur zwei Kategorien, nämlich Haushalte aus „kleinen“ und Haushalte aus „großen“ Gemeinden unterscheidet, so läßt sich das eben Gesagte wie folgt präzisieren: Die Grundgesamtheit umfaßt  $N$  Haushalte, von den  $N_1$  in kleinen und  $N_2$  in großen Gemeinden leben ( $N_1 + N_2 = N$ ). Wählt man uneingeschränkt zufällig eine vorgegebene Anzahl von  $n$  Haushalten aus, so erhält man in der Stichprobe  $n_1$  Haushalte aus kleinen und  $n_2$  Haushalten aus großen Gemeinden. Der Stichprobenanteil  $n_1/n$  der Haushalte aus kleinen Gemeinden ist eine Zufallsvariable mit dem Erwartungswert  $N_1/N$ . Entsprechend schwankt der Stichprobenanteil  $n_2/n$  um  $N_2/N$ . Sind also alle Haushalte antwortbereit, so unterscheidet die Stichprobenverteilung des Merkmals „Gemeindegröße“ nur zufallsbedingt von der Verteilung dieses Merkmals in der Grundgesamtheit.

Wir nehmen nun an, daß ein in die Stichprobe gelangter Haushalt mit der Wahrscheinlichkeit  $p_1$  bzw.  $p_2$  antwortet, je nachdem, ob der Haushalt in einer kleinen bzw. großen Gemeinde lebt. Die Zahl der antwortenden Haushalte aus kleinen bzw. großen Gemeinden wird mit  $n_1^*$  bzw.  $n_2^*$  bezeichnet. Beides sind Zufallsvariable.

Die Gesamtzahl der antwortenden Haushalte, also die Größe

$$(4.1.1) \quad n^* = n_1^* + n_2^*$$

ist jetzt ebenfalls zufallsabhängig. Bei gegebenem  $n_1$  besitzt  $n_1^*$  den bedingten Erwartungswert

$$n_1 p_1$$

und insgesamt den Erwartungswert

$$(4.1.2) \quad n p_1 N_1 / N.$$

Entsprechend ist der Erwartungswert von  $n_2^*$  gleich

$$(4.1.3) \quad n p_2 N_2 / N.$$

Die Summe (4.1.1), also der Umfang der „Nettostichprobe“, besitzt entsprechend den Erwartungswert

$$(4.1.4) \quad (n/N) (p_1 N_1 + p_2 N_2).$$

Bei hinreichend großem Stichprobenumfang ist der Erwartungswert des Stichprobenanteils  $n_1^*/n^*$  der Haushalte aus kleinen Gemeinden annähernd gleich

$$(4.1.5) \quad p_1 N_1 / (p_1 N_1 + p_2 N_2)$$

und entsprechend besitzt  $n_2^*/n^*$  näherungsweise den Erwartungswert

$$(4.1.6) \quad p_2 N_2 / (p_1 N_1 + p_2 N_2).$$

Sofern gleiche Antwortbereitschaft in kleinen und großen Gemeinden vorliegt ( $p_1 = p_2$ ), sind die Stichprobenanteile  $n_1^*/n^*$  und  $n_2^*/n^*$  unverzerrte Schätzungen für die Anteile  $N_1/N$  bzw.  $N_2/N$ . Unterscheiden sich aber die Antwortwahrscheinlichkeiten  $p_1$  und  $p_2$ , so liegt eine Verzerrung vor. Ist beispielsweise die Antwortbereitschaft in kleinen Gemeinden höher als in großen ( $p_1 > p_2$ ), so fällt der Stichprobenanteil  $n_1^*/n^*$  tendenziell zu hoch aus, d. h. der Erwartungswert von  $n_1^*/n^*$  ist größer als  $N_1/N$ .

Die Verzerrung bezüglich des Merkmals Gemeindegröße, die von der Abhängigkeit zwischen diesem Merkmal und der Antwortbereitschaft herrührt, führt auch zu Verzerrungen bezüglich derjenigen Verkehrsverhaltensmerkmale, die mit der Gemeindegröße korreliert sind. Wenn beispielsweise die mittlere Wegelänge in kleinen Gemeinden höher ist als in großen Gemeinden, so führt eine durch unterschiedliches Antwortverhalten bedingte „Überrepräsentierung“ der Haushalte aus kleinen Gemeinden zu einer systematischen Überschätzung der mittleren Wegelänge: Die Schätzung

$$(4.1.7) \quad \bar{y}^* = \sum_{h=1}^2 \bar{y}_h^* (n_h^*/n^*),$$

wobei  $\bar{y}_1^*$  bzw.  $\bar{y}_2^*$  die in der Nettostichprobe festgestellte mittlere Wegelänge für Haushalte aus kleinen bzw. großen Gemeinden bezeichnet, ist (nach oben) verzerrt.

Durch Gewichtung, genauer gesagt durch nachträgliche Schichtung bezüglich des Merkmals Gemeindegröße, läßt sich diese Verzerrung der Schätzung der mittleren Wegelänge korrigieren. Nach Einführung der Gewichtungsfaktoren

$$(4.1.8) \quad \frac{N_h/N}{n_h^*/n^*} \quad (h=1,2)$$

erhalten wir die gewichtete Schätzung

$$(4.1.9) \quad \begin{aligned} \bar{y}_{gew}^* &= \sum_{h=1}^2 \bar{y}_h^* (n_h^*/n) \frac{N_h/N}{n_h^*/n^*} \\ &= \sum_{h=1}^2 \bar{y}^* N_h/N \end{aligned}$$

die asymptotisch unverzerrt ist.

Da die Mittelwerte  $\bar{y}_1^*$  und  $\bar{y}_2^*$  nur auf den Angaben der *antwortenden* Haushalte aus kleinen bzw. großen Gemeinden beruhen, ist (4.1.9) nur dann als asymptotisch unverzerrt zu betrachten, wenn bei gegebener Gemeindegröße die mittlere Wegelänge von antwortenden und nicht antwortenden Haushalten dieselbe ist. Man kann diesen Sachverhalt auch wie folgt ausdrücken: Die gewichtete Schätzung (4.1.9) ist nur dann unverzerrt, wenn tatsächlich innerhalb jeder Gemeindegrößenklasse die Antwortwahrscheinlichkeit eines Haushalts nicht mit der mittleren Länge der Wege dieses Haushalts bzw. seiner Mitglieder im Zusammenhang steht. Wäre beispielsweise bei gegebener Gemeindegrößenklasse  $h$  die Antwortwahrscheinlichkeit eines Haushalts umso größer, je niedriger die mittlere Länge der Wege des Haushalts ist, so wäre der in (4.1.9) erscheinende Stichprobenmittelwert  $\bar{y}_h^*$  systematisch zu klein, da sich unter den antwortenden Haushalten vorzugsweise solche mit niedriger mittlerer Wegelänge befänden. Zusammenfassend kann man also feststellen:

1. Verzerrungen durch Antwortausfälle lassen sich durch nachträgliche Schichtung (Bildung von Hochrechnungsgruppen) teilweise korrigieren. Die Korrektur ist dabei umso wirksamer, je enger der Zusammenhang zwischen dem Verkehrsverhalten und den zur nachträglichen Schichtung benutzten Merkmalen ist.
2. Reduzierung ausfallbedingter Verzerrungen durch nachträgliche Schichtung gelingt nur dann, wenn innerhalb der verschiedenen Hochrechnungsgruppen die Antwortbereitschaft der Haushalte homogen ist. Variiert innerhalb der soziodemographischen Hochrechnungsgruppen die Antwortbereitschaft mit dem Verkehrsverhalten, so verliert die nachträgliche Schichtung an Wirksamkeit.

Da derzeit über die Abhängigkeit der Antwortbereitschaft von soziodemographischen und Verkehrsverhaltensmerkmalen wenig bekannt ist, kann über den Nutzen einer Gewichtung zur Korrektur ausfallbedingter Verzerrung vorwiegend nur spekuliert werden. Auf jeden Fall sollte aber im Rahmen einer nachträglichen Schichtung immer nur die gemeinsame Verteilung der Schichtungsmerkmale angepaßt werden. Wenn nämlich die Antwortbereitschaft zwischen den Hochrechnungsgruppen variiert, was wohl als Normalfall angenommen werden kann, so führt eine Gewichtung mit Hilfe marginaler Häufigkeiten zu verzerrten Schätzungen, auch wenn das Motiv für die Gewichtung gerade die Korrektur solcher Verzerrungen ist. (Vgl. hierzu auch Abschnitt 7.1.)

#### 4.2 Verzerrungskorrektur auf der Basis von Informationen über das Verhalten von Nichtantwortern

Idealerweise sollte die Korrektur von ausfallbedingten Verzerrungen auf der Grundlage von Daten über das Verkehrsverhalten der Nichtantworter vorgenommen werden. Solche Daten lassen sich mit Hilfe mehrphasiger Auswahlverfahren gewinnen: Aus den Non-Response-Fällen der schriftlichen Haushaltsbefragung (Phase 1) wird in der zweiten Phase eine Unterstichprobe gezogen und es wird für die dabei ausgewählten Haushalte mit allen zu Gebote stehenden Mitteln versucht, die benötigten Angaben einzuholen (STATISTISCHES BUNDESAMT, 1960). Wenn die Antwortbereitschaft innerhalb der einzelnen Hochrechnungsgruppen nicht konstant ist, sondern irgendwie mit Verkehrsverhaltensmerkmalen variiert, ist eine derartige Vorgehensweise die einzige Möglichkeit, ausfallbedingte Verzerrungen zu



korrigieren. Die Methode geht zurück auf *Hansen* und *Hurwitz* (1946) und ist insbesondere unter Einbeziehung von Kostengesichtspunkten stark verfeinert worden.

Das Grundprinzip der Schätzung bei Vorliegen solcher Zusatzinformationen kann wie folgt skizziert werden: Man geht von einem Response-Modell aus, nach welchem die Grundgesamtheit in zwei Klassen zerfällt:

1. Haushalte, bei denen in der schriftlichen Phase der Feldarbeit Daten erhoben werden können,
2. Alle übrigen Haushalte (Ausfälle).

Aus den Haushalten der Gruppe 2, die in die (Brutto-)Stichprobe gelangt sind und zunächst Ausfälle darstellen, wird eine Unterstichprobe  $U$  gezogen. Bei den Haushalten aus  $U$  wird dann eine Kurzversion des KONTIV-Fragebogens erhoben, die einige wenige Struktur- und Verhaltensmerkmale enthält. Die dabei gewonnenen Daten werden dann zur Gewichtung der Gesamtstichprobe herangezogen.

Zur Verdeutlichung betrachten wir folgendes Beispiel: Das Merkmal „mobil am Stichtag“ teilt die Grundgesamtheit der Personentage in zwei Kategorien:

Kategorie 1: mobil            Kategorie 2: nicht mobil.

Nach der schriftlichen Phase haben  $p \cdot 100$  Prozent ( $0 < p < 1$ ) der zur Bruttostichprobe gehörenden Personen geantwortet;  $q = 1 - p$  ist der Anteil der Non-Response-Fälle. Wir führen folgende Symbolik ein:

- $z$     Zahl der Personen, welche geantwortet haben
- $z_1$     Zahl der mobilen Personen
- $z_2$     Zahl der nicht mobilen Personen

Es gilt  $z = z_1 + z_2$ .

Bei der (beispielsweise telefonischen) Non-Response-Aktion können aus der Unterstichprobe die analog definierten Zahlen  $m$ ,  $m_1$  und  $m_2$  ( $m = m_1 + m_2$ ) ermittelt werden.

Der Anteil der mobilen Personen in der Grundgesamtheit wird nunmehr durch

$$r_1 = (z_1/z)p + (m_1/m)q$$

also durch das gewogene arithmetische Mittel der Verkehrsbeteiligungsquote von Personen aus antwortenden und (zunächst) nicht antwortenden Haushalten geschätzt. Analog ist

$$r_2 = (z_2/z)p + (m_2/m)q$$

ein Schätzwert für den Anteil der nicht mobilen Personen.

Korrekturen von ausfallbedingten Verzerrungen sollten sich stets, wie oben skizziert, auf empirische Daten über Haushalte, die die Teilnahme an der schriftlichen Befragung verweigert haben, stützen. Bedauerlicherweise gibt es bisher kaum irgendwelche veröffentlichten Ergebnisse über eventuelle echte Verhaltensunterschiede zwischen Antwortern und Nichtantwortern. Hierzu müßte das Verkehrsverhalten von Personen aus antwortenden und (zunächst) nicht antwortenden Haushalten bei gegebenen soziodemographischen Merkmalen miteinander verglichen werden. Nur so ist feststellbar, ob zur Korrektur ausfall-

bedingter Verzerrungen eine nachträgliche Schichtung ausreicht, oder ob zusätzliche Daten über das Verkehrsverhalten von Nichtantwortern herangezogen werden müssen.

An dieser Stelle muß noch darauf hingewiesen werden, daß die Nichtteilnahme an der schriftlichen Befragung keine absolute Antwortverweigerung darstellt. Wie bereits erwähnt, wurde nach dem am KONTIV-Design orientierten schweizerischen Mikrozensus „Verkehr und Umwelt“ im Jahr 1984 eine kleinere telefonische Kurzbefragung von Haushalten, die im Rahmen der schriftlichen Erhebung nicht geantwortet hatten, durchgeführt. Hierbei zeigte sich, daß rund zwei Drittel dieser „Antwortverweigerer“ am Telefon durchaus auskunftsbereit waren.

Schließlich ist auch noch zu beachten, daß es unter den Antwortern durchaus versteckte Antwortverweigerer geben kann. Es sind dies Haushalte, die sich der lästigen Aufgabe der Fragebogenausfüllung - insbesondere nach wiederholter Mahnung - mit möglichst geringem Aufwand zu entziehen versuchen (*Herz*, 1987). Diese Gruppe ist schwierig zu identifizieren und - wenn überhaupt - allenfalls an generell schlechter Ausfüllqualität der Erhebungsunterlagen zu erkennen. So führte im Zusammenhang mit einer Erhebung vom KONTIV-Typ eine telefonische Nachbefragung von rund 1.200 Haushalten in Berlin (West) mit mindestens einer Person mit leerem oder fehlendem Tagesprotokoll zu dem Resultat, daß rund drei Viertel der Personen mit leerem Tagesprotokoll am Stichtag tatsächlich mobil gewesen sind (*Kutter* und *Holz*, 1987).

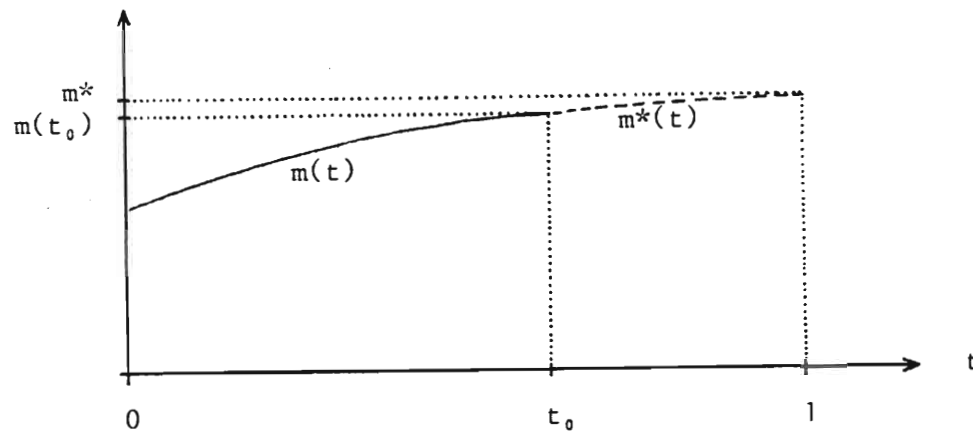
Im nächsten Abschnitt wird auf eine Extrapolationsmethode zur Non-Response-Korrektur eingegangen, die teilweise in sehr weitgehender Verfeinerung angewendet wurde (SOCIAL-DATA, 1984), aus heutiger Sicht aber als überaus problematisch eingestuft werden muß.

#### 4.3 Extrapolationsmethoden zur Verzerrungskorrektur: Darstellung und Kritik

Bei schriftlichen Befragungen gibt es manchmal einen Zusammenhang zwischen dem interessierenden Untersuchungsmerkmal und der Antwortgeschwindigkeit. Ist dies der Fall, so ändert sich der Schätzwert  $m$  für die zu schätzende statistische Maßzahl  $\mu$  ( $\mu$  kann beispielsweise ein Mittelwert oder eine Verhältniszahl sein) in ganz charakteristischer Weise, wenn man die Berechnung von  $m$  auf einen immer größeren Teil der Antwortenden stützt.

Es sei  $m(t)$  der Schätzwert für  $\mu$  berechnet auf der Basis der ersten  $100t$  Prozent der antwortenden Einheiten ( $0 \leq t \leq 1$ ). Gäbe es keine Antwortverweigerung, so könnte man für wachsende Werte von  $t$  jeweils  $m(t)$  berechnen. Solange  $t$  kleiner als Eins ist, hat der Schätzwert  $m(t)$  vorläufigen Charakter. Als „endgültigen“ Schätzwert  $m$  würde man selbstverständlich  $m(1)$  an der Stelle  $t = 1$  verwenden, d. h.  $m = m(1)$ . In Abbildung 4.1 ist dies nochmals graphisch veranschaulicht.

Abbildung 4.1  
Entwicklung eines vorläufigen Schätzwerts bei wachsendem Anteil antwortender Haushalte



Wenn nicht alle sondern nur  $100t_0$  Prozent der Befragten antworten, kann  $m(t)$  selbstverständlich nur für

$$0 \leq t \leq t_0 < 1$$

berechnet werden. Die Extrapolationsmethode besteht nun darin, den Verlauf von  $m(t)$  für  $0 \leq t \leq t_0$  empirisch zu ermitteln und danach eine Extrapolation für den Bereich  $t_0 < t \leq 1$  vorzunehmen. Die extrapolierten Werte kann man mit

$$m^*(t) \quad \text{für} \quad t_0 < t \leq 1$$

bezeichnen. Als Schätzwert für  $\mu$  wird dann nicht  $m(t_0)$  verwendet (dies ist der Schätzwert, der sich aus der Antworterstichprobe ergibt), sondern vielmehr der extrapolierte Wert

$$m^* = m^*(1).$$

Die Extrapolationsmethode ist zweifellos ein hilfreiches Instrument zur Korrektur von Verzerrungen durch Antwortverweigerung, wenn der Zusammenhang zwischen Antwortgeschwindigkeit und Untersuchungsgegenstand als gesichert betrachtet werden kann. In der stichprobentheoretischen Literatur wird das Extrapolationskonzept erstmals von *Hendricks* (1949) beschrieben. Im konkreten Fall der Haushaltsbefragungen vom KONTIV-Typ gibt es allerdings gewichtige Argumente gegen die Verwendung derartiger Extrapolationstechniken.

Vorliegende empirische Erkenntnisse deuten darauf hin, daß weniger das Verkehrsverhalten der Befragten als vielmehr die Ausfüllqualität der Fragebögen mit der Antwortgeschwindigkeit im Zusammenhang steht. So ist nachgewiesenermaßen der Anteil fehlender Angaben zu soziodemographischen Merkmalen bei den „Sofortantwortern“ und den „Spätantwortern“ überdurchschnittlich hoch. Dasselbe gilt für den Anteil der antwortenden Personen

mit leerem oder fehlendem Tagesprotokoll (*Kutter und Holz, 1987*). Wenn man in dieser Situation Personen mit leerem Tagesprotokoll als „am Stichtag immobil“ betrachtet, sind Fehlschlüsse bei Anwendung der Extrapolationsmethode im Zusammenhang mit der Schätzung der Verkehrsbeteiligungsquote unvermeidlich.

Da die Antwortgeschwindigkeit mit gewissen soziodemographischen Merkmalen des Haushalts und der darin lebenden Personen zusammenhängt, ist die Struktur des ersten, zweiten usw. Zehntels der Antworterstichprobe unterschiedlich. Dies wirkt sich auch auf Schätzungen bestimmter Kennziffern des Verkehrsverhaltens aus, wenn man diese getrennt nach den so gebildeten Teilstichproben berechnet. Eine Extrapolation ist hier vollkommen unangebracht, auch wenn sie durch ein „Abhängigkeitsmuster“ scheinbar nahegelegt wird. Es genügt hier die nachträgliche Schichtung der gesamten Antworterstichprobe (gemäß Abschnitt 4.1), gegebenenfalls ergänzt um Korrekturen auf der Basis von Zusatzinformationen aus einer Nacherhebung bei Nichtantwortern.

## 5. Korrektur von Verzerrungen durch Angabefehler

Stimmt der vom Befragten angegebene Wert eines Erhebungsmerkmals nicht mit dem wahren Wert überein, so spricht man von einem individuellen „Angabefehler“ (vgl. Abschnitt 2.2.5). Wir wollen nun der Frage nachgehen, unter welchen Bedingungen durch Angabefehler Verzerrungen entstehen können und welche Auswirkungen Angabefehler auf die Schätzgenauigkeit haben.

### 5.1 Antwortvariabilität ohne Antwortverzerrung

Das Auftreten von individuellen Angabefehlern führt nicht automatisch zu Verzerrungen der Hochrechnungsergebnisse. Keine Verzerrung ist zu befürchten, wenn bei jedem Befragten in der Stichprobe der angegebene Wert als eine Zufallsvariable betrachtet werden kann, deren Erwartungswert mit dem wahren Wert des Erhebungsmerkmals übereinstimmt. Eine solche Situation läge etwa dann vor, wenn die befragten Personen die Längen der von ihnen durchgeführten Wege zwar nicht genau angeben würden, wenn aber doch bei jeder Person eine subjektive Über- bzw. Unterschätzung der tatsächlichen Wegelänge gleich wahrscheinlich wäre.

In dieser Situation basiert die statistische Schätzung z. B. der mittleren Wegelänge zwar auf im Einzelfall falschen Wegelängenangaben, die Schätzung ist aber dennoch unverzerrt. Allerdings ist in einem solchen Fall von „Antwortvariabilität“ die Varianz dieser Schätzung größer als in dem Fall, wo alle Befragten die wahren Merkmalsausprägungen angeben. Die in der üblichen Weise (auf der Grundlage der angegebenen Werte) berechneten Konfidenzintervalle sind demnach unzulässig eng. Um das tatsächliche Ausmaß der stichproben- und angabebedingten Unsicherheit der Schätzung quantifizieren zu können, muß man zumindest einen Teil der Haushalte in geeigneter Weise mehrfach befragen, um so zu einer Schätzung für die durchschnittliche Antwortvarianz zu kommen. Näheres hierzu findet man bei *Strecker u. a. (1983)*.



### 5.2 Antwortverzerrung: Korrektur auf der Basis von Nacherhebungen

Wenn individuelle Antwortverzerrung vorliegt, d.h. wenn bei den Befragten der Erwartungswert der angegebenen Merkmalsausprägung nicht mit der wahren Merkmalsausprägung übereinstimmt, tritt zur Varianzvergrößerung noch eine Verzerrung der Hochrechnungsergebnisse. Dies wäre beispielsweise dann der Fall, wenn der Erwartungswert der angegebenen Wegelänge um einen konstanten oder auch von Person zu Person variierenden Faktor über der tatsächlichen Wegelänge liegen würde. Daß in dieser Situation unverzerrte Schätzungen nur möglich sind, wenn zumindest für einige Untersuchungseinheiten zusätzlich zu den angegebenen auch die wahren Merkmalsausprägungen vorliegen, ist bereits intuitiv klar.

Für einen bestimmten Stichtag mögen im einfachsten Fall  $n$  Haushalte nach der Ziehungsvorschrift mit Zurücklegen mit gleichen Auswahlwahrscheinlichkeiten gezogen werden. Mit

$$Y_k \quad (k = 1, \dots, n)$$

werde die angegebene Gesamtlänge der Wege sämtlicher Mitglieder des an  $k$ -ter Stelle gezogenen Haushalts bezeichnet. Nun wird angenommen, daß nach demselben Ziehungsverfahren weitere  $n^*$  Haushalte ausgewählt werden. Die angegebene Gesamtweglänge des an  $(n+i)$ ter Stelle ausgewählten Haushalts bezeichnen wir mit

$$Y_i^* \quad (i = 1, \dots, n^*).$$

Für die  $n^*$  zuletzt ausgewählten Haushalte werden durch eine Zusatzbefragung die wahren Wegelängen

$$X_i^* \quad (i = 1, \dots, n^*)$$

ermittelt. In dieser Situation ist

$$(5.2.1) \quad \frac{1}{n} \sum_{j=1}^n Y_j + \frac{1}{n^*} \sum_{i=1}^{n^*} (X_i^* - Y_i^*)$$

eine unverzerrte Schätzung für die Verkehrsleistung pro Haushalt am betreffenden Stichtag (km pro Haushalt und Tag). In (5.2.1) ist der erste Summand einfach der Stichprobenmittelwert der angegebenen Haushaltsverkehrsleistung (berechnet allerdings nur aus den Angaben der ersten  $n$  Haushalte), der zweite Summand dagegen hat den Charakter eines Korrekturfaktors; es handelt sich um das Stichprobenmittel (berechnet aus den Informationen über die restlichen  $n^*$  Haushalte) der Abweichung zwischen tatsächlicher und angegebener Haushaltsverkehrsleistung.

Die Varianz der unverzerrten Schätzfunktionen (5.2.1) kann auf einfache Weise geschätzt werden. Wenn man in (5.2.1) den ersten Summanden mit  $\bar{Y}$  und den zweiten mit  $(\bar{X}^* - \bar{Y}^*)$  bezeichnet, so lautet die Varianzschätzung

$$(5.2.2) \quad \frac{1}{n(n-1)} \sum (Y_j - \bar{Y})^2 + \frac{1}{n^*(n^*-1)} \sum [(X_i^* - Y_i^*) - (\bar{X}^* - \bar{Y}^*)]^2$$

Näheres siehe *Stenger* (1986).

### 5.3 Antwortverzerrung: Korrektur auf der Basis exogener Daten

Erhebungen des zuletzt beschriebenen Typs, bei denen durch Zusatzuntersuchungen für einen Teil der Stichprobeneinheiten die wahren Werte ermittelt werden, sind bisher nur in recht begrenztem Umfang durchgeführt worden. Entsprechend unsicher ist deshalb auch unser empirisches Wissen über Antwortvariabilität und Antwortverzerrung bei Haushaltsbefragungen vom KONTIV-Typ. Unter diesen Bedingungen ist gegenwärtig - wenn überhaupt - nur eine grobe Verzerrungskorrektur möglich. Hierbei müssen die Korrekturfaktoren aus den erwähnten kleinen Spezialerhebungen genommen und auf die Daten der hochzurechnenden Haushaltsbefragung übertragen werden.

Rechentchnisch läuft dies darauf hinaus, daß die in der Befragung festgestellten Merkmalsausprägungen mit gewissen Korrekturfaktoren aus Spezialerhebungen multipliziert werden. Es sei beispielsweise  $Y_{ki}$  die angegebene Länge des  $i$ -ten Weges in dem an  $k$ -ter Stelle ausgewählten Haushalt. Wenn die angegebene Wegelänge nach vorliegenden Erkenntnissen im Durchschnitt um etwa  $100 \alpha$  Prozent zu hoch wäre, könnte man

$$Y_{ki}^* = (1 + \alpha) Y_{ki}$$

verwenden, um eine Reduktion der Verzerrung der Hochrechnungsergebnisse zu erreichen. Wäre die Antwortverzerrung (subjektive Wegelängenüberschätzung) je nach Verkehrsmittel unterschiedlich, so müßte

$$Y_{ki}^* = (1 + \alpha_i) Y_{ki}$$

gesetzt werden, wobei  $\alpha_i$  davon abhängt, mit welchem Verkehrsmittel der  $i$ -te Weg durchgeführt worden ist.

Im folgenden wird nun auf die Größenordnung möglicher Antwortverzerrungen bei Haushaltsbefragungen vom KONTIV-Typ eingegangen. Über die Antwortvarianz, also die Varianz des angegebenen Werts (der ja als Zufallsvariable betrachtet werden kann), liegen nach unserem Kenntnisstand derzeit noch keine veröffentlichten Ergebnisse vor.

### 5.4 Verzerrungskorrektur bei Angabefehlern bezüglich der Zahl der Wege

Nach den Ausführungen des Abschnitts 2.2.5 sind als wichtigste Angabefehler bei Haushaltsbefragungen vom KONTIV-Typ anzusehen

- Nichtangabe von Wegen,
- Fehler bei der subjektiven Schätzung der Wegelänge und Wegedauer,
- Bevorzugung bestimmter Wegelängen- und Wegedauerwerte,
- Unterlassen von Verkehrsmittelnennungen, insbesondere der Nennung „zu Fuß“ bei Wegen mit Verkehrsmittelkombinationen,
- Nichtbeantwortung von einzelnen Fragen, d.h. fehlende Angaben zu bestimmten Erhebungsmerkmalen.

Die Nichtangabe von Wegen hat vor allem zwei Ursachen, wobei die erste mit der Validität des Meßinstruments (KONTIV-Fragebogen) und die zweite mit dem Antwortverhalten

zusammenhängt. Auf die Tendenz zur Nichtangabe von Wegen war bereits in Abschnitt 2.2.2 eingegangen worden. Bei *Wermuth* und *Maerschalk* (1981) findet man das Ergebnis einer Kontrollerhebung, wonach 14 Prozent der tatsächlich absolvierten Wege fehlen. Dieser Anteil (non-reported-trip-rate) variiert mit dem Wegezweck, dem benutzten Verkehrsmittel und der Wegelänge.

Wieviele dieser fehlenden Wege im Aufbereitungsprozeß „plausibel“ ergänzt werden können, läßt sich nicht genau quantifizieren. Bei einer speziellen 2-phasigen Erhebung wurde ein Anteil von 4,1 Prozent nichtangegebener Wege festgestellt, die auch bei „sorgfältiger Codierung“ nicht hätten gefunden werden können (SOCIALDATA, 1986).

Die Nichtangabe von Wegen hat eine weitere Ursache: Es gibt Grund zu der Annahme, daß bei insgesamt eher sinkender „Antwortmoral“ immer mehr Befragte vorgeben, am Stichtag keinen Weg unternommen zu haben, um sich die Mühe des Fragebogenausfüllens zu ersparen. Man hat es hier mit einer nicht unerheblichen Grauzone von „Pseudo-Häuslichkeit“ zu tun (*Herz*, 1987). Wie Auswertungen der KONTIV 82 zeigen, steigt mit wachsender Antwortdauer (also nach ggf. mehrmaliger Mahnung) der Anteil der Befragten, welche vollständig leere oder gar keine Tagesprotokolle zurücksenden, während der Anteil derer mit der Angabe „Zu Hause geblieben, weil ...“ praktisch konstant ist (*Kutter* und *Holz*, 1987). Behandelt man leere oder fehlende Tagesprotokolle wie berichtete Immobilität, so muß dies zu Verzerrungen der Verkehrsbeteiligungsquote nach unten führen. Auf diesen Aspekt des Antwortverhaltens ist bereits im Abschnitt 4.2 eingegangen worden.

#### 5.5 Verzerrungskorrektur bei Angabefehlern bezüglich Wegelänge und Wegedauer

Bei Erhebungen von KONTIV-Typ beruhen die Wegelänge- und Wegedauerangaben auf subjektiven Schätzungen der Befragten. Es versteht sich, daß diese subjektiven Schätzungen nicht genau mit den tatsächlichen Wegelängen und Wegedauerwerten übereinstimmen. Würden sich die individuellen Schätzfehler im Durchschnitt ausgleichen, und wären die individuellen Schätzfehler nicht mit irgendwelchen Erhebungsmerkmalen (z. B. benutztes Verkehrsmittel) korreliert, so ergäben sich hieraus für die Punktschätzung des Total- bzw. Mittelwerts der Wegelänge und Wegedauer keine gravierenden Probleme.

Aus einigen empirischen Studien geht nun allerdings hervor, daß Befragte dazu neigen, die Wegelänge und Wegedauer systematisch zu überschätzen, wobei die ermittelten Verzerrungen nach oben in der Größenordnung von rund 10 Prozent liegen; zum Teil wurde auch eine nach Verkehrsmitteln unterschiedliche Verzerrung festgestellt (*Brög* u. a., 1976; *Brög* und *Meyer*, 1978; *Morris*, 1983; *SWOV*, 1978). Eine Verwendung von entsprechenden Korrekturfaktoren bei Hochrechnungen kann jedoch kaum empfohlen werden, da der Grad der Zuverlässigkeit dieser Faktoren nicht hinreichend genau beurteilt werden kann.

Da die Befragten bestimmte „runde Zahlen“ bei Wegelängenangaben stark bevorzugen, kann es bei Tabellierungen von Wegelängenverteilungen zu recht beträchtlichen Verzerrungen kommen, da die bevorzugten Werte mit den Grenzen der gebräuchlichen Merkmalsklassen zusammenfallen. Die Definition der Wegelängensklassen („von ... bis unter ... km“ bzw. „über ... bis einschließlich ... km“) hat deshalb erheblichen Einfluß auf das Ergebnis. Im

Grundsatz gibt es, wenn man von aufwendigeren Glättungstechniken (*Hengst*, 1967) einmal absieht, drei Lösungsmöglichkeiten:

- Methode A: Man wählt Klasseneinteilungen derart, daß die bevorzugten Werte möglichst die Klassenmitte bilden.
- Methode B: Man teilt die Häufigkeit eines bevorzugten Werts, der zugleich eine Klassengrenze darstellt, auf die beiden benachbarten Klassen auf.
- Methode C: Man weist in der Tabelle die bevorzugten Werte und ihre Häufigkeiten gesondert aus.

Ein praktisches Beispiel hierfür findet man bei *Hautzinger* und *Tassaux* (1989).

#### 5.6 Verzerrungskorrektur bei der Schätzung von verkehrsmittelspezifischen Verkehrsleistungswerten

Die Neigung der Befragten, nicht alle tatsächlich benutzten Verkehrsmittel eines Weges anzugeben, führt bei Aufgliederungen des hochgerechneten Totals der Verkehrsleistung nach dem Verkehrsmittel zu teilweise gravierenden Verzerrungen, insbesondere im Zusammenhang mit der Verkehrsleistung von Fußgängern. Von *Hautzinger* und *Tassaux* (1989) wurde ein Verfahren zur Korrektur von derartigen Verzerrungen entwickelt und erprobt. Basis hierfür waren Daten einer speziell zu diesem Zweck durchgeführten Erhebung, bei welcher in einem 2-phasigen Verfahren Haushalte zunächst schriftlich und danach nochmals mündlich befragt wurden. Bei der mündlichen Befragung lag der Schwerpunkt auf der Exploration aller tatsächlich erforderlichen Verkehrsmittelnennungen pro Weg. Das Verfahren läßt sich wie folgt beschreiben.

Es sei  $m$  die Zahl der Ausprägungen des Merkmals Verkehrsmittel. Dann kann man  $m$  verkehrsmittelspezifische Wegelängenvariable definieren und die Gesamtlänge  $x$  eines Weges als Summe

$$(5.6.1) \quad x = x_1 + x_2 + \dots + x_m$$

darstellen, wobei  $x_i$  den auf das  $i$ -te Verkehrsmittel entfallenden Teil der Wegelänge bezeichnet. Für den Mittelwert der Wegelänge gilt selbstverständlich

$$(5.6.2) \quad \bar{x} = \bar{x}_1 + \bar{x}_2 + \dots + \bar{x}_m$$

und die Gliederungszahl

$$(5.6.3) \quad r_i = \bar{x}_i / \bar{x} \quad (i = 1, \dots, m)$$

gibt an, welcher Anteil am Wegelängentotal auf das  $i$ -te Verkehrsmittel entfällt.

Aus einer 2-phasigen Erhebung des oben beschriebenen Typs kann man empirische  $r_i$ -Werte differenziert nach dem in Phase 1 angegebenen (Haupt-)Verkehrsmittel  $j$  ( $j = 1, \dots, m$ ) berechnen. Wir bezeichnen mit

$$(5.6.4) \quad r(i|j) \quad (i, j = 1, \dots, m)$$

den im Durchschnitt auf das  $i$ -te Verkehrsmittel entfallenden Teil der Wegelänge derjenigen



Wege, für welche in Phase 1 das Verkehrsmittel  $j$  angegeben wurde. Beispielsweise ergab sich bei der erwähnten 2-phasigen Spezialerhebung

$$r(F|F) = 0,85$$

$$r(P|R) = 0,14$$

$$r(P|P) = 0,99$$

wobei F für „zu Fuß“, P für „Pkw-Fahrer“ und R für „Fahrrad“ steht. Der erste Zahlenwert (0,85) bedeutet, daß gemäß Phase 2 rund 15 Prozent der Gesamtlänge aller Wege, für die in Phase 1 „zu Fuß“ als Verkehrsbeteiligungsart angegeben wurde, tatsächlich mit einem Verkehrsmittel zurückgelegt worden ist.

Bei einer Schätzung des auf das  $i$ -te Verkehrsmittel entfallenden Teils der gesamten Verkehrsleistung können die aus der genannten Spezialerhebung stammenden  $r(i|j)$ -Werte zur Korrektur angabebedingter Verzerrungen verwendet werden. Wenn beispielsweise das Total der Fußgänger-Verkehrsleistung ( $i = 1$ ) zu schätzen ist, wird ein neues Wegemerkmal  $x_i$  („Weglänge zu Fuß“) definiert. Der Wert dieses Merkmals ergibt sich, indem die vom Befragten angegebene Weglänge  $x$  mit dem entsprechenden Faktor  $r(i|j)$  multipliziert wird, wobei  $j$  das beim betreffenden Weg angegebene (Haupt-)Verkehrsmittel ist. Das Total der Fußgänger-Verkehrsleistung wird nunmehr geschätzt durch die hochgerechnete Merkmalssumme des neuen Merkmals  $x_i$ .

Das Gesamttotal der Verkehrsleistung wird durch das beschriebene Verfahren nicht verändert. Die Korrektur bewirkt lediglich - im Vergleich zu einer nur am angegebenen (Haupt-)Verkehrsmittel orientierten Hochrechnung - strukturelle Verschiebungen zwischen den einzelnen Verkehrsmitteln. Die Auswirkungen dieser Korrektur sind beträchtlich. Gegenüber einer ausschließlich am „Hauptverkehrsmittel“ orientierten Verkehrsleistungsschätzung erhöht sich das Total der Fußgänger-Kilometer um über 40 Prozent. Bei der Schätzung des Totals der Verkehrsbeteiligungsdauer kann analog verfahren werden.

## 6. Genauigkeitsbeurteilung bei ungewichteten Daten

### 6.1 Varianzschätzung unter dem Design „Unabhängige Teilstichproben“

Wir gehen aus von einer Untergliederung der  $N$  Haushalte in  $G$  Gruppen und der  $T$  Tage des Untersuchungszeitraums in  $H$  Tagestypen. Eine Kombination  $(g, h)$  von Haushaltsgruppe und Tagestyp betrachten wir als eine Schicht von „Haushaltstagen“. Insgesamt wird also die Grundgesamtheit der  $NT$  Haushaltstage in  $GH$  Schichten zerlegt. Wir schreiben  $N_g$  bzw.  $T_h$  für die Zahl der Haushalte in Gruppe  $g$  bzw. für die Zahl der Tage des Typs  $h$ . Die Schicht  $(g, h)$  enthält demnach  $N_g T_h$  Haushaltstage.

Aus den  $N_g T_h$  Haushaltstagen der Schicht  $(g, h)$  werden  $n_{gh}$  Haushaltstage ausgewählt. Falls  $H = T$  (und dementsprechend  $T_h = 1$ ), so hat man es mit einem System von  $T$  unabhängigen „Tagesstichproben“ zu tun. Ein Beispiel für  $H < T$  wäre etwa der Fall, wo die Tage des Untersuchungszeitraums entsprechend dem Wochentagstyp (Werktag, Samstag, Sonn-/Feiertag) in  $H = 3$  Typen untergliedert werden; in dieser Situation werden Haushalte innerhalb der Gruppe  $g$  zufällig ausgewählt und ebenfalls nach einem Zufallsverfahren

einem bestimmten Tag (Befragungstichtag) des Typs  $h$  zugeordnet. Wir gehen davon aus, daß jeder ausgewählte Haushalt zum Verkehrsverhalten an einem bestimmten Stichtag befragt wird.

Wenn beim Ziehen mit Zurücklegen die Auswahlwahrscheinlichkeit eines Haushalts proportional zur Haushaltsgröße ist, was in der Praxis häufig auftritt, so ist beim Ziehungsverfahren für die Schicht  $(g, h)$

$$(6.1.1) \quad P_{ghi} = z_{gi} / Z_{gi} T_h$$

die Wahrscheinlichkeit, daß der  $i$ -te Haushalt der Gruppe  $g$  in einem bestimmten Zug gezogen und einem bestimmten Tag der insgesamt  $T_h$  Tage des Typs  $h$  zugeordnet wird. In (6.1.1) bezeichnet  $z_{gi}$  die Zahl der Personen im  $i$ -ten Haushalt der Gruppe  $g$  und  $Z_g$  die Gesamtzahl aller Personen in Haushalten der Gruppe  $g$ .

Bezeichnet man mit  $X_{ghi}$  die Ausprägung des Untersuchungsmerkmals (Verhaltensmerkmals) beim  $i$ -ten Stichproben-Haushalt der Schicht  $(g, h)$ , beispielsweise also die Länge aller Wege von Haushaltsmitgliedern am betreffenden Stichtag, so ist

$$(6.1.2) \quad \hat{X}_{gh} = (1/n_{gh}) \sum_{i=1}^{n_{gh}} X_{ghi} / P_{ghi}$$

eine unverzerrte Schätzfunktion für das Total des Untersuchungsmerkmals in Schicht  $(g, h)$ , beispielsweise also eine Schätzung für die Gesamtlänge aller Wege, die von Personen aus Haushalten der Gruppe  $g$  an Tagen des Typs  $h$  durchgeführt werden. Eine Schätzung des Gesamttotals ist dann durch

$$(6.1.3) \quad \hat{X} = \sum_g \sum_h \hat{X}_{gh}$$

gegeben.

Meist geht es bei Auswertungen allerdings um die Schätzung von Verhältniszahlen. Soll z. B. die mittlere Weglänge geschätzt werden, so ist analog zu  $X_{ghi}$  eine Zufallsgröße  $Y_{ghi}$  „Zahl der Wege (am betreffenden Stichtag) von Personen aus dem  $i$ -ten Stichprobenhalt in Schicht  $(g, h)$ “ zu definieren. Analog zu (6.1.2) erhält man eine Schätzfunktion  $\hat{Y}_{gh}$  für die Gesamtzahl aller Wege von Personen aus Haushalten der Gruppe  $g$  an Tagen des Typs  $h$ . Die Schätzfunktion für die Gesamtzahl aller Wege im Untersuchungszeitraum wird analog zu (6.1.3) mit  $\hat{Y}$  bezeichnet. Die mittlere Weglänge (Dimension „km/Weg“) wird dann durch

$$(6.1.4) \quad \hat{R} = \hat{X} / \hat{Y}$$

konsistent geschätzt.

Ein Konfidenzintervall für die zu schätzende Verhältniszahl  $r$  kann mit Hilfe der „relativen Varianz“ von  $\hat{R}$ , d. h. mittels der Größe

$$(6.1.5) \quad \text{var}(\hat{R}) / \hat{R}^2$$

auf einfache Weise angegeben werden. Man schätzt die relative Varianz von  $\hat{R}$  durch



$$(6.1.6) \quad \hat{V}_{\hat{R}}^2 = \hat{V}_{\hat{X}}^2 + \hat{V}_{\hat{Y}}^2 - 2\hat{V}_{\hat{X}\hat{Y}}$$

wobei

(6.1.7)

$$\hat{V}_{\hat{X}\hat{Y}} = \frac{1}{\hat{X}\hat{Y}} \sum_g \sum_h \frac{1}{n_{gh}(n_{gh}-1)} \sum_i \left( \frac{X_{ghi}}{P_{ghi}} - \hat{X}_{gh} \right) \left( \frac{Y_{ghi}}{P_{ghi}} - \hat{Y}_{gh} \right)$$

Es gilt ferner  $\hat{V}_{\hat{X}}^2 = \hat{V}_{\hat{X}\hat{X}}$  und  $\hat{V}_{\hat{Y}}^2 = \hat{V}_{\hat{Y}\hat{Y}}$ . Einzelheiten über Schätzfunktionen für Verhältniszahlen, Totalwerte oder Anteile im Fall von geschichteten Klumpenstichproben mit gleichen oder ungleichen Auswahlwahrscheinlichkeiten findet man z. B. bei Hansen u. a. (1953).

## 6.2 Praktische Beispiele für die Berechnung von Konfidenzintervallen

Bei Hautzinger und Tassaux (1989) findet man praktische Beispiele für die Berechnung von Konfidenzintervallen nach dem in Abschnitt 6.1 beschriebenen Standardverfahren. Hierbei wurde eine Haushaltsbefragung vom KONTIV-Typ zugrundegelegt, bei welcher G=10 Haushaltsgruppen unterschieden wurden. Der Untersuchungszeitraum umfaßte T=84 Tage und es wurden H=7 Tagestypen (nach dem Wochentag) gebildet. Die Zahl der Schichten betrug demnach GH=70.

Auf der Basis von n=4460 antwortenden Haushalten mit zusammen z=9640 Personen im Alter von 10 Jahren und mehr erhielt man folgende Schätzung:

Tabelle 6.1

Schätzwerte für ausgewählte Maßzahlen sowie zugehörige Variationskoeffizienten

Maßzahl	Schätzwert	Variationskoeffizient des Schätzwerts <sup>1</sup>
Gesamtzahl der Wege	16,6 · 10 <sup>6</sup> Wege	0,013
Gesamtzahl der Wege pro Person und Tag	2,66 Wege	0,008
Gesamtlänge der Wege	13,2 · 10 <sup>7</sup> km	0,049
Mittlere Wegelänge	7,97 km	0,047
Zurückgelegte Wegelänge pro Person und Tag	21,22 km	0,046
Gesamtdauer aller Wege	40,08 · 10 <sup>7</sup> min	0,029
Zeitaufwand pro Person und Tag	64,52 min	0,019
Anteil nichtmobiler Personen	24,5 %	0,028
Anteil Fußwege	33,0 %	0,024

1) Wurzel aus der relativen Varianz

Die relative Genauigkeit der einzelnen Schätzwerte differiert, wie man sieht, beträchtlich. Geht man davon aus, daß die Schätzfunktionen näherungsweise normal verteilt sind, so kann man Konfidenzintervalle für die Maßzahlen berechnen.

Ein 95-Prozent-Konfidenzintervall für eine zu schätzende Verhältniszahl r kann wie folgt geschrieben werden:

$$(6.2.1) \quad \hat{R} \pm 2 \hat{R} \hat{V}_{\hat{R}}$$

Für Totalwerte erhält man eine analoge Darstellung. Mit Hilfe von (6.2.1) kann man aus den Daten der Tabelle 6.1 leicht Konfidenzintervalle berechnen. Beispielsweise ergibt sich:

Tabelle 6.2

Konfidenzintervalle für ausgewählte Maßzahlen

Maßzahl	Schätzwert	95-Prozent-Konfidenzintervall
Zahl der Wege pro Person und Tag	2,66 Wege	2,62 bis 2,70
Zurückgelegte Wegelänge pro Person und Tag	21,2 km	19,3 bis 21,1

Obwohl im vorliegenden Fall der Stichprobenumfang groß ist (4.460 Haushalte mit 9.640 Personen und insgesamt 25.317 Wegen), sind einige Schätzwerte doch erstaunlich ungenau. Dies bedeutet, daß bei der Ergebnisinterpretation Vorsicht angezeigt ist.

Bei Hochrechnungen schätzt man beispielsweise die mittlere Wegelänge nicht nur für die Gesamtheit aller Haushalte während der gesamten Untersuchungsperiode. Vielmehr interessiert man sich normalerweise für die mittlere Wegelänge aufgliedert nach bestimmten Merkmalen des Weges (z. B. Wegezweck), der Person (z. B. Berufstätigkeit), des Tages (z. B. Wochentag) oder des Haushalts (z. B. Pkw-Besitz). Wenn man sich bei der Schätzung auf Wege beschränkt, die eine bestimmte Eigenschaft haben (z. B. Arbeitswege von Teilzeitbeschäftigten), so wird als Folge dieses „Untergruppeneffekts“ die Schätzung der mittleren Wegelänge zwangsläufig ungenauer.

Man kann dies leicht einsehen. Wir bezeichnen mit  $x_{ghi}$  die Gesamtlänge aller Arbeitswege von Teilzeitbeschäftigten aus dem i-ten Element (Haushaltstag) der Schicht (g, h). Für die meisten Elemente  $i$  ( $i=1, \dots, N_{gh}$ ) wird  $x_{ghi} = 0$  gelten, einfach deshalb, weil es in den meisten Haushalten gar keine Teilzeitbeschäftigten gibt. Für Elemente mit einem von 0 verschiedenen  $x_{ghi}$  wird diese Variable dagegen meist relativ große Werte annehmen. Als Folge hiervon ist die Variabilität des Merkmals „Länge der Arbeitswege von Teilzeitbeschäftigten“ relativ groß und die Schätzung entsprechend ungenau.

Tabelle 6.3 zeigt, wie bereits eine einfache Aufgliederung nach der Berufstätigkeit der Person zu einer ganz beträchtlichen Erhöhung der Variationskoeffizienten und damit zu einer Verbreiterung der Konfidenzintervalle führt.

Tabelle 6.3  
Geschätzter Variationskoeffizient der Wegelänge pro Person und Tag  
gegliedert nach Berufstätigkeit der Person

Berufstätigkeit der Person	Fallzahl in der Stichprobe		Wegelänge (km) pro Person und Tag	
	Personen	Wege	Schätz- wert	Variations- koeffizient
Hausfrau	911	2.555	16,7	0,182
Rentner	44	123	51,2	0,721
Lehrling	435	1.099	20,3	0,063
Schüler	1.363	3.544	17,4	0,319
Arbeitslose	206	597	21,2	0,170
Vollberufstätige	5.776	15.009	22,5	0,044
Teilberufstätige	354	1.035	17,3	0,117
keine Angaben	551	1.355	20,7	0,161
Insgesamt	9.640	25.317	21,2	0,046

Während für die Gesamtheit aller Personentage die Konfidenzschätzung (Sicherheitsgrad 95 Prozent) für die Kennziffer „Wegelänge pro Person und Tag“ durch

$$21,2 \pm 2,0 \quad (\text{km pro Person und Tag})$$

gegeben ist, lautet die entsprechende Schätzung für die Gruppe Hausfrauen

$$16,7 \pm 6,1 \quad (\text{km pro Person und Tag})$$

Die Genauigkeit der Schätzung hängt natürlich stark von der Variabilität des Untersuchungsmerkmals ab. Insofern ist es nicht überraschend, daß Schätzungen, die sich auf Wegelängen beziehen, am wenigsten genau sind.

Auch der Typ der Kennziffer hat Einfluß auf die Schätzungsgenauigkeit. Am höchsten ist die stichprobenbedingte Unsicherheit bei der Schätzung von Totalwerten. Verhältniszahlen (personen- oder wegebezogen) lassen sich mit höherer Genauigkeit schätzen. Bei *Hautzinger* und *Tassaux* (1989) findet man im Methoden-Anhang 6 umfangreiches Zahlenmaterial hierzu.

### 6.3 Quantifizierung des Design-Effekts

Zur Berechnung von Konfidenzintervallen nach den eben beschriebenen Verfahren ist spezielle Software erforderlich. Sofern nur Standard-Statistiksoftware verfügbar ist, kann man Konfidenzintervalle näherungsweise berechnen, falls man eine Vorstellung von der Größenordnung des „Design-Effekts“ hat. Unter dem Design-Effekt versteht man hierbei das Verhältnis der Varianz des betrachteten Schätzers unter dem Design „Geschichtete Auswahl von Haushaltstagen mit großenproportionalen Auswahlwahrscheinlichkeiten“ zur Varianz desselben Schätzers unter dem Design „Uneingeschränkt zufällige Auswahl von Personentagen bzw. Wegen“.

Tabelle 6.4 zeigt, daß die Konfidenzintervalle, die man erhält, wenn man der Einfachheit halber von uneingeschränkt zufälliger Auswahl (statt von Klumpenauswahl) ausgeht, teilweise erheblich zu eng sind. Das 95-Prozent-Konfidenzintervall für den durchschnittlichen Anteil immobiler Personen („Häuslichkeitsquote“) beträgt

23,7% bis 25,3% bei uneingeschränkter Zufallsauswahl

und

23,1% bis 25,9% bei Klumpenauswahl.

Diese Konfidenzintervalle ergeben sich auf der Basis von 9.640 antwortenden Personen und bei einem Schätzwert von 24,5 Prozent für den Anteil immobiler Personen.

Tabelle 6.4  
Standardabweichung der Schätzung des Anteils immobiler Personen

Berufs- tätigkeit	Anzahl Personen	Anteil immobiler Personen	Standardabweichung der Schätzung	
			Design A	Design B
Hausfrauen	911	0,496	0,016	0,017
Rentner	44	0,513	0,062	0,075
Lehrlinge	435	0,227	0,026	0,020
Schüler	1.363	0,299	0,018	0,012
Arbeitslose	206	0,466	0,036	0,035
Vollberufstätige	5.776	0,157	0,008	0,004
Teilberufstätige	354	0,183	0,027	0,021
keine Angaben	551	0,392	0,027	0,021
Insgesamt	9.640	0,245	0,007	0,004

Design A: Geschichtete Klumpenauswahl mit großenproportionalen Auswahlwahrscheinlichkeiten

Design B: Uneingeschränkt zufällige Auswahl von Personentagen



Ganz allgemein ist folgendes festzustellen (Hautzinger und Tassaux, 1989): Das Konfidenzintervall für die Verkehrsbeteiligungsquote von Personen des Typs  $k$  ist bei einer Klumpenauswahl von Haushalten umso größer,

- je kleiner die durchschnittliche Zahl  $m(k)$  der Personen des Typs  $k$  pro Haushalt ist, d. h. je seltener der betreffende Personentyp vorkommt,
- je gleichmäßiger die Personen des Typs  $k$  auf die Haushalte verteilt sind und
- je stärker die haushaltsspezifische Verkehrsbeteiligungsquote von Personen des Typs  $k$  zwischen den Haushalten streut.

Je nach betrachteter Personengruppe muß also von einem unterschiedlichen Design-Effekt bei der Schätzung der Verkehrsbeteiligungsquote ausgegangen werden. Für einstufige Klumpenstichproben und einer Typisierung der Personen nach der Berufstätigkeit liegt der Design-Effekt zwischen 1,4 und 3,4.

Für andere Kennziffern läßt sich der Design-Effekt (Klumpeneffekt) in entsprechender Weise quantifizieren. Sofern die Auswahl der Haushalte zweistufig erfolgt, was vor allem bei bundesweiten Erhebungen die Regel ist, kommt zum Klumpeneffekt noch ein Stufungseffekt. Näheres hierzu findet man bei Wermuth (1986).

Zusammenfassend kann man als Faustregel festhalten: Um Klumpen- und Stufungseffekte zu berücksichtigen, sollte man je nach Fragestellung den unter der Annahme „Uneingeschränkte Zufallsauswahl“ mit herkömmlicher Statistik-Software berechneten Standardfehler der Schätzung etwa verdoppeln bis verdreifachen, um zu einer realistischeren Konfidenzschätzung zu kommen.

## 7. Genauigkeitsbeurteilung bei gewichteten Daten

### 7.1 Einfluß der Gewichtung auf die Unverzerrtheit der Schätzung

Durch Gewichtung und Datenkorrektur soll, wie bereits ausgeführt, zum einen der Zufallsfehler der Schätzung verkleinert und zum anderen eine mögliche Verzerrung der Schätzung zumindest teilweise behoben werden. Wie die nachfolgenden Überlegungen zeigen, ist Gewichtung aber keineswegs gleichbedeutend mit Unverzerrtheit.

Wenn man nach Randhäufigkeiten gewichtet, um durch Einbringung von Zusatzinformationen die Schätzgenauigkeit zu erhöhen, so kommt es in der Praxis immer wieder vor, daß aus der amtlichen Statistik nicht die eigentlich interessierende mehrdimensionale Randverteilung bekannt ist, sondern lediglich mehrere eindimensionale Randverteilungen vorliegen. In einer Arbeit von Stenger (1986) wird gezeigt, daß – uneingeschränkte Zufallsauswahl von Haushalten und vollständige Antwortbereitschaft vorausgesetzt – das Gewichten nach mehreren eindimensionalen Randverteilungen zu asymptotisch unverzerrten Schätzfunktionen führt. Auch das Gewichten nach einer mehrdimensionalen Randverteilung liefert im übrigen nur asymptotisch unverzerrte Schätzungen.

Wenn man gewichtet, um ausfallbedingte Verzerrungen zu korrigieren, so ist Vorsicht angezeigt, da es in dieser Situation wesentlich darauf ankommt, ob in den einzelnen Hochrechnungsgruppen gleiche Antwortbereitschaft vorliegt. Die oben angesprochene asymptotische

Unverzerrtheit der beiden gewichteten Schätzungen gilt nämlich nur unter der zusätzlichen Voraussetzung gleicher Antwortbereitschaft in den Hochrechnungsgruppen. Wenn dagegen – was in der Praxis wohl die Regel ist – die Antwortquoten systematisch zwischen den Gruppen variieren, so kommt es bei Gewichtung nach zwei oder mehreren eindimensionalen Randverteilungen zu Verzerrungen: Der Erwartungswert der Schätzfunktion stimmt auch asymptotisch nicht mit der zu schätzenden Kennziffer des Verkehrsverhaltens überein. Man hat dann die paradoxe Situation, daß wegen der unterschiedlichen Antwortbereitschaft auch die gewichtete Schätzung verzerrt ist, obwohl die Gewichtung gerade mit der unterschiedlichen Antwortbereitschaft motiviert wird.

Eine zusätzliche Schwierigkeit entsteht, wenn die Auswahl der Haushalte nicht uneingeschränkt zufällig, sondern mit variierenden Auswahlwahrscheinlichkeiten erfolgt, was in der Erhebungspraxis häufig der Fall ist. Bei Stenger (1985) wird an einem Beispiel gezeigt, daß dann, wenn aus methodischen Gründen nach der Haushaltsgröße gewichtet werden muß (etwa weil Haushalte aus einer Personendatei ausgewählt wurden), eine zusätzliche Gewichtung nach der Zahl der Pkw im Haushalt zu Verzerrungen führen kann. Es wird dabei deutlich, daß die Gewichtungen, mit denen zufallsbedingte Unvollkommenheiten behoben werden sollen, auf die Gewichtung abzustimmen sind, die aus methodischen Gründen erforderlich ist.

### 7.2 Einfluß der Gewichtung auf den Standardfehler der Schätzung

Wenn ausfallbedingte Verzerrungen vernachlässigt werden können und die gewichteten Schätzungen demnach asymptotisch unverzerrt sind, so wird z. B. von Cochran (1977) gezeigt, daß die Varianz einer Schätzfunktion mit Gewichtung nach einer mehrdimensionalen Randverteilung asymptotisch kleiner ist als die Varianz der entsprechenden Schätzfunktion ohne Gewichtung. Die Einbeziehung von Zusatzinformationen in Form der mehrdimensionalen Zensus-Verteilung bringt den erwünschten Genauigkeitsgewinn. Bei hinreichend großem Stichprobenumfang innerhalb jeder Hochrechnungsgruppe (etwa  $n > 20$ ) erzielt man fast denselben Effekt wie durch Schichtung mit proportionaler Aufteilung des Gesamtstichprobenumfangs.

Wird dagegen gleichzeitig nach mehreren eindimensionalen Randverteilungen gewichtet, so ist die oben beschriebene Varianzverkleinerung nicht sichergestellt. Wie bei Stenger (1986) gezeigt wird, kann eine derartige Gewichtung, in die drei oder mehr Merkmale einbezogen werden, sogar zu einer Varianzvergrößerung (verglichen mit der ungewichteten Schätzung) führen. Dieser unerwünschte Effekt tritt aber nur bei extremer Abhängigkeit der zur Gewichtung benutzten Merkmale ein. Sind die drei oder mehr Gewichtungsmerkmale unabhängig, so ist die Varianz der gewichteten Schätzung im allgemeinen kleiner als die Varianz der Schätzung ohne Verwendung von Zusatzinformationen.

### 7.3 Vereinfachte Fehlerrechnung bei komplexen Stichprobendesigns

Im Abschnitt 6 war gezeigt worden, wie Varianzschätzungen durchgeführt werden können, wenn unter dem Design „Unabhängige Teilstichproben“ das Auswahlverfahren zur Klasse der Standardverfahren der Stichprobentheorie gehört. Die dort behandelte Schätztechnik ist

bereits - theoretisch und praktisch - relativ kompliziert. Die Varianzschätzung wird weiter erschwert, wenn es sich um mehrstufige Verfahren der Haushaltsauswahl handelt oder wenn das Design „Partielle Rotation“ vorliegt. Die angesprochenen Erschwerungen treten wohl-gemerkt bereits dann ein, wenn die Erhebung vollkommen nach Plan verläuft, d. h. insbesondere keine Ausfälle und Angabefehler zu berücksichtigen sind.

Falls Gewichtungen und/oder Datenkorrekturen notwendig oder zumindest zweckmäßig sind, wird eine Varianzschätzung im klassischen Sinn rasch unmöglich. Dies gilt ebenso in dem Fall, wo - aus welchen Gründen auch immer - das verwendete Auswahlverfahren nicht zum Standardrepertoire der Stichprobentheorie gehört. Ein typisches Beispiel hierfür ist etwa die KONTIV 82, bei der man sich von den stichprobentheoretischen Standardverfahren so weit entfernte, daß eine konventionelle statistische Effizienzbeurteilung allenfalls noch in grober Näherung möglich ist.

In Situationen der eben beschriebenen Art ist eine vereinfachte Fehlerrechnung möglich, wenn das Stichprobenverfahren von vornherein replikativ oder doch zumindest quasi-replikativ gestaltet wird. Das Konzept der replikativen Stichproben wird ausführlich bei *Deming* (1960) sowie bei *McCarthy* (1966 und 1969) beschrieben. Als replikativ werden Stichprobenverfahren bezeichnet, bei denen ein beliebiges zufälliges Auswahlverfahren mehrfach, sagen wir  $K$ -mal, unabhängig durchgeführt und im Anschluß an jede Durchführung eine unverzerrte Schätzung nach demselben Prinzip berechnet wird (*Stenger*, 1985). Die so gewonnenen (unverzerrten) Schätzungen

$$(7.3.1) \quad \hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_K$$

sind identisch verteilt und unabhängig.

Bei der Konstruktion der Schätzfunktionen können ohne weiteres auch Gewichtungen vorgenommen werden. Wichtig ist hierbei allerdings, daß die Gewichte für jede Teilstichprobe (Replikation) neu berechnet werden. Die Bildung vieler Hochrechnungsgruppen für die nachträgliche Schichtung mit entsprechend geringen Fallzahlen pro Gruppe beeinträchtigt die Wirksamkeit des replikativen Verfahrens nicht.

Liegen unabhängige Schätzungen gemäß (7.3.1) vor, so ist

$$(7.3.2) \quad \hat{Y} = (1/K) \sum \hat{Y}_i$$

die endgültige Schätzung für die interessierende Kennziffer. Die Varianz der Schätzung  $\hat{Y}$  kann durch

$$(7.3.3) \quad \frac{1}{K(K-1)} \sum (\hat{Y}_i - \hat{Y})^2$$

geschätzt werden.

Die Punktschätzung  $\hat{Y}$  hängt nicht wesentlich davon ab, wieviele Teilstichproben gebildet wurden. Dagegen unterliegt die Schätzung der Varianz von  $\hat{Y}$  um so größeren Zufallsschwankungen, je kleiner die Zahl  $K$  der Replikationen ist. Man sollte deshalb normalerweise mindestens fünf Teilstichproben bilden ( $K \geq 5$ ). Prinzipiell kommt aber auch der Fall  $K=2$  in Betracht.

Mit Hilfe von (7.3.2) und (7.3.3) lassen sich Konfidenzintervalle in der üblichen Weise konstruieren, falls zumindest eine der beiden Bedingungen

- $K$  hinreichend groß
- $\hat{Y}_1, \dots, \hat{Y}_K$  näherungsweise normalverteilt

erfüllt ist. Im Fall  $K=2$  hat (annähernde Normalverteilung vorausgesetzt) das Konfidenzintervall zum Sicherheitsgrad 95 Prozent die besonders einfache Form

$$(7.3.4) \quad \hat{Y} \pm |\hat{Y}_1 - \hat{Y}_2|$$

wobei  $\hat{Y} = (\hat{Y}_1 + \hat{Y}_2)/2$

In der Praxis kann es unter Umständen günstiger sein, „quasi-replikativ“ vorzugehen. In diesem Fall wird zunächst eine Gesamtstichprobe gezogen. Diese Stichprobe zerlegt man dann uneingeschränkt zufällig in  $K$  Teilstichproben von jeweils gleichem Umfang. Wie zuvor berechnet man nun wieder (gegebenenfalls nach Gewichtung getrennt für die Teilstichproben) vorläufige Schätzungen  $\hat{Y}_1, \dots, \hat{Y}_K$ . Diese werden gemäß (7.3.2) zu einer endgültigen Schätzung  $\hat{Y}$  zusammengeführt. Die Varianz von  $\hat{Y}$  kann auch beim quasi-replikativen Design durch (7.3.3) geschätzt werden. Der Erwartungswert der Schätzfunktion (7.3.3) ist im Fall einer quasi-replikativen Vorgehensweise geringfügig größer als beim replikativen Verfahren. Konfidenzintervalle fallen beim quasi-replikativen Vorgehen also etwas größer aus als bei replikativen Stichproben.

Die Varianzschätzung auf der Basis replikativer oder quasi-replikativer Stichproben ist nach Lage der Dinge bei Haushaltsbefragungen vom KONTIV-Typ gegenwärtig das einzig realistische Konzept für methodisch fundierte Genauigkeitsbeurteilungen.

#### 7.4 Ein Anwendungsbeispiel

Ein erstes Anwendungsbeispiel für Varianzschätzungen im Rahmen einer replikativen Haushaltsbefragung vom KONTIV-Typ wird bei *Hautzinger* und *Tassaux* (1989), Anhang 8, beschrieben. Bei dieser Befragung wurden zunächst Gemeindegrößenklassen gebildet und innerhalb jeder Gemeindegrößenklasse eine vorgegebene gerade Anzahl von Gemeinden (Primäreinheiten) uneingeschränkt zufällig ausgewählt. Innerhalb einer jeden Gemeindegrößenklasse wurde in jeder Stichprobengemeinde dieselbe (fest vorgegebene) Anzahl von Adressen gezogen. Die Zahl der auszuwählenden Gemeinden pro Größenklasse und die Zahl der zu befragenden Haushalte pro Gemeinde wurde dabei so festgelegt, daß der Gesamtstichprobenumfang (Haushalte) proportional auf die Gemeindegrößenklassen verteilt wurde.

Für die Hochrechnung wurde getrennt nach Gemeindegrößenklassen die Menge der Stichprobengemeinden nach einem reinen Zufallsverfahren in zwei gleich große Teilmengen zerlegt. Das Vorgehen war also quasi-replikativ mit  $K=2$ . Für beide so entstandenen Teilstichproben von Haushalten bzw. Haushaltstagen wurde jeweils eine nachträgliche Schichtung nach Haushaltsgröße, Gemeindegröße und Wochentagstyp (dreidimensionale Verteilung) und auf der Personenebene eine Anpassung an eine dreidimensionale Randverteilung



(Altersgruppe, Geschlecht und Wochentagstyp) vorgenommen. Die Hochrechnungsfaktoren wurden so gewählt, daß sich die Schätzwerte (z. B. für die Gesamtzahl aller Wege) auf den Zeitraum von einer Woche beziehen. Die Erhebung selbst fand im wesentlichen im Monat Dezember statt. Verwertbare Angaben konnten von 811 Haushalten (bereinigte Antwortquote 47 Prozent) erhalten werden.

Die Vorgehensweise bei der Varianzschätzung soll am Beispiel der Kennziffern „Gesamtzahl der Wege“ und „Wege pro Person und Tag“ (Personen ab 10 Jahren) dargestellt werden. Es ergaben sich folgende Schätzwerte:

Replikation	Wege pro Woche (in Mio.)	Wege pro Person und Tag
1	1.055,2	2,91
2	1.084,9	2,99
Mittelwert	1.070,1	2,95

Hieraus erhält man unter Verwendung von (7.3.4) folgende Konfidenzintervalle jeweils zum Sicherheitsgrad 95 Prozent:

Wege pro Woche (in Mio.)  
1070,1  $\pm$  29,7 oder (1040,4; 1099,8)

Wege pro Person und Tag  
2,95  $\pm$  0,08 oder ( 2,87; 3,03)

Die Tabelle 7.1 zeigt Ergebnisse von Konfidenzschätzungen, welche sich auf die Verkehrsmittelbenutzung bzw. auf Wegezwecke beziehen. Man erkennt, daß die Schätzung der Verkehrsmittelanteile generell weniger genau ausfällt als die Schätzung der Anteile der einzelnen Wegezwecke.

Die Rechnung hat wegen  $K=2$  aber doch mehr exemplarischen Charakter. Bei drei Wegezwecken (Wirtschaftsweg, Inanspruchnahme von Dienstleistungen, Einkauf) unterscheiden sich die Schätzwerte aus den beiden Teilstichproben jeweils erst auf der zweiten Nachkommastelle, was zu sehr engen Konfidenzintervallen führt. Hier ist nochmals auf die bei kleiner Anzahl von Replikationen gegebene Unsicherheit der Varianzschätzung hinzuweisen.

Tabelle 7.1

Ergebnisse von Konfidenzschätzungen bei quasi-replikativem Design ( $K=2$ )

Wegetyp	Schätzung der rel. Hfkt. des Wegetyps (%)			
	Repl. 1	Repl. 2	Insg.	95-Prozent-KI
<u>Verkehrsmittel*</u>				
zu Fuß	33,4	31,1	32,2	30,0 bis 34,6
Fahrrad	6,1	9,5	7,8	4,4 bis 11,2
Mofa/Moped/ Motorrad	1,6	0,9	1,3	0,6 bis 2,0
Pkw-Fahrer	39,8	38,1	39,0	37,3 bis 40,7
Pkw-Mitfahrer	11,4	11,8	11,6	11,2 bis 12,0
Bus	3,9	6,3	5,1	2,7 bis 7,5
Übriger öV	2,7	1,5	2,1	0,9 bis 3,3
Sonstiges	1,0	0,8	0,9	0,7 bis 1,1
Summe	100,0	100,0	100,0	
<u>Verkehrszweck</u>				
zur Arbeit	11,6	11,9	11,8	11,5 bis 12,1
dienstlich	2,8	2,7	2,8	2,7 bis 2,9
zur Ausbildung	4,5	4,4	4,5	4,4 bis 4,6
zum Einkauf	14,7	14,7	14,7	14,7 bis 14,7
Wirtschaftsweg	0,1	0,1	0,1	0,1 bis 0,1
Serviceweg	3,4	2,8	3,1	2,5 bis 3,7
Freizeitweg	17,3	17,1	17,2	17,0 bis 17,4
Dienstleistung	3,2	3,2	3,2	3,2 bis 3,2
nach Hause	42,4	43,0	42,7	42,1 bis 43,3
Zweck unbekannt	—	0,1	0,1	—
Summe	100,0	100,0	100,0	

\*) Erste Nennung

Ergänzend sei noch darauf hingewiesen, daß die praktische Durchführung der Hochrechnung und Genauigkeitsbeurteilung nach dem eben beschriebenen Verfahren vergleichsweise unkompliziert ist. EDV-technisch kann, sofern jedem Datensatz der Auswertungsdatei die Replikationsnummer hinzugefügt wurde, die Hochrechnung der Teilstichproben mit her-

kömmlichen Statistik-Softwarepaketen (z.B. SPSSX oder SAS) problemlos durchgeführt werden.

Wegen der Einfachheit des replikativen Verfahrens können die Varianzen der Schätzungen – und damit die Konfidenzintervalle – mit Computerprogrammen berechnet werden, die vom Anwender leicht selbst erstellt werden können.

## 8. Resümee und Ausblick

Haushaltsbefragungen vom KONTIV-Typ haben sich in der Vergangenheit als Instrument der empirischen Verkehrsforschung in überzeugender Weise bewährt. Durch eine Vielzahl von Forschungsanstrengungen konnten in den letzten Jahren auch die statistischen Grundlagen für

- Hochrechnung
- Gewichtung und
- Genauigkeitsbeurteilung

von derartigen Verkehrserhebungen geschaffen werden. Zu allen diesen Aspekten der Stichprobenauswertung wurden inzwischen sowohl theoretische Erkenntnisse als auch praktische Erfahrungen gewonnen.

Mit der vorliegenden Arbeit wurde versucht, eine zusammenfassende Darstellung und Bewertung des erreichten Forschungsstands zu geben. Man kann durchaus sagen, daß die Entwicklung einer statistischen Methodik für Haushaltsbefragungen vom KONTIV-Typ einen gewissen vorläufigen Abschluß gefunden hat. Dies gilt zumindest für Erhebungen, die nach dem Design „Unabhängige Teilstichproben“ durchgeführt werden. Theoretische Weiterentwicklungen sind hier wohl nur im Zusammenhang mit neuen empirischen Erkenntnissen über strukturelle Unterschiede zwischen Antwortern und Nichtantwortern sowie mit Erkenntnissen über Art und Ausmaß von Angabefehlern der Befragten zu erwarten. Insofern müssen weitere Forschungsaktivitäten zunächst vorzugsweise in diese Richtung gehen.

Natürlich kann das KONTIV-Design als solches – also die schriftlich-postalische Haushaltsbefragung unter Verwendung von Tagesprotokollen – nicht pauschal als „Königsweg“ oder ultima ratio der empirischen Verkehrsforschung angesehen werden. Verbesserungen der Erhebungstechnik scheinen möglich, etwa durch Kombination von schriftlicher und telefonischer Befragung. Ein solcher Übergang zu einem Methoden-Mix brächte allerdings auch wieder neue statistische Aufgabenstellungen in den Bereichen Hochrechnung, Gewichtung und Genauigkeitsbeurteilung mit sich.

## Literatur

- Barnett, V.: Elements of Sampling Theory, Hodder and Stoughton, London, 1984  
 Bölken, F.: Auswahlverfahren, Teubner-Verlag, Stuttgart, 1976  
 Brög, W., Meyer, N.: From Price Elasticity to Price Susceptibility, Socialdata, München, 1978  
 Brög, W., Heuwinkel, D., Neumann, K.H.: Psychological Determinants of User Behaviour, 34th Round Table on Transport Economics, CEMT, Paris, 1976  
 Cochran, W.G.: *Sampling Techniques*, Wiley, New York, 1977  
 Deming, W.E.: Sampling Design in Business Research, Wiley, New York, 1960  
 Hansen, M.H., Hurwitz, W.N.: The Problem of Non-Response in Sample Surveys, Journal of the American Statistical Association, Vol. 41, 1946, S. 517-529  
 Hansen, M.H., Hurwitz, W.N., Madow, W.G.: Sample Survey Methods and Theory, Bd. I und II, Wiley, New York, 1953  
 Hautzinger, H.: Pkw-Besitz der deutschen Wohnbevölkerung: Ein kritischer Vergleich von EVS- und KONTIV-Ergebnissen, Internationales Verkehrswesen, 38. Jg., Heft 3, 1986, S. 184-188  
 Hautzinger, H.: Statistische Designprinzipien für kontinuierliche Verkehrserhebungen, Schriftenreihe der DVWG, Band B97, Bergisch Gladbach, 1987, S. 5-25  
 Hautzinger, H., Schorer, B.: Statistik-Programmsystem zur Schätzung von Merkmalssummen und Verhältniszahlen sowie deren Varianzen und Variationskoeffizienten (1985), in: *Hautzinger und Tassaux* (1989), Methoden-anhang  
 Hautzinger, H., Stenger, H.: Hochrechnungsverfahren zur Infratest-Erhebung „Reiseausgaben im Ausland“, Gutachten, Heilbronn, 1986  
 Hautzinger, H., Tassaux, B.: Verkehrsmobilität und Unfallrisiko in der Bundesrepublik Deutschland, Schlußbericht zum BAST-Forschungsprojekt FP 8303, Forschungsberichte der BAST, Band 195, Bergisch Gladbach, 1989  
 Hendricks, W.A.: Adjustment for Bias by Non-Response in Mailed Surveys, Agr. Econ. Res., Vol. 1, 1949, S. 52-56  
 Hengst, M.: Einführung in die mathematische Statistik und ihre Anwendung, Mannheim, 1967  
 Herz, R.: Diskussionspapier zur Projektgruppensitzung „Verbesserung der Gewichtung und Hochrechnung von Haushaltsbefragungen vom KONTIV-Typ“, unveröffentlichtes Manuskript, Karlsruhe, 1987  
 Krug, W., Nourney, M.: Wirtschafts- und Sozialstatistik: Gewinnung von Daten, Oldenbourg-Verlag, München, 1982  
 Kutter, E., Holz, R.: Diskussionspapier zur Projektgruppensitzung „Verbesserung der Gewichtung und Hochrechnung von Haushaltsbefragungen vom KONTIV-Typ“, unveröffentlichtes Manuskript, Berlin, 1987  
 McCarthy, P.J.: Replication: An Approach to the Analysis of Data from Complex Surveys, Nation Center for Health Statistics, Washington, D.C., Series 2, 14, 1966  
 McCarthy, P.J.: Pseudo Replication: Half-Samples, Rev. Int. Stat. Inst., 37, 1969, S. 239-264  
 Morris, N.: The Accuracy of Commuter's Reported Walking Times: A Comparison of First World and Third World Commuters. Beitrag zur Konferenz "New Survey Methods in Transport", Hungerford Hill (Australien), 1983  
 Roth, E.: Sozialwissenschaftliche Methoden, Oldenbourg-Verlag, München, 1987  
 Socialdata: Marktanalyse des Verkehrsverhaltens im Personenfernverkehr, Expertenband, München, 1981  
 Socialdata: KONTIV 82 - Bericht, München, 1984  
 Socialdata: Erhebungen zur Ermittlung von Fußwegen und "non-reported trips", Untersuchung im Auftrag des BMV, München, 1986  
 Sozialforschung Brög: KONTIV - Methoden, Band I, München, 1975  
 Statistisches Bundesamt (Hrsg.): Stichproben in der Amtlichen Statistik, Kohlhammer-Verlag, Stuttgart, 1960  
 Stenger, H.: Varianzschätzung bei komplexen Stichprobenerhebungen, Schriftenreihe der DVWG, Band B81, Bergisch Gladbach, 1985, S. 143-150  
 Stenger, H.: Stichproben, Physica-Verlag, Heidelberg/Wien, 1986  
 Stenger, H.: Zur heutigen Gewichtungspraxis in Markt- und Meinungsforschung, Manuskript (erscheint demnächst), 1986  
 Strecker, H. u.a.: Messung der Antwortvariabilität auf Grund von Erhebungsmodellen mit Wiederholungszählungen, Vandenhoeck & Ruprecht, Göttingen, 1983  
 SWOV (Hrsg.): The Extent to Which People Can Estimate Their Trip Distances, Voorburg (Niederlande), 1978  
 Wermuth, M.: Datenerhebung und Datenanalyse im Verkehrswesen, Schriftenreihe der DVWG, Band B66, 1983, S. 13-43



- Wermuth, M.: Schriftliche Haushaltsbefragungen zum Verkehrsverhalten - Stichprobenauswahl und Nonresponse-problematik, Schriftenreihe der DVWG, Band B81, Bergisch Gladbach, 1985, S. 151-175
- Wermuth, M.: Theoretische Anforderungen in der Stichprobenplanung, Schriftenreihe der DVWG, Band B85, Bergisch Gladbach, 1986, S. 137-163
- Wermuth, M., Maerschalk, G.: Zur Repräsentanz schriftlicher Haushaltsbefragungen zum Verkehrsverhalten, Untersuchung im Auftrag des BMV, München, 1981
- Wermuth, M., Maerschalk, G.: Entwicklung von Verfahren für Verkehrserhebungen aus der Sicht der Empirischen Sozialforschung und der Ingenieurwissenschaften, Untersuchung im Auftrag des BMV, München, 1983

### Summary

The paper contains a deep and comprehensive description and assessment of the statistical methodology for KONTIV-type household travel surveys. After a characterization of KONTIV-type surveys from the viewpoint of sampling theory, the different sources of errors (sampling error, procedural errors) are discussed in detail.

Subsequently, measures for reducing the sampling error and weighting procedures for correcting biases are presented. Such biases may occur due to nonresponse and/or individual response errors.

Finally, the problem of measuring the accuracy of a KONTIV-type survey is treated for the case of unweighted and weighted data. In this context it is shown, how approximative confidence intervals can be calculated.

v. st. e.

## Die Entwicklung des Güterverkehrsaufkommens nach Verkehrszweig und Güterart

VON HANS JOSEF KAYSER UND FRANZ-JOSEF HOLZMÜLLER

### 1. Einleitung

In der Diskussion über die künftige Ausrichtung der Verkehrspolitik werden Investitionen in die Verkehrswege zunehmend kritisch beurteilt. Von dieser Einschätzung, die u. a. auf ein verstärktes Umweltbewußtsein zurückgeführt wird, ist insbesondere der Straßenverkehr betroffen, der durch die Inanspruchnahme natürlicher Ressourcen (Flächenbedarf, Energieverbrauch) sowie die Emissionen von Lärm und Schadstoffen die natürliche Umwelt belastet. So wird der Verzicht auf Neu- und Ausbaumaßnahmen im Straßennetz und eine möglichst weitgehende Verlagerung des Verkehrsaufkommens von der Straße auf die Schiene gefordert, wobei der Güterverkehr zunehmend in den Blickpunkt des öffentlichen Interesses gerät. Es ist daher von Interesse, das Geschehen und die Entwicklung auf dem Güterverkehrsmarkt näher zu betrachten.

### 2. Datengrundlage

Als wesentliche Grundlage für die folgende Darstellung dienen die aus der amtlichen Statistik des Statistischen Bundesamtes zusammengestellten Angaben zum Güterverkehr.<sup>1)</sup> Dort wird das Verkehrsaufkommen (t/a) ausgewiesen nach Verkehrszweig, Hauptverkehrsbeziehungen, Güterart, Verkehrsbezirken und Entfernungsstufen.

Die Besonderheiten der verschiedenen Verkehrszweige, die neben den technologisch-organisatorischen Maßnahmen der verschiedenen Verkehrsbetriebe auch von den ordnungspolitischen Rahmenbedingungen abhängig sind, wirken sich auf die Art der Erfassung aus. Die für die Auswertung benötigten Angaben (z. B. Güterart, Versand- und Empfangsverkehrsbezirk) werden im Regelfall aus dem Frachtpapier erhoben. Transporte, für die kein Frachtpapier ausgestellt wird, sind in der Statistik nicht oder nur durch Schätzwerte enthalten (z. B. Straßengüternahverkehr, Stückgutverkehr, Teile des Rohrleitungsverkehrs).<sup>2)</sup>

Durch die unterschiedliche Art der Erfassung sind die Angaben für die verschiedenen Verkehrszweige nur bedingt vergleichbar. Während in den Werten für den Straßenverkehr der

#### *Anschrift der Verfasser:*

Universitätsprofessor Dr.-Ing. Hans Josef Kayser  
Dipl.-Ing., Dipl.-Wirt. Ing. Franz-Josef Holzmüller  
Lehrstuhl und Institut für Straßenwesen, Erd- und Tunnelbau  
der Rheinisch-Westfälischen Technischen Hochschule Aachen  
Mies-van-der-Rohe-Straße 1  
D-5100 Aachen

1) Vgl. Statistisches Bundesamt: Verkehr; Fachserie 8, Reihe 1, Güterverkehr der Verkehrszweige 1987, Stuttgart, Mainz 1988.

2) Ebenda, S. 7.