

Sonderdruck aus:

# Mitteilungen aus der Arbeitsmarkt- und Berufsforschung

Werner Karr

Die G-Kartei als statistisches Instrument in der  
Arbeitsmarkt- und Berufsforschung

3. Jg./1970

**1**

## Mitteilungen aus der Arbeitsmarkt- und Berufsforschung (MittAB)

Die MittAB verstehen sich als Forum der Arbeitsmarkt- und Berufsforschung. Es werden Arbeiten aus all den Wissenschaftsdisziplinen veröffentlicht, die sich mit den Themen Arbeit, Arbeitsmarkt, Beruf und Qualifikation befassen. Die Veröffentlichungen in dieser Zeitschrift sollen methodisch, theoretisch und insbesondere auch empirisch zum Erkenntnisgewinn sowie zur Beratung von Öffentlichkeit und Politik beitragen. Etwa einmal jährlich erscheint ein „Schwerpunktheft“, bei dem Herausgeber und Redaktion zu einem ausgewählten Themenbereich gezielt Beiträge akquirieren.

### *Hinweise für Autorinnen und Autoren*

Das Manuskript ist in dreifacher Ausfertigung an die federführende Herausgeberin Frau Prof. Jutta Allmendinger, Ph. D. Institut für Arbeitsmarkt- und Berufsforschung 90478 Nürnberg, Regensburger Straße 104 zu senden.

Die Manuskripte können in deutscher oder englischer Sprache eingereicht werden, sie werden durch mindestens zwei Referees begutachtet und dürfen nicht bereits an anderer Stelle veröffentlicht oder zur Veröffentlichung vorgesehen sein.

Autorenhinweise und Angaben zur formalen Gestaltung der Manuskripte können im Internet abgerufen werden unter [http://doku.iab.de/mittab/hinweise\\_mittab.pdf](http://doku.iab.de/mittab/hinweise_mittab.pdf). Im IAB kann ein entsprechendes Merkblatt angefordert werden (Tel.: 09 11/1 79 30 23, Fax: 09 11/1 79 59 99; E-Mail: [ursula.wagner@iab.de](mailto:ursula.wagner@iab.de)).

### Herausgeber

Jutta Allmendinger, Ph. D., Direktorin des IAB, Professorin für Soziologie, München (federführende Herausgeberin)  
Dr. Friedrich Buttler, Professor, International Labour Office, Regionaldirektor für Europa und Zentralasien, Genf, ehem. Direktor des IAB  
Dr. Wolfgang Franz, Professor für Volkswirtschaftslehre, Mannheim  
Dr. Knut Gerlach, Professor für Politische Wirtschaftslehre und Arbeitsökonomie, Hannover  
Florian Gerster, Vorstandsvorsitzender der Bundesanstalt für Arbeit  
Dr. Christof Helberger, Professor für Volkswirtschaftslehre, TU Berlin  
Dr. Reinhard Hujer, Professor für Statistik und Ökonometrie (Empirische Wirtschaftsforschung), Frankfurt/M.  
Dr. Gerhard Kleinhenz, Professor für Volkswirtschaftslehre, Passau  
Bernhard Jagoda, Präsident a.D. der Bundesanstalt für Arbeit  
Dr. Dieter Sadowski, Professor für Betriebswirtschaftslehre, Trier

### Begründer und frühere Mitherausgeber

Prof. Dr. Dieter Mertens, Prof. Dr. Dr. h.c. mult. Karl Martin Bolte, Dr. Hans Büttner, Prof. Dr. Dr. Theodor Ellinger, Heinrich Franke, Prof. Dr. Harald Gerfin, Prof. Dr. Hans Kettner, Prof. Dr. Karl-August Schäffer, Dr. h.c. Josef Stingl

### Redaktion

Ulrike Kress, Gerd Peters, Ursula Wagner, in: Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB), 90478 Nürnberg, Regensburger Str. 104, Telefon (09 11) 1 79 30 19, E-Mail: [ulrike.kress@iab.de](mailto:ulrike.kress@iab.de); (09 11) 1 79 30 16, E-Mail: [gerd.peters@iab.de](mailto:gerd.peters@iab.de); (09 11) 1 79 30 23, E-Mail: [ursula.wagner@iab.de](mailto:ursula.wagner@iab.de); Telefax (09 11) 1 79 59 99.

### Rechte

Nachdruck, auch auszugsweise, nur mit Genehmigung der Redaktion und unter genauer Quellenangabe gestattet. Es ist ohne ausdrückliche Genehmigung des Verlages nicht gestattet, fotografische Vervielfältigungen, Mikrofilme, Mikrofotos u.ä. von den Zeitschriftenheften, von einzelnen Beiträgen oder von Teilen daraus herzustellen.

### Herstellung

Satz und Druck: Tümmels Buchdruckerei und Verlag GmbH, Gundelfinger Straße 20, 90451 Nürnberg

### Verlag

W. Kohlhammer GmbH, Postanschrift: 70549 Stuttgart; Lieferanschrift: Heßbrühlstraße 69, 70565 Stuttgart; Telefon 07 11/78 63-0; Telefax 07 11/78 63-84 30; E-Mail: [waltraud.metzger@kohlhammer.de](mailto:waltraud.metzger@kohlhammer.de), Postscheckkonto Stuttgart 163 30. Girokonto Städtische Girokasse Stuttgart 2 022 309. ISSN 0340-3254

### Bezugsbedingungen

Die „Mitteilungen aus der Arbeitsmarkt- und Berufsforschung“ erscheinen viermal jährlich. Bezugspreis: Jahresabonnement 52,- € inklusive Versandkosten; Einzelheft 14,- € zuzüglich Versandkosten. Für Studenten, Wehr- und Ersatzdienstleistende wird der Preis um 20 % ermäßigt. Bestellungen durch den Buchhandel oder direkt beim Verlag. Abbestellungen sind nur bis 3 Monate vor Jahresende möglich.

### Zitierweise:

MittAB = „Mitteilungen aus der Arbeitsmarkt- und Berufsforschung“ (ab 1970)  
Mitt(IAB) = „Mitteilungen“ (1968 und 1969)  
In den Jahren 1968 und 1969 erschienen die „Mitteilungen aus der Arbeitsmarkt- und Berufsforschung“ unter dem Titel „Mitteilungen“, herausgegeben vom Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit.

**Internet:** <http://www.iab.de>

# Die G-Kartei als statistisches Instrument in der Arbeitsmarkt- und Berufsforschung

Werner Karr

Die G-Kartei, eine nach dem Buchstabenprinzip gezogene Stichprobe aus der ehemaligen Totalkartei aller abhängig beschäftigten Arbeiter und Angestellten bei den Arbeitsämtern, sollte als Grundlage einer kurzfristigen Beschäftigtenstatistik der Bundesanstalt für Arbeit dienen. Sie wird seit 1. Januar 1964 geführt. Die aus ihr ermittelten Zahlen für den genannten Zweck schienen jedoch so unplausibel, daß sie nie veröffentlicht wurden.

Es stellte sich nun die Frage, ob die G-Kartei andere statistische Erfordernisse, insbesondere im Rahmen der Arbeitsmarkt- und Berufsforschung, erfüllen kann. In der folgenden Untersuchung wird die Qualität der G-Kartei-Zahlen mit diesen unterschiedlichen statistischen Anforderungen verglichen. Dabei wird insbesondere geprüft, ob die zum Teil bekannten schlechten Ergebnisse auf das Stichproben-Auswahlverfahren zurückzuführen sind oder ob dafür auch andere Gründe in Frage kommen.

Die Untersuchungen stützen sich einerseits auf Vergleiche der G-Kartei mit der ehemaligen Totalkartei, andererseits auf Vergleiche zwischen G-Kartei, Mikrozensus und Industriebericht-erstattung.

Die im Ergebnis gefundenen signifikanten Unterschiede zwischen den hochgerechneten G-Kartei-Zahlen und den Zahlen der ehemaligen Totalkartei, des Mikrozensus und der Industriebericht-erstattung sind zum Teil so beträchtlich, daß die G-Kartei als statistische Grundlage für nur wenige Forschungszwecke in Frage kommt.

Wenn am Schluß der Untersuchung trotzdem für die Weiterführung einer Unterstichprobe aus der G-Kartei plädiert wird, so deshalb, weil gegenwärtig und in absehbarer Zukunft kein vergleichbares statistisches Instrument vorhanden ist bzw. geschaffen wird und weil der Hauptfehler der G-Kartei, nämlich ihr ungenügender Bearbeitungsstand, weitgehend beseitigt werden kann.

Die Untersuchung wurde im Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB), Erlangen, durchgeführt.

## Gliederung

### A. Einleitung

### B. Die Qualität der G-Stichprobe

#### I. Allgemeine Bemerkungen über Fehlermöglichkeiten

#### II. Güte und Verwendungszweck

##### 1. Konjunkturbeobachtung und Konjunkturforschung

##### 2. Strukturforschung

###### a) Querschnittsanalysen, komparativ statische Analysen

###### b) Längsschnittuntersuchungen

##### 3. Auswahlgrundlage

#### III. Die Untersuchungen im einzelnen

##### 1. Untersuchung auf wirtschaftsfachliche Repräsentanz

###### a) Die $\chi^2$ -Methode

###### b) Entwicklung zwischen den Zähltagen 1962 und 1963

###### c) Untersuchung einzelner Wirtschaftszweige

###### d) Ergebnisse und Folgerungen

##### 2. Vergleich der G-Kartei mit anderen Statistiken

###### a) Bestandsvergleich G-Kartei — Mikrozensus für April 1968

###### b) Bestandsvergleich G-Kartei

###### — Industriebericht für April 1968

###### c) Bestandsvergleich im Wirtschaftszweig 45: Deutsche Bundesbahn

###### d) Veränderungsvergleich G-Kartei — Mikrozensus 1966—68

###### e) Veränderungsvergleich G-Kartei — Industriebericht 1966—68

###### f) Altersverteilung der Beschäftigten nach der G-Kartei und dem Mikrozensus im April 1968

###### g) Ergebnisse

##### 3. Die Altersverteilung des ersten Beschäftigungsverhältnisses

##### 4. Über die Berufsstruktur

#### IV. Die G-Kartei als Grundlage für Panel-Untersuchungen

### C. Zusammenfassung

#### A. Einleitung

Die G-Kartei ist die Kartei einer Stichprobe aus der Gesamtheit der deutschen Arbeiter und Angestellten, die nach dem sogenannten Buchstabenauswahlverfahren gezogen wurde. Theoretisch müssen alle Beschäftigten dieses Personenkreises, deren Familienname mit dem Buchstaben G beginnt, in ihr enthalten sein.

Die Grundgesamtheit, der diese Stichprobe entnommen wurde, war die ehemalige Totalkartei der beschäftigten deutschen Arbeitnehmer, die bis 31. 12. 1963 bei den Arbeitsämtern geführt wurde. Diese Totalkartei, die die Basis für eine vierteljährliche Beschäftigtenstatistik gebildet hatte, war aufgrund des mangelhaften Fortschreibungsmechanismus im Laufe der Zeit immer mehr „versandet“, so daß die aus ihr gewonnenen statistischen Ergebnisse sowohl ihrer absoluten Höhe nach als auch hinsichtlich ihrer Strukturen den Anforderungen nicht mehr genügten. Im Jahre 1950 betrug die Überhöhung der aus der Totalkartei gewonnenen Zahlen gegenüber jenen des Statistischen Bundesamtes noch rund 152000, im Jahre 1963 bereits über 1 Million.

Diese Totalkartei hatte inzwischen einen Umfang von über 22 Millionen Karten erreicht, so daß die Vermutung nahe lag, daß neben dem unzureichenden Fortschreibungsmechanismus, der auf §53 des ehemaligen AVAVG<sup>1)</sup> beruhte, auch beträchtliche Bearbeitungsschwierigkeiten zu den schlechten Ergebnissen beitrugen. Dies führte zu der Überlegung, eine Teilkartei einzurichten, von der man annahm, daß sie schon wegen ihres vergleichsweise geringen Umfanges leichter auf dem laufenden zu halten sei. Daneben sollte die Fortschreibung verbessert werden (beispielsweise über Meldungen der Einwohnermeldeämter, Standesämter usw.).

Der Auswahlatz der Stichprobe, die nach Arbeitsamtsbezirken geschichtet gezogen wurde, sollte 5 v. H. betragen. Bei Anwendung des Buchstabenauswahlverfahrens erreichte man diesen v. H.-Satz am ehesten durch Verwendung des Buchstabens G (Bundesdurchschnitt 5,18 v. H.).

Die G-Kartei wurde zum 1.1.1964 eingerichtet, die Führung der Totalkartei zum gleichen Zeitpunkt eingestellt.

Allerdings hatten Probezählungen, die man unter anderem am 31. 3. 1962 und 30. 9. 1963 durchführte (an beiden Stichtagen wurde sowohl die Totalkartei als auch die Stichprobe nach verschiedenen Merkmalen ausgezählt), ergeben, daß die hochgerechneten Stichprobenwerte von jenen der Totalkartei abwichen. Doch begründete man diese Abweichungen zunächst mit dem allgemein schlechten Karteizustand. Man hoffte, nach einer großangelegten Bereinigungsaktion mit der G-Stichprobe ein statistisches Instrument zu haben, das aktuelle und nach ausgesuchten Wirtschaftszweigen und Regionen hinreichend genaue Beschäftigtenzahlen liefern würde.

<sup>1)</sup> Gesetz über Arbeitsvermittlung und Arbeitslosenversicherung.

<sup>2)</sup> Vergleiche: Stichproben in der amtlichen Statistik, Stuttgart und Mainz 1960, S. 30.

<sup>3)</sup> Vergleiche Anderson, Oskar: Probleme der statistischen Methodenlehre in den Sozialwissenschaften; 5. Auflage, Würzburg 1965, S. 83 ff.

Leider hat sich dieser Wunsch nicht erfüllt: die ersten Auszählungen im Jahr 1964 zeigten erhebliche Abweichungen zwischen den Zahlen der G-Kartei und denen anderer Beschäftigtenstatistiken, z. B. des Mikrozensus. Die Zahlen wurden deshalb nie veröffentlicht. Die Auszählungen wurden bis Juni 1966 fortgesetzt und dann eingestellt. Die G-Kartei selbst wurde jedoch weitergeführt, da sie möglicherweise als statistisches Instrument für das im Frühjahr 1967 gegründete Institut für Arbeitsmarkt- und Berufsforschung dienen konnte. Eine der ersten Aufgaben des Instituts war es deshalb, die Brauchbarkeit der G-Kartei für seine Zwecke zu untersuchen.

## B. Die Qualität der G-Stichprobe

### I. Allgemeine Bemerkungen über Fehlermöglichkeiten

Die Anwendung des Buchstabenverfahrens bei der Ziehung einer Stichprobe ergibt theoretisch keine uneingeschränkte Zufallsauswahl. Vielmehr handelt es sich dabei um ein Ersatzverfahren, auf das jedoch die Methoden der Wahrscheinlichkeitsrechnung voll anwendbar sind<sup>2)</sup>.

Die Größe des *Zufallsfehlers*, mit dem Stichprobenergebnisse im allgemeinen behaftet sind, läßt sich mithin abschätzen.

Neben Zufallsfehlern können statistische Daten jedoch auch noch *systematische Fehler* enthalten<sup>3)</sup>. Ein solcher systematischer Fehler läge in der G-Stichprobe beispielsweise dann vor, wenn eine Abhängigkeit zwischen dem Namensanfang G und einem zu erhebenden Merkmal, z. B. wirtschaftsfachliche Zugehörigkeit, bestünde. Die Strukturen zwischen G-Kartei und Grundgesamtheit würden dann nicht nur zufallsbedingt voneinander abweichen. Mit anderen Worten: die G-Kartei würde die Grundgesamtheit nicht repräsentieren.

Der Repräsentationsfehler der Stichprobenergebnisse setzt sich also aus dem Zufallsfehler und dem möglichen systematischen Fehler zusammen.

Von diesen beiden statistischen Fehlerquellen im engeren Sinne ist noch eine dritte Fehlerquelle zu unterscheiden: es handelt sich dabei um die „Sauberheit“ der Kartei. Ist z. B. der definierte Personenkreis nicht richtig erfaßt, sind also verschiedene Personen in der Kartei gar nicht, andere dagegen doppelt oder mehrfach enthalten, so dürfte dies zu falschen Ergebnissen sowohl hinsichtlich des Niveaus als auch der Struktur führen.

Es ist also zu prüfen, ob die G-Kartei neben dem Zufallsfehler einen durch das Auswahlverfahren und/oder einen durch den Erfassungs- und Fortschreibungsmechanismus bedingten Fehler enthält.

## II. Güte und Verwendungszweck

Für *jeden* Verwendungszweck sind Stichprobenergebnisse, die *lediglich* mit dem unvermeidlichen, aber abschätzbaren Zufallsfehler behaftet sind, der Idealfall, obwohl sich auch hier bereits Unzulänglichkeiten ergeben können. Trotzdem gibt es Verwendungszwecke, bei denen ein zusätzlicher Fehler der oben geschilderten Art keine schwerwiegende Rolle spielt. In Anlehnung an den im IAB entwickelten erwerbsstatistischen Anforderungskatalog und die Systematisierung der Anforderungsbereiche<sup>4)</sup> könnte die G-Kartei theoretisch für folgende Zwecke verwendet werden, für die die Anforderung an die Genauigkeit der Ergebnisse angegeben werden:

### 1. Konjunkturbeobachtung und Konjunkturforschung

Voraussetzung für eine Statistik, die der Konjunkturbeobachtung und Konjunkturforschung dienen soll, ist eine kurze Periodizität und hohe Aktualität. Beide Forderungen könnten mit der G-Kartei erfüllt werden.

Darüber hinaus ist jedoch eine hohe Genauigkeit der Zahlen sowohl hinsichtlich des Niveaus in den einzelnen Wirtschaftszweigen als auch insbesondere hinsichtlich der Veränderungen erforderlich.

Was das Niveau anbetrifft, bedeutet dies die Forderung nach einer „sauberen“ Kartei einerseits und ausreichender wirtschaftsfachlicher Repräsentanz andererseits. Für die Veränderungen zwischen den Stichtagen genügt nicht nur die verlässliche Bestimmung der *Veränderungsrichtung*- diese muß auch in ihrer absoluten Höhe ermittelt werden können.

### 2. Strukturforschung

#### a) Querschnittsanalysen; komparativ statische Analysen

Aus der G-Kartei könnten nach Art und Zahl der in ihr enthaltenen Merkmale Strukturdaten gewonnen werden. Mehrere Erhebungen in regelmäßigen Abständen und nach gleichen Merkmalen und gleicher Gliederungstiefe würden die Beobachtung von strukturellen Änderungen und eventuell deren Trendbestimmung ermöglichen. Hierbei ist insbesondere zu prüfen, inwieweit die G-Kartei die wichtigsten Strukturmerk-

male (wirtschaftsfachliche Zugehörigkeit, Beruf, Region und Alter) repräsentiert.

Dabei ist es allerdings nicht erforderlich, wenn auch wünschenswert, daß alle Merkmale zusammen richtig repräsentiert werden. Eine nur wirtschaftsfachliche Verzerrung beispielsweise würde immer noch eine Menge Untersuchungen, wie etwa Altersaufbau der Beschäftigten und dessen Veränderung je Wirtschaftszweig, erlauben.

#### b) Längsschnittuntersuchungen

Längsschnittuntersuchungen, die in neuester Zeit immer mehr an Bedeutung gewinnen, könnten mit der G-Kartei durchgeführt werden, da in ihr individuelle Verläufe (Berufswege, Fluktuationen, Arbeitslosigkeit etc.) aufgezeichnet sind. Der Repräsentationsgrad für die Grundgesamtheit hat hier die geringste Bedeutung, da es nicht mehr auf die Gewinnung eines Strukturbildes für einen Stichtag, sondern auf typische Verläufe für eine bestimmte Personengruppe ankommt.

In erster Linie sind deshalb die Zuverlässigkeit und Vollständigkeit der Eintragungen von Bedeutung.

### 3. Auswahlgrundlage

Die G-Kartei kann zusätzlich als Auswahlgrundlage für Sonderuntersuchungen verwendet werden. Dabei hängt es allerdings von dem Untersuchungsziel ab, welche Ansprüche an die G-Kartei gestellt werden müssen.

## III. Die Untersuchungen im einzelnen

### 1. Untersuchung auf wirtschaftsfachliche Repräsentanz

#### a) die $\chi^2$ -Methode

Die G-Stichprobe mit dem Merkmal wirtschaftsfachliche Zugehörigkeit hat theoretisch eine mehrdimensionale hypergeometrische Verteilung. Unter Vernachlässigung der Endlichkeit der Grundgesamtheit, aus welcher sie gezogen ist, kann man sie als sogenannte Polynomial- oder Multinomialverteilung betrachten. Der adäquate Test zur Prüfung einer solchen Stichprobe auf Strukturgleichheit mit der Grundgesamtheit ist der  $\chi^2$ -Test von *Pearson*<sup>5)</sup>. Ein solcher Vergleichstest zur Grundgesamtheit ist für den Zeitraum bis 1964 möglich, da ja erst dann die Totalkartei eingestellt wurde.

Die ersten Berechnungen zur Beurteilung der wirtschaftsfachlichen Struktur der G-Kartei wurden *ohne* Berücksichtigung der Schichtung nach der Formel

$$\chi^2 = \sum_i \frac{[n_i - E(n_i)]^2}{E(n_i)}$$

vorgenommen<sup>6)</sup>.

Dabei wurden für die Auszählungen am 31. 3. 1962 und am 30. 9. 1963 die in den folgenden Übersichten aufgeführten Ergebnisse ermittelt.

<sup>4)</sup> Vergleiche „Mitteilungen“ des IAB, Nr. 2, Mai 1968, S. 23.

<sup>5)</sup> Vergleiche *Morgenstern, Dietrich*: Einführung in die Wahrscheinlichkeitsrechnung und Mathematische Statistik; Berlin — Göttingen — Heidelberg, 1964, S. 72 ff.

<sup>6)</sup> Symbole:

$n_i$  = Beschäftigte mit Namensanfang G im i-ten Wirtschaftszweig

$E(n_i)$  = Erwartungswert von  $n_i = np_i = \frac{N_i}{N} n$

$n$  = Gesamtzahl der Beschäftigten mit Namensanfang G

$N_i$  = Gesamtzahl der Beschäftigten im i-ten Wirtschaftszweig

$N$  = Gesamtzahl der Beschäftigten

Die empirischen  $\chi^2$ -Werte wurden den bei einer Irrtumswahrscheinlichkeit von 5 v. H. höchstzulässigen theoretischen  $\chi^2$ -Werten gegenübergestellt.

**Übersicht 1:**

**Test für die Auszählung am 31. 3. 1962**

Bereich	$\chi^2 = \sum_i \frac{[n_i - E(n_i)]^2}{E(n_i)}$	$\chi^2_{\alpha} = 0,05$
Bundesgebiet (ohne Berlin)	559,77	119,87
Schleswig-Holstein	152,96	95,08
Niedersachsen	250,05	99,62
Nordrhein-Westfalen	554,54	99,62
Hessen	210,45	99,62
Rheinland-Pfalz - Saarland	188,24	97,35
Baden-Württemberg	156,82	97,35
Nordbayern	154,97	99,62
Südbayern	162,93	97,35
Berlin	163,45	101,88
Hamburg	187,12	100,75
Bremen	67,40	81,38

Wegen der regional stark streuenden G-Quoten<sup>7)</sup> und der Wahrscheinlichkeit, daß die Wirtschaftszweige ebenfalls eine unterschiedliche regionale Verteilung aufweisen, war es in einem weiteren Schritt erforderlich, die empirischen  $\chi^2$ -Werte unter Berücksichtigung der Schichtung zu ermitteln.

Ist z. B. ein Industriezweig in Regionen mit kleinen G-Quoten konzentriert, so ist er in der G-Stichprobe insgesamt, also ohne Berücksichtigung der Schichtung, unterrepräsentiert.

Wie aus den beiden letzten Tabellen ersichtlich ist, sind die Unterschiede der empirischen  $\chi^2$ -Werte für die beiden Stichtage 31. 3. 1962 und 30. 9. 1963 gering, so daß die sehr umfangreichen Berechnungen auf die Zahlen des 31. 3. 1962 beschränkt werden konnten.

Die empirischen  $\chi^2$ -Werte wurden dabei nach der Formel

$$\chi^2 = \sum_i \frac{\left( n_i - \sum_h n_h \frac{N_{hi}}{N_h} \right)^2}{\sum_h n_h \frac{N_{hi}}{N_h}}$$

ermittelt<sup>8)</sup>.

Es wurden folgende empirischen  $\chi^2$ -Werte festgestellt, denen wiederum die bei einer Irrtumswahrscheinlichkeit von 5 v. H. höchstzulässigen theoretischen Werte gegenübergestellt wurden.

**Übersicht 2:**

**Test für die Auszählung am 30. 9. 1963**

Bereich	$\chi^2 = \sum_i \frac{[n_i - E(n_i)]^2}{E(n_i)}$	$\chi^2_{\alpha} = 0,05$
Bundesgebiet (ohne Berlin)	548,85	118,75
Schleswig-Holstein	176,70	107,52
Niedersachsen	195,69	115,39
Nordrhein-Westfalen	623,76	114,27
Hessen	345,79	109,77
Rheinland-Pfalz - Saarland	215,01	107,52
Baden-Württemberg	199,50	109,77
Nordbayern	158,87	105,27
Südbayern	197,88	112,02
Berlin	101,94	99,02
Hamburg	119,80	99,62
Bremen	88,94	99,62

**Übersicht 3:**

**Test für die Auszählung am 31. 3. 1962 unter Berücksichtigung der Schichtung**

Bereich	$\chi^2 = \sum_i \frac{\left( n_i - \sum_h n_h \frac{N_{hi}}{N_h} \right)^2}{\sum_h n_h \frac{N_{hi}}{N_h}}$	$\chi^2_{\alpha} = 0,05$
Bundesgebiet (ohne Berlin und Hamburg)	283,42	99,62
Schleswig-Holstein	173,08	99,62
Niedersachsen	183,60	99,62
Nordrhein-Westfalen	292,35	99,62
Hessen	157,88	99,62
Rheinland-Pfalz - Saarland	185,11	99,62
Baden-Württemberg	157,41	99,62
Nordbayern	130,30	99,62
Südbayern	156,83	99,62
Berlin	163,45	101,88
Hamburg	187,12	100,75
Bremen	75,51	99,62

7) Vergleiche Amtliche Nachrichten der Bundesanstalt für Arbeit, 5/62, S. 254 ff.

8) Symbole wie in Fußnote 4), der Index h numeriert die Arbeitsamtsbezirke.

**Übersicht 4:**  
Unterschiede in den  $\chi^2$ -Werten mit und ohne Berücksichtigung der Schichtung

Bereich	$\chi^2$ geschichtet	$\chi^2$ ungeschichtet
Bundesgebiet (ohne Berlin und Hamburg)	283,42	459,40
Schleswig-Holstein	173,08	152,86
Niedersachsen	183,60	250,05
Nordrhein-Westfalen	292,35	554,54
Hessen	157,88	210,45
Rheinland-Pfalz - Saarland	185,11	188,24
Baden-Württemberg	157,41	156,83
Nordbayern	130,30	157,85
Südbayern	156,83	162,49
Berlin	163,45	163,45
Hamburg	187,12	187,12
Bremen	75,51	64,73

Die hier aufgeführten  $\chi^2$ -Werte ohne Berücksichtigung der Schichtung weichen zum Teil von den Werten der Übersicht 2 ab. Dies liegt daran, daß die Berechnungen auf einer anderen Aggregationsebene durchgeführt wurden.

**b) Entwicklung zwischen den Zähltagen 1962 und 1963**

Der oben (Übersicht 1—3) bereits sichtbar gewordene signifikante Unterschied zwischen den wirtschaftsfachlichen Strukturen der Totalkartei und der G-Ausschnitte für beide Zähltage impliziert nicht unbedingt einen Unterschied in der Entwicklung zwischen Grundgesamtheit und Stichprobe. Ein solcher Unterschied würde, neben der Abhängigkeit zwischen Namensanfang G und wirtschaftsfachlicher Zugehörigkeit, eine Abhängigkeit zwischen Namensanfang G und Fluktuation bedeuten.

Um dies zu prüfen, wurden die Erwartungswerte der Veränderungen je Wirtschaftszweig für den G-Ausschnitt ermittelt und den tatsächlichen Veränderungen zwischen dem 31. 3. 1962 und dem 30. 9. 1963 gegenübergestellt.

Der ermittelte  $\chi^2$ -Wert beträgt 7683. Bei einer Irrtumswahrscheinlichkeit von  $\alpha = 0,05$  hätte er den Wert von 116,51 nicht überschreiten dürfen.

Um diese fast unglaubliche Abweichung plausibler zu machen, sind in der folgenden Übersicht die Erwartungswerte der G-Veränderung  $E(\Delta n_i)$  den tatsächlichen Werten  $\Delta n_i$  für einige ausgewählte Wirtschaftszweige gegenübergestellt.

Darüber hinaus ist in Übersicht 6 die Entwicklung des Wirtschaftszweiges 52a (Bildungswesen, Kirche) innerhalb der Landesarbeitsamtsbezirke

dargestellt. In diesem Wirtschaftszweig steht einer Zunahme von 14 005 Beschäftigten innerhalb der Totalkartei ( $\Delta N_i$ ) eine Abnahme um 64 im G-Ausschnitt ( $\Delta n_i$ ) gegenüber.

**Übersicht 5:**  
Vergleich zwischen  $E(\Delta n_i)$  und  $\Delta n_i$  für ausgewählte Wirtschaftszweige (Bundesgebiet ohne Berlin) 1962—1963

Wirtschaftszweig	$E(\Delta n_i)$	$\Delta n_i$
7a Gewinnung und Bearbeitung von natürl. Steinen und Erden	+ 251	+ 101
11b Eisen- u. Stahlgießerei	— 391	— 282
13a Eisen-, Stahl- und Metallwarenindustrie	— 333	— 473
17d/1 Straßenfahrzeugbau	+ 2 894	+ 3 110
18 Elektroindustrie	+ 1 451	+ 958
23a Druck- und Vervielfältigungsgewerbe	+ 518	+ 397
27a Tischlerei	— 226	— 607
41/42 Handel und Handelshilfsgewerbe	+ 9 320	+ 10 198
39d/2 Gebäudeinnenreinigung	+ 395	+ 555
48 Gaststättenwesen	+ 1 731	+ 2 250
58 Häusliche Dienste	— 1 262	— 1 582
49 Verwaltung	+ 2 254	+ 1 539
52a Bildungswesen, Kirche	+ 729	— 64

**Übersicht 6:**  
Vergleich zwischen  $\Delta N_i$  und  $\Delta n_i$  innerhalb der Landesarbeitsamtsbezirke für Wirtschaftszweig 52a (Bildungswesen, Kirche) 1962—1963

Bereich	$\Delta N_i$	$\Delta n_i$
Schleswig-Holstein	+ 664	+ 60
Hamburg	+ 286	+ 48
Niedersachsen	+ 2 092	— 107
Bremen	— 72	— 7
Nordrhein-Westfalen	+ 2 307	— 8
Baden-Württemberg	+ 3 398	— 17
Nordbayern	+ 1 312	— 2
Südbayern	+ 1 747	— 7
Hessen	+ 1 355	— 63
Rheinland-Pfalz - Saarland	+ 916	+ 39
Bundesgebiet (ohne Berlin)	+ 14 005	— 64

**c) Untersuchung einzelner Wirtschaftszweige**  
Einige der Wirtschaftszweige, deren Anteil in der G-Stichprobe von dem in der Grundgesamtheit besonders stark abwich (oder: deren G-Quote stark von der Gesamtquote 0,0521 abwich), wur-

den daraufhin untersucht, ob diese Abweichungen durch Zählfehler in einem oder mehreren Arbeitsamtsbezirken entstanden sein konnten oder ob eine systematische Abweichung in einer Richtung über alle Arbeitsamtsbezirke erkennbar ist. Diese Untersuchung konnte deshalb durchgeführt werden, weil die G-Stichprobe nach Arbeitsamtsbezirken geschichtet vorliegt, für jeden Arbeitsamtsbezirk also unabhängige Zahlen vorhanden sind.

Es wurden dabei jene Wirtschaftszweige ausgewählt, die neben der hohen Abweichung auch relativ hohe Besetzungszahlen aufwiesen. In die Untersuchung wurden einbezogen:

Wirtschaftszweig	18	Elektroindustrie
	39b	Hoch- und Tiefbau
	41/42	Handel und Handelshilfsgewerbe
	44	Deutsche Bundespost
	45	Deutsche Bundesbahn

Dabei wurden zunächst für jeden der ausgewählten Wirtschaftszweige die Schwankungsintervalle bestimmt, innerhalb derer die relative

Häufigkeit  $\frac{n_i}{n}$  mit einer Wahrscheinlichkeit von 68 v. H. bzw. 95 v. H. zu erwarten war. Diese Intervalle berechnen sich zu

$$P_i \pm \sqrt{\frac{P_i(1-P_i)}{n}} \text{ bzw. } P_i \pm 1,96 \sqrt{\frac{P_i(1-P_i)}{n}}$$

Bei einem Vergleich der relativen Häufigkeit

$P_i = \frac{n_i}{n}$  der Stichprobe mit diesen Schwankungsintervallen konnte angegeben werden, in welcher Richtung und in welchem Grade die Anteile in der Stichprobe von denen in der Grundgesamtheit abweichen; mit anderen Worten: ob und wie stark der betreffende Wirtschaftszweig in der Stichprobe über- bzw. unterrepräsentiert ist.

Diese Berechnungen bezogen sich auf die Zahlen des Bundesgebietes (ohne Berlin). Anschließend wurden die Anteile je ausgewählten Wirtschaftszweig und Arbeitsamtsbezirk aus der Stichprobe und der Grundgesamtheit miteinander verglichen und ihre Abweichungen mit + bzw. — angeschrieben.

Unter der Annahme, die G-Stichprobe verhalte sich wie eine Zufallsauswahl, wären die Abweichungen über die einzelnen Arbeitsamtsbezirke binomial verteilt mit der Wahrscheinlichkeit

$$p = \frac{1}{2}. \text{ (Vorzeichentest)}$$

Ergebnisse:

Wirtschaftszweig 18 — Elektroindustrie:

Der Anteilswert der Elektroindustrie innerhalb der G-Stichprobe liegt (für das gesamte Bundesgebiet) *oberhalb* des doppelten Fehlerberei-

ches. Bei 98 von 174 Arbeitsamtsbezirken liegt eine positive Abweichung vor. Dies entspricht einer Wahrscheinlichkeit von (nur) 0,0643. Die Hypothese auf Zufälligkeit kann bei  $\alpha = 0,05$  nicht verworfen werden.

Wirtschaftszweig 39b — Hoch- und Tiefbau:

Der Anteilswert des Hoch- und Tiefbaus in der G-Stichprobe liegt (für das gesamte Bundesgebiet) *oberhalb* des einfachen, jedoch innerhalb des doppelten Fehlerbereiches. Bei 96 von 174 Arbeitsamtsbezirken liegt eine positive Abweichung vor. Dies entspricht einer Wahrscheinlichkeit von (nur) 0,0869. Die Hypothese auf Zufälligkeit kann bei  $\alpha = 0,05$  nicht verworfen werden.

Wirtschaftszweig 41/42 — Handel und Handelshilfsgewerbe:

Der Anteilswert des Handels und Handelshilfsgewerbes in der G-Stichprobe liegt *unterhalb* des doppelten Fehlerbereiches. Bei 92 von 174 Arbeitsamtsbezirken liegt eine negative Abweichung vor. Dies entspricht einer Wahrscheinlichkeit von 0,2236. Die Hypothese auf Zufälligkeit kann bei  $\alpha = 0,05$  nicht verworfen werden.

Wirtschaftszweig 44 — Deutsche Bundespost:

Der Anteilswert der Deutschen Bundespost in der G-Stichprobe liegt *unterhalb* des doppelten Fehlerbereiches. Bei 107 von 174 Arbeitsamtsbezirken liegt eine negative Abweichung vor. Dies entspricht einer Wahrscheinlichkeit von 0,00122. Die Hypothese auf Zufälligkeit muß bei  $\alpha = 0,05$  verworfen werden.

Wirtschaftszweig 45 — Deutsche Bundesbahn:

Der Anteilswert der Deutschen Bundesbahn in der G-Stichprobe liegt *unterhalb* des doppelten Fehlerbereiches. Bei 100 von 174 Arbeitsamtsbezirken liegt eine negative Abweichung vor. Dies entspricht einer Wahrscheinlichkeit von 0,0244. Die Hypothese auf Zufälligkeit muß bei  $\alpha = 0,05$  verworfen werden.

#### d) Ergebnisse und Folgerungen

(1) Der Unterschied zwischen der wirtschaftsfachlichen Struktur der G-Kartei und der der Totalkartei ist gesichert. Dies gilt für die Auszählungen am 31. 3. 1962 und am 30. 9. 1963 (Übersicht 1 und 2).

Die Berücksichtigung der Schichtung bringt zwar teilweise eine erhebliche Reduzierung der Abweichungen mit sich, doch liegen auch hier die ermittelten Werte für das Bundesgebiet wie auch für die Landesarbeitsamtsbezirke (mit Ausnahme des Landesarbeitsamts Bremen) noch über der zulässigen Schranke (Übersicht 3).

(2) Neben dem hier festgestellten statischen Unterschied beider Strukturen läßt sich auch ein Unterschied in der Entwicklung zwischen den Zähltagen nachweisen.

(3) Von den fünf einzeln untersuchten Wirtschaftszweigen kann bei zweien (Bahn und Post) die Annahme der Zufälligkeit der Abweichungen nicht aufrechterhalten werden. In beiden Fällen wird die sich aus den Gesamtzahlen ergebende Unterrepräsentation in der Stichprobe durch die Abweichungen in den Arbeitsamtsbezirken belegt.

In zwei weiteren Wirtschaftszweigen (Elektroindustrie und Bau) wurde zwar die Hypothese der Zufälligkeit der Abweichungen bei  $\alpha = 0,05$  aufrechterhalten, doch kann nicht übersehen werden, daß die Wahrscheinlichkeiten beider Vorzeichenverteilungen von 0,0643 bzw. 0,0869 nur wenig von der Signifikanzschranke abweichen. Die Verteilungen weisen zumindest in die Richtung der Überrepräsentation, wie sie sich aus den Gesamtzahlen ergibt.

Nur der Wirtschaftszweig Handel weist eine Vorzeichenverteilung auf, die durchaus als zufällig bezeichnet werden kann. Die Verteilung innerhalb der Landesarbeitsamtsbezirke ist jedoch teilweise ebenfalls extrem.

Nach diesen Ergebnissen dürfte es schwer fallen, die Abweichungen der G-Stichprobe von der Totalkartei mit Zählfehlern in einigen Arbeitsamtsbezirken zu begründen.

Die hier ermittelten signifikanten Unter- bzw. Überrepräsentationen sind nicht ohne Einfluß auf die Anteile anderer Wirtschaftszweige. Es kann deshalb angenommen werden, daß sich weitere entsprechende Ergebnisse finden ließen.

(4) Aus den unter (1) bis (3) genannten Ergebnissen folgt, daß sowohl eine Abhängigkeit zwischen Namensanfang G und wirtschaftsfachlicher Zugehörigkeit als auch zwischen Namensanfang G und Fluktuation — beides in hohem Ausmaß — besteht. Dafür ist keine plausible Erklärung zu finden. Vielmehr drängt sich der Verdacht auf, daß die G-Kartei innerhalb der Totalkartei eine eigene Entwicklung genommen hat und beide Strukturen mithin überhaupt nicht vergleichbar sind.

Freilich ist es schwierig zu sagen, wann diese Entwicklung begann und ob sie bereits die Zähl-ergebnisse vom 31. 3. 1962 beeinflusste.

Diese Möglichkeit ist jedoch nicht auszuschließen: die Diskussion um die Umstellung der Totalkartei auf die G-Kartei war schon seit längerer Zeit im Gange; dabei wurden auch Probleme der Bereinigung der G-Kartei und ihre „Laufendhaltung“ erörtert. Ende September und Ende Dezember 1961 wurden vorbereitende Zählungen (Total und G) durchgeführt. Für diese Auszählungen sollte die Kartei besonders sorgfältig bereinigt werden. Möglicherweise konzentrierte sich die Bereinigung auf den G-Ausschnitt.

<sup>9)</sup> Zu diesem Vergleich wurde das arithmetische Mittel der Beschäftigtenzahlen vom 31. 3. 1961 und 30. 9. 1961 verwendet.

Da die ehemalige Totalkartei neben der absoluten Überhöhung von ca. 1,2 Millionen Beschäftigten gegenüber der Volkszählung 1961 auch die Tendenz hatte, in wachsenden Wirtschaftszweigen *relativ* zu wenig Beschäftigte auszuweisen, in schrumpfenden oder stagnierenden Zweigen jedoch absolut und relativ zu viele, ist anzunehmen, daß eine bessere Führung der G-Kartei innerhalb der Totalkartei eine Verschiebung der G-Struktur in umgekehrter Richtung mit sich bringen müßte. Mit anderen Worten: die bessere Bereinigung und Bearbeitung der G-Kartei würde tendenziell zu einer Anpassung der Struktur der G-Kartei an die der Volkszählung 1961 führen; die in der Totalkartei übererfaßten Wirtschaftszweige wären in der G-Kartei unterrepräsentiert und umgekehrt.

Um dies nachzuprüfen, wurden alle stärker besetzten Wirtschaftszweige, für die ein Vergleich mit der Volkszählungs-Systematik ohne größere Zusammenfassungen möglich war, ausgewählt.

Übersicht 7:

	Wirtschaftszweige	Totalkartei gegenüber Volkszählung 1961 <sup>9)</sup>	G-Kartei gegenüber Totalkartei 31. 3. 1962
3	Steinkohlenbergbau	+	+
18	Elektroindustrie	—	+
20a	Chemische Industrie	—	+
21a	Textilgewerbe	+	—
23a	Druck- und Vervielfältigungsgewerbe	—	+
26	Sägeindustrie	+	+
27a, 39c2	Möbelindustrie	—	+
30b, 31a	Bäckereigewerbe Süßwarenherstellung	+	—
36	Schneiderei, Kleiderherstellung	—	—
37	Schuhindustrie	+	—
43a	Geldinstitute	—	—
45	Bundesbahn	+	—
47	Schiffahrt	+	—
58	Häusliche Dienste	+	—

Von den vierzehn Wirtschaftszweigen haben vier gleiche Vorzeichen und widersprechen somit der Hypothese. Dagegen sind die Vorzeichen von zehn Wirtschaftszweigen verschieden und weisen in die angegebene Richtung.

Ebenfalls zehn von den vierzehn angeführten Wirtschaftszweigen belegen die Hypothese im Hinblick auf die Veränderung zwischen den Zähltagen: gegenüber den in der Totalkartei übererfaßten Wirtschaftszweigen nahmen die

Zahlen der G-Kartei stärker ab bzw. weniger zu und umgekehrt. Beispielsweise betrug die Veränderungen im Wirtschaftszweig 58 (Häusliche Dienste), der gegenüber der Volkszählung 1961 um 177400 Beschäftigte überhöht war, in der Totalkartei —5,17 v. H., in der G-Kartei — 6,48 v. H. Im Wirtschaftszweig 20a (Chemische Industrie), der gegenüber der Volkszählung 1961 um 4300 Beschäftigte<sup>10)</sup> untererfaßt war, betrug die Zunahme in der Totalkartei 1,94 v. H., in der G-Kartei 2,23 v. H.

Die teilweise größer gewordenen Fehler zwischen Totalkartei und G-Kartei am 30.9. 1963 deuten ebenfalls in diese Richtung.

Ein weiterer Anhaltspunkt, der die oben genannte Hypothese stützt, ergab sich aus der G-Auszählung im April 1968. Bei dieser Auszählung wurde unter anderem das Jahr des 1. Beschäftigungsverhältnisses erhoben. Dieses Datum ist weithin identisch mit der Anlage einer Karteikarte überhaupt.

Die folgenden Zahlen sprechen für sich:

Jahr	Zahl der Eintragungen	
1958	34 193	
1959	34 339	+ 146
1960	34 379	+ 40
1961	35 578	+ 1 199
1962	37 712	+ 2 134
1963	45 675	+ 7 963
1964	47 538	+ 1 863

Während die Zahlen der Eintragungen zwischen 1958 und 1960 praktisch gleichbleiben, nehmen sie ab 1961 zu, und zwar mit zunehmender Beschleunigung. Es kann angenommen werden, daß dies hauptsächlich mit der besonderen Bearbeitung der G-Kartei zusammenhängt.

Nach diesen Ergebnissen ist ein Unterschied im Bearbeitungsstand zwischen Stichprobe und Grundgesamtheit sehr wahrscheinlich, und zwar auch schon vor dem 31. 3. 1962.

Dann aber ergibt sich, daß die sehr umfangreichen und detaillierten Berechnungen mit völlig unzureichendem Zahlenmaterial durchgeführt wurden. *Der im Ergebnis gefundene signifikante Unterschied zwischen G-Kartei und Totalkartei deutet weniger auf einen systematischen, durch das Auswahlverfahren bedingten Fehler in der Stichprobe hin als auf die unterschiedliche Entwicklung beider Karteien.*

*Damit entfällt aber auch gleichzeitig die Möglichkeit, einen (möglichen) systematischen Fehler der G-Stichprobe nachzuweisen.*

Man wird, um die Qualität der aus der G-Kartei gewonnenen Zahlen beurteilen zu können, auf

<sup>10)</sup> Die relative Untererfassung betrug 0,12 v. H., das entspricht etwa 28000 Beschäftigten.

den Vergleich mit anderen Statistiken ausweichen müssen. Die Aussagen dürften entsprechend schwächer sein, da das Gebiet exakter Berechnungen gegen jenes plausibler Schätzungen getauscht wird.

## 2. Vergleich der G-Kartei mit anderen Statistiken

### a) Bestandsvergleich G-Kartei — Mikrozensus für April 1968

Die auch für das Projekt „Beschäftigungstendenzen älterer Arbeitnehmer“ des IAB erforderliche Auszählung der G-Kartei wurde zeitlich so gelegt, daß erstmals ein Vergleich der G-Kartei-Ergebnisse mit jenen des Mikrozensus ohne Interpolation möglich wurde. Neben dieser zeitlichen Abstimmung war eine sachliche Abgrenzung erforderlich: die Mikrozensuszahlen mußten um die ausländischen Erwerbstätigen, Beamte, Heimarbeiter und eine bestimmte Gruppe von Teilzeitarbeitern bereinigt werden. Zu diesem Zweck wurde vom Statistischen Bundesamt eine Sonderaufbereitung des Mikrozensus für den gewünschten Personenkreis vorgenommen.

Die ermittelten Zahlen aus der G-Kartei liegen mit 1,97 Millionen oder 10,85 v. H. über den Zahlen des Mikrozensus (Mikrozensus = 100).

Unterstellt man, daß die Mikrozensuszahlen insgesamt zumindest der Größenordnung nach richtig sind, so weist die G-Kartei einen Niveaufehler auf, der weit über dem der ehemaligen Totalkartei (1961 = ca. 1,2 Millionen Überhöhung) liegt. Dieser Niveaufehler von durchschnittlich + 11 v. H. ist auch, mit einer gewissen Streuung, über alle Regionen und über die Wirtschaftsbereiche hinweg zu beobachten. Besonders kraß tritt er im Bereich Handel und Verkehr mit + 26,1 v. H. hervor. Auf die möglichen Ursachen dieser erstaunlichen Verzerrung wird später noch eingegangen werden.

Auch in der Kombination Wirtschaftsbereich X Region sind, von ganz wenigen Ausnahmen abgesehen, Überhöhungen festzustellen. Freilich dürften sich bei dieser Untergliederung — im Gegensatz zu den Gesamtzahlen — auch die Stichprobenfehler beider Erhebungen auswirken.

Die unterschiedliche Überhöhung in den einzelnen Wirtschaftsbereichen (die Proportionen des Mikrozensus müssen wiederum als etwa richtig unterstellt werden) zeigt neben dem Niveaufehler auch einen wirtschaftsfachlichen Strukturfehler. Dieser Strukturfehler hat eine Größenordnung, die durch einen Mangel im Auswahlverfahren wohl schwerlich begründet werden kann.

Die Abweichungen in den einzelnen Wirtschaftsbereichen können, wie die entsprechenden Berechnungen ergaben, auch nicht als Zufallsfehler gedeutet werden.

b) *Bestandsvergleich G-Kartei — Industriebericht für April 1968*

Die möglicherweise den oben gemachten Ausführungen entgegenzuhaltenden Argumente, der Mikrozensus sei wegen seiner oft umstrittenen Ergebnisse keine geeignete Vergleichsbasis zur Beurteilung der G-Kartei, lassen sich mit den folgenden Zahlen weitgehend entkräften. Für eine ganze Reihe von Wirtschaftszweigen war ein *direkter* Vergleich zwischen den Beschäftigten der Industrieberichterstattung — die als eine der besten Beschäftigtenstatistiken überhaupt gilt — und denen der G-Kartei möglich. Allerdings mußten auch hier Schätzungen vorgenommen werden. So war es erforderlich, die Zahlen der Industrieberichterstattung, die ungefähr 98 v. H. der Industriebeschäftigten repräsentieren, für die infrage kommenden Wirtschaftszweige auf 100 v. H. hochzuschätzen. Darüber hinaus mußten die Zahlen, um zu einer vertretbaren Abgrenzung der Personenkreise zu gelangen, um die ausländischen Beschäftigten bereinigt und um die Beschäftigten des Handwerks ergänzt werden.

Diese Operationen ließen sich mit dem vorhandenen statistischen Material jedoch relativ einfach durchführen.

Für den Vergleich konnten ca. 99 v. H. der abhängig Beschäftigten aus der Industrieberichterstattung herangezogen werden; die diesem Personenkreis entsprechenden Zahlen aus der G-Kartei lagen um 10,4 v. H. höher!

Die Streuung der Abweichungen in den einzelnen Wirtschaftszweigen ist groß. Dabei dürften auch Abgrenzungsschwierigkeiten beider Systematiken eine gewisse Rolle gespielt haben. Die Tendenz der Überhöhung, die in den Ingesamtzahlen zum Ausdruck kommt, ist jedoch auch bei Betrachtung der einzelnen Wirtschaftszweige nicht zu verkennen.

Von 39 Aggregaten weisen 35 positive und 4 negative Abweichungen zwischen G-Kartei und Industrieberichterstattung auf.

Ein beträchtlicher Teil der Abweichung könnte, da die Aggregate oft nur schwach besetzt sind, auf Zufallsfehlern innerhalb der G-Kartei beruhen. Entsprechende Berechnungen ergaben jedoch, daß bei Berücksichtigung des doppelten Standardfehlers die Abweichungen von nur 8 Wirtschaftszweigen, bei Berücksichtigung des einfachen Standardfehlers sogar von nur 6 Wirtschaftszweigen mit Zufallsfehlern begründet werden können.

c) *Bestandsvergleich im Wirtschaftszweig 45: Deutsche Bundesbahn*

Die Deutsche Bundesbahn ist im Verzeichnis der Wirtschaftszweige für die Arbeitsstatistik ein selbständiger Wirtschaftszweig (45).

Da die Bundesbahn ihre Sozialversicherungspflichtigen Beschäftigten in Personalkarteien bei den Sozialversicherungsstellen führt, war es möglich, Vergleichszahlen für den Wirtschaftszweig 45 ohne großen Aufwand zu erhalten.

Die Sozialversicherungsstellen der Deutschen Bundesbahn wurden gebeten, ihre Versicherungspflichtigen Beschäftigten mit Namensanfang G (ohne Ausländer) zum 24.4. 1968 auszu zählen.

Das Ergebnis enthält die folgende Übersicht.

**Übersicht 8:**

Versicherungspflichtige Beschäftigte mit Namensanfang G (ohne Ausländer) am 24. 4. 1968	G-Fälle der Arbeitsämter Wirtschaftszweig 45 (Deutsche Bundesbahn) am 24. 4. 1968
9 920	12 193

Diese Zahlen, die theoretisch völlig miteinander übereinstimmen müßten, da ein Stichprobenfehler hier nicht möglich ist, beweisen eindrucksvoll die „Unsauberkeit“ der Kartei. Die Abweichung beträgt 2273, was auf der Basis der Bundesbahnzahlen 22,91 v. H. ausmacht.

Ein regionaler Vergleich konnte wegen der unterschiedlichen Abgrenzung der Bundesbahnbezirke leider nicht durchgeführt werden. Auch war es nicht möglich, von allen Bundesbahn-Dienststellen die Ingesamtzahlen der Versicherungspflichtigen Beschäftigten am Stichtag zu erfahren, um damit die Güte der G-Quoten zu testen.

d) *Veränderungsvergleich G-Kartei — Mikrozensus 1966—1968*

Obwohl, wie weiter oben schon erwähnt, ein exakter Bestandsvergleich zwischen Mikrozensus und G-Kartei wegen der unterschiedlichen Stichtage früher nicht möglich war, soll doch versucht werden, die Veränderung beider Statistiken für einen etwa identischen Zeitraum einander gegenüberzustellen. Dazu wurden für den Mikrozensus die Aprilzählung 1966, für die G-Kartei die vorletzte Auszählung vom Juni 1966, herangezogen. Die Zahlen wurden jeweils nach den vier Wirtschaftsbereichen und nach Bundesländern gegliedert. Es wurde festgestellt, daß zwar die Veränderungsrichtung der *Ingesamtzahlen* stimmt, das Ausmaß der Ingesamt-Veränderung jedoch unterschiedlich groß ist. Die Mikrozensuszahlen nahmen im Beobachtungszeitraum um rund 7 v. H. ab, die Zahlen aus der G-Kartei dagegen nur um 0,42 v. H.

Bei der geringsten Differenzierung jedoch stimmen nicht einmal mehr die Richtungen überein.

#### e) *Veränderungsvergleich G-Kartei — Industriebericht für 1966—1968*

Wegen der Art der früheren Erhebungsmethoden bei der Auszählung der G-Kartei konnten die Veränderungen leider nur in 6 Wirtschaftszweigen beobachtet werden. Die Ergebnisse fallen nicht ganz so kraß aus wie jene, die auf der Basis des Mikrozensus gefunden wurden. Sie haben jedoch die gleiche Tendenz.

In den 6 (zusammengefaßten) Wirtschaftszweigen weist die Industrieberichterstattung als Folge der Rezession im beobachteten Zeitraum (Juni 1966 bis April 1968) reduzierte Beschäftigtenbestände aus.

In der G-Kartei ist dieser Rückgang nur in 4 Wirtschaftszweigen festzustellen. Dazu kommt noch, daß das Ausmaß des Rückganges fast ausnahmslos geringer ist als in der Industrieberichterstattung.

Auch die Verhältniszahlen<sup>11)</sup> bestätigen die beim Mikrozensusvergleich gefundenen Ergebnisse: von einer Ausnahme abgesehen, liegen alle Relationen für das Jahr 1968 über denen des Jahres 1966. Sie weisen also ebenfalls in die Richtung der *zunehmenden* Überhöhung der G-Kartei.

#### f) *Altersverteilung der Beschäftigten nach der G-Kartei und dem Mikrozensus im April 1968*

Von den in der G-Kartei erfaßten Merkmalen erscheint das Alter als das „stabilste“. Das Geburtsdatum ist keinen Veränderungen unterworfen wie etwa Wirtschaftszweig, Beruf oder Wohnort und bietet somit kaum Gelegenheit für Eintragungsfehler.

Die Frage ist nur, ob neben der oben bereits sichtbar gewordenen absoluten Überhöhung und wirtschaftsfachlichen und regionalen Verzerrung der Kartei auch eine Verzerrung der Altersstruktur der aus ihr gewonnenen Beschäftigtenzahlen vorhanden ist.

Der Vergleich wurde nach 5er Altersgruppen, Wirtschaftsabteilungen bzw. -bereichen und Geschlecht durchgeführt.

Die Ergebnisse zeigen, daß die Verteilungen der Beschäftigten nach Altersgruppen absolut und relativ zwar voneinander abweichen, jedoch in unterschiedlicher Art je Altersgruppe.

Trotz der beträchtlichen Überhöhung der G-Karteizahlen gegenüber denen des Mikrozensus sind beispielsweise die Beschäftigten der Altersgruppe bis 20 Jahre in der G-Kartei *absolut untererfaßt*. Dies gilt für die Ingesamtszahlen und für die nach dem Geschlecht gegliederten Zahlen. Bei der Gliederung nach Wirtschaftsabtei-

lungen weist in dieser Altersgruppe nur die Abteilung 3 (Baugewerbe) in der G-Kartei eine höhere Erfassung gegenüber dem Mikrozensus aus. In der weiteren Untergliederung nach dem Geschlecht sind es bei den Männern die Abteilungen 3 (Baugewerbe) und 9 (Gebietskörperschaften und Sozialversicherung), bei den Frauen die Abteilung 2 (Verarbeitendes Gewerbe), die höhere Bestände nach der G-Kartei ausweisen. Alle anderen Abteilungen liegen — wie die Gesamtzahl dieser Altersgruppe — unter den Mikrozensusergebnissen.

Ganz anders sieht es aus bei den folgenden Altersgruppen: die Ingesamtszahlen der G-Kartei in den Altersgruppen 20 bis unter 25, 25 bis unter 30... usw. bis 40 bis unter 45 liegen zum Teil erheblich über denen des Mikrozensus. Dies kehrt sich ab der Altersgruppe 45 bis unter 50 Jahre (mit Ausnahme der Gruppe 60 bis unter 65 Jahre) um. Bei den Relativzahlen kommt diese Tendenz auch zum Ausdruck, jedoch mit erheblich geringerer Ausprägung.

Generell kann gesagt werden:

1. die Abweichungen sind in den jüngeren und mittleren Altersgruppen stärker als in den älteren,
2. die Abweichungen sind in den Abteilungen der Dienstleistungen (im weiteren Sinne) und der Landwirtschaft größer als anderswo,
3. die Abweichungen sind bei den Frauen größer als bei den Männern.

Die Erwartungen, die man in bezug auf die Zuverlässigkeit der aus der G-Kartei ermittelten Altersstruktur hatte, haben sich trotzdem in den meisten Punkten erfüllt: Die Abweichungen der Strukturen beider Statistiken bei den Männern sind so gering, daß sie durchaus mit Zufallsschwankungen erklärt werden können. Bei den Frauen sind die Abweichungen in den meisten Altersgruppen geringer als 1 v. H. Sie weichen jedoch bei den Altersgruppen bis 30 Jahren relativ stark ab. Darauf wurde bereits bei der ersten Auswertung der G-Karteizahlen hingewiesen<sup>12)</sup>. In wirtschaftsfachlicher Gliederung sind die Abweichungen im Dienstleistungsbereich am stärksten.

#### g) *Ergebnisse*

(1) Die aus der G-Kartei ermittelten Beschäftigtenzahlen für den 24.4. 1968 liegen um über 10 v. H. über den Zahlen vergleichbarer Statistiken (10,9 v. H. über Mikrozensus; 10,4 v. H. über Industrieberichterstattung), wobei die Überhöhung bei den Frauen prozentual größer ist als bei den Männern.

(2) Neben diesem Niveaufehler sind Verzerrungen in wirtschaftsfachlicher und regionaler Hinsicht vorhanden. Beim Mikrozensusvergleich fällt

<sup>11)</sup>  $\frac{\text{G-Kartei}}{\text{Industrieberichterstattung}} \cdot 100$

<sup>12)</sup> Vergleiche „Mitteilungen“ des IAB, Nr. 8, Juni 1969, S. 620.

dabei insbesondere der Wirtschaftsbereich Handel und Verkehr auf, bei dem die Abweichung 26,1 v. H. beträgt. Die Abweichungen bei den Wirtschaftszweigen der Industrieberichterstattung streuen ebenfalls beträchtlich.

Die Überhöhung im Wirtschaftszweig Deutsche Bundesbahn beträgt + 22,9 v. H. Theoretisch müßten hier beide Zahlen völlig übereinstimmen.

(3) Die festgestellte Überhöhung wird im Zeitablauf größer. Sowohl gegenüber dem Mikrozensus als auch gegenüber der Industrieberichterstattung liegen die Verhältniszahlen nahezu aller relevanten Kombinationen für das Jahr 1968 über denen des Jahres 1966. Diese Tendenz überlagert häufig gegenläufige reale Bewegungen, so daß *Veränderungsrichtungen* mit der G-Kartei nicht zuverlässig erfaßt werden können.

(4) Der Vergleich der Altersschichtung zeigt, daß die Unterschiede der prozentualen Strukturen in den meisten Fällen gering sind. Ausnahmen hiervon treten bei einigen Altersgruppen der Frauen und vor allem im Dienstleistungsbereich auf.

Diese Ergebnisse bestätigen die bereits an früherer Stelle geäußerten Vermutungen, daß die für die meisten Merkmale mangelnde Qualität der G-Karteizahlen in erster Linie nicht auf einem falschen oder unzureichenden (Stichproben-) Auswahlprinzip, sondern auf dem schlechten Bearbeitungsstand der Kartei beruht.

Darauf weist insbesondere der Umstand hin, daß die hochgerechneten Zahlen nicht einfach einen (gleichbleibend) schlechten Zustand anzeigen, sondern einen sich ständig verschlechternden. Dieser Trend ist nur durch die Schwächen des Fortschreibungsmechanismus erklärbar.

Dieser Fortschreibungsmechanismus, der auf § 53 des alten AVAVG bzw. § 10 (1) AFG basiert, ist zwar theoretisch lückenlos, doch hat die Praxis gezeigt, daß eine Rechtsgrundlage allein ohne Sanktionen für ihre Durchsetzung nicht genügt.

Insbesondere die Angaben über Einstellungen und Entlassungen, die der Arbeitgeber *direkt* an das zuständige Arbeitsamt zu melden hat — dies betrifft im wesentlichen den Personenkreis der bei den Ersatzkassen pflichtversicherten Beschäftigten — scheinen unvollständig zu sein.

Bei den Einstellungen würde dies tendenziell auf eine Untererfassung schließen lassen, die zu den gefundenen Ergebnissen zunächst in Widerspruch zu stehen scheint.

Diese Vermutung läßt sich allerdings belegen: in dem Altersvergleich Mikrozensus — G-Kartei

liegt die Altersgruppe der unter 20jährigen, bei denen die Einstellungen die Entlassungen übertreffen dürften, absolut und relativ unter den Ergebnissen des Mikrozensus.

Genau umgekehrt aber ist es bei den Frauen der Altersgruppen 20 bis unter 25 und 25 bis unter 30 Jahren, deren Zahlen absolut und relativ erheblich über den entsprechenden Werten des Mikrozensus liegen. Hier schlägt sich das Ausscheiden aus dem Erwerbsleben in der sogenannten zweiten Phase nieder. Die „Entlassungen“ bei dieser Personengruppe übersteigen die Einstellungen und werden offenbar ebenfalls nicht alle registriert.

Wirtschaftsfachlich funktioniert das Meldeverfahren am schlechtesten im Dienstleistungsbereich (Handel, Verkehr, Kredit- und Nachrichtenwesen usw.), da hier ein sehr hoher Prozentsatz der Beschäftigten im Angestelltenverhältnis und mithin auch bei Ersatzkassen versichert ist. Die Abweichungen der G-Kartei in diesem Wirtschaftsbereich von den entsprechenden Mikrozensuszahlen belegen dies.

Diese Mängel des Meldesystems führen zwangsläufig per Saldo zu einer Überhöhung der Kartei. Hinzu kommt noch, daß für verschiedene G-Beschäftigte Doppelkarten bzw. Mehrfachkarten vorhanden sind. Die Gefahr dafür besteht immer dann, wenn ein Arbeitsplatzwechsel mit einem Wohnungswechsel zusammenfällt. Auch dies trägt zur Überhöhung in den Altersgruppen der 20- bis unter 25- und 25- bis unter 30jährigen, bei denen die Fluktuation am größten ist<sup>13)</sup>, bei.

Die hier gefundenen Ergebnisse und ihre Interpretation decken sich auch mit den Ergebnissen von zwei Untersuchungen der Bundesanstalt

aa) Im August 1965 wurde auf repräsentativer Basis geprüft, wieviel Doppel- bzw. Mehrfachkarten die G-Kartei enthält. Dabei stellte sich heraus, daß bei etwa 2 v. H. der Männer und 1,5 v. H. der Frauen mehrere Karten vorhanden waren. Es wurde weiterhin festgestellt, daß der Anteil der Mehrfachkarten bei den 20- bis 30jährigen Beschäftigten besonders groß war.

bb) In Zusammenarbeit mit dem Statistischen Bundesamt wurde für einige Zählbezirke die Übereinstimmung der vom Mikrozensus und von der G-Kartei erfaßten Personen für Ende April 1965 überprüft.

Nach dem Mikrozensus und der G-Kartei *zusammen* waren 310 Fälle zu untersuchen. 90 davon konnten jedoch nicht verglichen werden, da sie entweder nur im Mikrozensus oder nur in der G-Kartei erfaßt waren. Von 220 Fällen stimmten 188 in beiden Statistiken überein. Von den restlichen 32 waren lt. Mikrozensus 24 nicht erwerbstätig bzw. nicht als Arbeiter oder Angestellte erwerbstätig. Das *sind auf der Basis der*

<sup>13)</sup> Vergleiche: Der zwischenbetriebliche Arbeitsplatzwechsel im Bundesgebiet einschließlich Berlin (West) im Jahre 1965, in: Amtliche Nachrichten der Bundesanstalt für Arbeitsvermittlung und Arbeitslosenversicherung Heft Nr. 12/1966 und 7/1967.

220 vergleichbaren Fälle ca. 10 v. H. Die übrigen 8 erfüllten in anderen Punkten die Kriterien der G-Kartei nicht.

### 3. Die Altersverteilung des ersten Beschäftigungsverhältnisses

Will man die G-Kartei nicht nur als Auswahlgrundlage, sondern deren Inhalt unmittelbar zu Längsschnittanalysen heranziehen, so interessiert unter anderem, wie weit die Karteieintragen zeitlich im Mittel zurückreichen.

Aus diesem Grund wurde in der Auszählung vom April 1968 auch nach dem Datum des ersten Beschäftigungsverhältnisses gefragt. Dieses Datum ist nicht unbedingt identisch mit dem Alter der entsprechenden Karte, da ja manche Eintragungen sozusagen „nachgeholt“ werden, doch dürfte es in der überwiegenden Zahl der Fälle mit ihm identisch sein.

Die Altersverteilung des ersten Beschäftigungsverhältnisses hat ein arithmetisches Mittel von 16,7 Jahren bei einer Streuung von 11,7 Jahren.

Der Median liegt bei 14,77 Jahren. Diese Zeit, die in der Hälfte aller Fälle beobachtbar ist, dürfte für Längsschnittbeobachtungen schon brauchbar sein. Fraglich ist jedoch, ob die Eintragungen auf der Karte, insbesondere über den Beruf bzw. den Berufswechsel, den Anforderungen genügen. Die im Jahr 1970 ablaufende Großerhebung der Arbeitsämter, bei der G-Karteiangaben mit persönlichen Auskünften von Probanden verglichen werden können, dürfte darüber einige Auskünfte vermitteln.

### 4. Ober die Berufsstruktur

Leider ist es nicht gelungen, für die sich aus der G-Auszählung vom April 1968 ergebende Berufsstruktur eine Vergleichsbasis zu finden.

Trotzdem sind Aussagen über die Verteilung der Berufe nach der G-Kartei möglich:

Da die Berufe in den einzelnen Wirtschaftszweigen nicht identisch verteilt sind<sup>14)</sup>, die Wirtschaftszweige wiederum unterschiedlich repräsentiert sind, ergibt sich daraus eine Verzerrung der Berufsstruktur insgesamt. Diese Aussage ist *sicher*, sie ist das Implikat der gefundenen Ergebnisse.

Weiterhin ist *wahrscheinlich*, daß auch die gefundenen Abweichungen in der G-Kartei sowohl hinsichtlich des Alters als auch hinsichtlich des Geschlechts nicht ohne Einfluß auf die Berufsstruktur sind.

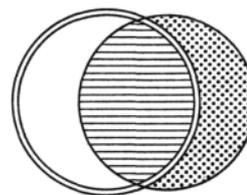
Es bleibt zu überlegen, inwieweit die Berufsstruktur für eine bestimmte Kombination (z. B. Berufsstruktur je Wirtschaftszweig) hinlänglich

genau ermittelt werden kann. Hier gilt tendenziell das gleiche, was oben bereits ausgeführt wurde: wären nur die Wirtschaftszweige (um beim Beispiel zu bleiben) untereinander verzerrt, so hätte das auf die Berufsstruktur pro Wirtschaftszweig vermutlich keinen Einfluß. Da jedoch Alter und Geschlecht, die beide als nicht unabhängig von der Berufsstruktur betrachtet werden müssen, in einzelnen Wirtschaftszweigen ebenfalls verzerrt auftreten, kann die G-Kartei auch in Kombinationen von Merkmalen keine einwandfreie Berufsstruktur liefern.

Die bisherigen Überlegungen sind Ableitungen aus den gefundenen Ergebnissen. Unerörtert blieb die Qualität der Eintragungen. Sieht man einmal davon ab, daß die Eintragungen selbst mit Fehlern behaftet sein können, wie dies auch bei den anderen Merkmalen zu erwarten ist, über deren Größenordnung jedoch keine Vorstellungen bestehen, so bleibt immer noch eine Schwäche im Fortschreibungssystem, die sicher nicht ohne Einfluß auf die Resultate zur Berufsstruktur ist: In der Regel werden *Berufswechselfälle* immer nur dann bekannt, wenn ein zwischenbetrieblicher Arbeitsplatzwechsel stattfindet. Ein innerbetrieblicher Arbeitsplatz- und/oder Berufswechsel wird dem Arbeitsamt nicht gemeldet. Dies ist bekannt und auch der Grund dafür, warum von der Bundesanstalt früher nie Beschäftigtenstatistiken aus der Totalkartei nach Berufen erstellt worden sind.

### IV. Die G-Kartei als Grundlage für Panel-Untersuchungen

Der Zustand der G-Kartei läßt sich an folgendem Diagramm veranschaulichen:



Der doppelt ausgezogene Kreis stellt die theoretisch einwandfreie G-Kartei dar; der einfach gezogene Kreis den tatsächlichen Bestand. Die Überschneidung beider Kreise (schraffierte Fläche) zeigt den in der G-Kartei richtig erfaßten Bestand; die leere Fläche repräsentiert die in der G-Kartei nicht erfaßten, die gepunktete Fläche die übererfaßten Personen.

Es stellt sich nun die Frage, ob eine Kartei in diesem Zustand den Anforderungen als Auswahlgrundlage für ein Panel genügt. Dies kann zunächst nur mit Nein beantwortet werden. Dabei ist aber unberücksichtigt gelassen, daß die Kartei — zunächst wenigstens einseitig — bereinigt werden kann.

<sup>14)</sup> Vergleiche dazu F. Stooß: Die Veränderungen der beruflichen Gliederung der Erwerbspersonen nach Wirtschaftszweigen in der Bundesrepublik 1950—1961, in: „Mitteilungen“ des IAB, Heft 4, Oktober 1968, S. 248 ff.

Nach Ziehung einer Unterstichprobe, gedacht ist an einen Umfang von etwa 100000 Personen, können, nach entsprechender Bearbeitung, alle Doppel- und Mehrfachkarten sowie alle definitiv nicht dem Personenkreis der G-Kartei zuzuordnenden Karten entfernt werden. Diese Unterstichprobe hätte dann nur noch den Fehler der Untererfassung, der auch kurzfristig nicht beseitigt werden kann. Aber auch dieser Fehler ließe sich in absehbarer Zeit korrigieren: mit Einführung der neuen Beschäftigtenstatistik, die ja *alle* rentenversicherungspflichtigen Beschäftigten enthalten soll, könnte der bisher nicht erfaßte Personenkreis an Hand des Stichprobenplanes leicht identifiziert werden. Mit dieser Stichprobe würde die Arbeitsmarkt- und Berufsforschung über ein Instrument verfügen, das insbesondere für soziologische Untersuchungen wie Ausbildungen, Berufsverläufe, Fluktuationen und ähnliche individuell zu erhebende Merkmale in Frage käme.

Ein weiterer Vorteil dieser Kartei ist, daß sie zur Hälfte Vergangenheitsinformationen von mindestens 15 Jahren enthält. Wenn diese Information auch nicht ohne weiteres verwertet werden kann, so dient sie doch als Gedächtnisstütze bzw. Kontrolle bei Befragungen. Zuletzt wäre noch zu nennen, daß mehrfache Befragungen unterschiedlichen Inhalts am gleichen Personenkreis zu einer Fülle von Merkmalen und damit auch Kombinationen führen, die anderweitig nicht zu erreichen sind. Zusammenfassend sprechen also drei gewichtige Gründe für die Ziehung einer Unterstichprobe aus der G-Kartei:

1. es ist gegenwärtig kein statistisches Instrument vorhanden und in absehbarer Zukunft keines zu erwarten, das, trotz der aufgezeigten Mängel, bessere Möglichkeiten für Verlaufsuntersuchungen bieten würde;
2. bei jeder anderen Stichprobe würde von vornherein auf eine ganze Menge Vergangenheitsinformationen verzichtet. Echte Längsschnittuntersuchungen könnten erst nach einer Reihe von Jahren durchgeführt werden;
3. die Unterstichprobe ergäbe die Basis für die auf dem Gebiet der Berufsforschung wünschenswerten Panel-Untersuchungen.

Umfang und Stichprobenplan sollen nach Vorliegen der Ergebnisse aus der Untersuchung des IAB über Berufsverläufe festgelegt werden. Diese sind im Herbst 1970 zu erwarten.

## C. Zusammenfassung

Wie die Ausführungen gezeigt haben, befriedigt die Qualität der G-Kartei kaum eines der vielfältigen statistischen Bedürfnisse in der Arbeitsmarkt- und Berufsforschung: die weiter oben an erster Stelle genannte Konjunkturbeobachtung, für die in der Bundesrepublik Deutschland sowohl hinsichtlich der Aktualität als auch der Genauigkeit keine ausreichenden Beschäftigtenstatistiken vorhanden sind, kann mit Zahlen aus der G-Kartei nicht einmal näherungsweise verfolgt werden.

Von den für Strukturforschungszwecke zu erhebenden Merkmalen erscheint nur das Alter als ausreichend genau repräsentiert. Die große Lücke, die besonders in der Berufsstatistik vorhanden ist, kann auch mit der G-Kartei nicht geschlossen werden.

Berücksichtigt man, daß für Längsschnittuntersuchungen der Repräsentationsgrad die geringste Rolle spielt, so scheint die G-Kartei hier am ehesten als statistisches Instrument geeignet. Für diesen Zweck ist jedoch eine Stichprobe dieses Umfanges keinesfalls erforderlich. Es ist deshalb auch vorgesehen, eine Unterstichprobe zu ziehen, deren Umfang auf etwa 100 000 Karten beschränkt bleibt. Diese Unterstichprobe muß bereinigt werden und anschließend weit intensiver bearbeitet werden als die jetzige G-Kartei. Sie wird in absehbarer Zeit (etwa 2—3 Jahre) in Verbindung mit der neuen Beschäftigtenstatistik auf einen völlig einwandfreien Stand gebracht und könnte dann nur noch den theoretisch zwar möglichen, doch bisher nicht nachweisbaren, durch das Auswahlverfahren bedingten systematischen Fehler haben.

Dieses Problem sollte zum Schluß nochmals aufgegriffen werden: die sehr schlechten Ergebnisse, die man bisher mit Daten aus der G-Kartei erzielte, wurden fast durchweg auf diesen systematischen Fehler zurückgeführt. Die hier vorgenommenen Untersuchungen haben jedoch ergeben, daß sie vorwiegend auf anderen Ursachen beruhen. Ob daneben tatsächlich ein systematischer Stichprobenfehler vorhanden ist, konnte leider nicht mehr festgestellt werden. Er würde sich jedoch wahrscheinlich in einer Größenordnung bewegen, die für die meisten der genannten statistischen Zwecke vertretbar wäre.