

# WIRTSCHAFT UND STATISTIK

- Sozialrechnungsmatrix für Deutschland • Maße für die faktische Anonymität von Mikrodaten • Geheimhaltungsmethoden auf dem Prüfstand – eine Analyse anhand der Umsatzsteuerstatistik
- Arbeitsmarkt in Deutschland • Mikrozensus • Unternehmen der Binnenschifffahrt • Hedonische Preismessung bei EDV-Investitionsgütern



**6/2004**

**Statistisches Bundesamt**

Herausgeber: Statistisches Bundesamt, Wiesbaden

Schriftleitung: Johann Hahlen  
Präsident des Statistischen Bundesamtes  
Verantwortlich für den Inhalt:  
Brigitte Reimann,  
65180 Wiesbaden

- Telefon: + 49 (0) 6 11/75-20 86
- E-Mail: [wirtschaft-und-statistik@destatis.de](mailto:wirtschaft-und-statistik@destatis.de)

Vertriebspartner: SFG Servicecenter Fachverlage  
Part of the Elsevier Group  
Postfach 43 43  
72774 Reutlingen  
Telefon: + 49 (0) 70 71/93-53 50  
Telefax: + 49 (0) 70 71/93-53 35  
E-Mail: [destatis@s-f-g.com](mailto:destatis@s-f-g.com)

Druck: Kern & Birner, Frankfurt am Main

Erscheinungsfolge: monatlich

Erschienen im Juli 2004

Einzelpreis: EUR 7,50 [D]

Jahresbezugspreis der Printversion: EUR 121,- [D]

zuzüglich Versandkosten

Bestellnummer: 1010200-04106-2 – ISSN 1619-2907



Allgemeine Informationen über das Statistische Bundesamt und sein Datenangebot erhalten Sie:

- im Internet: [www.destatis.de](http://www.destatis.de)

oder bei unserem Informationsservice  
65180 Wiesbaden

- Telefon: + 49 (0) 6 11/75-24 05
- Telefax: + 49 (0) 6 11/75-33 30
- E-Mail: [info@destatis.de](mailto:info@destatis.de)

#### Abkürzungen

WiSta	=	Wirtschaft und Statistik
MD	=	Monatsdurchschnitt
VjD	=	Vierteljahresdurchschnitt
HjD	=	Halbjahresdurchschnitt
JD	=	Jahresdurchschnitt
D	=	Durchschnitt (bei nicht addierfähigen Größen)
Vj	=	Vierteljahr
Hj	=	Halbjahr
a. n. g.	=	anderweitig nicht genannt
o. a. S.	=	ohne ausgeprägten Schwerpunkt
St	=	Stück
Mill.	=	Million
Mrd.	=	Milliarde

#### Zeichenerklärung

p	=	vorläufige Zahl
r	=	berichtigte Zahl
s	=	geschätzte Zahl
–	=	nichts vorhanden
0	=	weniger als die Hälfte von 1 in der letzten besetzten Stelle, jedoch mehr als nichts
.	=	Zahlenwert unbekannt oder geheim zu halten
...	=	Angabe fällt später an
X	=	Tabellenfach gesperrt, weil Aussage nicht sinnvoll
l oder —	=	grundsätzliche Änderung innerhalb einer Reihe, die den zeitlichen Vergleich beeinträchtigt
/	=	keine Angaben, da Zahlenwert nicht sicher genug
()	=	Aussagewert eingeschränkt, da der Zahlenwert statistisch relativ unsicher ist

Abweichungen in den Summen ergeben sich durch Runden der Zahlen.

© Statistisches Bundesamt, Wiesbaden 2004

Für nichtgewerbliche Zwecke sind Vervielfältigung und unentgeltliche Verbreitung, auch auszugsweise, mit Quellenangabe gestattet. Die Verbreitung, auch auszugsweise, über elektronische Systeme/Datenträger bedarf der vorherigen Zustimmung. Alle übrigen Rechte bleiben vorbehalten.

<b>Inhalt</b>		<b>Seite</b>
Kurznachrichten		595
<b>Textteil</b>		
<i>Manfred Klose, Alexander Opitz, Norbert Schwarz</i>	Sozialrechnungsmatrix für Deutschland	605
<i>Dr. Rainer Lenz, Roland Sturm, Dr. Daniel Vorgrimler</i>	Maße für die faktische Anonymität von Mikrodaten	621
<i>Dr. Rainer Lenz, Dr. Daniel Vorgrimler</i>	Geheimhaltungsmethoden auf dem Prüfstand – eine Analyse anhand der Umsatzsteuerstatistik	639
<i>Joachim Fritz, Dr. Josef Hartmann, Matthias Sacher</i>	Arbeitsmarkt in Deutschland – Analysen zu Ausfallprozessen in der Piloterhebung zum ILO-Erwerbsstatus	649
<i>Holger Breiholz</i>	Ergebnisse des Mikrozensus 2003	663
<i>Uwe Reim</i>	Unternehmen der Binnenschifffahrt 2002	673
<i>Dr. Stefan Linz, Timm Behrmann, Ulf Becker</i>	Hedonische Preismessung bei EDV-Investitionsgütern	682
<i>Gudrun Eckert</i>	Preise im Mai 2004	690
Übersicht über die im laufenden Jahr erschienenen Textbeiträge		695
<b>Tabellenteil</b>		
Inhalt		1*
Statistische Monatszahlen		2*

Für die Zeit vor dem 1. Januar 2002 ermittelte DM-Beträge wurden zum amtlich festgelegten Umrechnungskurs 1 Euro = 1,95583 DM in Euro umgerechnet. Aufgrund der kaufmännischen Rundung kann es bei der Summenbildung zu geringfügigen Abweichungen kommen. Auch vor dem 1. Januar 2002 aus DM-Werten errechnete Zuwachsraten und Anteile können aus diesem Grund geringfügig von den in Euro dargestellten Werten abweichen.

Angaben für die Bundesrepublik Deutschland nach dem Gebietsstand seit dem 3. 10. 1990. Die Angaben für das „frühere Bundesgebiet“ beziehen sich auf die Bundesrepublik Deutschland nach dem Gebietsstand bis zum 3. 10. 1990; sie schließen Berlin-West ein. Die Angaben für die „neuen Länder und Berlin-Ost“ beziehen sich auf die Länder Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, Thüringen sowie auf Berlin-Ost.

**Contents**

**Page**

News in brief

595

**Texts**

<i>Manfred Klose, Alexander Opitz, Norbert Schwarz</i>	Social accounting matrix for Germany	605
<i>Dr. Rainer Lenz, Roland Sturm, Dr. Daniel Vorgrimler</i>	Measures for the de-facto anonymity of microdata	621
<i>Dr. Rainer Lenz, Dr. Daniel Vorgrimler</i>	Confidentiality methods on the proving stand – an analysis taking the example of turnover tax statistics	639
<i>Joachim Fritz, Dr. Josef Hartmann, Matthias Sacher</i>	Labour market in Germany – analyses of nonresponse in the pilot survey on the ILO activity status	649
<i>Holger Breiholz</i>	Results of the microcensus, 2003	663
<i>Uwe Reim</i>	Enterprises engaged in inland water transport, 2003	673
<i>Dr. Stefan Linz, Timm Behrmann, Ulf Becker</i>	Hedonic price measurement for EDP investment goods	682
<i>Gudrun Eckert</i>	Prices in May 2004	690
	List of the contributions published in the current year	695

**Tables**

Summary	1*
Monthly statistical figures	2*

**Table des matières**

**Page**

Informations sommaires

595

**Textes**

<i>Manfred Klose, Alexander Opitz, Norbert Schwarz</i>	Matrice de comptabilité sociale pour l'Allemagne	605
<i>Dr. Rainer Lenz, Roland Sturm, Dr. Daniel Vorgrimler</i>	Mesures pour l'anonymité de fait des micro-données	621
<i>Dr. Rainer Lenz, Dr. Daniel Vorgrimler</i>	Méthodes de confidentialité sur le banc d'essai – une analyse à l'aide de la statistique des impôts sur le chiffre d'affaires	639
<i>Joachim Fritz, Dr. Josef Hartmann, Matthias Sacher</i>	Marché du travail en Allemagne – analyses concernant les processus de non-réponse dans l'enquête pilote de l'OIT sur le statut en matière d'emploi	649
<i>Holger Breiholz</i>	Résultats du microrecensement de 2003	663
<i>Uwe Reim</i>	Entreprises de la navigation intérieure en 2003	673
<i>Dr. Stefan Linz, Timm Behrmann, Ulf Becker</i>	Evaluation hédonique des prix de biens d'investissement informatiques	682
<i>Gudrun Eckert</i>	Prix en mai 2004	690
	Liste des contributions publiées dans l'année en cours	695

**Tableaux**

Résumé	1*
Chiffres statistiques mensuels	2*

The data for the Federal Republic of Germany relate to its territory since 3 October 1990. The data for the "former territory of the Federal Republic" relate to the territory of the Federal Republic of Germany before 3 October 1990; they include Berlin-West. The data for the "new Länder and Berlin-East" relate to the Länder of Brandenburg, Mecklenburg-Western Pomerania, Saxony, Saxony-Anhalt, Thuringia as well as to Berlin-East. Données pour la République fédérale d'Allemagne selon le territoire depuis le 3 octobre 1990. Les données pour «l'ancien territoire fédéral» se réfèrent à la République fédérale d'Allemagne, territoire jusqu'au 3 octobre 1990; Berlin-Ouest y est inclus. Les données pour les «nouveaux Länder et Berlin-Est» se réfèrent aux Länder Brandebourg, Mecklembourg-Poméranie occidentale, Saxe, Saxe-Anhalt, Thuringe ainsi qu'à Berlin-Est.

# Kurznachrichten

## Aus Europa

### 53. Sitzung des Ausschusses für das Statistische Programm (ASP)/ 42. EWR-Konferenz

Unter Beteiligung der Leiter der Statistischen Zentralämter der Europäischen Union (EU) fand am 27. Mai 2004 die 53. Sitzung des Ausschusses für das Statistische Programm (ASP)/42. EWR-Konferenz statt.

Dem ASP wurden drei Verordnungsentwürfe der Kommission und ein Entwurf einer Entscheidung der Kommission zur Abstimmung vorgelegt. Die Mehrheit des ASP stimmte dem Verordnungsentwurf der Kommission zur Durchführung einer Entscheidung des Europäischen Parlaments und des Rates hinsichtlich der Innovationsstatistik zu. Die Verordnung, die im Zusammenhang mit der Lissabonner Strategie über mehr Wachstum, Beschäftigung und sozialen Zusammenhalt in der EU aus dem Jahr 2000 begründet wird, enthält grundlegende Bestimmungen und einen technischen Anhang über zu erhebende Variablen. Zustimmung fand auch der Entwurf einer Kommissionsverordnung für die Erstellung jährlicher Statistiken über die Stahlindustrie für die Berichtsjahre 2003 bis 2009. Die Verordnung bezieht sich auf die Abänderung der Definition von Merkmalen aus den Bereichen Schrott- und Gussbruchwirtschaft, Elektroenergiebilanz und Investitionen und regelt das technische Format der Übermittlung von Daten. Den Mitgliedstaaten Belgien, Portugal, Griechenland und Zypern werden mit einer weiteren Kommissionsverordnung Abweichungen von den Bestimmungen der EG-Verordnung Nr. 2150/2002 über

Abfallstatistiken genehmigt. Damit sollen den betreffenden Mitgliedstaaten Übergangsfristen für die Anpassung der nationalen Datenerhebungssysteme eingeräumt werden. Die vom ASP verabschiedete dritte Entscheidung zur Durchführung einer Richtlinie des Rates über die statistische Erfassung des Güter- und Personenseeverkehrs aktualisiert die Liste der Häfen und die so genannten „Codes“ in den vier Systematiken Ladungsart, Küstengebiete, Nationalität der Flagge und Schiffstyp.

Das Statistische Amt der Europäischen Gemeinschaften (Eurostat) legte dem ASP einen überarbeiteten Entwurf einer Verordnung zu gemeinschaftlichen Statistiken über die Struktur und Tätigkeit von Auslandsunternehmenseinheiten (FATS) zur Stellungnahme vor. Mehrere Mitgliedstaaten – darunter auch Deutschland – halten zur Durchführung der FATS-Verordnung eine klare Regelung hinsichtlich des möglichen Austausches von Einzeldaten der einzelnen Mitgliedstaaten untereinander und die Gewähr der Vertraulichkeit dieser Daten für erforderlich. Mit dieser Thematik wird sich die von der ASP-Partnerschaftsgruppe eingerichtete Task Force „Rechts- und Verfassungsfragen“ befassen.

Dem ASP wurde ein erster Entwurf des Statistischen Jahresprogramms 2005 vorgelegt. Die Mehrheit der ASP-Mitglieder, darunter auch die deutschen Vertreter, sahen es als notwendig an, den vorliegenden Entwurf zu überarbeiten. Erforderlich sei eine klare Prioritätensetzung unter Beteiligung der Mitgliedstaaten und der europäischen Nutzer. Eurostat bat die Mitgliedstaaten um konkrete Vorschläge zur Reduzierung des Programmentwurfs bis Ende Juni 2004.

Der ASP nahm den Überblick über den Stand der Umsetzung der Konjunkturstatistikverordnung des Rates in den

einzelnen Mitgliedstaaten zur Kenntnis. Die Situation für Deutschland ist vergleichsweise sehr gut, dennoch brachte der deutsche Vertreter zum Ausdruck, dass die Anforderungen der Konjunkturstatistik insgesamt sehr hoch seien und einer Prioritätensetzung unterzogen werden sollten, bevor an eine Ausweitung der Konjunkturstatistik gedacht wird. Eurostat zieht zur vollständigen Durchführung der Verordnung neben bilateralen Verhandlungen auch die Einleitung von Vertragsverletzungsverfahren in Betracht.

Ferner begrüßte der ASP das vorgeschlagene Arbeitsprogramm zur Festlegung einer neuen Strategie 2005 im Bereich der Finanzhilfen, mit dem Eurostat die Konsequenzen aus den im vergangenen Jahr zu Tage getretenen inkorrekten Haushaltspraktiken bei Eurostat und einigen Eurostat Data Shops ziehen möchte. Der ASP beschließt die Einrichtung einer Task Force unter Teilnahme der EU-Mitgliedstaaten Deutschland, Finnland, Irland, der Niederlande, Österreich, Schweden, Slowenien und Spanien, die sich mit Detailfragen zur neuen Strategie befassen soll.

Der ASP begrüßte den von der Task Force vorgelegten „Bericht für eine verbesserte Prioritätensetzung im Rahmen der Statistischen Programmplanung“, der unter Mitwirkung Deutschlands entstanden ist. Eurostat sagte zu, die Vorschläge der Task Force ab dem Jahresprogramm 2006 zu berücksichtigen. Aus deutscher Sicht sind vor allem folgende Vorschläge von Bedeutung: Erstellung von Kostenkalkulationen durch Eurostat und die Mitgliedstaaten für Projekte und Vorhaben, stärkere Einbeziehung der Mitgliedstaaten und der europäischen Nutzer, Erstellung eines Inventars der europäischen Anforderungen und Rechtsgrundlagen durch Eurostat, zeitliche Befristung von Rechtsakten und bereichsweise genaue Durchforstung des Statistikprogramms, beginnend mit den Wirtschaftsstatistiken ab dem Jahr 2006.

Der ASP dankte Eurostat für den vorgelegten Fortschrittsbericht zum Thema Erfassung selbst genutzter Eigentümerwohnungen im Harmonisierten Verbraucherpreisindex (HVPI) und stimmte dem vorgeschlagenen Arbeitsprogramm zu. Der Bericht schlägt eine Ausweitung der bereits durchgeführten Pilotstudien auf weitere Mitgliedstaaten auf eine Erfassung der Kosten im Zusammenhang mit Bestandsimmobilien – bisher wurden nur Neubauten betrachtet – und auf die Entwicklung von Preisindizes für Bauland vor.

Der ASP begrüßte den Fortschrittsbericht über den Stand der Umsetzung der neuen Verbreitungspolitik von Eurostat. Kernpunkte des Papiers sind die kostenlose Bereitstellung von Daten durch Eurostat ab dem 1. Oktober 2004, der Aufbau von Beratungszentren durch die Nationalen Statistischen Ämter mit finanzieller Unterstützung Eurostats, die Überarbeitung der Eurostat-Internetseiten und die weitere Harmonisierung der Dokumentationen. Die deutsche Delegation wies darauf hin, dass vor der kostenlosen Bereitstellung von Regionaldaten eine nationale Abstimmung mit den Statistischen Ämtern der Länder erfolgen muss. Deutschland bat Eurostat um den Aufbau eines Eurostat-internen Prüfungssystems, um künftig Fehler bei der Verbreitung von Regionaldaten in der Datenbank New Cronos zu vermeiden.

Unter einem weiteren Tagesordnungspunkt berichtete die deutsche Delegation über die bisherigen Arbeiten der Task Force „Centres of Excellence“, in der neben Deutschland

Eurostat, Frankreich, Italien, die Niederlande, Österreich, Spanien und Schweden vertreten sind. Bislang hat die Task Force verschiedene Formen der Kooperation von Mitgliedstaaten im Rahmen von Netzwerken auf europäischer Ebene analysiert und mehrere Modelle für die so genannten „Centres of Excellence“ (Kompetenzzentren) entwickelt. Die Modelle reichen von einem Kompetenzzentrum, das ein Projekt nach einem Ausschreibungsverfahren durchführt, bis hin zur Einrichtung von Agenturen unter dem Dach der Kommission. Eurostat regte an, die künftigen Arbeiten der Task Force auf die einfacheren Modelle zu konzentrieren und die Einrichtung von Agenturen nicht weiter zu verfolgen.

Die niederländische Delegation stellte den Entwurf eines Mandats der Arbeitsgruppe „Statistik“ des Rates vor, dem der ASP im Grundsatz zustimmte. Die Arbeitsgruppe soll sich nicht nur mit statistikrelevanten Rechtsakten befassen, sondern generell auch über statistikrelevante Themen in anderen Ratsarbeitsgruppen informiert werden – zum Beispiel dadurch, dass sie mit besonders bedeutsamen statistikrelevanten Berichten befasst wird. Die Niederlande haben im zweiten Halbjahr 2004 die Ratspräsidentschaft inne und werden sich um eine Umsetzung der Vorschläge bemühen.

## Aus dem Inland

### Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik

Am 4. und 5. Juni 2004 fand in Wiesbaden die 2. Konferenz für Sozial- und Wirtschaftsdaten statt. Ziel der Konferenz, die vom Bundesministerium für Bildung und Forschung (BMBF) finanziert wurde, war die Weiterentwicklung der Dateninfrastruktur für die Sozial- und Wirtschaftswissenschaften in Deutschland. Veranstalter waren der Gründungsausschuss des Rates für Sozial- und Wirtschaftsdaten und die beiden Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder.

Viele Konferenzteilnehmerinnen und -teilnehmer nutzten die Möglichkeit zum Dialog mit den Mitarbeiterinnen und Mitarbeitern der Forschungsdaten- und Datenservicezentren. Diese Zentren arbeiten an der Verbesserung des Zugangs der empirisch arbeitenden Wissenschaft zu Daten der amtlichen Statistik und weiterer wichtiger Datenproduzenten im öffentlichen Bereich und ermöglichen empirisches Arbeiten im Bereich der Sozial- und Wirtschaftswissenschaften nach modernen technologischen Standards. Der Verbesserung des Datenzugangs kommt auch bei der Reformdiskussion in Deutschland im Hinblick auf die Bewertung unterschiedlicher Lösungsansätze eine wichtige Rolle zu. Auf der Agenda der Wiesbadener Konferenz standen daher auch inhaltliche, methodische und rechtliche Themen zum Datenbedarf und zu Datenproblemen, zum Beispiel in den Bereichen Soziale Sicherung, Gesundheit und Pflege, Bildung und Arbeit.

Neben dem wissenschaftlichen Programm des Kongresses wurden auch die Mitglieder des neu eingerichteten Rates

für Sozial- und Wirtschaftsdaten (RatSWD) gewählt. Rund 600 Wissenschaftlerinnen und Wissenschaftler aus den einschlägigen Fachdisziplinen haben die Chance zur Wahl genutzt. Weitere Informationen zur Konferenz und zum Rat für Sozial- und Wirtschaftsdaten finden sich im Internet unter <http://www.RatSWD.de>.

## Einladung zum ZUMA-Workshop

Unter dem Titel „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Arbeiten mit den Daten“ veranstaltet das Zentrum für Umfragen, Methoden und Analysen (ZUMA), Mannheim, vom 29. bis 30. September 2004 in enger Kooperation mit dem Statistischen Bundesamt (Gruppe VIII C – Mikrozensus, Arbeitskräftestichprobe, Haushalte und Familie) unter der Leitung von Helga Christians einen Workshop für die Nutzer des Mikrozensus-Scientific Use Files.

Neben den gängigen großen sozialwissenschaftlichen Erhebungen wie der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS), dem Sozio-oekonomischen Panel (SOEP) und dem Wohlfahrtssurvey stehen der Wissenschaft seit einiger Zeit mit den Scientific Use Files des Mikrozensus (70%-Substichproben) auch Daten der größten laufenden amtlichen Haushaltsstichprobe zur Verfügung. Diese faktisch anonymisierten Daten enthalten für über 500 000 Personen und rund 230 000 Haushalte vielfältige Informationen über die soziale und wirtschaftliche Lage der Bevölkerung, der Haushalte und Familien. Allerdings unterscheidet sich der Mikrozensus in Konzeption, Inhalt und Zugangsbedingungen von den typischen sozialwissenschaftlichen Erhebungen.

Der Workshop legt den Schwerpunkt auf das praktische Arbeiten mit den Daten des Mikrozensus. Die Vorträge des ersten Tages dienen sowohl der Einführung in den Mikrozensus als auch der thematischen Vorbereitung auf die praktischen Übungen. Am zweiten Tag wird in parallel stattfindenden Sitzungen anhand ausgewählter Themen eine praktische Einführung in das Arbeiten mit den Daten des Mikrozensus gegeben. Für methodisch interessierte Wissenschaftler und Studierende wird darüber hinaus ein Kurs zum Stichprobenfehler und zur Gewichtung im Mikrozensus angeboten.

Im Zentrum steht die Vermittlung von anwendungsorientiertem Basiswissen sowohl für interessierte Wissenschaftler und Studierende, die bisher nicht oder kaum mit den Daten des Mikrozensus gearbeitet haben, als auch für Forscher mit ersten Erfahrungen. Voraussetzung sind Grundkenntnisse in SPSS.

Der Teilnahmebeitrag beläuft sich auf 120,- Euro (Studierende: 80,- Euro). Die Teilnehmerzahl ist auf 16 begrenzt, wobei pro Themenblock maximal 8 Personen vorgesehen werden können.

Interessenten werden gebeten, sich beim ZUMA-Tagungssekretariat ([workshop@zuma-mannheim.de](mailto:workshop@zuma-mannheim.de), Telefon 06 21 / 1 24 62 21) anzumelden. Anmeldeschluss ist der 16. August 2004.

## Neuerscheinungen

### Im Blickpunkt: Energie und Umwelt

Aus Anlass der „Internationalen Konferenz für Erneuerbare Energien“, die Anfang Juni 2004 in Bonn stattgefunden hat, ist vom Statistischen Bundesamt in der Reihe „Im Blickpunkt“ ein neues Heft mit dem Titel „Energie und Umwelt“ in englischer Sprache herausgebracht worden. Darin wird auf 40 Seiten mit Texten, Grafiken und Tabellen ein statistischer Überblick zum Thema Umwelt und Energiewirtschaft in Deutschland gegeben.

Ein zentrales Ziel der Energiepolitik in Deutschland ist, dass die Energie sicher, ökonomisch und ökologisch verträglich produziert und verwandt wird. Dabei wird die natürliche Umwelt vielfältig genutzt: Materialien werden für die Energiegewinnung gefördert, die Landschaft wird als Fläche für die Produktion und Verteilung der Energie benötigt und die Natur wird als „Senke“ für entstandene Rückstände genutzt. Das Prinzip der Nachhaltigkeit verlangt, dass die Natur nur in dem Maße in Anspruch genommen wird, dass auch künftige Generationen sich an einer intakten Umwelt erfreuen können.

Normalerweise werden Energie- und Umweltfragen getrennt untersucht. Der neue Blickpunktband bringt jetzt Daten zusammen, die über Umweltaspekte von Energieangebot und -nachfrage berichten. So wird u. a. den Fragen nachgegangen, wie sich seit Beginn der 1990er-Jahre die Energieeffizienz entwickelt hat, wie sich Energiegewinnung und -verbrauch auf den Ausstoß von Luftschadstoffen sowie die Wassergewinnung und Abwasserbeseitigung ausgewirkt haben und welche Rolle die Abfallwirtschaft im Rahmen der Energiewirtschaft spielt. Die Darstellung geht auch auf die zunehmende Bedeutung der erneuerbaren Energien ein.

Die Buchausgabe „In the Spotlight: Energy and Environment“ (ISBN 3-8246-0723-9) ist zum Preis von 10,- Euro im Buchhandel und über den Statistik-Shop auf der Internetseite des Statistischen Bundesamtes (<http://www.destatis.de/shop>) erhältlich. Der Band steht hier außerdem zum Preis von 7,50 Euro im PDF-Format als Download zur Verfügung; eine Leseprobe der Veröffentlichung „In the Spotlight: Energy and Environment“ kann als PDF-Dokument kostenfrei heruntergeladen werden. Journalisten erhalten Rezensionsexemplare der Publikation bei der Pressestelle des Statistischen Bundesamtes (Telefon 06 11 / 75-34 44, E-Mail: [presse@destatis.de](mailto:presse@destatis.de)).

Weitere Informationen erteilt  
Dr. Bernd Becker, Telefon 0 18 88 / 6 44-82 00,  
E-Mail: [umwelt@destatis.de](mailto:umwelt@destatis.de).

### Mikrodatenfiles der Sozialhilfempfeänger für die Berichtsjahre 2001 und 2002 auf CD-ROM erhältlich

Gemäß § 132 Abs. 2 des Bundessozialhilfegesetzes stellen die Statistischen Ämter der Länder dem Statistischen Bundesamt jedes Jahr u. a. Einzelangaben aus einer Zufallsstichprobe mit einem Auswahlatz von 25% der Empfänger

von laufender Hilfe zum Lebensunterhalt zur Verfügung. Nach Anonymisierung und nutzerfreundlicher Aufbereitung werden die Mikrodaten anschließend vom Statistischen Bundesamt auf CD-ROM als so genanntes „Public Use File“ veröffentlicht.

Insgesamt ergeben sich dadurch für Politik, Verwaltung, Wissenschaft und interessierte Öffentlichkeit neue und von Forschung/Lehre gewünschte Auswertungsmöglichkeiten, die – im Vergleich zur Standardaufbereitung der Sozialhilfestatistik – weitaus tieferegehende Analysen zum Sozialhilfebezug in Deutschland ermöglichen.

Die Mikrodatensätze zum Jahresende 2001 bzw. 2002 umfassen jeweils knapp 700 000 Empfänger/-innen von laufender Hilfe zum Lebensunterhalt außerhalb von Einrichtungen (= sog. „Sozialhilfe im engeren Sinne“). Es sind sowohl personen- als auch haushaltsbezogene Daten auf den CD-ROMs enthalten. Insgesamt 61 Variablen (Berichtsjahr 2001) bzw. 62 Variablen (Berichtsjahr 2002), wie zum Beispiel Geburtsjahr, Geschlecht, Staatsangehörigkeit, Erwerbsstatus, höchster Schul- bzw. Berufsausbildungsabschluss, Haushaltstyp, Anspruch und Bruttobedarf der Bedarfsgemeinschaften je Monat, vorhandene Einkommensarten usw. ermöglichen eine Vielzahl von Auswertungsmöglichkeiten.

Auch methodisch gesehen ist die 25%-Stichprobe eine interessante Datenquelle: Sie bietet dem Nutzer einerseits eine sehr große Anzahl von Einzeldatensätzen und erfüllt somit hinsichtlich der Repräsentativität hohe Ansprüche. Darüber hinaus kann der Nutzer seine errechneten Stichprobenergebnisse mit den aggregierten Gesamtzahlen der Totalerhebung abgleichen, welche beispielsweise in den Fachserien oder Arbeitsunterlagen des Statistischen Bundesamtes veröffentlicht sind.

Die Daten stehen dem Nutzer auf den CD-ROMs als ASCII-Files sowie als SPSS- und SAS-Dateien zur Verfügung. Ergänzend enthalten die Datenträger Hilfsdateien, die das Konvertieren des ASCII-Files in SPSS- oder SAS-Dateien ermöglichen. Die genauen Datensatzbeschreibungen sowie allgemeine Informationen zur Konzeption und Qualität der Stichproben runden den Lieferumfang ab.

Die CD-ROMs können zum Preis von jeweils 199,- Euro zuzüglich Versandkosten beim Vertriebspartner des Statistischen Bundesamtes (nähere Angaben siehe Impressum) sowie über den Statistik-Shop (<http://www.destatis.de/shop>) angefordert werden. Sofern Interesse an der 2001er- und der 2002er-Ausgabe besteht, können beide CD-ROMs zum Setpreis von 298,- Euro zuzüglich Versandkosten bezogen werden. Ausgewählte Nutzergruppen erhalten einen Rabatt.

Auch die entsprechenden CD-ROMs mit den Mikrodaten zum Stichtag 31.12.1997 bis 31.12.2000 können beim Statistischen Bundesamt zum Preis von je 199,- Euro zuzüglich Versandkosten bezogen werden.

Weitere Auskünfte erteilen

Tim Weber, Telefon 0 1888 / 6 44-81 40,  
E-Mail: [tim.weber@destatis.de](mailto:tim.weber@destatis.de) sowie  
Sascha Krieger, Telefon 0 1888 / 6 44-81 59,  
E-Mail: [sascha.krieger@destatis.de](mailto:sascha.krieger@destatis.de).

## Campus File zur Sozialhilfestatistik 1998

Das Forschungsdatenzentrum des Statistischen Bundesamtes bietet erstmals ein Campus File zur Sozialhilfestatistik (Stichtag 31.12.1998) an. Das Campus File ist ein Public Use File, das speziell für Lehrende und Studierende erstellt wird. Damit soll die praktische Statistikausbildung mit amtlichen Einzeldaten angereichert und den Hochschulen ein neues Werkzeug für eine fundierte empirische Lehre zur Verfügung gestellt werden. Das Campus File zur Sozialhilfestatistik 1998 steht als kostenfreier Download im Internet unter <http://www.destatis.de/fdz/leistungen/campusfiles.htm> bereit.

## Aktuelle Informationen zum Stand der Raumentwicklung in Deutschland und Europa

Über den aktuellen Stand der Raumentwicklung in Deutschland und Europa unterrichten jährlich der vom Bundesamt für Bauwesen und Raumordnung in Kooperation mit den Statistischen Ämtern des Bundes und der Länder herausgegebene Bericht „Aktuelle Daten zur Entwicklung der Städte, Kreise und Gemeinden“ sowie die CD-ROM „INKAR – Indikatoren und Karten zur Raumentwicklung“. Die Vielzahl und Vielfalt der angebotenen regionalstatistischen Indikatoren ermöglichen umfassende, raum- und zeitvergleichende Analysen von Lebensbedingungen im Bundesgebiet und in den Regionen der Europäischen Union.

Die Ausgaben 2003 des Berichtes und der CD-ROM bieten rund 200 bundesweit bis auf die Ebene von Kreisen und kreisfreien Städten berechnete Indikatoren zu den Themen: Bevölkerung und Siedlungsstruktur, Natürliche Entwicklung, Wanderungen, Alters- und Sozialstruktur, Wirtschaft, Beschäftigung und Arbeitslosigkeit, Bildung, Medizinische Versorgung, Verkehr und Erreichbarkeit, Baulandmarkt und Bautätigkeit sowie Gebäude- und Wohnungsbestand.

Das Tabellenprogramm der Vorjahre wurde zum größten Teil fortgeschrieben; neue Indikatoren werden insbesondere zu den Themenbereichen „Siedlungsstruktur“ und „Erreichbarkeitsverhältnisse“ angeboten. Ergänzt werden die auf Deutschland bezogenen kreisscharf abgegrenzten Informationen durch 25 Indikatoren zu den Themen Bevölkerung, Beschäftigung, Arbeitslosigkeit, Wohnungen und Finanzsituation der Gemeinden mit mehr als 20 000 Einwohnern.

Zur Darstellung der Raumentwicklung in Europa werden 40 Indikatoren zur Bevölkerungs- und Altersstruktur, zu Erwerbstätigkeit und Arbeitslosigkeit angeboten. Damit kann die Situation einzelner Kreise, Gemeinden und Regionen auch vor dem Hintergrund der Regionalentwicklung in Europa analysiert werden. Die europäischen Indikatoren werden auf den EU-NUTS-Ebenen 0, 1 und 2 nachgewiesen. Ergänzt werden sie um eine Auswahl zentraler Indikatoren für die mehr als 1 000 Regionaleinheiten der NUTS-Ebene 3.

Mit der CD-ROM INKAR 2003 kann der Nutzer eigenständig Tabellen, Karten und Diagramme zum Stand und zur Entwicklung der Lebensverhältnisse in Deutschland und Europa erstellen. Ein einfach zu bedienendes Datenretrievalprogramm bietet die Möglichkeit zur beliebigen Kombination von Indikatoren aus verschiedenen Beobachtungsbereichen und Raumbezügen in eigenen Tabellen. Diese können sehr komfortabel in alle gängigen Formate exportiert werden. Mit dem ebenfalls enthaltenen Kartographieprogramm lassen sich alle Indikatoren mit nur wenigen Mausklicks als thematische Karten auf der Ebene von Stadt- und Landkreisen, Regionen, Bundesländern oder den Regionen der Europäischen Union darstellen. Die Karten können entweder alle Raumeinheiten umfassen und somit Deutschland bzw. die Europäische Union komplett abdecken oder auch nur einzelne ausgewählte Raumeinheiten darstellen (geographische Selektion). Hierzu steht neben einer Suchroutine eine sehr komfortable, stufenlose Zoomfunktion zur Verfügung.

INKAR 2003 bietet ferner die Möglichkeit, sämtliche Indikatorenwerte auf den verschiedenen räumlichen Einheiten graphisch als Balken-, Linien- oder Kreisdiagramm darzustellen. Diese Graphiken können selbstverständlich ausgedruckt werden; der Anwender kann sie aber auch wie die Karten exportieren, um sie in andere Anwendungen oder Dokumente zu integrieren.

Der Bericht „Aktuelle Daten zur Entwicklung der Städte, Kreise und Gemeinden, Ausgabe 2003“ (ISBN: 3-87994-067-3, Preis: 20,- Euro zuzüglich Versandkosten) und die CD-ROM „INKAR – Indikatoren und Karten zur Raumentwicklung“, CD-ROM zu Band 17 der Schriftenreihe „Berichte“, Ausgabe 2003 (ISBN: 3-87994-924-7, Preis: 25,- Euro zuzüglich Versandkosten) sind über den Selbstverlag des Bundesamtes für Bauwesen und Raumordnung (Telefon: 01888/401-2209, Telefax: 01888/401-2292, E-Mail: [selbstverlag@bbr.bund.de](mailto:selbstverlag@bbr.bund.de)) zu beziehen.

## Kompakt

### Lebenserwartung in Deutschland hat erneut zugenommen

Die durchschnittliche Lebenserwartung in Deutschland beträgt nach der aktuellen Sterbetafel 2000/2002 für einen neugeborenen Jungen 75,4 Jahre und für ein Mädchen 81,2 Jahre. In der vorhergehenden Sterbetafel 1999/2001 waren hierfür 75,1 bzw. 81,1 Jahre errechnet worden; die Lebenserwartung für Jungen hat also stärker zugenommen als die für Mädchen.

Auch für ältere Personen ist die Lebenserwartung gestiegen. So kann nach der aktuellen Sterbetafel 2000/2002 ein 60-jähriger Mann noch mit einer ferneren Lebenserwartung von im Durchschnitt 19,7 Jahren rechnen gegenüber 19,5 Jahren nach der vorhergehenden Sterbetafel. Für eine gleichaltrige Frau ergeben sich statistisch noch 23,8 weitere Lebensjahre gegenüber zuvor 23,7 Jahren.

Die Sterbetafeln basieren auf den Angaben über die Gestorbenen und die Bevölkerung der letzten drei Jahre.

Die durchschnittliche Lebenserwartung gibt damit die Zahl der weiteren Lebensjahre an, die eine Person in einem bestimmten Alter nach den gegenwärtigen Sterblichkeitsverhältnissen voraussichtlich noch leben könnte. Eine mögliche Veränderung in den kommenden Jahren wird dabei nicht berücksichtigt.

Im internationalen Vergleich nimmt Deutschland keine Spitzenstellung ein. Unter den Mitgliedstaaten der Europäischen Union weisen einige Länder deutlich höhere Lebenserwartungen als Deutschland auf. So betrug die durchschnittliche Lebenserwartung im Jahr 2002 für neugeborene Jungen in Schweden 77,7 Jahre und für neugeborene Mädchen in Spanien 83,1 Jahre, das sind 2,3 bzw. 1,9 Jahre mehr als in Deutschland. Gegenüber der durchschnittlichen Lebenserwartung der Europäischen Union (25 Länder) kann Deutschland für neugeborene Jungen eine um 0,6 Jahre und für neugeborene Mädchen eine um 0,1 Jahre höhere Lebenserwartung verzeichnen. International gehört Japan zu den Ländern mit der höchsten durchschnittlichen Lebenserwartung, für neugeborene Jungen lag diese dort im Jahr 2002 bei 78,3 Jahren, für Mädchen bei 85,2 Jahren.

Die aktuelle Sterbetafel 2000/2002 für Deutschland steht als Excel-Tabelle im Internetangebot des Statistischen Bundesamtes unter <http://www.destatis.de/download/d/bevoe/sterbetafel02.xls> zur Verfügung.

Weitere Auskünfte erteilt  
Matthias Eisenmenger, Telefon 06 11 / 75-32 75,  
E-Mail: [matthias.eisenmenger@destatis.de](mailto:matthias.eisenmenger@destatis.de).

## Rauchgewohnheiten der Bevölkerung in Deutschland und ihre Folgen

### Fast drei Viertel der Bevölkerung sind Nichtraucher

Im Mai 2003 waren 73% der Bevölkerung im Alter ab 15 Jahren Nichtraucher: 54% hatten nie geraucht, 19% waren ehemalige Raucher. Insgesamt waren 78% der Frauen Nichtraucherinnen und 67% der Männer Nichtraucher. Das zeigen die Ergebnisse der Mikrozensus-Zusatzbefragung, bei der alle vier Jahre knapp 0,5% der Bevölkerung (370 000 Personen) zu gesundheitsbezogenen Themen befragt werden, darunter auch zu ihren Rauchgewohnheiten.

Der Anteil der Nichtraucher wächst mit zunehmendem Alter der Befragten. Der niedrigste Wert von 59% ist bei den 20- bis unter 25-Jährigen zu finden. Bei den 50- bis unter 55-Jährigen sind es 70%. Die Ausstiegsquote, das heißt der Anteil der ehemaligen Raucher an der Gesamtbevölkerung, ist bei Männern mit 24% größer als bei Frauen mit 13%. Das liegt daran, dass 65% der Frauen nie geraucht haben – bei Männern liegt dieser Anteil bei 43%.

Zum Rauchen bekannten sich im Mai 2003 dagegen 27% der Bevölkerung im Alter ab 15 Jahren, 1999 waren es 28% gewesen. Der Anteil der Raucher betrug 2003 bei den Männern 33%, bei den Frauen 22%. In jeder Altersgruppe rauchten Frauen weniger häufig als Männer.

### 363 Mill. Zigaretten werden pro Tag geraucht

Pro Tag gehen in Deutschland durchschnittlich 363 Mill. Zigaretten in Rauch auf. Dieser Wert wurde auf Grundlage der Zahl der versteuerten Zigaretten im Kalenderjahr 2003 ermittelt. Daneben werden täglich durchschnittlich noch 9 Mill. Zigarren und Zigarillos, 51 t Feinschnitt und 2 t Pfeifentabak geraucht. Rein rechnerisch kommen damit auf jeden Einwohner – ob Raucher oder Nichtraucher, Erwachsener oder Kind – vier Zigaretten pro Tag und jeden zehnten Tag eine Zigarre oder ein Zigarillo.

Verglichen mit dem Vorjahr konsumierten die Raucher im Jahr 2003 weniger Zigaretten (-9%), aber deutlich mehr Feinschnitt (+20%).

Von dem Tabakkonsum profitiert auch die Bundeskasse, die im Kalenderjahr 2003 hierfür 14,1 Mrd. Euro (39 Mill. Euro pro Tag) Steuern eingenommen hat. Auf jeden Einwohner entfällt damit rechnerisch ein Tabaksteueraufkommen von 47 Cent pro Tag.

### Lungenkrebs ist fünfthäufigste Todesursache

An Erkrankungen, die in Zusammenhang mit dem Konsum von Tabakprodukten gebracht werden können, starben im Jahr 2002 in Deutschland 40 715 Personen. Neben Lungenkrebs (39 105 Sterbefälle) werden auch Kehlkopfkrebs (1 583 Sterbefälle) und Luftröhrenkrebs (27 Sterbefälle) zu den Folgeerkrankungen des Rauchens gezählt. An den genannten Todesursachen starben dreimal mehr Männer als Frauen. Insgesamt waren im Jahr 2002 4,8% aller Sterbefälle auf eine für Raucher symptomatische Erkrankung zurückzuführen.

Allein der Lungenkrebs stellt in Deutschland die fünfthäufigste Todesursache dar; bei den Männern war er sogar – wie bereits im Jahr 2001 – die dritthäufigste Todesursache. Bei Frauen war Lungenkrebs im Jahr 2001 die neunthäufigste Todesursache, im Jahr 2002 rangierte er auf Platz 10. Das durchschnittliche Alter der an Lungen-, Kehlkopf- und Luftröhrenkrebs Gestorbenen lag 2002 bei 68,9 Jahren.

Auch unter den im Krankenhaus im Jahr 2001 aufgrund bösartiger Neubildungen der Lunge, des Kehlkopfes und der Luftröhre stationär behandelten Fällen (179 950) dominierten die Männer mit 75,2%.

Bösartige Neubildungen der zuvor genannten Art waren der vierthäufigste Grund für einen stationären Krankenhausaufenthalt von Männern, bei Frauen war nur jeder 53. Krankenhausaufenthalt darauf zurückzuführen.

### Computer in der Freizeit von jungen Männern bevorzugt

Männliche Jugendliche zwischen 15 und 17 Jahren verbringen täglich nahezu eineinhalb Stunden ihrer Freizeit mit dem Computer – so lange wie keine Bevölkerungsgruppe

sonst. Vor allem Computerspiele stehen dabei mit mehr als einer Stunde Zeitaufwand im Vordergrund. Junge Frauen im gleichen Alter beschäftigen sich dagegen weniger als eine halbe Stunde ihrer täglichen Freizeit mit dem Computer (darunter 11 Minuten mit Computerspielen).

Betrachtet man die Altersgruppe der 10- bis 14-jährigen Kinder und Jugendlichen, so sitzen die Jungen täglich im Durchschnitt eine Stunde während ihrer Freizeit am Computer. Die gleichaltrigen Mädchen nutzen den Computer hingegen nur knapp eine halbe Stunde. Mit zunehmendem Alter verliert der Computer als Ort der Freizeittätigkeit an Bedeutung. Bereits in der Altersgruppe der 18- bis 24-jährigen Männer nimmt die Beschäftigung mit dem Computer in der Freizeit deutlich ab, auf rund eine Stunde am Tag. Der Rückgang verstärkt sich bei den 25- bis 39-jährigen noch weiter. Männer verbringen dann täglich weniger als eine halbe Stunde ihrer Freizeit am Computer, Frauen 8 Minuten. Seniorinnen und Senioren ab 65 Jahren verwenden den Computer als Freizeitbeschäftigung im Durchschnitt nur noch 6 Minuten täglich.

Diese Verhaltensweisen zeigen Ergebnisse der Zeitbudgeterhebung, einer der umfassendsten Studien zur Zeitverwendung in Deutschland, die im Auftrag des Bundesministeriums für Familie, Senioren, Frauen und Jugend 2001/02 zum zweiten Mal nach 1991/92 vom Statistischen Bundesamt durchgeführt wurde.

Weitere Auskünfte erteilt

Erlend Holz, Telefon 0 18 88 / 6 44-88 02,  
E-Mail: [erlend.holz@destatis.de](mailto:erlend.holz@destatis.de).

### Einfuhren aus Südostasien gingen im Jahr 2003 weiter zurück

Im Jahr 2003 wurden von Deutschland Waren im Wert von 11,8 Mrd. Euro in die zur ASEAN (Association of Southeast Asian Nations) gehörenden südostasiatischen Staaten Brunei Darussalam, Indonesien, Laos, Kambodscha, Malaysia, Myanmar, Philippinen, Singapur, Thailand und Vietnam exportiert. Gegenüber dem Vorjahr nahmen die deutschen Ausfuhren um 0,2% zu, nachdem sie 2002 gegenüber 2001 um 3,0% zurückgegangen waren.

Die wichtigsten Handelspartner in der Region sind Singapur und Malaysia. In diese beiden Staaten gingen im Jahr 2003 mit 7,0 Mrd. Euro alleine 59,3% des in die ASEAN-Staaten ausgeführten Wertes. Der Anteil der Ausfuhren in die ASEAN-Staaten an den gesamten deutschen Ausfuhren betrug im Jahr 2003 1,8%, gegenüber einem Anteil von 2,8%, den die ASEAN-Lieferungen an den gesamten deutschen Einfuhren ausmachten.

Bei den Einfuhren setzte sich der Negativtrend der letzten Jahre fort, gegenüber 2002 war im Jahr 2003 ein Rückgang um 3,3% auf 14,9 Mrd. Euro zu verzeichnen. Seit dem Jahr 2000 beträgt der Rückgang bei den Einfuhren aus den ASEAN-Ländern somit 8,2%. Für die Einfuhren aus den beiden wichtigsten ASEAN-Lieferländern Malaysia und Sin-

gapur wurde 2003 gegenüber 2002 ein Plus von 2,1% bzw. ein Rückgang von 7,2% verzeichnet.

Zu den wichtigsten Exportartikeln im Warenhandel mit den ASEAN-Staaten zählten 2003 – wie schon in den Vorjahren – Waren der Nachrichtentechnik, Rundfunk- und Fernsehgeräte sowie elektronische Bauelemente mit einem Ausfuhrwert von 2,9 Mrd. Euro und Maschinen mit einem Ausfuhrwert von 2,4 Mrd. Euro. Bei den Einfuhren dominierten, ebenfalls traditionell, Erzeugnisse der Nachrichtentechnik, Rundfunk- und Fernsehgeräte und elektronische Bauelemente (Einfuhrwert: 3,9 Mrd. Euro) sowie Büromaschinen, Datenverarbeitungsgeräte und -einrichtungen (3,5 Mrd. Euro).

Weitere Auskünfte erteilt

Konrad Schemer, Telefon 06 11 / 75-23 12,

E-Mail: [konrad.schemer@destatis.de](mailto:konrad.schemer@destatis.de).

## Hochschulausgaben 2002

Die öffentlichen und privaten Hochschulen in Deutschland gaben im Jahr 2002 insgesamt 30,4 Mrd. Euro aus. Damit erhöhten sich die Hochschulausgaben gegenüber 2001 um 6,0%, die Studierendenzahlen stiegen im gleichen Zeitraum um 3,8%.

Hinter dem durchschnittlichen Anstieg von 6% – deutlich mehr als in den Vorjahren (2001: +4,1%; 2000: +1,6%) – stehen je nach Bundesland sehr unterschiedliche Entwicklungen. In den meisten Ländern lagen die Zuwachsraten der Ausgaben unter 3%, so in Sachsen (+2,6% gegenüber 2001), Rheinland-Pfalz (+2,5%), Berlin (+2,3%), Bayern (+1,8%), Schleswig-Holstein (+0,9%) sowie dem Saarland (+0,6%). In Hessen blieben sie mit –0,3% praktisch unverändert und in Hamburg ergab sich vor allem wegen außergewöhnlicher Ausgaben für Baumaßnahmen im Jahr 2001 ein Rückgang um 13,5%. Die hohe Steigerungsrate von 32,4% in Bremen wurde durch die Neugründung einer privaten Universität verursacht. Bei Thüringen geht die Steigerung um 21,0% primär auf größere Baumaßnahmen im Berichtsjahr zurück. Die überdurchschnittliche Ausgabeentwicklung in Nordrhein-Westfalen (+15,3%) und Niedersachsen (+13,3%) hat unter anderem ihren Grund in der Einführung von Mietzahlungen der Hochschulen an das Land.

Mit 18,2 Mrd. Euro, das entspricht 60,1% aller Ausgaben, waren die Personalaufwendungen im Jahr 2002 der größte Ausgabeposten der Hochschulen. Der übrige laufende Sachaufwand betrug 8,9 Mrd. Euro. Die Investitionen gingen leicht auf 3,2 Mrd. Euro zurück.

Auf die Universitäten (ohne medizinische Einrichtungen) entfielen 13,4 Mrd. Euro der Gesamtausgaben, 6,8% mehr als 2001. Die medizinischen Einrichtungen der Universitäten gaben 13,9 Mrd. Euro (+4,9%) und die Fach- und Verwaltungsfachhochschulen 3,1 Mrd. Euro (+8,0%) aus. 307 von 391 Hochschulen befanden sich in öffentlicher Trägerschaft. Das Ausgabevolumen der insgesamt 84 privaten Hochschulen in Deutschland lag bei 379,9 Mill. Euro. Auf die medizinischen Fachbereiche und Einrichtungen der Hochschulen entfielen 45,8% aller Hochschulausgaben; sie wurden sowohl für Lehre und Forschung als auch für Kran-

kenbehandlung aufgewendet. Für die Fächergruppe Mathematik und Naturwissenschaften gaben die Hochschulen 3,3 Mrd. Euro aus; für die Ingenieurwissenschaften 2,7 Mrd. Euro. In den Rechts-, Wirtschafts- und Sozialwissenschaften lagen die Ausgaben bei 1,6 Mrd. Euro.

Die Hochschulen erzielten 2002 Einnahmen in Höhe von 13,0 Mrd. Euro (+6,2%). Davon waren 9,3 Mrd. Euro Verwaltungseinnahmen (+3,7%), die zu 94,4% von den medizinischen Einrichtungen erzielt wurden, überwiegend als Entgelte für Krankenbehandlungen. Darüber hinaus warben die Hochschulen 3,3 Mrd. Euro an Drittmitteln bei privaten und öffentlichen Stellen ein (+7,4%). Der größte Teil dieser Drittmittel, 2,4 Mrd. Euro, floss an die Universitäten (ohne medizinische Einrichtungen). Deren Drittmittelleinnahmen stiegen gegenüber 2001 um 11,9%. Wichtigste Drittmittelgeber der Hochschulen waren die Deutsche Forschungsgemeinschaft (1,0 Mrd. Euro), die Unternehmen (854 Mill. Euro) und der Bund (822 Mill. Euro).

Ein Indikator für den Beitrag der Hochschulträger zum Unterhalt der Hochschulen sind die Grundmittel. Sie werden ermittelt, indem die Verwaltungs- und Drittmittelleinnahmen von den Hochschulausgaben abgezogen werden. Die auf diese Weise berechneten Grundmittel der Hochschulen lagen 2002 bei 17,8 Mrd. Euro, 7,0% mehr als 2001. Auf jeden Studierenden kamen im Jahr 2002 rechnerisch 7 510 Euro Grundmittel für laufende Zwecke.

Detaillierte Informationen erhalten Sie kostenlos als Download unter <http://www.statistik-shop.de>.

Weitere Auskünfte erteilt

Heike Franziska Haug, Telefon 06 11 / 75-41 13,

E-Mail: [bildungsausgaben@destatis.de](mailto:bildungsausgaben@destatis.de).

## Erste Ergebnisse der Erbschaft- und Schenkungsteuerstatistik 2002

Erste Eckzahlen aus der Erbschaft- und Schenkungsteuerstatistik 2002, die zum ersten Mal seit 1978 wieder erhoben wurde, liegen jetzt vor. Danach beliefen sich die Erwerbe von Todes wegen (Erbschaften, Vermächtnisse usw.), für die im Jahr 2002 erstmals Steuern festgesetzt wurden, auf insgesamt 12,1 Mrd. Euro und die Schenkungen auf 4,6 Mrd. Euro. Dafür fielen insgesamt 2,2 Mrd. Euro Erbschaftsteuer und 0,6 Mrd. Euro Schenkungsteuer an.

Der nachrichtlich ermittelte Reinnachlass (Nachlassvermögen abzüglich Nachlassverbindlichkeiten, ungeachtet einer Steuerpflicht) betrug 15,4 Mrd. Euro.

30% des Wertes aller Nachlassgegenstände bestand aus Grundvermögen, 62% aus übrigem Vermögen und 8% aus Betriebsvermögen.

Die meisten steuerpflichtigen Erwerbe von Todes wegen, nämlich 68%, lagen im Wert unter 50 000 Euro. Aus diesen Fällen flossen jedoch nur 9% der gesamten Erbschaftsteuer. Dagegen errechneten sich für die 0,1% der Fälle, in denen jeweils mehr als 5 Mill. Euro erworben wurden, 18% der angefallenen Erbschaftsteuer.

Bei den Schenkungen blieben gut 64% der Erwerbe im Wert unter 50000 Euro. Diese trugen 11% zur gesamten Schenkungsteuer bei. Dagegen machten die 0,3% der Fälle, in denen mehr als 5 Mill. Euro erworben wurden, 17% der gesamten Schenkungsteuer aus.

Die Statistik kann keine Informationen über *alle* Vermögensübergänge des Jahres 2002 liefern, da zum einen die meisten Erwerbe von Todes wegen und Schenkungen innerhalb der Freibeträge, für die keine Steuer festgesetzt wird, liegen; sie tauchen folglich in der Statistik nicht auf. Zum anderen können die Finanzämter Steuern für Erwerbe von Todes wegen oder Schenkungen des Jahres 2002 auch noch in Zukunft festsetzen – auch in den Zahlen für 2002 sind Vermögensübergänge enthalten, die mehrere Jahre zurücklagen, bei denen die Steuer aber 2002 erst festgesetzt wurde.

Die Erbschaft- und Schenkungsteuerstatistik soll in Zukunft alle fünf Jahre durchgeführt werden, die nächste Erhebung ist entsprechend für das Berichtsjahr 2007 vorgesehen.

Weitere Auskünfte erteilt  
Elke Eckert, Telefon 06 11 / 75-41 26,  
E-Mail: [elke.eckert@destatis.de](mailto:elke.eckert@destatis.de).

## 57 Mrd. Euro Einnahmen aus „Umweltsteuern“ im Jahr 2003

Die Einnahmen aus „umweltbezogenen Steuern“ betragen nach Angaben der Umweltökonomischen Gesamtrechnungen im Jahr 2003 rund 57 Mrd. Euro. Davon entfielen 43,2 Mrd. Euro auf die Mineralölsteuer, 7,3 Mrd. Euro auf die Kraftfahrzeugsteuer und 6,5 Mrd. Euro auf die Stromsteuer. Im Vergleich zum Jahr 1991 haben sich die Einnahmen aus den umweltbezogenen Steuern damit fast verdoppelt (+91%). Der Anteil dieser Einnahmen an den gesamten Steuereinnahmen der öffentlichen Haushalte (im Jahr 2003 insgesamt 479,6 Mrd. Euro) hat sich in diesem Zeitraum von 8,8 auf 11,9% erhöht.

Die Zusammenfassung der drei genannten Steuerarten zu „umweltbezogenen Steuern“ folgt einer auf internationaler Ebene entwickelten Abgrenzung, die alle Steuern zusammenfasst, die den Energieverbrauch, die Emissionen, den Verkehr oder schädliche Stoffausbringungen (Pestizide o.Ä.) besteuern, unabhängig von den Beweggründen für die Einführung der Steuer oder von der Verwendung der Einnahmen. Entsprechend umfassen die so abgegrenzten Umweltsteuern zum Beispiel die gesamten Mineralölsteuereinnahmen und nicht nur den Anteil, der sich durch die zum 1. April 1999 in Kraft getretenen Gesetze zur ökologischen Steuerreform (Ökosteuern) ergeben hat.

Gegenüber 1998, dem Jahr vor Einführung der „Ökosteuern“, sind diese Steuereinnahmen um rund 36% gestiegen. Dahinter stehen eine schrittweise Erhöhung der Steuersätze, aber auch eine Verringerung der besteuerten und damit der verbrauchten Mengen. Von 1998 bis 2003 ist beispielsweise die Menge an versteuertem Benzin von 40,8 Mill. m<sup>3</sup> auf 34,9 Mill. m<sup>3</sup> und damit um knapp 15% zurückgegangen. Bei den Dieselmotorkraftstoffen war in diesem Zeitraum zunächst ein Anstieg von 32,5 Mill. m<sup>3</sup> auf 34,3

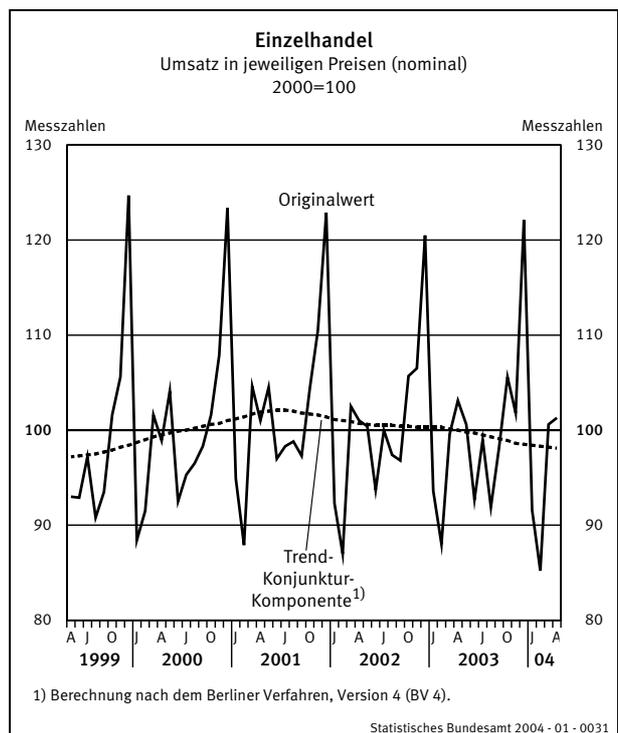
Mill. m<sup>3</sup> im Jahr 2001 zu verzeichnen, dem ein Rückgang auf 33,1 Mill. m<sup>3</sup> folgte. Seit 2003 liegt der Mineralölsteuersatz für unverbleites und schwefelfreies Benzin bei 65,45 Cent je Liter, für schwefelfreies Dieselmotorkraftstoff bei 47,04 Cent.

Weitere Auskünfte erteilt  
Ursula Lauber, Telefon 06 11 / 75-27 37,  
E-Mail: [ursula.lauber@destatis.de](mailto:ursula.lauber@destatis.de).

## Weitere wichtige Monatszahlen

### Einzelhandel

Nach vorläufigen Ergebnissen setzte der Einzelhandel in Deutschland im *April 2004* nominal 1,7% und real 1,8% weniger als im April 2003 um. Beide Monate hatten jeweils 24 Verkaufstage. Das vorläufige Ergebnis wurde aus Daten von sechs Bundesländern berechnet, in denen 81% des Gesamtumsatzes im deutschen Einzelhandel getätigt werden. Nach Kalender- und Saisonbereinigung der Daten wurde im Vergleich zum März 2004 nominal 1,0% und real 0,6% mehr abgesetzt.



Im Einzelhandel mit Lebensmitteln, Getränken und Tabakwaren lag der Umsatz nominal genauso hoch ( $\pm 0,0\%$ ) und real um 0,5% niedriger als im April 2003. Die Umsatzentwicklung in den beiden Teilbereichen des Lebensmittel-einzelhandels war unterschiedlich: In den Lebensmittelgeschäften mit einem breiten Sortiment (Supermärkte, SB-Warenhäuser und Verbrauchermärkte) stiegen die Umsätze um nominal 0,3% (real -0,2%), im Facheinzel-

handel mit Lebensmitteln – dazu gehören zum Beispiel die Getränkemärkte und Fischgeschäfte – lagen die Umsätze nominal um 4,0% und real um 5,8% unter den Umsatzwerten des Vorjahresmonats.

Im Einzelhandel mit Nicht-Lebensmitteln (dazu gehört der Einzelhandel mit Gebrauchs- und Verbrauchsgütern) wurde das Ergebnis des Vorjahresmonats nominal und real nicht erreicht (nominal -3,0% , real -2,7%). Nur der Facheinzelhandel mit Textilien, Bekleidung, Schuhen und Lederwaren (nominal +0,7%, real +1,0%) erzielte nominal und real höhere Umsätze als im April 2003. Der Facheinzelhandel mit kosmetischen, pharmazeutischen und medizinischen Produkten erzielte bei einem nominalen Umsatzminus (-0,7%) real genauso hohe Umsätze (±0,0%) wie im Vorjahresmonat. Alle anderen Branchen verbuchten nominale und reale Umsatzrückgänge gegenüber dem Vorjahresmonat: der Facheinzelhandel mit Einrichtungsgegenständen, Haushaltsgeräten und Baubedarf (nominal -1,2%, real -1,0%), der sonstige Facheinzelhandel (z. B. Bücher, Zeitschriften, Schmuck, Sportartikel) (nominal -5,9%, real -4,6%), der sonstige Einzelhandel mit Waren verschiedener Art, zu dem die Waren- und Kaufhäuser gehören (nominal -6,6%, real -6,3%) und der Versandhandel (nominal -9,5%, real -8,9%).

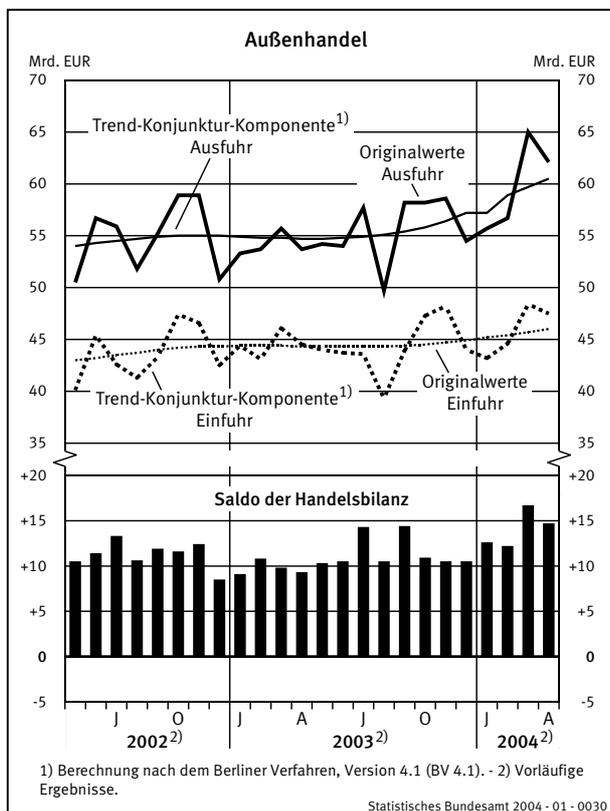
zustellen. Die Preise der Einfuhren erhöhten sich im April 2004 gegenüber dem Vorjahresniveau um 0,4%. Werden Erdöl und Mineralölerzeugnisse nicht berücksichtigt, war ein Rückgang von -0,5% zu verzeichnen. Auch die Preise der Ausfuhren nahmen gegenüber April 2003 leicht zu (+0,2%).

Der Überschuss der Außenhandelsbilanz stieg gegenüber dem gleichen Vorjahresmonat (+9,2 Mrd. Euro) auf 14,6 Mrd. Euro. Zusammen mit den Salden der Dienstleistungsbilanz (-3,2 Mrd. Euro), der Bilanz der laufenden Übertragungen (-1,7 Mrd. Euro), der Bilanz der Ergänzungen zum Warenverkehr (-0,3 Mrd. Euro), der Bilanz der Erwerbs- und Vermögenseinkommen (+0,8 Mrd. Euro) resultiert nach vorläufigen Berechnungen der Deutschen Bundesbank daraus im April 2004 ein Leistungsbilanzsaldo von +10,2 Mrd. Euro. Im gleichen Vorjahresmonat betrug der Überschuss der Leistungsbilanz +3,4 Mrd. Euro.

Gegenüber März 2004 sanken die deutschen Ausfuhren nominal um 4,3%, die Importe gingen nominal um 1,8% zurück. Dagegen sind im April die Exporte saisonbereinigt gegenüber März 2004 um 5,2% gestiegen. Auch die Einfuhren haben sich saisonbereinigt erhöht (um 4,1%). [u](#)

## Außenhandel

Im April 2004 erhöhten sich die Ausfuhren aus Deutschland im Vorjahresvergleich um 15,7% auf 62,1 Mrd. Euro, die Gütereinfuhren stiegen um 6,8% auf 47,5 Mrd. Euro. Leichte Anstiege waren bei den Außenhandelspreisen fest-





Betriebswirt (grad.) Manfred Klose, Dipl.-Volkswirt Alexander Opitz, Dipl.-Volkswirt Norbert Schwarz

# Sozialrechnungsmatrix für Deutschland

## Konzepte und erste Ergebnisse für das Jahr 2000

*Eine alternde Gesellschaft und ein verschärfter globaler Wettbewerb sind Phänomene, die heute und in den nächsten Jahrzehnten die wirtschaftliche und soziale Entwicklung in Deutschland beeinflussen werden. Dabei gilt, dass es zwischen ökonomischer und sozialer Entwicklung eine Vielzahl von Interdependenzen gibt. Vor dem Hintergrund der Finanzierung sozialer Sicherungssysteme und der Belastung des Faktors Arbeit mit Sozialabgaben steht in der öffentlichen Diskussion zumeist die Last, die der Sozialstaat für die wirtschaftliche Entwicklung darstellt, im Vordergrund. Weitere Beispiele für Interdependenzen zwischen ökonomischer und sozialer Entwicklung sind die Auswirkungen einer veränderten Haushalts- und Altersstruktur oder die Veränderungen in der Einkommensverteilung auf die Nachfrage privater Haushalte nach Konsumgütern. Auch die Notwendigkeit, sich gegen soziale Risiken wie Krankheit und Alter vermehrt privat abzusichern, beeinflusst die Konsumausgaben privater Haushalte.*

*Solche sozioökonomischen Entwicklungen und die daraus resultierenden Verhaltensänderungen privater Haushalte schlagen sich in den Makrogrößen der Volkswirtschaftlichen Gesamtrechnungen nur mittelbar nieder. So sind die seit zwei Jahren real rückläufigen Konsumausgaben privater Haushalte vermutlich nicht nur auf die konjunkturelle Schwäche zurückzuführen, sondern es zeigen sich möglicherweise von den privaten Haushalten antizipierte Reaktionen auf erwartete zukünftige gesellschaftliche und ökonomische Entwicklungen. Dafür spricht beispielsweise der Anstieg der Sparquote privater Haushalte von 9,8% im Jahr 2000 auf 10,8% im Jahr 2003, der als ein Ausdruck erhöhten „Sicherheitsdenkens“ in „unsicheren“ Zeiten gewertet werden kann.*

*Der Einfluss sozioökonomischer Entwicklungen konnte bisher nicht im Detail mit Ergebnissen aus den Volkswirtschaftlichen Gesamtrechnungen beleuchtet werden. Um beispielsweise Haushalts- und Familienstrukturen in Analysen einzubeziehen, mussten Primärstatistiken wie der Mikrozensus und die Einkommens- und Verbrauchsstichprobe herangezogen werden. Dies hatte zur Folge, dass sowohl konzeptionell als auch empirisch zum Teil deutliche Abweichungen zu den Größen der Volkswirtschaftlichen Gesamtrechnungen in Kauf genommen werden mussten. Mit dem hier vorgestellten Instrumentarium einer Sozialrechnungsmatrix existiert ein Rahmen, in dem zum einen auf die Volkswirtschaftlichen Gesamtrechnungen abgestimmte Ergebnisse in tiefer sozioökonomischer Gliederung präsentiert und zum anderen Analysen im Kreislaufzusammenhang der Volkswirtschaftlichen Gesamtrechnungen vorgenommen werden können.*

*In dem Aufsatz werden erstmals Ergebnisse der Volkswirtschaftlichen Gesamtrechnungen in Form einer Sozialrechnungsmatrix gezeigt. Sie beziehen sich auf das Jahr 2000, das aktuellste Jahr, für das alle benötigten Daten vorliegen. Mit dem Beitrag werden zwei Ziele angestrebt: Zum einen soll einer breiteren Öffentlichkeit das Konzept und der Aufbau einer Sozialrechnungsmatrix vorgestellt werden und zum anderen werden mögliche Verknüpfungen zwischen Volkswirtschaftlichen Gesamtrechnungen und Sozialstatistik aufgezeigt. Damit stellt eine Sozialrechnungsmatrix auch einen Verbindungsrahmen zwischen Volkswirtschaftlichen Gesamtrechnungen und dem im Aufbau befindlichen sozioökonomischen Berichtssystem dar.<sup>1)</sup> Weitere, tiefer gehende Ergebnisse der Sozialrechnungsmatrix sollen Ende dieses Jahres veröffentlicht werden.*

1) Informationen zum sozioökonomischen Berichtssystem siehe Stahmer, C.: „Aufbau eines sozio-ökonomischen Berichtssystems für eine nachhaltige Gesellschaft“ in „Sozialer Wandel“, Band 41 der Schriftenreihe „Forum der Bundesstatistik“, Wiesbaden 2003, S. 72 ff.

## 1 Konzept und Aufbau einer Sozialrechnungsmatrix

Die Konzepte für eine Sozialrechnungsmatrix, abgekürzt SAM von dem englischen Begriff Social Accounting Matrix, gehen vor allem auf Richard Stone zurück, der schon in den 50er-Jahren des vergangenen Jahrhunderts vorgeschlagen hatte, die Ergebnisse der Volkswirtschaftlichen Gesamtrechnungen in Matrixform darzustellen.<sup>2)</sup> In den 1980er- und 1990er-Jahren wurden die Arbeiten zur Erstellung einer SAM insbesondere im Statistischen Amt der Niederlande (Centraal Bureau voor de Statistiek) vorangetrieben. Die Matrixdarstellung der Volkswirtschaftlichen Gesamtrechnungen ist auch im weltweiten System of National Accounts 1993 (SNA 93) der Vereinten Nationen und im Europäischen System Volkswirtschaftlicher Gesamtrechnungen von 1995 (ESVG 95) erläutert. Zusätzlich veröffentlichte das Statistische Amt der Europäischen Gemeinschaften (Eurostat) im Jahr 2003 ein Handbuch zur Erstellung einer SAM.<sup>3)</sup> Neben dem Aufbau einer SAM ist dort die Verbindung zu Arbeitsmarktstatistiken ausführlich dargestellt.

Ausgehend von den makroökonomischen Größen der Volkswirtschaftlichen Gesamtrechnungen ermöglicht eine SAM eine umfassende und konsistente Verbindung zu Informationen aus den Sozialstatistiken. Damit lassen sich sozioökonomische Fragestellungen in den Rahmen der volkswirtschaftlichen Makrogrößen einbinden und Wechselwirkungen können aufgezeigt werden. Die Matrixdarstellung bietet den Vorteil, dass verschiedene Darstellungseinheiten, wie zum Beispiel Sektoren, Einkommensarten oder Gütergruppen, miteinander kombiniert werden können. Auch ist es möglich, die monetären Angaben aus den Volkswirtschaftlichen Gesamtrechnungen mit nichtmonetären Größen, wie beispielsweise Bevölkerungszahlen, zu verknüpfen.

Entscheidendes Kriterium für eine konsistente Darstellung einer SAM ist die Anknüpfung an die Konzepte des ESGV 95. Daher sind die originären Ergebnisse aus den Sozialstatistiken sowohl konzeptionell als auch quantitativ auf die Volkswirtschaftlichen Gesamtrechnungen abzustimmen. Eine SAM erweitert das bisherige Spektrum der Volkswirtschaftlichen Gesamtrechnungen um zusätzliche Analysemöglichkeiten. Beispielsweise kann untersucht werden, wie sich Änderungen in der Güterproduktion auf die Einkommen privater Haushalte und in der Folge auf deren Konsumverhalten auswirken.

Für die Aufstellung einer SAM werden in einem ersten Schritt Ergebnisse aus den verschiedenen Rechenwerken der Volkswirtschaftlichen Gesamtrechnungen (VGR) zusammengeführt. Aus den Ergebnissen der Input-Output-Rechnung, der Einkommensrechnung für volkswirtschaftliche Sektoren und der Finanzierungsrechnung entsteht so ein umfassendes Gesamtbild in Form einer Kreuztabelle (Matrix). Streng genommen handelt es sich dabei noch

nicht um eine SAM, sondern um eine VGR-Matrix (NAM von dem englischen Begriff National Accounting Matrix), da in den Zellen ausschließlich vorhandene Ergebnisse aus den Volkswirtschaftlichen Gesamtrechnungen abgebildet sind. Erst durch die anschließenden Verknüpfungen mit sozioökonomischen Informationen wird hieraus eine Sozialrechnungsmatrix im eigentlichen Sinn. Aus Vereinfachungsgründen wird im Folgenden jedoch durchgängig der Begriff SAM verwendet. Den Grundaufbau einer SAM zeigt Übersicht 1.

Kernstück einer SAM sind die Einkommenstransaktionen, angefangen von der Einkommensentstehung über die primäre und sekundäre Einkommensverteilung bis hin zur Einkommensverwendung. Durch die Matrixform wird jede Transaktion durch einen Eintrag dargestellt, wie dies in der Input-Output-Rechnung üblich ist. Dagegen erfolgt die Darstellung in der Einkommens- und in der Finanzierungsrechnung üblicherweise in Kontenform mit jeweils einer Buchung auf der Aufkommens- und Verwendungsseite bzw. Aktiv- und Passivseite des Kontos. Gerade bei der Darstellung der Einkommenstransaktionen zeigen sich die Vorteile einer Matrix. Durch die Matrixform wird unmittelbar ersichtlich, von wem an wen eine Transaktion erfolgt, das heißt es wird beispielsweise aufgezeigt, von welchem volkswirtschaftlichen Sektor das Einkommen stammt und welcher der empfangende Sektor ist.

Jede in Übersicht 1 aufgeführte Zelle stellt eine eigene, in sich geschlossene Matrix dar, sodass eine SAM aus einer Vielzahl von Untermatrizen besteht. Ergebnisse werden jedoch nur in den namentlich benannten Zellen aufgeführt. Zur besseren Orientierung sind im „Schema einer Sozialrechnungsmatrix“ (siehe Übersicht 1) die Zellennummern angegeben. An erster Stelle steht die laufende Nummer für die Zeilengliederung, an zweiter Stelle die laufende Nummer der Spaltengliederung. In den folgenden Erläuterungen wird hierauf Bezug genommen. Die Grundstruktur einer SAM lässt sich folgendermaßen umschreiben:

- Zeile 1 und Spalte 1 der SAM zeigen die Güterströme innerhalb der Volkswirtschaft und im Außenverhältnis zur übrigen Welt.
- Die Einkommensbetrachtung beginnt mit der Matrix des Einkommensentstehungskontos (Zelle 3/2). Dort wird die Einkommensentstehung nach Einkommensarten und Wirtschaftsbereichen aufgezeigt. Die folgenden Matrizen zeigen die Primäreinkommen aus der Produktionstätigkeit (Zelle 4/3), die Vermögenseinkommen (Zelle 4/4), die Verteilung der Primäreinkommen (Zelle 5/4) und die Umverteilung der Einkommen durch Steuern, Sozialbeiträge und -transfers sowie sonstige Transfers (Zelle 5/5). Die verfügbaren Einkommen sind in der Zelle 6/5 dargestellt.
- In den Zeilen bzw. Spalten 7 bis 9 werden dann vor allem Vorgänge der Sach- und der Geldvermögensbildung dargestellt.

2) Siehe hierzu u. a. Stone, R.: „Input-Output and the Social Accounts“ in University of Pisa (Hrsg.): „The Structural Interdependence of the Economy, Proceedings of an International Conference on Input-Output Analysis“, New York 1955, S. 155 ff., und Stone, R./Croft-Murray, G.: „Social Accounting and Economic Models“, London 1959. Ausführlichere Erläuterungen zur Entstehungsgeschichte der SAM sind zu finden in Stahmer, C.: „Das unbekannteste Meisterwerk – Sir Richard Stone und sein System of Social and Demographic Statistics“ in Hartard, S./Stahmer, C. (Hrsg.): „Magische Dreiecke – Berichte für eine nachhaltige Gesellschaft“, Band 3, Sozioökonomische Berichtssysteme, Marburg 2003, S. 13 ff.

3) European Commission: „Handbook on Social Accounting Matrices and Labour Accounts“, Luxemburg 2003.

Übersicht 1: Schema einer Sozialrechnungsmatrix (SAM)

	1. Waren und Dienstleistungen (Gütergruppen)	2. Produktion (Wirtschaftsbereiche)	3. Einkommensentstehung (Einkommenskategorien)	4. Primäre Einkommensverteilung (volkswirtschaftliche Sektoren)	5. Sekundäre Einkommensverteilung (volkswirtschaftliche Sektoren)	6. Einkommensverwendung (volkswirtschaftliche Sektoren)	7. Vermögensänderungen (volkswirtschaftliche Sektoren)	8. Bruttoanlageinvestitionen (Wirtschaftsbereiche)	9. Finanzierung (finanzielle Kategorien)	10. Übrige Welt: Güter- und Einkommens-transaktionen	11. Übrige Welt: Vermögens-transfers	Insgesamt
1. Waren und Dienstleistungen (Gütergruppen)	Zelle 1/1 Handels- und Transportspannen	Zelle 1/2 Intermediäre Verwendung (Vorleistungen)				Zelle 1/6 Konsum	Zelle 1/7 Vorratsveränderung, Zugang an Wert-sachen	Zelle 1/8 Bruttoanlageinvestitionen		Zelle 1/10 Export von Waren und Dienstleistungen		Gesamte Güterverwendung
2. Produktion (Wirtschaftsbereiche)	Zelle 2/1 Produktionswerte zu Herstellungspreisen											
3. Einkommensentstehung (Einkommenskategorien)		Zelle 3/2 Einkommensentstehung								Zelle 3/10 Arbeitnehmerentgelt aus d.ü.W.		
4. Primäre Einkommensverteilung (volkswirtschaftliche Sektoren)	Zelle 4/1 Gütersteuern minus Gütersubventionen		Zelle 4/3 Primäreinkommen aus Produktionstätigkeit	Zelle 4/4 Vermögenseinkommen						Zelle 4/10 Vermögenseink., Prod.-Steu. minus Subv. aus der ü.W.		
5. Sekundäre Einkommensverteilung (volkswirtschaftliche Sektoren)				Zelle 5/4 Verteilung der Primäreinkommen	Zelle 5/5 Umverteilung durch Lfd. Transfers					Zelle 5/10 Lfd. Transfers aus der ü.W.		
6. Einkommensverwendung (volkswirtschaftliche Sektoren)					Zelle 6/5 Verfügbare Einkommen	Zelle 6/6 Änderung betrieblicher Versorgungsansprüche				Zelle 6/10 Änderung betrieblicher Versorgungsansprüche an die ü.W.		
7. Vermögensänderungen (volkswirtschaftliche Sektoren)						Zelle 7/6 Sparen	Zelle 7/7 Vermögens-transfers	Zelle 7/9 Nettoanstieg der Passiva <sup>1)</sup>		Zelle 7/11 Vermögens-transfers aus der ü.W.		
8. Bruttoanlageinvestitionen (Wirtschaftsbereiche)		Zelle 8/2 Abschreibungen					Zelle 8/7 Nettoanlageinvestitionen					
9. Finanzierung (finanzielle Kategorien)							Zelle 9/7 Nettoerwerb von finanziellen Aktiva <sup>2)</sup>			Zelle 9/11 Finanzierungssaldo zur ü.W.		
10. Übrige Welt: Güter- und Einkommens-transaktionen	Zelle 10/1 Importe von Waren und Dienstleistungen		Zelle 10/3 Arbeitnehmerentgelte an die ü.W.	Zelle 10/4 Vermögenseinkommen an die ü.W.	Zelle 10/5 Lfd. Transfers an die ü.W.	Zelle 10/6 Änderung betrieblicher Versorgungsansprüche aus der ü.W.						
11. Übrige Welt: Vermögens-transfers							Zelle 11/7 Vermögens-transfers an die ü.W.			Zelle 11/10 Saldo lfd. Transaktionen mit der ü.W.		
Insgesamt	Güteraufkommen zu Anschaffungspreisen											

ü.W. = übrige Welt. – 1) Außenfinanzierung. – 2) Geldvermögensbildung.

Durch die vollständige Darstellung einer SAM wird der Zusammenhang zwischen den einzelnen Transaktionen – auch im Verhältnis zur übrigen Welt – in den verschiedenen Untermatrizen ersichtlich. So erfolgt die Darstellung der durch Produktionstätigkeiten entstandenen Einkommen (Zelle 3/2) nach dem Inlandskonzept, während der primären Einkommensverteilung aus Produktionstätigkeit (Zelle 4/3) das Inländerkonzept zugrunde liegt. Der Übergang vom Inlands- zum Inländerkonzept und damit die Beziehungen zur übrigen Welt werden in den Zeilen bzw. Spalten 10 und 11 gezeigt. Die vom Inland an die übrige Welt fließenden Güter- und Einkommenstransaktionen sind in Zeile 10 und die Vermögenstransfers in Zeile 11 dargestellt, die aus dem Ausland stammenden Güter- und Einkommenstransaktionen bzw. Vermögenstransfers in den Spalten 10 und 11.

## 2 Darstellungseinheiten der Sozialrechnungsmatrix

Darstellungseinheiten in der hier gezeigten Grundversion einer SAM sind Gütergruppen, Wirtschaftsbereiche, Einkommenskategorien, volkswirtschaftliche Sektoren und Finanzierungsarten. In den Spalten und Zeilen erfolgt die Darstellung immer spiegelbildlich. In Zeile und Spalte 1 sind jeweils Gütergruppen, in den Zeilen und Spalten 2 und 8 Wirtschaftsbereiche, in Zeile und Spalte 3 Einkommensarten, in den Zeilen und Spalten 4 bis 7 volkswirtschaftliche Sektoren und in Zeile und Spalte 9 Finanzierungsarten dargestellt. In den Zeilen und Spalten 10 und 11 werden die Transaktionen mit der übrigen Welt gezeigt. Abschließend wird der Gesamtwert der Zeilen bzw. Spalten ausgewiesen. Der spiegelbildliche Vergleich der jeweiligen Gesamtwerte zeigt, ob die SAM konsistente Ergebnisse beinhaltet.

Die Darstellung der Güter erfolgt nach der europäischen Klassifikation für *Gütergruppen*.<sup>4)</sup> Nach dieser Klassifikation können die Güter auf verschiedenen Aggregationsniveaus zu homogenen Gütergruppen zusammengefasst werden.

Die Gliederung nach *Wirtschaftsbereichen* baut auf der Klassifikation der Wirtschaftszweige, Ausgabe 1993 in der Fassung für die Volkswirtschaftlichen Gesamtrechnungen, auf.<sup>5)</sup> Die Zuordnung zu den Wirtschaftsbereichen erfolgt nach dem wirtschaftlichen Schwerpunkt der wirtschaftlichen Einheiten, die in der Regel den Unternehmen als institutionelle Einheiten entsprechen. Abweichend von der schwerpunktmäßigen Zuordnung wird die gesamte Wohnungsvermietung einschließlich unterstellter Produktion aus der Eigennutzung von Wohnimmobilien im Wirtschaftsbereich „Grundstücks- und Wohnungswesen“ nachgewiesen, unabhängig davon, ob sie als Haupt- oder Nebentätigkeit einer wirtschaftlichen Einheit ausgeübt wird.

Dritte Darstellungseinheit in der Sozialrechnungsmatrix sind die entstandenen Einkommen. An *Einkommensar-*

*ten* werden das Arbeitnehmerentgelt, das Selbstständigeneinkommen, der Nettobetriebsüberschuss und der Saldo der sonstigen Produktionsabgaben abzüglich der sonstigen Subventionen unterschieden. Das Arbeitnehmerentgelt umfasst die Bruttolöhne und -gehälter der Arbeitnehmer sowie die Sozialbeiträge der Arbeitgeber. Das Selbstständigeneinkommen und der Nettobetriebsüberschuss errechnen sich als Saldo, indem von den entstandenen Einkommen insgesamt, die der Nettowertschöpfung entsprechen, das Arbeitnehmerentgelt und die sonstigen Produktionsabgaben abgezogen sowie die sonstigen Subventionen hinzugerechnet werden. Auf die Unterschiede zwischen Selbstständigeneinkommen und Nettobetriebsüberschuss wird bei der Erläuterung der volkswirtschaftlichen Sektoren eingegangen. Zu den sonstigen Produktionsabgaben zählen sämtliche Steuern, die im Zusammenhang mit der Produktionstätigkeit zu entrichten sind und keine Gütersteuern darstellen. Sie sind unabhängig von der Menge und dem Wert der produzierten Güter zu zahlen (z. B. Grundsteuer, Kraftfahrzeugsteuer der Unternehmen). Entsprechend zählen zu den sonstigen Subventionen alle Subventionen, die nicht Gütersubventionen sind (z. B. Forschungsförderung, Zinszuschüsse).

Eine weitere Darstellungseinheit ist die Gliederung nach *volkswirtschaftlichen Sektoren*. Unterschieden werden gemäß ESVG 95 Nichtfinanzielle Kapitalgesellschaften, Finanzielle Kapitalgesellschaften, Staat einschließlich Sozialversicherungen und Private Haushalte einschließlich Privater Organisationen ohne Erwerbszweck<sup>6)</sup>. Zu den Kapitalgesellschaften gehören neben den Unternehmen mit eigener Rechtspersönlichkeit auch Personengesellschaften (z. B. Offene Handelsgesellschaft, Kommanditgesellschaft), die als so genannte Quasi-Kapitalgesellschaften in die Sektoren Nichtfinanzielle und Finanzielle Kapitalgesellschaften einbezogen sind. Einzelunternehmen und sonstige Selbstständige, wie die freiberuflich Tätigen, werden dem Sektor Private Haushalte zugerechnet. Dies ist für die begriffliche Unterscheidung von Selbstständigeneinkommen und Nettobetriebsüberschuss wichtig. Die Einkommen der Selbstständigen im Bereich der privaten Haushalte werden als Selbstständigeneinkommen bezeichnet und die Einkommen aus Produktionstätigkeit der anderen Sektoren als Nettobetriebsüberschuss. Beim Nettobetriebsüberschuss privater Haushalte handelt es sich nur um unterstellte Einkommen aus der Eigennutzung von Wohnimmobilien und um den Nettobetriebsüberschuss privater Organisationen ohne Erwerbszweck.

Schließlich werden in der Sozialrechnungsmatrix sieben *finanzielle Kategorien* unterschieden, die einerseits der Geldvermögensbildung dienen und andererseits die verschiedenen Möglichkeiten zur Außenfinanzierung darstellen. Die Begriffe sind weit gefasst. Sie umfassen außer unmittelbaren Gläubiger-Schuldnerverhältnissen, wie Bargeld und Sichteinlagen, Wertpapieren und Finanzderiva-

4) Die Abgrenzung entspricht der Statistischen Güterklassifikation in Verbindung mit den Wirtschaftszweigen in der Europäischen Wirtschaftsgemeinschaft (CPA: Classification of Products by Activity).

5) Die Klassifikation der Wirtschaftszweige ist aus der Statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft (NACE Rev. 1) abgeleitet.

6) Aufgrund unzureichender Ausgangsdaten ist eine Trennung privater Haushalte und privater Organisationen ohne Erwerbszweck nicht vollständig möglich. Es erfolgt daher eine zusammengefasste Darstellung als Sektor Private Haushalte einschließlich Privater Organisationen ohne Erwerbszweck.

ten sowie Krediten, auch Finanzbeziehungen in Form von Aktien und anderen Beteiligungen (Anteilsrechte), Ansprüche gegenüber Versicherungen und Forderungen gegenüber der übrigen Welt in Form von Währungsgold und Sonderziehungsrechten. Dazu kommen sonstige Forderungen und sonstige Verbindlichkeiten u. a. aufgrund von Pensionsrückstellungen. Hier nicht berücksichtigt werden alle anderen zur Geldanlage dienenden Sachgüter wie Kunstgegenstände und Sammlungen.

### 3 Erläuterung der Ergebnisse in der Sozialrechnungsmatrix

Ziel des Kapitels ist es, eine SAM mit ihren Untermatrizen und den Übergängen von einer zur nächsten Untermatrix zu erläutern. Aus Gründen der Anschaulichkeit beschränkt sich die Darstellung auf das in Übersicht 2 auf den Seiten 610 und 611 gezeigte hohe Aggregationsniveau. Für eine Vielzahl von Untermatrizen existieren differenziertere Ergebnisse in den Volkswirtschaftlichen Gesamtrechnungen, auf die im Folgenden nur verwiesen wird. Da die Untermatrizen in Übersicht 2 inhaltlich den Zellen der SAM gemäß dem Grundschemata der Übersicht 1 entsprechen, werden sie weiterhin durch Zeilen- und Spaltennummern gekennzeichnet und als Zellen bezeichnet. Der Verweis auf einzelne Werte erfolgt dagegen durch die zusätzliche Angabe eines Buchstabens für die Zeile und Spalte der jeweiligen Untermatrix.

Die Erläuterung der Ergebnisse erfolgt in drei Schritten. Zuerst wird die Darstellung der Güterströme aufgezeigt. Danach werden die Einkommenstransaktionen und abschließend die Finanzierungsströme dargestellt. Neben der Interpretation der Ergebnisse wird damit die Konzeption der SAM in praktischer Weise näher erläutert.

#### 3.1 Darstellung der Güterströme

Die gütermäßige Verflechtung innerhalb der Volkswirtschaft und mit der übrigen Welt wird in der SAM unter der Rubrik Waren und Dienstleistungen aufgezeigt, die in der ersten Zeile und der ersten Spalte zu finden ist (siehe Übersicht 2). In der Zeile 1 (Zellen 1/1 bis 1/10) ist die Verwendung nach Gütergruppen dargestellt und in der Spalte 1 (Zellen 1/1 bis 10/1) das Güteraufkommen. Aus didaktischen Gründen beschränkt sich die nachfolgende Darstellung auf sechs hochaggregierte Gütergruppen.<sup>7)</sup>

Bei den gütermäßigen Angaben handelt es sich um Ergebnisse der Input-Output-Rechnung und der Investitionsrechnung im Rahmen der Volkswirtschaftlichen Gesamtrechnungen. So ist das Aufkommen der Güter und ihre Verwendung in so genannten Aufkommens- und Verwendungstabellen nachgewiesen, die im Rahmen der Input-Output-Rechnung erstellt werden.<sup>8)</sup> Die Informationen der Investitionsrech-

nung werden genutzt, um die Angaben der Verwendungstabelle zu den Bruttoanlageinvestitionen und Vorratsveränderungen weiter nach Wirtschaftsbereichen bzw. Sektoren aufzuteilen.

Zwischen SAM und Aufkommens- und Verwendungstabelle gibt es konzeptionelle Unterschiede, sodass Umrechnungen notwendig sind. Quantitativ am wichtigsten ist, dass in der SAM der Konsum der privaten Haushalte (Zelle 1/6) nach dem Inländerkonzept darzustellen ist, der Verwendungstabelle jedoch das Inlandskonzept zugrunde liegt. Die Umrechnung vom Inlands- auf das Inländerkonzept erfolgt anhand von Angaben aus der Zahlungsbilanz und Informationen zu Reiseausgaben privater Haushalte aus der Einkommens- und Verbrauchsstichprobe und den laufenden Wirtschaftsrechnungen. Weitere konzeptionelle Unterschiede betreffen die Abgrenzung und die Bewertung von Importen und Exporten.<sup>9)</sup>

Die Darstellung der Güterströme in der SAM lässt sich anhand der Übersicht 2 verdeutlichen. Ausgangspunkt ist das Güteraufkommen im Jahr 2000, das in der ersten Spalte (1a bis 1f) der SAM für sechs aggregierte Gütergruppen gezeigt wird. Die Güter können aus der inländischen Produktion (Zelle 2/1) oder aus der Einfuhr (Zelle 10/1) stammen. In die Importe sind die Käufe privater Haushalte im Ausland eingerechnet. Die inländische Produktion in der Zelle 2/1 ist nicht nur nach Gütergruppen unterteilt, sondern auch nach den sie produzierenden Wirtschaftsbereichen untergliedert. Da die hier gezeigten Produktionswerte zu Herstellungspreisen nachgewiesen sind, müssen für eine Bewertung zu Anschaffungspreisen, die von den Käufern auf der Verwendungsseite aufzuwenden sind, noch die Handels- und Transportleistungen (Zelle 1/1) sowie die Gütersteuern abzüglich Gütersubventionen (Zelle 4/1) hinzugefügt werden. Das gesamte Aufkommen für jede Gütergruppe zu Anschaffungspreisen errechnet sich durch Addition der inländischen Produktion zu Anschaffungspreisen und der Importe. Es ist in der letzten Zeile von Spalte 1 dargestellt.

Die Güterverwendung im Jahr 2000 wird in der ersten Zeile (1a bis 1f) von Übersicht 2 nachgewiesen. Da Aufkommen und Verwendung für jede Gütergruppe identisch sind, enthält die Spalte insgesamt von Zeile 1 die gleichen Güterwerte wie die Zeile insgesamt von Spalte 1. Die gesamte Verwendung von Gütern ergibt sich durch Addition der einzelnen Verwendungskategorien. Zunächst zeigt die Matrix der intermediären Verwendung (Zelle 1/2) zeilenweise die Vorleistungen nach Gütergruppen in den Wirtschaftsbereichen. Der Konsum nach dem Inländerkonzept (Zelle 1/6) ist, ebenso wie die Vorratsveränderungen (Zelle 1/7), nach volkswirtschaftlichen Sektoren untergliedert. Die Bruttoanlageinvestitionen (Ausrüstungen und Bauten) sind in Zelle 1/8 in kombinierter Gliederung nach Gütergruppen und Wirtschaftsbereichen dargestellt. Die Ausfuhr (Zelle 1/10) enthält auch die Käufe von Gebietsfremden in Deutschland.

<sup>7)</sup> Im Rahmen der Input-Output-Rechnung ist eine maximale Gliederungstiefe von 71 Gütergruppen möglich.

<sup>8)</sup> Für ausführliche Informationen zur Input-Output-Rechnung siehe u. a. Fachserie 18 „Volkswirtschaftliche Gesamtrechnungen“, Reihe 2 „Input-Output-Rechnung 2000“, Wiesbaden 2004, und Bleses, P./Stahmer, C.: „Input-Output-Rechnung 1995 und 1997“ in WiSta 12/2000, S. 901 ff.

<sup>9)</sup> Im Unterschied zur Aufkommenstabelle umfassen die Im- und Exporte von Waren in der SAM auch den Lagerverkehr (Warenverkehr über Freihafen und Zolllager) auf inländische Rechnung und sind fob (free on board) dargestellt, d. h. ohne Fracht- und Versicherungsleistungen zwischen den Grenzen. Die Importe in der Aufkommenstabelle sind demgegenüber cif (cost, insurance, freight), d. h. frei Einfuhrgrenze, angegeben.

Übersicht 2: Sozialrechnungsmatrix  
Mrd.

Klassifikation		1. Waren und Dienstleistungen						2. Produktion						3. Einkommensentstehung				
		1a	1b	1c	1d	1e	1f	2a	2b	2c	2d	2e	2f	3a	3b	3c	3d	
1. Waren und Dienstleistungen	Erzeugnisse der Land- und Forstwirtschaft, Fischerei	1a						2	42	0	2	1	1					
	Bergbau, Verarbeitendes Gewerbe, Wasser und Energie	1b						16	640	85	81	22	61					
	Erzeugnisse des Baugewerbes	1c						0	9	4	5	22	9					
	Handelsleistung, Dienstleistungen des Gastgewerbes, Verkehr und Nachrichtenübermittlung	1d	14	333		-347		1	33	2	98	11	15					
	Kredit- und Versicherungsgewerbe, Grundstücks- und Wohnungswesen, Dienstleistungen für Unternehmen	1e						5	146	33	85	259	48					
	Öffentliche Verwaltung, Verteidigung, soziale und andere Dienstleistungen	1f						2	20	2	12	19	36					
2. Produktion	Land- und Forstwirtschaft, Fischerei	2a	46	0	0	0	0											
	Bergbau, Verarbeitendes Gewerbe, Wasser und Energie	2b		1 278	5	43	23	4										
	Baugewerbe	2c		6	215	2	0											
	Handel, Gastgewerbe, Verkehr und Nachrichtenübermittlung	2d	1	7	2	597	11	1										
	Kredit- und Versicherungsgewerbe, Grundstücks- und Wohnungswesen, Dienstleistungen für Unternehmen	2e		1	0	0	831											
	Öffentliche Verwaltung, Verteidigung, soziale und andere Dienstleistungen	2f		0	2	1	0	572										
3. Einkommensentstehung	Arbeitnehmerentgelt	3a						10	333	65	219	166	307					
	Selbstständigeneinkommen	3b						7	10	16	20	99	28					
	Nettobetriebsüberschuss	3c						0	45	10	43	109	24					
	Sonstige Produktionsabgaben minus sonstige Subventionen	3d						-2	4	1	8	8	-8					
4. Primäre Einkommensverteilung	Nichtfinanzielle Kapitalgesellschaften	4a														226		
	Finanzielle Kapitalgesellschaften	4b														-49		
	Staat	4c	-1	140	21	11	27	8								-3	12	
	Private Haushalte einschl. Privater Organisationen ohne Erwerbszweck	4d												1 099	181	57		
5. Sekundäre Einkommensverteilung	Nichtfinanzielle Kapitalgesellschaften	5a																
	Finanzielle Kapitalgesellschaften	5b																
	Staat	5c																
	Private Haushalte einschl. Privater Organisationen ohne Erwerbszweck	5d																
6. Einkommensverwendung	Nichtfinanzielle Kapitalgesellschaften	6a																
	Finanzielle Kapitalgesellschaften	6b																
	Staat	6c																
	Private Haushalte einschl. Privater Organisationen ohne Erwerbszweck	6d																
7. Vermögensänderungen	Nichtfinanzielle Kapitalgesellschaften	7a																
	Finanzielle Kapitalgesellschaften	7b																
	Staat	7c																
	Private Haushalte einschl. Privater Organisationen ohne Erwerbszweck	7d																
8. Bruttoanlageinvestitionen	Land- und Forstwirtschaft, Fischerei	8a						8										
	Bergbau, Verarbeitendes Gewerbe, Wasser und Energie	8b							73									
	Baugewerbe	8c								6								
	Handel, Gastgewerbe, Verkehr und Nachrichtenübermittlung	8d									46							
	Kreditinstitute und Versicherungen, Wohnungswesen, Dienstleistungen für Unternehmen	8e										116						
	Öffentliche Verwaltung, Verteidigung, soziale und andere Dienstleistungen	8f											54					
	Käufe abzüglich Verkäufe von gebrauchten Anlagen	8g																
		8h																
9. Finanzierung	Währungsgold und Sonderziehungsrechte	9a																
	Bargeld und Sichteinlagen	9b																
	Wertpapiere (ohne Anteilsrechte) und Finanzderivate	9c																
	Kredite	9d																
	Anteilsrechte	9e																
	Versicherungstechnische Rückstellungen	9f																
	Sonstige Forderungen/Verbindlichkeiten	9g																
10. Übrige Welt	Güter- und Einkommenstransaktionen	10	18	548	4	63	34	12						5			-1	
11. Übrige Welt	Vermögenstransfers	11																
Insgesamt			78	2 313	249	372	927	597	47	1 353	223	619	833	575	1 104	181	230	10



Eine nähere Betrachtung des gesamten Aufkommens zeigt, dass die Gütergruppe Bergbau, Verarbeitendes Gewerbe, Wasser und Energie – im Folgenden „Erzeugnisse des Produzierenden Gewerbes“ genannt<sup>10)</sup> – mit 2 313 Mrd. Euro die quantitativ bedeutsamste ist (Zeile insgesamt, Spalte 1b); sie soll daher zur Veranschaulichung der geschilderten Zusammenhänge dienen.

Wie Spalte 1b der Zelle 2/1 zeigt, erfolgt die inländische Produktion der Erzeugnisse des Produzierenden Gewerbes hauptsächlich im Wirtschaftsbereich Bergbau, Verarbeitendes Gewerbe, Wasser und Energie – kurz Produzierendes Gewerbe – (1 278 Mrd. Euro). In weit geringerem Umfang wird diese Gütergruppe darüber hinaus als Nebentätigkeit in anderen Wirtschaftsbereichen erstellt (insgesamt 14 Mrd. Euro). Um das inländische Einkommen dieser Gütergruppe zu Anschaffungspreisen zu erhalten, müssen noch Handels- und Transportleistungen (Zeile 1d, Spalte 1b) und Gütersteuern abzüglich Gütersubventionen (Zeile 4c, Spalte 1b) addiert werden. Außerdem wird mit 548 Mrd. Euro knapp ein Viertel des Einkommens dieser Gütergruppe aus der übrigen Welt bezogen (Zeile 10, Spalte 1b).

Eine nähere Betrachtung der Verwendung der Güter des Produzierenden Gewerbes zeigt bei der intermediären Verwendung (siehe Zeile 1b, Spalten 2a bis f), dass diese Erzeugnisse größtenteils als Vorleistung des Wirtschaftsbereichs, der diese Güter schwerpunktmäßig produziert, verwendet werden (640 Mrd. Euro). Daneben werden sie auch in der Land- und Forstwirtschaft, Fischerei (16 Mrd. Euro), im Baugewerbe (85 Mrd. Euro) sowie in den verbleibenden Dienstleistungsbereichen (zusammen 164 Mrd. Euro) verbraucht.

Konsumiert werden die Erzeugnisse des Produzierenden Gewerbes überwiegend von den privaten Haushalten (579 Mrd. Euro), wie Zeile 1b, Spalte 6d zeigt. Zusätzlich werden 28 Mrd. Euro für Konsumausgaben des Staates verbraucht (Zeile 1b, Spalte 6c). Dies sind vor allem Arzneimittel und andere medizinische Hilfsmittel, die als soziale Sachleistungen der gesetzlichen Krankenkassen im Sektor Staat gebucht werden.

Die Vorratsveränderungen (Zeile 1b, Spalten 7a bis d) fallen bei dieser Gütergruppe gegenüber den Bruttoanlageinvestitionen (Zeile 1b, Spalten 8a bis g) quantitativ nicht ins Gewicht. Wichtige Bereiche, die Erzeugnisse des Produzierenden Gewerbes als Investitionsgüter nutzen, sind neben dem entsprechenden Wirtschaftsbereich selbst (63 Mrd. Euro) der zusammengefasste Bereich Kredit- und Versicherungsgewerbe, Grundstücks- und Wohnungswesen und Dienstleistungen überwiegend für Unternehmen (71 Mrd. Euro) sowie der Bereich Handel, Gastgewerbe, Verkehr und Nachrichtenübermittlung (39 Mrd. Euro). Der Saldo der Käufe abzüglich der Verkäufe von gebrauchten Anlagen in Höhe von – 11 Mrd. Euro (Zeile 1b, Spalte 8g) kann aus statistischen Gründen nicht auf Wirtschaftsbereiche aufgeteilt werden.

Güter des Produzierenden Gewerbes im Wert von 603 Mrd. Euro gingen schließlich als Exporte in die übrige Welt (Zeile 1b, Spalte 10). Die Verwendung insgesamt (siehe Zeile 1b, Spalte insgesamt) entspricht mit 2 313 Mrd. Euro wiederum dem gesamten Einkommen (Spalte 1b, Zeile insgesamt).

### 3.2 Die Einkommensströme

Wie im Kontensystem der Volkswirtschaftlichen Gesamtrechnungen werden in der SAM die Einkommensströme ausgehend von der Einkommensentstehung über die Einkommensverteilung und -umverteilung bis hin zur Einkommensverwendung unterteilt nach volkswirtschaftlichen Sektoren gezeigt.<sup>11)</sup> Vorteil einer Darstellung der Ergebnisse in Matrixform gegenüber einer Darstellung im Kontensystem ist, dass die Ströme zwischen den Sektoren transparent werden, das heißt man sieht, von welchen Sektoren bzw. an welche Sektoren die Einkommensströme fließen.

Konzeptionell gibt es zwei Unterschiede zwischen dem Kontensystem der Volkswirtschaftlichen Gesamtrechnungen und der SAM, die beide bedingt sind durch die Darstellung in der Input-Output-Rechnung. Dies betrifft die Buchung der Gütersteuern und Gütersubventionen sowie der so genannten unterstellten Bankgebühr.<sup>12)</sup> In der SAM wird die unterstellte Bankgebühr – wie in der Input-Output-Rechnung – sowohl beim Güteraufkommen (Zelle 2/1) als auch bei der intermediären Verwendung (Zelle 1/2) den Kreditinstituten zugeordnet. Auf der Einkommenseite hat dies dann zur Folge, dass der nachgewiesene Nettobetriebsüberschuss der finanziellen Kapitalgesellschaften in Zelle 4/3 zu niedrig und der Nettobetriebsüberschuss der anderen Sektoren bzw. die Selbstständigeneinkommen privater Haushalte zu hoch ausgewiesen werden. Um dieses Problem zu umgehen, erfolgt im Kontensystem die verwendungsseitige Buchung erst im Zusammenhang mit den Vermögenseinkommen in Zelle 4/4. Die danach folgenden Untermatrizen der SAM entsprechen dann wieder der Darstellung im Kontensystem. Der zweite Konzeptunterschied betrifft die Buchung der Gütersteuern abzüglich Gütersubventionen (Zelle 4/1), die in der SAM zunächst nach dem Inlandskonzept erfolgt. Der Übergang auf das Inländerkonzept, das heißt der Abzug der an die übrige Welt (Europäische Union) zu leistenden Gütersteuern und das Hinzufügen der von dort empfangenen Gütersubventionen, erfolgt durch eine Korrekturbuchung in Zelle 4/10, sodass die Zeilensumme der Zellen 4/1 und 4/10 für den Sektor Staat den gleichen Wert wie im Kontensystem ergibt.

Der enge Zusammenhang zwischen Güterproduktion und Einkommensentstehung wird in Spalte 2 der SAM deutlich (siehe Übersicht 2). Als Bestandteile des Produktionswertes werden dort neben dem Vorleistungsverbrauch (Zelle 1/2) und den Abschreibungen (Zelle 8/2) die Einkommensentstehung in den einzelnen Wirtschaftsbereichen nach Einkommensarten in Zelle 3/2 dargestellt. Am Beispiel des

10) In der üblichen Darstellung in den Volkswirtschaftlichen Gesamtrechnungen umfasst das Produzierende Gewerbe auch das Baugewerbe. In der SAM wird hiervon abgewichen, da es deutliche strukturelle Unterschiede zwischen dem Baugewerbe und dem sonstigen Produzierenden Gewerbe gibt (siehe dazu auch Abschnitt 4.1).

11) Zur Einkommensrechnung nach volkswirtschaftlichen Sektoren siehe Fachserie 18 „Volkswirtschaftliche Gesamtrechnungen“, Reihe 1.3 „Konten und Standardtabellen – Hauptbericht“, Erläuterungen zu 3.4, Tabellen für Sektoren, Wiesbaden 2003.

12) Zurzeit wird als unterstellte Bankgebühr die Differenz zwischen den empfangenen Vermögenseinkommen und den gezahlten Zinsen der Kreditinstitute angesehen.

Produzierenden Gewerbes sei dies verdeutlicht. Im Produzierenden Gewerbe (Spalte 2b in Zelle 3/2) werden Arbeitnehmerentgelte in Höhe von 333 Mrd. Euro gezahlt, es entsteht ein Selbstständigeneinkommen von 10 Mrd. Euro, ein Nettobetriebsüberschuss von 45 Mrd. Euro und es fallen per saldo sonstige Produktionsabgaben minus sonstige Subventionen in Höhe von 4 Mrd. Euro an, die an den Staat gehen. Die Summe dieser Einkommensarten plus der Vorleistungen (Spalte 2b in Zelle 1/2) und der Abschreibungen (Spalte 2b in Zelle 8/2) ergibt den Produktionswert des Produzierenden Gewerbes in Höhe von 1 353 Mrd. Euro (Zeile insgesamt, Spalte 2b). Spiegelbildlich kann der Produktionswert auch im Produktionskonto abgelesen werden. So ergibt die Summe über die Zeile 2b in Zelle 2/1 ebenfalls 1 353 Mrd. Euro (Spalte insgesamt, Zeile 2b).

In der folgenden Zelle 4/3 wird die Verteilung der Primäreinkommen aus der Produktionstätigkeit auf die volkswirtschaftlichen Sektoren gezeigt. Der Sektor Private Haushalte bezieht beispielsweise ein Arbeitnehmerentgelt in Höhe von 1 099 Mrd. Euro. Dieses empfangene Arbeitnehmerentgelt errechnet sich, indem das im Rahmen der inländischen Produktion geleistete Arbeitnehmerentgelt in Höhe von 1 100 Mrd. Euro (Zeile 3a, Zelle 3/2) um das Entgelt für Einpendler (5 Mrd. Euro) vermindert und um das Entgelt für Arbeitnehmertätigkeit in der übrigen Welt (Auspendler; 4 Mrd. Euro) erhöht wird. Darüber hinaus erzielen private Haushalte ein Selbstständigeneinkommen in Höhe von 181 Mrd. Euro und einen Nettobetriebsüberschuss von 57 Mrd. Euro. Beim Nettobetriebsüberschuss privater Haushalte handelt es sich um unterstellte Einkommen aus selbstgenutztem Wohneigentum sowie um den Nettobetriebsüberschuss privater Organisationen ohne Erwerbszweck. Insgesamt beläuft sich das Einkommen privater Haushalte aus Produktionstätigkeiten somit auf 1 337 Mrd. Euro. Auch beim Sektor Staat entstehen Einkommen aus Produktionstätigkeiten, nämlich einerseits sonstige Produktionsabgaben (minus sonstige Subventionen) in Höhe von 12 Mrd. Euro und andererseits der Nettobetriebsüberschuss aus Nebentätigkeiten (– 3 Mrd. Euro). Bei der Interpretation des Nettobetriebsüberschusses der finanziellen und nichtfinanziellen Kapitalgesellschaften ist die oben beschriebene Buchung der unterstellten Bankgebühr zu beachten.

Die in den einzelnen Sektoren verbleibenden Primäreinkommen (Markteinkommen der inländischen Sektoren) sind in Zelle 5/4 ersichtlich. Sie ergeben sich wie folgt:

Einkommen aus Produktionstätigkeit nach dem Inländerkonzept (Zelle 4/3)

- +/- Empfangene und geleistete Vermögenseinkommen zwischen den inländischen Sektoren (Zelle 4/4)
- + Vermögenseinkommen aus der übrigen Welt (Zelle 4/10)<sup>13)</sup>
- Vermögenseinkommen an die übrige Welt (Zelle 10/4)

+ Geleistete Gütersteuern (minus empfangene Gütersubventionen) inländischer Sektoren an den Staat (Zelle 4/1)

= Primäreinkommen der inländischen Sektoren

Bei den Vermögenseinkommen handelt es sich um empfangene und geleistete Zinsen, Ausschüttungen und Entnahmen von den Sektoren Nichtfinanzielle und Finanzielle Kapitalgesellschaften, Pachteinkommen, Vermögenseinkommen aus Versicherungsverträgen und reinvestierte Gewinne von der und an die übrige Welt.

An den Sektoren Private Haushalte und Nichtfinanzielle Kapitalgesellschaften sei die Buchung erläutert. Der Nettobetriebsüberschuss nichtfinanzieller Kapitalgesellschaften in Höhe von 226 Mrd. Euro (Zelle 4/3) erhöht sich durch empfangene Vermögenseinkommen von anderen inländischen Sektoren (+ 33 Mrd. Euro, Zeile 4a in Zelle 4/4) und Vermögenseinkommen aus der übrigen Welt (+ 28 Mrd. Euro, Zeile 4a in Zelle 4/10) um 61 Mrd. Euro. Dem stehen geleistete Vermögenseinkommen in Höhe von insgesamt 297 Mrd. Euro gegenüber. Davon gehen 276 Mrd. Euro an inländische Sektoren (Spalte 4a in Zelle 4/4) und 21 Mrd. Euro (Spalte 4a in Zelle 10/4) an die übrige Welt. Im Ergebnis ergibt sich ein Primäreinkommen von – 11 Mrd. Euro. Daraus können jedoch keine Rückschlüsse auf die Höhe der Unternehmensgewinne gezogen werden, da es sich bei den Primäreinkommen um Einkommen nach geleisteten Ausschüttungen und Entnahmen handelt.

Private Haushalte beziehen ein Primäreinkommen aus Produktionstätigkeit in Höhe von 1 337 Mrd. Euro. Sie leisten Vermögenseinkommen – vor allem in Form von Zinszahlungen – in Höhe von 100 Mrd. Euro (Spalte 4d in Zelle 4/4) und empfangen Vermögenseinkommen aller Art in Höhe von 299 Mrd. Euro von inländischen Sektoren (Zeile 4d in Zelle 4/4) und 10 Mrd. Euro aus der übrigen Welt (Zelle 4/10). Damit ergibt sich per saldo ein Primäreinkommen der privaten Haushalte von 1 545 Mrd. Euro (Zeile 5d, Spalte 4d).

Um zum verfügbaren Einkommen der Sektoren in Zelle 6/5 zu gelangen, sind die Primäreinkommen um die empfangenen und geleisteten laufenden Transfers im Inland und gegenüber der übrigen Welt zu bereinigen. Diesen Übergang zeigt Zelle 5/5. Der Sektor Private Haushalte beispielsweise leistet Transfers in Form von Steuern, Sozialbeiträgen, Nettoprämien für Schadensversicherungen und andere Transfers in Höhe von insgesamt 696 Mrd. Euro an inländische Sektoren (Spalte 5d in Zelle 5/5) und 12 Mrd. Euro an die übrige Welt (Spalte 5d in Zelle 10/5), die von den empfangenen Primäreinkommen (1 545 Mrd. Euro) abzuziehen sind. Mit 600 Mrd. Euro geht der weitaus größte Teil der Transfers an den Sektor Staat, der auch die gesetzlichen Sozialversicherungen einschließt. Von inländischen Sektoren empfangen private Haushalte insgesamt 470 Mrd. Euro (Zeile 5d in Zelle 5/5) und aus dem Ausland 3 Mrd. Euro (Zeile 5d in Zelle 5/10); hierbei handelt es sich vor allem um empfangene Transfers vom Staat, wie Rentenzahlun-

<sup>13)</sup> Einschließlich des Saldos aus Gütersteuern minus Gütersubventionen zur übrigen Welt: Dabei handelt es sich um Gütersteuern abzüglich Gütersubventionen aus der Europäischen Union in Höhe von – 8 Mrd. Euro (Zeile 4c in Zelle 4/10). Aufgrund konzeptioneller Unterschiede zwischen Input-Output-Rechnung und Kontensystem ist diese Korrekturbuchung notwendig.

gen, Kranken- oder Arbeitslosengeld. Daraus ergibt sich ein negativer Transfersaldo von 235 Mrd. Euro und ein verfügbares Einkommen privater Haushalte von 1 310 Mrd. Euro (Zeile 6d, Spalte 5d).

In der Spalte 6 der SAM ist die Aufteilung des verfügbaren Einkommens auf Konsum und Sparen dargestellt. So zeigt Zelle 1/6 den Staatskonsum (Spalte 6c) und den Konsum privater Haushalte einschließlich privater Organisationen ohne Erwerbszweck (Spalte 6d) nach Gütergruppen. Das Sparen der Sektoren ist in Zelle 7/6 aufgezeigt. Das Sparen privater Haushalte einschließlich privater Organisationen ohne Erwerbszweck errechnet sich aus deren verfügbarem Einkommen in Höhe von 1 310 Mrd. Euro minus deren Konsum von 1 196 Mrd. Euro plus den gestiegenen betrieblichen Versorgungsansprüchen privater Haushalte in Höhe von 15 Mrd. Euro (Zelle 6/6), sodass sich ein Sparen von 129 Mrd. Euro ergibt (Zeile 7d, Spalte 6d).

Neben dem Sparen aus dem laufenden Einkommen steht den privaten Haushalten auch noch der Saldo aus empfangenen minus geleisteten Vermögenstransfers (Zeile 7d abzüglich Spalte 7d in Zelle 7/7) in Höhe von 16 Mrd. Euro zur Vermögensbildung zur Verfügung. Vermögenstransfers empfangen private Haushalte beispielsweise in Form von Investitionszuschüssen für den Wohnungsbau oder als Sparprämien vom Staat. Zu den geleisteten Vermögenstransfers zählen die Erbschaft- und Schenkungsteuer sowie Vermächtnisse und Schenkungen. Sparen und per saldo empfangene Vermögenstransfers verwenden private Haushalte einerseits für Nettoanlageinvestitionen (Zelle 8/7) – Bruttoanlageinvestitionen (Zelle 1/8) minus Abschreibungen (Zelle 8/2) –, wobei sich dahinter vor allem Investitionen in den Wohnungsbau verbergen, und andererseits für die Geldvermögensbildung, die im nächsten Abschnitt erläutert wird.

### 3.3 Finanzielle Transaktionen

Finanzielle Transaktionen werden unterschieden in die Geldvermögensbildung und die Außenfinanzierung. Die Außenfinanzierung bzw. der Nettoanstieg der Passiva (Zelle 7/9) zeigt die Veränderung der Verbindlichkeiten und die Geldvermögensbildung bzw. der Nettoerwerb von finanziellen Aktiva (Zelle 9/7) die Änderungen in der Anlage von Finanztiteln nach Sektoren auf. Zugänge an Forderungen oder Verbindlichkeiten einer bestimmten Art werden mit den betreffenden Abgängen in der Periode aufgerechnet. Die Nettogeldvermögensbildung nach Sektoren ergibt sich aus der Differenz von Geldvermögensbildung minus Außenfinanzierung. Zur Bewertung der finanziellen Ströme sind grundsätzlich die tatsächlichen Transaktionspreise heranzuziehen. Falls dies nicht möglich ist, erfolgt die Bewertung anhand von Bestandinformationen mit den jeweils dazugehörigen Marktpreisen.

Definitionen und Konzepte der Finanzierungsrechnung sind Bestandteil der Volkswirtschaftlichen Gesamtrechnungen

und im ESVG 95 erläutert. Ermittelt werden die Daten im Rahmen der Finanzierungsrechnung der Deutschen Bundesbank.<sup>14)</sup>

Auch hier sei am Beispiel der privaten Haushalte die Buchung erläutert (siehe Übersicht 2). Private Haushalte einschließlich privater Organisationen ohne Erwerbszweck nahmen Kredite in Höhe von 43 Mrd. Euro auf und reduzierten ihre sonstigen Verbindlichkeiten um 1 Mrd. Euro (Zeile 7d in Zelle 7/9). Bei der Geldvermögensbildung privater Haushalte (Spalte 7d in Zelle 9/7) zeigt sich folgendes Bild: Bargeld und Sichteinlagen wurden um 31 Mrd. Euro reduziert, dagegen erhöhten private Haushalte im Jahr des auslaufenden Aktienbooms ihre Anteilsrechte in Form von Aktien, Investmentzertifikaten und sonstigen Beteiligungen um 75 Mrd. Euro. Auch die Ansprüche privater Haushalte gegenüber Versicherungen erhöhten sich um 56 Mrd. Euro. Wertpapiere und Finanzderivate erfuhren einen Zuwachs von 6 Mrd. Euro und die sonstigen Forderungen, wobei es sich vor allem um betriebliche Pensionsrückstellungen handelt, stiegen um 11 Mrd. Euro.

Abschließend sei noch auf zwei Besonderheiten hingewiesen, die sich empirisch im spiegelbildlichen Vergleich der Zeile 7 (7a bis 7d) und Spalte 7 (7a bis 7d) in den jeweiligen Insgesamtwerten zeigen. So ist dort unschwer zu erkennen, dass die spiegelbildlichen Insgesamtwerte für die Sektoren Nichtfinanzielle und Finanzielle Kapitalgesellschaften sowie für den Sektor Staat nicht übereinstimmen. Dies hat zwei Gründe: Während die Angaben aus den Volkswirtschaftlichen Gesamtrechnungen des Statistischen Bundesamtes auf dem Rechenstand Sommer 2003 basieren, liegt den Ergebnissen der Finanzierungsrechnung der Rechenstand Februar 2003 zugrunde. Zudem kommt es aufgrund der eigenständigen Berechnungen von nichtfinanziellen Strömen und finanziellen Strömen zu einer statistischen Differenz zwischen der in der Finanzierungsrechnung ermittelten Nettogeldvermögensbildung und dem Finanzierungssaldo nach Sektoren, der im Rahmen der nichtfinanziellen Sektorrechnung ermittelt wird. Diese statistische Differenz in Höhe von 32 Mrd. Euro wird im Sektor Nichtfinanzielle Kapitalgesellschaften gebucht.

## 4 Sozioökonomische Erweiterungen der Volkswirtschaftlichen Gesamtrechnungen

Bei den in Kapitel 3 aufgezeigten Ergebnissen handelt es sich – wie schon angesprochen – im Grunde nicht um eine SAM, sondern um eine VGR-Matrix (NAM), da lediglich Ergebnisse der Volkswirtschaftlichen Gesamtrechnungen in Matrixform dargestellt werden. Erst Erweiterungen bzw. Vertiefungen dieser Ergebnisse führen zu einer SAM im eigentlichen Sinn. An den drei folgenden Beispielen soll dies verdeutlicht werden.

<sup>14)</sup> Siehe hierzu auch Deutsche Bundesbank: „Ergebnisse der gesamtwirtschaftlichen Finanzierungsrechnung für Deutschland 1991 bis 2002“, Statistische Sonderveröffentlichung 4, 2003.

## 4.1 Qualifikationsstruktur der Erwerbstätigen

Die erste Erweiterung knüpft an die Arbeitnehmerentgelte nach Wirtschaftsbereichen an. Diese Angaben sind in Zelle 3/2 der SAM zu finden, also in der Zeilengliederung bei der Einkommensentstehung bzw. in der Spaltengliederung bei der Produktion. Die Zusatzinformationen enthalten Angaben zur Struktur der Arbeitnehmer nach Wirtschaftsbereichen in der Gliederung nach Alter, Geschlecht und Qualifikationsniveau auf Basis der ISCED 97.<sup>15)</sup> Angaben zur Zahl der Arbeitnehmer nach Wirtschaftsbereichen gehören zum Standardveröffentlichungsprogramm der Volkswirtschaftlichen Gesamtrechnungen und dienen als Basis für die zusätzlichen Untergliederungen. Damit ist die Konsistenz zur Erwerbstätigenrechnung in den Volkswirtschaftlichen Gesamtrechnungen gewahrt. Die Zahl der Arbeitnehmer der einzelnen Wirtschaftsbereiche entspricht also den veröffentlichten Ergebnissen der Volkswirtschaftlichen Gesamtrechnungen. Die weiteren Aufteilungen wurden anhand von Strukturinformationen des Mikrozensus 2000 vorgenommen.

In Schaubild 1 wird die Zusammensetzung der weiblichen und männlichen Arbeitnehmer nach deren Qualifikations-

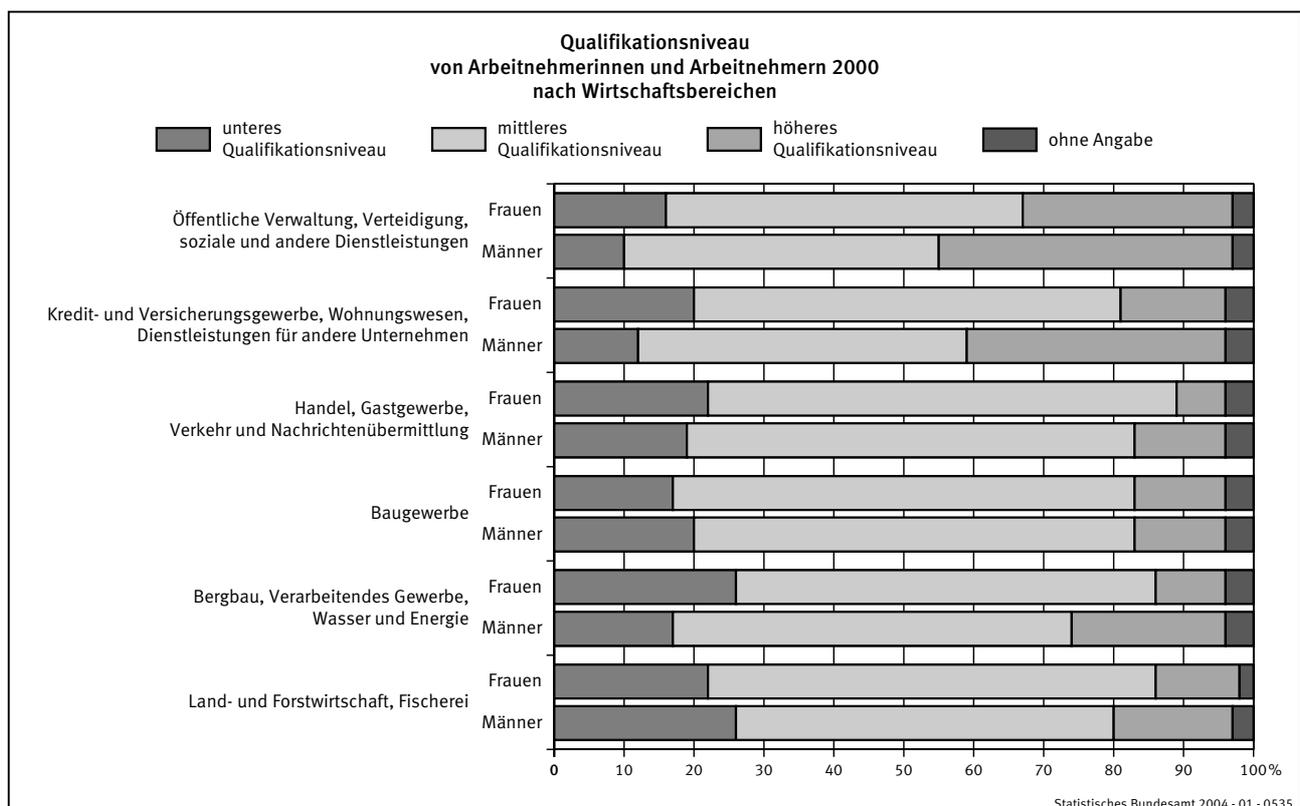
niveau in den einzelnen Wirtschaftsbereichen gezeigt. Die grafische Darstellung veranschaulicht, dass zwischen den Wirtschaftsbereichen deutliche Unterschiede hinsichtlich des Qualifikationsniveaus der Arbeitnehmer bestehen.

Während im Bereich Öffentliche Verwaltung, Verteidigung, soziale und andere Dienstleistungen gut ein Drittel aller beschäftigten Arbeitnehmer einen höheren Bildungsabschluss angegeben hat, liegt der entsprechende Anteil im Baugewerbe bei etwa einem Achtel. Im Bereich Handel, Gastgewerbe, Verkehr und Nachrichtenübermittlung verfügt nur jeder zehnte Arbeitnehmer über einen höheren Bildungsabschluss. Zudem zeigt sich, dass der Anteil der Arbeitnehmer mit höherem Qualifikationsniveau bei männlichen Arbeitnehmern – mit Ausnahme des untypischen Baugewerbes – über dem der Frauen liegt. Unabhängig von Geschlecht und Wirtschaftsbereich besitzen die meisten Arbeitnehmer einen mittleren Bildungsabschluss.

## 4.2 Einkommen privater Haushalte nach sozioökonomischen Haushaltsgruppen

Ein zentraler Baustein zur Erstellung der SAM ist die Unterteilung der Einkommen privater Haushalte nach sozioöko-

Schaubild 1



15) Die ISCED (International Standard Classification of Education) ist die international üblicherweise verwendete Klassifikation für Ausbildungsabschlüsse. Zum *unteren* Qualifikationsniveau (ISCED-Level 1 und 2) zählen Personen ohne allgemeinen schulischen oder beruflichen Abschluss sowie Personen mit Haupt- oder Realschulabschluss, die keine Lehrausbildung abgeschlossen haben. Zum *mittleren* Qualifikationsniveau (ISCED-Level 3 und 4) zählen Personen, die eine abgeschlossene Lehrausbildung oder einen berufsqualifizierenden Abschluss an einer Berufsfachschule bzw. Kollegschule haben oder die Fachhochschulreife/allgemeine Hochschulreife besitzen. Zum *höheren* Qualifikationsniveau (ISCED-Level 5 und 6) zählen Personen mit abgeschlossener Meister-/Technikerausbildung oder einem gleichwertigen Fachschulabschluss, Personen mit abgeschlossenem Hochschul- oder Fachhochschulstudium und Personen mit Promotion.

nomischen Haushaltsgruppen. Die Grundlagen für den Aufbau eines Rechensystems zur Einkommensverteilung nach Haushaltsgruppen innerhalb der Volkswirtschaftlichen Gesamtrechnungen wurden schon in den 70er- und 80er-Jahren des letzten Jahrhunderts gelegt. Seitdem wurden in unregelmäßigen Abständen Ergebnisse veröffentlicht.<sup>16)</sup> In Anlehnung an die Empfehlungen zur Bildung von Teilsektoren des Sektors Private Haushalte im ESVG 95 werden die Erwerbstätigenhaushalte nach Haushalten von Selbstständigen, Beamten, Angestellten und Arbeitern sowie die Nichterwerbstätigenhaushalte nach Haushalten von Arbeitslosengeld-/hilfeempfängern, Rentnern, Pensionären, Sozialhilfeempfängern und sonstigen Nichterwerbstätigen unterschieden. Die Zuordnung der Haushalte zu den einzelnen Haushaltsgruppen richtet sich nach dem überwiegenden Einkommen der Bezugsperson des Haushalts. Bezugsperson ist diejenige Person, die am meisten zum Lebensunterhalt des Haushaltes beiträgt. Zum Haushaltseinkommen gehören die Einkommen aller Haushaltsmitglieder aus verschiedenen Quellen. Zu beachten ist, dass die Einkommen einzelner Haushalte erheblich von den dargestellten Durchschnittseinkommen je Haushalt abweichen können.

Wie sich die Einkommen der Haushalte aus Arbeitnehmerentgelten, Unternehmens- und Vermögenseinkommen und empfangenen laufenden Transfers zusammensetzen und was nach Abzug der geleisteten laufenden Transfers an verfügbarem Einkommen für Konsumzwecke und das Sparen übrig bleibt, zeigt Tabelle 1. Die Berechnungen schließen unmittelbar an die in der SAM dargestellten Ergebnisse

an. So ist beispielsweise in Zelle 5/4 das Primäreinkommen und in der Zelle 5/5 der SAM die gesamtwirtschaftliche Umverteilung in Form von empfangenen und geleisteten Transfers aufgezeigt. Diese Makrogrößen dienen als Basis für die Berechnung der Einkommen nach sozioökonomischen Haushaltsgruppen.

Addiert man zum Primäreinkommen den Saldo der empfangenen abzüglich der geleisteten laufenden Transfers, ergibt sich das verfügbare Einkommen (siehe Tabelle 1). Haushalte, die ihr Einkommen überwiegend aus Erwerbsarbeit beziehen, haben einen deutlich negativen Transfersaldo, Nichterwerbstätigenhaushalte bestreiten ihren Lebensunterhalt dagegen überwiegend aus Transferzahlungen. Bei empfangenen Transfers handelt es sich vor allem um Zahlungen der Sozialversicherungen und um andere staatliche Transfers wie Kindergeld, Sozialhilfe und Wohngeld. Die geleisteten Transfers sind insbesondere direkte Steuern und Sozialbeiträge.

Aus den dargestellten Ergebnissen geht deutlich hervor, dass es sich bei den Haushaltseinkommen um Mischeinkommen handelt. Bei einem Vergleich der Einkommen einzelner Haushaltsgruppen müssen einige spezifische Aspekte der Volkswirtschaftlichen Gesamtrechnungen beachtet werden, deren Auswirkungen bei den einzelnen Haushaltsgruppen sehr unterschiedlich sind: So reduzieren bei Arbeitern und Angestellten die Beiträge zur Sozialversicherung deren verfügbares Einkommen, da sie als geleistete laufende Transfers abgezogen werden.<sup>17)</sup> Selbstständige müssen dagegen ihre private Altersvorsorge im

Tabelle 1: Einkommen und laufende Transfers nach Haushaltsgruppen im Jahr 2000

Haushaltsgruppe	Empfangene Einkommen und Transfers				Geleistete laufende Transfers	Verfügbares Einkommen
	insgesamt	Arbeitnehmerentgelte	Unternehmens- und Vermögenseinkommen	Empfangene laufende Transfers		
EUR je Haushalt						
Privathaushalte insgesamt .....	53 100	29 100	11 500	12 500	19 300	33 800
darunter:						
Selbstständigenhaushalte .....	127 600	12 900	107 100	7 600	32 100	95 500
Beamtenhaushalte .....	83 300	68 100	5 400	9 700	35 800	47 500
Angestelltenhaushalte .....	70 400	61 600	4 200	4 700	32 700	37 700
Arbeiterhaushalte .....	54 900	47 200	2 700	5 000	24 300	30 600
Nichterwerbstätigenhaushalte mit überwiegendem Lebensunterhalt der Bezugsperson aus ...						
Arbeitslosengeld/-hilfe .....	29 800	7 400	2 500	19 800	10 000	19 700
Rente .....	28 800	3 500	4 900	20 300	6 000	22 700
%						
Privathaushalte insgesamt .....	100	55	22	23	36	64
darunter:						
Selbstständigenhaushalte .....	100	10	84	6	25	75
Beamtenhaushalte .....	100	82	6	12	43	57
Angestelltenhaushalte .....	100	87	6	7	46	54
Arbeiterhaushalte .....	100	86	5	9	44	56
Nichterwerbstätigenhaushalte mit überwiegendem Lebensunterhalt der Bezugsperson aus ...						
Arbeitslosengeld/-hilfe .....	100	25	8	67	34	66
Rente .....	100	12	17	71	21	79

16) Siehe hierzu Statistisches Bundesamt (Hrsg.): „Einkommensverteilung nach Haushaltsgruppen und Einkommensarten 1991 bis 2002“, Wiesbaden 2003, sowie Spies, V.: „Verfügbares Einkommen nach Haushaltsgruppen“ in WiSta 7/1992, S. 418 ff.

17) Bei Beamten werden unterstellte Sozialbeiträge für die Alterssicherung gebucht, die einerseits im Arbeitnehmerentgelt enthalten sind und andererseits in den geleisteten laufenden Transfers.

Tabelle 2: Verfügbares Einkommen nach Haushaltsgruppen im Jahr 2000  
EUR je Haushalt

Haushaltsgruppe	Alle Haushalte	Einpersonenhaushalte	Zweipersonenhaushalte	Dreipersonenhaushalte	Vierpersonenhaushalte	Haushalte mit fünf oder mehr Personen
Privathaushalte insgesamt .....	33 800	18 900	35 400	45 200	52 600	56 800
darunter:						
Selbstständigenhaushalte .....	95 500	63 500	97 700	100 900	109 400	121 200
Beamtenhaushalte .....	47 500	26 600	50 200	53 700	58 600	62 600
Angestelltenhaushalte .....	37 700	21 000	41 100	45 700	51 900	54 500
Arbeiterhaushalte .....	30 600	16 100	29 200	34 900	38 700	42 000
Nichterwerbstätigenhaushalte mit überwiegendem Lebensunterhalt der Bezugsperson aus ...						
Arbeitslosengeld/-hilfe .....	19 700	9 900	21 900	27 500	29 500	30 200
Rente .....	22 700	15 300	27 100	39 300	45 100	50 300

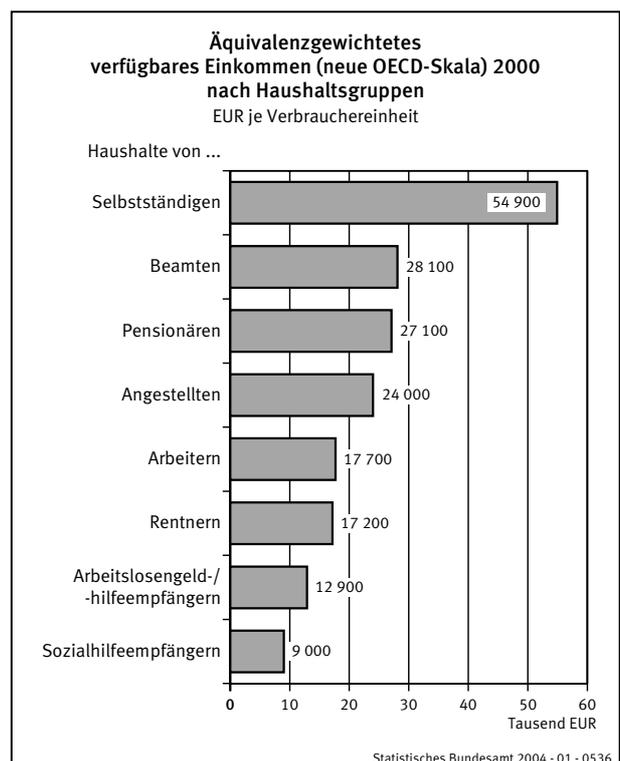
Allgemeines aus ihrem verfügbaren Einkommen leisten, soweit es sich nicht um Zahlungen an die gesetzliche Rentenversicherung handelt. Auch die Art der Krankenversicherung beeinflusst die Höhe des verfügbaren Einkommens. Erstattungen privater Krankenkassen und Beihilfezahlungen an Beamte und Pensionäre erhöhen deren verfügbares Einkommen gegenüber gesetzlich Versicherten, da Sachleistungen der gesetzlichen Krankenkassen nicht als empfangene Transfers gebucht werden und somit nicht im verfügbaren Einkommen enthalten sind. Der im Vergleich zu den anderen Erwerbstätigenhaushalten höhere Anteil an empfangenen Transfers bei den Beamtenhaushalten ist zum Großteil auf Beihilfezahlungen und Erstattungen privater Krankenversicherungen zurückzuführen. Erstattungen privater Krankenversicherungen sind daneben auch bei Selbstständigenhaushalten von überdurchschnittlichem Gewicht. Diese spezifischen Aspekte der Volkswirtschaftlichen Gesamtrechnungen beeinträchtigen den Vergleich der Einkommen nach Haushaltsgruppen, da sie das verfügbare Einkommen von Selbstständigen- und Beamtenhaushalten überzeichnen.

Konzeptionell unterscheidet sich die SAM von der Berechnung des Einkommens nach Haushaltsgruppen dadurch, dass die innersektorale Umverteilung zwischen den einzelnen Haushaltsgruppen bzw. zwischen privaten Haushalten und privaten Organisationen ohne Erwerbszweck in der gesamtwirtschaftlichen Darstellung in der SAM – mit Ausnahme der in Zusammenhang mit Arbeitnehmerentgelten stehenden sozialen Leistungen und Sozialbeiträgen – nicht enthalten ist. Die Berechnungen zur Einkommensverteilung nach Haushaltsgruppen berücksichtigen auch innersektorale Umverteilungsströme, wie zum Beispiel Unterhaltsleistungen von geschiedenen Ehepartnern, gezahlte Kirchensteuer und Vereinsbeiträge.

Tabelle 2 zeigt die Verteilung der verfügbaren Einkommen anknüpfend an Zelle 6/5 nach Haushaltsgruppen und der Haushaltsgröße. Mit zunehmender Haushaltsgröße steigt in allen Haushaltsgruppen auch das verfügbare Haushaltseinkommen an. Dies ist vor allem auf einen Anstieg der Primäreinkommen zurückzuführen, da sich mit steigender Haushaltsgröße auch die Zahl der potenziellen Einkommensbezieher erhöht.

Eine Betrachtung der Einkommen je Haushalt allein ist für Wohlstandsuntersuchungen wenig aussagekräftig. Deshalb sollte man für diese Untersuchungen auch die Zahl der Haushaltsmitglieder, die von dem verfügbaren Einkommen leben, einbeziehen. Häufig wird deshalb das Einkommen je Haushaltsmitglied angegeben. Dies berücksichtigt allerdings nicht die Kostenvorteile, die sich ergeben, wenn mehrere Personen gemeinsam einen Haushalt führen. Deshalb werden auch Einkommen je Verbrauchereinheit berechnet, wobei den Personen im Haushalt unterschiedliche Gewichte zugerechnet werden. Diese Einkommen werden als Äquivalenzeinkommen bezeichnet. In Schaubild 2 werden die äquivalenzgewichteten verfügbaren Einkommen der einzelnen Haushaltsgruppen gezeigt. International üblich ist die Darstellung nach der alten oder neuen Äquivalenzskala der Organisation für wirtschaftliche Zusammenar-

Schaubild 2



beit und Entwicklung (OECD). Bei der hier gewählten neuen OECD-Skala wird die erste Person im Haushalt mit 1,0, jede weitere Person im Haushalt ab 15 Jahren mit 0,5 und Kinder unter 15 Jahren mit 0,3 gewichtet. Äquivalenzeinkommen sind aussagekräftiger für einen Einkommensvergleich zwischen Haushalten unterschiedlicher Größe und Zusammensetzung (Erwachsene, Kinder) als eine einfache Pro-Kopf-Betrachtung.

### 4.3 Konsum privater Haushalte nach sozioökonomischen Haushaltsgruppen

Im Anschluss an die Ergebnisse zum verfügbaren Einkommen nach Haushaltsgruppen rückt jetzt die Verwendung des verfügbaren Einkommens für Konsumausgaben und Sparen in derselben sozioökonomischen Gliederung ins Zentrum der Betrachtung. Die Ergebnisse knüpfen an die in Zelle 1/6 der SAM dargestellte letzte Verwendung nach Gütergruppen für private Haushalte (Spalte 6d in Übersicht 2) an. Allerdings wird der Konsum hier nicht nach Gütergruppen, sondern nach Verwendungszwecken gemäß dem Systematischen Verzeichnis der Einnahmen und Ausgaben der privaten Haushalte in der derzeit gültigen Fassung von 1998 (SEA 98) dargestellt, da Verwendungszwecke für die Analyse von Ausgabenstrukturen nach Haushaltsgruppen aussagekräftiger sind.<sup>18)</sup> Das Sparen (einschließlich Zunahme betrieblicher Versorgungsansprüche) der privaten Haushalte – in der SAM in Zelle 7/6 Spalte 6d aufgeführt – ergibt sich als Differenz zwischen dem verfügbaren Einkommen und dem Konsum. Bei den gezeigten Angaben handelt es sich um erste, vorläufige Ergebnisse. Weitere methodische Arbeiten mit dem Ziel, tiefer disaggregierte Ergebnisse zu ermitteln, sind für die nächsten Monate geplant.

Ausgangspunkt der Berechnungen sind die in den Volkswirtschaftlichen Gesamtrechnungen für private Haushalte insgesamt ausgewiesenen Konsumausgaben nach Verwendungszwecken im Inland, die anhand von Angaben zu Reiseausgaben in das Inländerkonzept überführt werden. Grundlage für die weitere Aufteilung dieser Angaben auf Haushaltsgruppen sind die Ergebnisse der Einkommens- und Verbrauchsstichprobe 1998 (EVS 98). Da es sowohl konzeptionell als auch in der Berechnungsmethodik Unterschiede zwischen EVS und Volkswirtschaftlichen Gesamtrechnungen gibt, waren die folgenden Anpassungen notwendig:<sup>19)</sup>

- Die Ausgangsdaten der EVS wurden zunächst auf der Einkommens- und Ausgabenseite an die Konzepte des ESVG 95 angepasst. Hierzu gehören beispielsweise die Beiträge für private Schadensversicherungen, deren geleistete Bruttoprämien in der EVS erfasst sind, während nach dem ESVG 95 nur die hierin enthaltenen Dienstleistungsentgelte zum Konsum zählen. Daher war es notwendig, die in der EVS nachgewiesenen Bruttoprämien anhand von Zusatzinformationen aus den Volkswirtschaftlichen Gesamtrechnungen zu Versicherungsunternehmen in Nettoprämien und Dienstleistungsentgelte aufzuteilen.

- In den Angaben zum Konsum privater Haushalte in den Volkswirtschaftlichen Gesamtrechnungen ist auch die in Anstalten wie Pflege- und Kinderheimen lebende Bevölkerung eingeschlossen. In der EVS ist diese so genannte Anstaltsbevölkerung nicht befragt worden. Für diese Bevölkerungsgruppe wurden anhand von Angaben aus dem Mikrozensus über die Zahl der in Anstalten lebenden Personen Abschläge am Konsum privater Haushalte gemäß den Volkswirtschaftlichen Gesamtrechnungen vorgenommen. Auch Haushalte mit einem monatlichen Haushaltsnettoeinkommen von über 35 000 DM (17 895 Euro) sind in der EVS 1998 nicht erfasst. Für diese Haushalte wurden Zuschätzungen auf die Ergebnisse der EVS vorgenommen, sodass diese Haushalte im Gegensatz zur Anstaltsbevölkerung in den weiter unten gezeigten Angaben enthalten sind.

- Da die Untergliederung der Ausgaben in der EVS nicht überall den Abgrenzungen des SEA 98 folgt, mussten einige Positionen neu zugeordnet bzw. aufgeteilt werden.

Die so bereinigten Ergebnisse der EVS wurden anschließend mit der Zahl der Haushalte hochgerechnet und auf die Eckgrößen aus den Volkswirtschaftlichen Gesamtrechnungen für das Jahr 1998 abgestimmt. Anschließend erfolgte eine Fortschreibung auf das Berichtsjahr 2000 anhand der Ausgangswerte aus den Volkswirtschaftlichen Gesamtrechnungen.

Tabelle 3 zeigt die Ergebnisse zur Verwendung des verfügbaren Einkommens nach acht Verwendungszwecken und dem als Rest ermittelten Sparen für die in Abschnitt 4.2 näher erläuterten sechs sozioökonomischen Haushaltsgruppen. Bei den Angaben handelt es sich um Durchschnittswerte für die jeweiligen Haushaltstypen, die je nach individueller Lebenslage deutlich von den Werten für einzelne Haushalte abweichen können. In der ersten Zeile sind die verfügbaren Einkommen je Haushalt im Jahr 2000 angegeben, darunter der Konsum nach Verwendungszwecken in Prozent des verfügbaren Einkommens und abschließend die haushaltsspezifischen Sparquoten.

Bei der Verwendung des Einkommens für konsumtive Zwecke stellen bei allen Haushaltsgruppen die Ausgaben für das Wohnen einschließlich der Wohnungsnebenkosten den größten Posten dar. Haushaltsgruppen mit geringem Durchschnittseinkommen, wie etwa Empfänger von Arbeitslosengeld oder -hilfe, müssen dabei einen vergleichsweise hohen Anteil ihres verfügbaren Einkommens (30%) für das Wohnen aufwenden. Lediglich bei den aufgeführten Nichterwerbstätigenhaushalten liegen die Ausgaben für Nahrungsmittel, Getränke und Tabakwaren über denen für Verkehr und Nachrichtenübermittlung. Bei Erwerbstätigenhaushalten liegen dagegen die Ausgaben für Verkehr und Nachrichtenübermittlung an zweiter Stelle.

Bei den Ausgaben für Gesundheitspflege ist zu beachten, dass laut ESVG 95 privat abgerechnete Krankheitskosten – wie bei privat Versicherten üblich – einerseits Konsum

18) In den Volkswirtschaftlichen Gesamtrechnungen wird der Konsum sowohl nach Verwendungszwecken als auch nach Gütergruppen berechnet.

19) Zu den Konzepten der EVS 98 siehe Fachserie 15 „Wirtschaftsrechnungen“, Heft 7 „Aufgabe, Methode und Durchführung der EVS 1998“, Wiesbaden 2002, S. 11 ff.

Tabelle 3: Verwendung des verfügbaren Einkommens im Jahr 2000 nach Haushaltsgruppen

Einkommensverwendung	Alle Haushalte	Darunter:					
		Selbstständigenhaushalte	Beamtenhaushalte	Angestelltenhaushalte	Arbeiterhaushalte	Nichtwerbstätigenhaushalte mit überwiegender Lebensunterhalt der Bezugsperson aus ...	
						Arbeitslosengeld/-hilfe	Rente
EUR je Haushalt und Jahr							
Verfügbares Einkommen .....	33 800	95 500	47 500	37 700	30 600	19 700	22 700
Verwendung des verfügbaren Einkommens in %							
Nahrungsmittel, Getränke und Tabakwaren .....	14	10	12	13	17	21	17
Einrichtungsgegenstände, Bekleidung und Schuhe .....	12	10	12	13	12	12	13
Wohnung, Wasser, Strom, Gas u.a. Brennstoffe .....	22	16	17	19	22	30	28
Gesundheitspflege <sup>1)</sup> .....	3	3	8	2	2	2	4
Verkehr und Nachrichtenübermittlung .....	15	13	16	18	17	16	14
Freizeit, Unterhaltung und Kultur .....	9	7	9	9	9	10	10
Beherbungs- und Gaststättendienstleistungen .....	6	5	6	6	5	5	6
Sonstige Waren und Dienstleistungen <sup>2)</sup> .....	9	9	9	9	9	10	10
Sparen .....	10	27	11	11	7	-6	-2

1) Ohne soziale Sachleistungen der gesetzlichen Krankenversicherung, aber einschließlich privat abgerechneter Leistungen. – 2) Einschl. Bildungswesen.

sind und andererseits die dazugehörigen Erstattungen privater Krankenkassen oder auch Beihilfezahlungen an Beamte Einkommen darstellen (siehe hierzu auch die Erläuterungen in Abschnitt 4.2). Beiträge für gesetzliche Krankenkassen und soziale Sachleistungen für gesetzlich Krankenversicherte sind dagegen weder Bestandteil des verfügbaren Einkommens noch im Konsum privater Haushalte enthalten, sondern Teil des Staatskonsums. Bei der Betrachtung von Beamtenhaushalten ist beispielsweise zu beachten, dass sowohl deren verfügbares Einkommen als auch spiegelbildlich deren Ausgaben für Gesundheitspflege im Vergleich zu anderen Arbeitnehmerhaushalten überhöht sind.

Die höchste Sparquote weisen mit 27% die Selbstständigenhaushalte auf. Dies kann einerseits auf deren vergleichsweise hohes verfügbares Einkommen (95 500 Euro) zurückgeführt werden und andererseits müssen Selbstständige im Gegensatz zu Arbeitnehmern ihre Alterssicherung

zumeist privat aus ihrem verfügbaren Einkommen bestreiten. Auch dies begründet die hohe Sparquote der Selbstständigenhaushalte. Arbeitnehmerhaushalten verbleibt ein deutlich geringerer Teil ihres verfügbaren Einkommens – zwischen 7 und 11% – für das Sparen. Ganz anders stellt sich die Situation bei Nichtwerbstätigenhaushalten dar. Arbeitslosen- und Rentnerhaushalte finanzieren – im Durchschnitt – einen Teil ihrer Konsumausgaben durch die Auflösung von Ersparnissen oder die Aufnahme von Schulden. Bei den Angaben zum Sparen ist zu berücksichtigen, dass sie nur im Rahmen einer Restrechnung ermittelt werden können. Daraus ergeben sich vergleichsweise hohe statistische Unsicherheiten.

Tabelle 4 zeigt, wie sich Konsum und Sparen in Abhängigkeit von der Haushaltsgröße entwickeln. Bei den Ergebnissen fällt auf, dass die durchschnittlichen relativen Kosten für das Wohnen bei zunehmender Anzahl von Haushalts-

Tabelle 4: Verwendung des verfügbaren Einkommens im Jahr 2000 nach Haushaltsgröße

Einkommensverwendung	Alle Haushalte	Einpersonenhaushalte	Zweipersonenhaushalte	Dreipersonenhaushalte	Vierpersonenhaushalte	Haushalte mit fünf oder mehr Personen
EUR je Haushalt und Jahr						
Verfügbares Einkommen .....	33 800	18 900	35 400	45 200	52 600	56 800
Verwendung des verfügbaren Einkommens in %						
Nahrungsmittel, Getränke und Tabakwaren .....	14	12	14	15	15	16
Einrichtungsgegenstände, Bekleidung und Schuhe .....	12	11	13	12	12	11
Wohnung, Wasser, Strom, Gas u.a. Brennstoffe .....	22	26	21	20	19	19
Gesundheitspflege <sup>1)</sup> .....	3	4	4	3	3	3
Verkehr und Nachrichtenübermittlung .....	15	15	16	16	15	14
Freizeit, Unterhaltung und Kultur .....	9	10	9	8	8	8
Beherbungs- und Gaststättendienstleistungen .....	6	6	6	5	5	4
Sonstige Waren und Dienstleistungen <sup>2)</sup> .....	9	10	9	9	9	9
Sparen .....	10	6	8	12	14	16

1) Ohne soziale Sachleistungen der gesetzlichen Krankenversicherung, aber einschließlich privat abgerechneter Leistungen. – 2) Einschl. Bildungswesen.

mitgliedern fallen. Ursachen hierfür dürften neben dem vergleichsweise höheren Einkommen größerer Haushalte auch fallende Kosten je Quadratmeter bei steigender Wohnungsgröße und die geringere Wohnfläche pro Kopf sein. Zugleich steigt mit zunehmender Haushaltsgröße die Sparquote. Für diese auf den ersten Blick etwas verblüffende Entwicklung können mehrerer Ursachen maßgebend sein. So steigt mit der Zahl potenzieller Einkommensempfänger das durchschnittliche Haushaltseinkommen und die Vorteile des gemeinsamen Wirtschaftens (z. B. durch gemeinsame Nutzung von Gebrauchsgütern) senken tendenziell den Bedarf an Konsumgütern pro Kopf. Daneben gibt es noch weitere Einflussgrößen auf das Sparverhalten, die in diesen Ergebnissen nicht unmittelbar zum Ausdruck kommen, wie beispielsweise Alter oder Familientyp. So sind beispielsweise unter den Ein- und Zweipersonenhaushalten überproportional häufig Rentnerhaushalte vertreten, die eine vergleichsweise niedrige Sparquote aufweisen.

## 5 Ausblick

Mit den hier vorgelegten ersten Ergebnissen für eine Sozialrechnungsmatrix in Deutschland soll vor allem die Konzeption einer SAM aufgezeigt werden. Dazu musste sich die Darstellung auf ein hochaggregiertes Niveau beschränken. Für die meisten hier aufgezeigten Angaben liegen jedoch erheblich differenziertere Ergebnisse aus den Volkswirtschaftlichen Gesamtrechnungen vor. Die gütermäßige Verflechtung der Volkswirtschaft, die in der Input-Output-Rechnung gezeigt wird, liegt für das Jahr 2000 in einer Darstellungstiefe von 71 Wirtschaftsbereichen und 71 Gütergruppen vor. Aus dem Kontensystem der Volkswirtschaftlichen Gesamtrechnungen können die Einkommensströme nach volkswirtschaftlichen Sektoren in ihren Komponenten tief gegliedert nachvollzogen werden und die Finanzierungsrechnung zeigt die finanziellen Transaktionen detaillierter, als dies in einem Überblicksaufsatz wie diesem möglich ist.

Eine Erkenntnis ist, dass der konzeptionelle und empirische Zusammenhang zwischen den Rechenwerken der Volkswirtschaftlichen Gesamtrechnungen durch die Zusammenführung zu einem Gesamtbild in Form einer SAM deutlich wird. So ist es beispielsweise möglich, mit einer SAM nachzuvollziehen, welche Auswirkungen Änderungen in der Güterproduktion auf die Einkommensentstehung und in der Folge auf die Einkommensverteilung und das verfügbare Einkommen haben. Daraus wiederum ergeben sich Auswirkungen auf die letzte Verwendung und die für Investitionen in Sach- und Geldvermögen verbleibenden finanziellen Mittel.

Die Vorteile einer SAM kommen aber erst dann zur Geltung, wenn (sozio-)ökonomische Wechselwirkungen in tiefer gehenden Strukturen untersucht werden. Dabei stehen vor allem die privaten Haushalte in ihrer Funktion als Anbieter von Arbeit und Kapital und als Nachfrager von Konsumgütern im Mittelpunkt. Wesentliches Ziel einer SAM ist es daher, die privaten Haushalte in ihrer Struktur tiefer zu beleuchten. Die hier dargestellten Ergebnisse sind ein erster Schritt dazu. In den nächsten Monaten ist geplant, weitere und detailliertere Ergebnisse für private Haushalte und

die dort lebende Bevölkerung zu erstellen. Einerseits soll die Struktur der Erwerbstätigen tiefer beleuchtet werden und andererseits sollen differenziertere Ergebnisse zum Konsum privater Haushalte nach sozioökonomischen Gruppen berechnet werden. Die daraus hervorgehenden Ergebnisse werden sowohl konzeptionell als auch empirisch auf die Angaben aus den Volkswirtschaftlichen Gesamtrechnungen abgestimmt. Diese Ergebnisse sollen dann in der Schriftenreihe zum Sozioökonomischen Berichtssystem veröffentlicht werden. [u](#)

Dr. Rainer Lenz, Dipl.-Volkswirt Roland Sturm, Dr. Daniel Vorgrimler

# Maße für die faktische Anonymität von Mikrodaten

*Kann eine faktische Anonymisierung von Unternehmens- und Betriebsdaten gelingen, auch wenn diese bei Wirtschaftsdaten als wesentlich problematischer angesehen wird als im Bereich der Haushalts- und Personenerhebungen? Mit dieser Frage beschäftigt sich – vom Bundesministerium für Bildung und Forschung (BMBF) gefördert – das Projekt „Faktische Anonymisierung wirtschaftsstatistischer Einzeldaten“, worüber in dieser Zeitschrift bereits berichtet wurde<sup>1)</sup>. Gemeinsam mit der Wissenschaft wollen die statistischen Ämter das so genannte Wissenschaftsprivileg des Bundesstatistikgesetzes (BStatG)<sup>2)</sup> umsetzen, nach dem für die Durchführung wissenschaftlicher Vorhaben faktisch anonymisierte Mikrodaten an Hochschulen und sonstige Einrichtungen mit der Aufgabe unabhängiger wissenschaftlicher Forschung übermittelt werden dürfen. In einem Workshop zur faktischen Anonymität von Unternehmens- und Betriebsdaten am 23. April 2004, an dem sowohl Fachstatistiker als auch Mitarbeiter der Forschungsdatenzentren der statistischen Ämter teilnahmen, wurde das im Projekt erarbeitete Schutzwirkungskonzept diskutiert.<sup>3)</sup> Im vorliegenden Beitrag werden die Inhalte des Workshops und Vorschläge zum künftigen Vorgehen bei der faktischen Anonymisierung vorgestellt.*

## Vorbemerkung

Faktisch anonymisierte Einzeldaten müssen zwei gleichrangigen Ansprüchen gerecht werden: Einerseits verlangt das

BStatG, dass ein ausreichender Schutz der Einzelangaben gewährleistet ist; andererseits müssen die Analysemöglichkeiten der anonymisierten Daten so weit wie möglich erhalten bleiben. Beide Ziele sind sowohl für die Nutzer als auch für die statistischen Ämter von großer Bedeutung.

Die Sicherung der Vertraulichkeit von Einzelangaben ist nicht nur eine den statistischen Ämtern vom Gesetzgeber vorgegebene Aufgabe. Der amtlichen Statistik kann auch bereits ohne tatsächlichen Missbrauch der Daten Schaden entstehen, da bereits der Verdacht des leichtfertigen Umgangs der statistischen Ämter mit Einzelangaben erhebliche Gefahren für die Auskunftsbereitschaft der Befragten birgt. Der erkennbar verantwortungsvolle Umgang mit der statistischen Geheimhaltung ist daher wichtig für die Qualität der Statistik insgesamt.

Um das Wissenschaftsprivileg des § 16 Absatz 6 BStatG beim Datenzugang nicht unnötig einzuschränken, darf eine faktische Anonymisierung andererseits nur auf realistische Gefährdungsszenarien abstellen und nicht dazu führen, dass unnötig hohe Schutzmaßnahmen aufgebaut werden. Den Nutzern faktisch anonymisierter Daten wird von Seiten der statistischen Ämter generell keine Missbrauchsabsicht unterstellt.

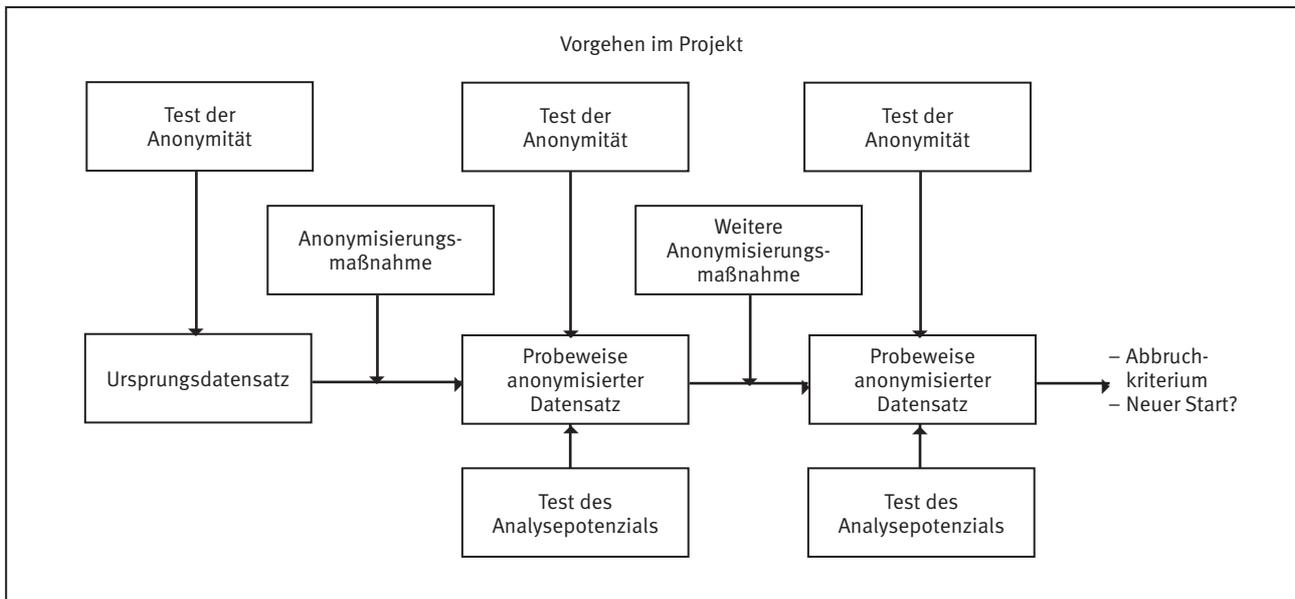
Während mit den künftigen Nutzern von Einzeldaten im Rahmen des Projektes bereits intensive Diskussionen bei zwei Veranstaltungen in Tübingen geführt wurden, widmete sich

1) Siehe Sturm, R.: „Wirtschaftsstatistische Einzeldaten für die Wissenschaft“ in WiSta 2/2002, S. 101 ff.

2) Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) vom 22. Januar 1987 (BGBl. I S. 462, 565), zuletzt geändert durch Artikel 16 des Gesetzes vom 21. August 2002 (BGBl. I S. 3322).

3) Zum Schutzwirkungskonzept siehe Höhne, J./Sturm, R./Vorgrimler, D.: „Konzept zur Beurteilung der Schutzwirkung von faktischer Anonymisierung“ in WiSta 4/2003, S. 287 ff., zum Workshop siehe WiSta 5/2004, S. 501 f.

Schaubild 1



der Workshop in Wiesbaden der Schutzwirkung von Anonymisierungsmaßnahmen.

Im Projekt werden die beiden Aspekte „Schutzwirkung“ und „Analysepotenzial“ umfassend untersucht. Schaubild 1 zeigt schematisch das Vorgehen bei der Anonymisierung einer Datei. Zunächst werden die Ursprungsdaten auf ihre Schutzbedürftigkeit hin untersucht. Dateninkompatibilitäten zu externem Zusatzwissen, die Beschränktheit des Zusatzwissens auf wenige Überschneidungsmerkmale und Eigenschaften der jeweiligen Statistik (Stichprobe, Erhebungsjahr der Daten) wirken bereits als Schutz gegen Deanonimisierung, ohne dass die Originalwerte behandelt werden müssten. Allerdings reicht diese „natürliche“ Schutzwirkung nicht aus, um die Daten aller betroffenen Einheiten einer Erhebung ausreichend zu schützen. Daher werden als nächster Schritt Anonymisierungsmaßnahmen auf die Daten angewendet. Neben den bereits bisher in den statistischen Ämtern eingesetzten Anonymisierungsmaßnahmen werden im Projekt auch Auswirkungen datenverändernder Verfahren untersucht.<sup>4)</sup> Im schematischen Vorgehen wird zunächst eine schwach dosierte Anonymisierungsmaßnahme gewählt, um möglichst viel des Analysepotenzials in den Daten zu erhalten. Die so gewonnene probeweise anonymisierte Datei wird im Statistischen Bundesamt auf ihre Schutzwirkung und vom Institut für Angewandte Wirtschaftsforschung (IAW) auf ihr Analysepotenzial untersucht. Abhängig vom Ergebnis kann ein weiterer Anonymisierungsschritt (z. B. stärkere Dosierung) erfolgen und wiederum auf seine Auswirkung geprüft werden. Bei der Steigerung der Dosierung der Anonymisierungsmaßnahmen ist zu erwarten, dass der erreichte Schutz steigen und das Analysepotenzial abnehmen wird. In den statistischen Ämtern muss entschieden werden, wann der durch die Anonymisierungs-

maßnahmen erzeugte Schutz ausreicht, um die vom BStatG geforderte faktische Anonymität zu gewährleisten.

Das Wissenschaftsprivileg für den Zugang zu faktisch anonymisiertem Einzelmaterial der statistischen Ämter ist im Bundesstatistikgesetz (BStatG) formuliert. Gemäß dem Gebot des § 16 Absatz 6 BStatG können Einzeldaten dann an wissenschaftliche Einrichtungen übermittelt werden, wenn sie nur mit einem „*unverhältnismäßig hohen Aufwand*“ zugeordnet (reidentifiziert) werden können. Dieses Unverhältnismäßigkeitsgebot ist gemeint, wenn von faktischer anstelle absoluter Anonymität gesprochen wird.

Im Rahmen des Projektes „Faktische Anonymisierung wirtschaftsstatistischer Einzeldaten“ wurde ein „Konzept zur Beurteilung der Schutzwirkung von faktischer Anonymisierung“ entwickelt. Durch einen Beitrag in Heft 4/2003 dieser Zeitschrift<sup>5)</sup> wurde die Diskussion über dieses Konzept angestoßen. Um zu einer Operationalisierung zu gelangen, was bei der Reidentifikation von Einzeldaten ein „*unverhältnismäßig hoher Aufwand*“ ist, werden im genannten Konzept realitätskonforme Annahmen für einen Datenangriff zugrunde gelegt. Aus dem Rationalkalkül eines potenziellen Datenangreifers<sup>6)</sup> kann abgeleitet werden, dass ein Deanonimisierungsversuch nur dann erfolgreich wird, wenn eine mögliche Reidentifikation für einen Datenangreifer als nutzbringend einzustufen ist. Der Versuch einer Reidentifikation wäre für einen Datenangreifer dann erwägenswert, wenn er abschätzen kann, dass der (erwartete) Ertrag eines Reidentifikationsversuchs dessen (erwarteten) Aufwand übersteigt. Kosten und Nutzen sind davon abhängig, wie „sicher“ ein Datenangreifer brauchbare Informationen enthüllen kann. In der Praxis wird sich ein Datenangreifer mit zahlreichen Problemen konfrontiert sehen:

4) Für eine Übersicht über Anonymisierungsmethoden und die Auswahl der im Projekt untersuchten Methoden siehe Höhne, J.: „Methoden zur Anonymisierung wirtschaftsstatistischer Einzeldaten“ in Gnoss, R./Ronning, G. (Hrsg.): „Anonymisierung wirtschaftsstatistischer Einzeldaten“, Band 42 der Schriftenreihe „Forum der Bundesstatistik“, Wiesbaden 2003, S. 69 ff.

5) Siehe Fußnote 3.

6) Zum Konzept des Datenangreifers und des Datenangriffs siehe Fußnote 3, S. 288 f.

- mit der Existenz von Dateninkompatibilitäten zwischen Datensätzen aus dem Zusatzwissen und amtlichen Datensätzen<sup>7)</sup>,
- mit der Unkenntnis, ob eine gesuchte Einheit in der amtlichen Erhebung enthalten ist (insbesondere im Fall von Stichprobenerhebungen),
- mit Unsicherheiten über die Richtigkeit von Zuordnungsversuchen,
- mit Unsicherheiten über die Qualität von enthüllten Informationen.

Im Folgenden sollen Berechnungen anhand probeweiser anonymisierter Daten der Kostenstrukturerhebung (KSE) im Verarbeitenden Gewerbe 1999 und der Umsatzsteuerstatistik (USt) 2000 vorgestellt werden.<sup>8)</sup> Als Zusatzwissen dient unter anderem die Datenbank MARKUS (Marketinguntersuchungen), die Informationen zu ausgewählten Unternehmen der Creditreform enthält.<sup>9)</sup> Sie ist im Handel frei erhältlich und wird vierteljährlich herausgegeben. Die Berechnungen zeigen, wie die genannten Unsicherheiten – zum einen auf der Datensatzebene die Unsicherheiten über die Richtigkeit von Zuordnungsversuchen, zum anderen auf der Merkmalsebene die Unsicherheiten über die Qualität von enthüllten Informationen – wirken und wie sie gemessen werden können.<sup>10)</sup>

## 1 Elemente des Schutzes von Mikrodaten: Ebene der Datensätze

### 1.1 Dateninkompatibilitäten zwischen Zusatzwissen und amtlichen Datensätzen

Ein wichtiger Einflussfaktor auf die Datenqualität ist der Weg, auf welchem die Daten gewonnen werden. Die Abhängigkeit vom datengenerierenden Prozess führt dazu, dass Merkmale zum gleichen Sachverhalt, die auf unterschiedlichen Wegen gewonnen werden, unterschiedliche Werte annehmen können. Dies kann zum Beispiel darin begründet liegen, dass ein Sachverhalt zwar gleich deklariert wird, sich die exakten Definitionen hinter den Begriffen jedoch unterscheiden, oder darin, dass die Personen, die an den unterschiedlichen Erhebungen beteiligt sind, manche Begrifflichkeiten jeweils verschieden auffassen und damit unterschiedlich ermitteln. Je unterschiedlicher der Prozess, desto größere „Dateninkompatibilitäten“ können auftreten. Bei den hier untersuchten Daten handelt es sich

einmal um primärstatistische Daten, die durch eine spezielle statistische Erhebung gewonnen werden (Daten der Kostenstrukturerhebung), einmal um sekundärstatistisch aus Verwaltungsquellen übernommene Daten (Daten der Umsatzsteuerstatistik) und einmal um eine Mischform aus primär- und sekundärstatistischen Daten (MARKUS-Datenbank). Die datengenerierenden Prozesse unterscheiden sich demnach erheblich, weshalb Dateninkompatibilitäten unvermeidlich sind. Abweichungsanalysen haben ergeben, dass sich bei Unternehmen des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden mit 20 Beschäftigten und mehr die Wirtschaftszweiguordnung auf Zweistellerebene zwischen der Kostenstrukturerhebung und der Umsatzsteuerstatistik in 25% der Fälle unterscheidet, zwischen der Umsatzsteuerstatistik und der MARKUS-Datenbank in fast 29% der Fälle und zwischen der Kostenstrukturerhebung und der MARKUS-Datenbank in 23% der Fälle. Die Umsatzangaben in Umsatzsteuerstatistik und Kostenstrukturerhebung stimmten verhältnismäßig gut überein (bei lediglich 18,8% der Unternehmen traten Abweichungen von mehr als 10% auf). Dagegen waren starke Abweichungen zwischen Umsatzsteuerstatistik und MARKUS-Datenbank zu verzeichnen (über 60% der Unternehmen wiesen beim Umsatz Abweichungen von mehr als 10% auf). Zwischen der Kostenstrukturerhebung und der MARKUS-Datenbank weichen bei fast 50% der Unternehmen die Ausprägungen bei den Merkmalen Umsatz und Anzahl der Beschäftigten um mehr als 10% voneinander ab. Selbst bei der Regionalerkennung treten zwischen jeweils zwei betrachteten Erhebungen beinahe 2% unterschiedliche Ausprägungen in der Regionalangabe [bezogen auf den siedlungsstrukturellen Regionstyp<sup>11)</sup>] auf.

Der Vergleich insbesondere der Umsatzangaben zeigt, dass die Abweichungen zwischen den amtlichen Erhebungen deutlich geringer sind als etwa zwischen der Umsatzsteuerstatistik und der MARKUS-Datenbank. Solche Dateninkompatibilitäten wirken als natürlicher Schutz gegenüber Reidentifikationsversuchen.<sup>12)</sup>

### 1.2 Unsicherheit über die Möglichkeit einer Zuordnung

Um einen Merkmalsträger einwandfrei in einer Erhebung identifizieren zu können, muss ein Datenangreifer Kenntnis darüber besitzen, dass der gesuchte Merkmalsträger an der Erhebung teilgenommen hat (Teilnahmekennntnis). Ohne diese Kenntnis ist jede erfolgte eindeutige Zuordnung mit der zusätzlichen Unsicherheit behaftet, ob diese Zuordnung zustande kam, obwohl der gesuchte Merkmalsträger nicht

7) Unter einem Datensatz (Record) wird im Folgenden die Menge der einem statistischen Objekt (Unternehmen, Betrieb) zugehörigen Ausprägungen verschiedener Merkmale (Variable) verstanden.

8) Zur Beschreibung der Kostenstrukturerhebung und möglicher Anonymisierungen siehe Rosemann, M./Vorgrimler, D./Lenz, R.: „Erste Ergebnisse faktischer Anonymisierung wirtschaftsstatistischer Einzeldaten“ in Allgemeines Statistisches Archiv 1/2004, S. 73 ff., und zur Umsatzsteuerstatistik siehe Lenz, R./Vorgrimler, D.: „Geheimhaltungsmethoden auf dem Prüfstand – eine Analyse anhand der Umsatzsteuerstatistik“ in diesem Heft auf S. 639 ff.

9) Informationen zur Datenbank MARKUS finden sich auf der Internetseite [http://www.branchen-adressen.de/html/markus\\_marketingdatenbank.html](http://www.branchen-adressen.de/html/markus_marketingdatenbank.html).

10) Die verschiedenen Risiken für einen Datenangreifer werden durch Anonymisierungsmaßnahmen verstärkt, bestehen teilweise aber bereits unabhängig davon. Zur Wirkungsweise von Anonymisierungsmaßnahmen siehe ausführlicher Höhne, J.: „Methoden zur Anonymisierung wirtschaftsstatistischer Einzeldaten“, Fußnote 4.

11) Zur Definition der nichtadministrativen Gebietsschlüssel des Bundesamtes für Bauwesen und Raumordnung siehe Internetseite <http://www.bbr.bund.de/raumordnung/raumbearbeitung/gebietstypen.htm>.

12) Siehe ausführlich zum natürlichen Schutz Lenz, R./Vorgrimler, D., Fußnote 8.

in der Erhebung enthalten ist. Ist dies der Fall, kann es sich nur um eine Fehlzuordnung handeln.

Bei der Kostenstrukturserhebung 1999 wurden alle Unternehmen mit mehr als 500 Beschäftigten vollständig erfasst. Die mittleren Unternehmen (250 bis 499 Beschäftigte) wurden zu 73% und die kleineren Unternehmen (20 bis 249 Beschäftigte) zu 38% erfasst. Demnach kann ein Datenangreifer bei der Suche nach großen Unternehmen von der Teilnahme des gesuchten Unternehmens an der Erhebung ausgehen. Dagegen muss er bei kleinen und mittleren Unternehmen die Unsicherheit in Kauf nehmen, dass das gesuchte Unternehmen nicht an der Erhebung teilgenommen hat.

Bei der Umsatzsteuerstatistik 2000 wurden alle Unternehmen, die Umsatzsteuer-Voranmeldungen abgeben, mit einem Jahresumsatz (ohne Umsatzsteuer) von über 32 500 DM (16 617 Euro) erfasst. Damit sind rund 2,9 Mill. Unternehmen enthalten. Unter diesen Bedingungen ist die Umsatzsteuerstatistik mit wenigen Ausnahmen eine Vollerhebung und der Datenangreifer besitzt nahezu gesicherte Kenntnis darüber, dass das gesuchte Unternehmen im Datenbestand enthalten ist.

### 1.3 Unsicherheit über die Richtigkeit von Zuordnungsversuchen

Im Zusammenhang mit Deanonymisierungsversuchen werden Informationen, die einem Datenangreifer zusätzlich zum anonymisierten Datensatz vorliegen oder beschaffbar sind und die zur Deanonymisierung eingesetzt werden können, als Zusatzwissen bezeichnet. Unter einer Reidentifikation wird verstanden, ein bestimmtes Zusatzwissen einem Datensatz der anonymisierten Erhebung eindeutig und richtig zuzuordnen. Auf diese Weise können diesem Datensatz direkte Identifikatoren (Name, Anschrift) zugewiesen werden. Ein Zuordnungsversuch kann in der Regel nur mittels der Überschneidungsmerkmale vorgenommen werden, also der Merkmale, über die Informationen sowohl im anonymisierten Zieldatensatz als auch im Zusatzwissen vorliegen.<sup>13)</sup> Die Reidentifikation (richtige und eindeutige Zuordnung eines gesuchten Merkmalsträgers) eröffnet einem Datenangreifer Kenntnisse über sensible Sachverhalte von ihm interessierenden Merkmalsträgern, die aus anderen Quellen nicht beschaffbar sind (Zielmerkmale).

Schaubild 2

Externe Daten (Zusatzwissen)		
Identifikatoren (Name, Anschrift)	Überschneidungsmerkmale (z. B. Umsatz)	
	Überschneidungsmerkmale (z. B. Umsatz)	Zielmerkmale (z. B. Kostenstrukturen)
Zieldaten (anonymisierte Daten)		

Ein Schutzeffekt von Anonymisierungsmaßnahmen besteht darin, die Verwendbarkeit der Überschneidungsmerkmale im Zieldatensatz für eine Zuordnung zu stören. Ein Zuordnungsversuch verursacht für einen Datenangreifer Aufwand und damit Kosten. Es ist wichtig, sich vor Augen zu halten, dass ein Datenangreifer vor dem Problem steht, korrekte von falschen Zuordnungen zu unterscheiden. Da er die Richtigkeit seiner vorgenommenen Zuordnungen nicht prüfen kann, wird ein rationaler Datenangreifer die Unsicherheit, ob eine von ihm angenommene Zuordnung zutrifft, bei der Bewertung seines Nutzens berücksichtigen.

Zuordnungsexperimente

	Angriffstechnik	Zusatzwissen	Zieldaten
Experiment 1	Massenfischzug	Originaldaten der Kostenstrukturserhebung	formal und durch verschiedene Arten der Mikroaggregation anonymisierte Daten der Kostenstrukturserhebung
Experiment 2	Massenfischzug	MARKUS-Datenbank	formal und durch verschiedene Arten der Mikroaggregation anonymisierte Daten der Kostenstrukturserhebung
Experiment 3	Massenfischzug	Originaldaten der Umsatzsteuerstatistik	formal und durch verschiedene Arten der Mikroaggregation anonymisierte Daten der Kostenstrukturserhebung
Experiment 4	Massenfischzug	Originaldaten der Kostenstrukturserhebung ergänzt um das Merkmal Rechtsform <sup>1)</sup>	formal, traditionell und durch verschiedene Arten der Mikroaggregation anonymisierte Daten der Umsatzsteuerstatistik
Experiment 5	Massenfischzug	MARKUS-Datenbank	durch Mikroaggregation anonymisierte probeanonymisierte Daten der Umsatzsteuerstatistik
Experiment 6	Einzelangriffe	Datenbankrecherche	formal und durch verschiedene Arten der Mikroaggregation anonymisierte Daten der Umsatzsteuerstatistik

1) In der Kostenstrukturserhebung ist das Merkmal „Rechtsform“ nicht erfasst. Für die Untersuchung wurde zu einem früheren Zeitpunkt als Ersatzlösung für noch nicht verfügbare externe Datenbanken ein Zusatzwissen erstellt, das aus den Merkmalen der Kostenstrukturserhebung und dem dazugespielten Merkmal Rechtsform besteht. Die Verwendung dieser Datei hat für die hier dargestellte Untersuchung den Vorteil, dass man die Zuordnungsexperimente besser miteinander vergleichen kann, da in allen drei Fällen die gleichen Überschneidungsmerkmale verwendet werden.

13) Ein – allerdings wenig überzeugender – Ansatz über Data Mining auch ohne Überschneidungsmerkmale Zusatzwissen einem Merkmalsträger zuzuordnen findet sich in Torra, V./Domingo-Ferrer, J./Torres, A.: "Data Mining Methods for Linking Data Coming from Several Sources", Invited Paper for Joint ECE/Eurostat work session on statistical data confidentiality, Luxemburg 2003; Internetseite: <http://neon.vb.cbs.nl/casc/default.htm>.

Aufgrund seines Rationalkalküls wird ein Datenangreifer von dem Versuch einer Reidentifikation absehen, wenn er davon ausgehen muss, dass der Anteil der falsch zuordenbaren an allen zuordenbaren Datensätzen über einer kritischen Schwelle liegt, sein Zuordnungsrisiko also zu hoch ist.

Im Folgenden soll das Zuordnungsrisiko eines Datenangreifers untersucht werden. Hierbei werden sechs unterschiedliche Experimente betrachtet. Drei davon beschäftigen sich in Abschnitt 1.3.1 mit (auf fünf unterschiedliche Weisen anonymisierten) Daten der Kostenstrukturerhebung und drei in Abschnitt 1.3.2 mit der Umsatzsteuerstatistik.

### 1.3.1 Zuordnungsversuche bei der Kostenstrukturerhebung

In den Experimenten 1 bis 3 werden fünf Anonymisierungsvarianten verwendet, die im Zwischenbericht des Projektes „Faktische Anonymisierung wirtschaftsstatistischer Einzeldaten“<sup>14)</sup> detailliert beschrieben sind:

- Formale Anonymisierung, das heißt Löschung direkter Identifikatoren wie Name und Anschrift (kurz: Formal)
- Mikroaggregation über 1 Gruppe (kurz: MA1G)
- Mikroaggregation über 8 Gruppen (MA8G)
- Mikroaggregation über 11 Gruppen (MA11G)
- Mikroaggregation über 30 Gruppen (MA30G)

Bei den Mikroaggregationsvarianten wurden zunächst die metrischen Merkmale in Gruppen unterteilt, wobei hoch korrelierte Merkmale zusammengruppiert wurden. Innerhalb der Gruppen wurden die Merkmale standardisiert und für jeden Merkmalsträger aufsummiert, sodass die Merkmalsträger nach diesen so genannten Z-scores geordnet werden können. Im nächsten Schritt wurden die Merkmalsträger mit dem größten bzw. kleinsten Z-score mit ihren zwei nächsten Nachbarn (bei mehrdimensionaler Mikroaggregation bezüglich der euklidischen Norm) zusammengelegt und bei den so entstandenen Tripeln merkmalsweise die Werte gemittelt. Mit den Anonymisierungsmaßnahmen der Mikroaggregation über 1 Gruppe und 30 Gruppen wurden die stärkste und die schwächste Variante der Mikroaggregation auf die Zieldaten angewendet. Während bei der schwächsten Variante jedes metrische Merkmal seine eigene Gruppe definiert, werden bei der stärksten Variante alle metrischen Merkmale zusammen gruppiert, sodass Tripel von Merkmalsträgern erzeugt werden, welche in allen metrischen Merkmalen übereinstimmen und sich lediglich in den kategorialen Merkmalen Wirtschaftszweigzuordnung und Regionalkennung unterscheiden können. Eine detaillierte Beschreibung der Gruppenbildung bei den Varianten MA8G (acht Gruppen von Merkmalen) und MA11G (elf Gruppen von Merkmalen) findet sich im Zwischenbericht des Projektes (siehe Fußnote 14, S. 68 f.).

Mit der Mikroaggregation über 1 Gruppe und der über 30 Gruppen wurden die stärkste und die schwächste Variante der Mikroaggregation als Anonymisierungsmaßnahmen auf die Zieldaten angewendet. Während bei der schwächsten Variante jedes metrische Merkmal seine eigene Gruppe definiert, werden bei der stärksten Variante alle metrischen Merkmale zusammen gruppiert, sodass Tripel von Merkmalsträgern erzeugt werden, welche in allen metrischen Merkmalen übereinstimmen und sich lediglich in den kategorialen Merkmalen *Wirtschaftszweigzuordnung* und *Regionalkennung* unterscheiden können.

#### Experiment 1

In diesem Experiment wird eine Obergrenze für das Reidentifikationsrisiko bestimmt. Mit den Originaldaten wird dem potenziellen Datenangreifer ein unrealistisch gutes Zusatzwissen unterstellt. Die Originaldaten repräsentieren eine perfekte externe Datenbank, zwischen deren Merkmalen und dem Zieldatensatz keine Dateninkompatibilitäten auftreten.

In einem solchen unrealistischen „Worst-Case“-Szenario stünden dem Datenangreifer sämtliche Originalwerte für die in der Kostenstrukturerhebung vorhandenen Merkmale als Überschneidungsmerkmale zur Verfügung. Im vorliegenden Fall werden jedoch vorrangig solche Kombinationen von Überschneidungsmerkmalen (und damit ein geringer Ausschnitt aus dem Kanon der Originalmerkmale) betrachtet, welche in realistischen Szenarien zu erwarten sind. Recherchen über externe Datenbanken haben ergeben, dass einem Datenangreifer als Überschneidungsmerkmale in der Regel eine Teilmenge der nachfolgenden fünf Merkmale zur Verfügung steht:

- Wirtschaftszweig auf Ebene der Zweisteller der Klassifikation der Wirtschaftszweige
- Siedlungsstruktureller Kreistyp BBR9<sup>15)</sup>
- Umsatz
- Zahl der Beschäftigten
- Aufwand für Forschung und Entwicklung (F&E)

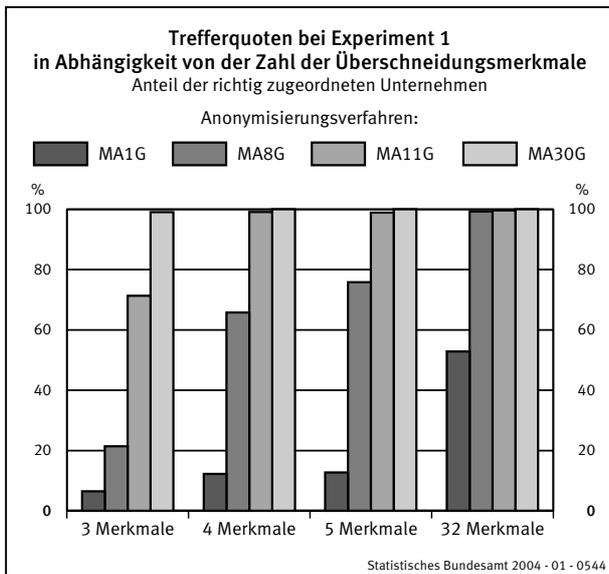
Schaubild 3 enthält die Trefferquoten des Massenfischzug-szenarios unter Verwendung von drei Überschneidungsmerkmalen (*Wirtschaftszweigzuordnung*, *Regionalkennung* und *Umsatz*), vier Überschneidungsmerkmalen (*Wirtschaftszweigzuordnung*, *Regionalkennung*, *Umsatz* und *Zahl der Beschäftigten*), fünf Überschneidungsmerkmalen (*Wirtschaftszweigzuordnung*, *Regionalkennung*, *Umsatz*, *Zahl der Beschäftigten* und *Aufwand für Forschung und Entwicklung*) und allen 32 Merkmalen der probeanonymisierten Kostenstrukturerhebung.

Erwartungsgemäß steigt das Reidentifikationsrisiko mit der Anzahl der Überschneidungsmerkmale. Selbst bei dem hier unterstellten guten Zusatzwissen ist ein beachtlicher Schutz

14) Ronning, G./Gnoss, R. u. a.: „Faktische Anonymisierung wirtschaftsstatistischer Einzeldaten“, Zwischenbericht an das Bundesministerium für Bildung und Forschung 2003.

15) Zur Definition der nichtadministrativen Gebietsschlüssel des Bundesamtes für Bauwesen und Raumordnung siehe Fußnote 11, sowie Lenz, R./Vorglimmer, D., Fußnote 8.

Schaubild 3

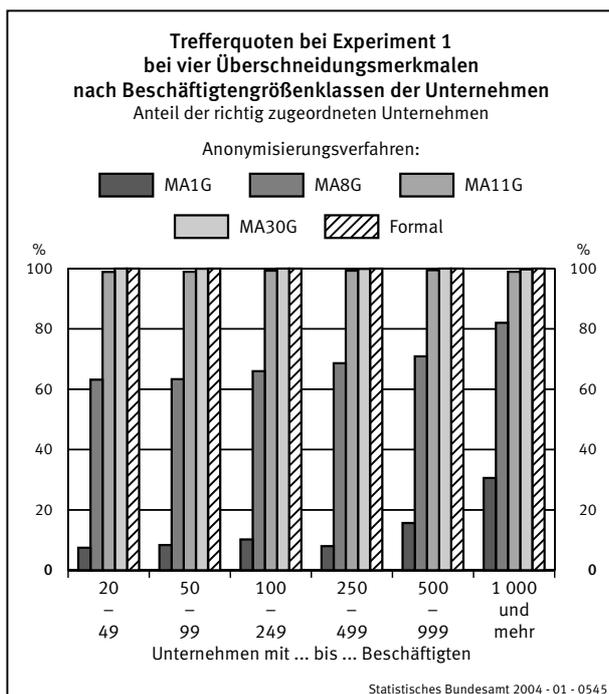


durch die Anonymisierungsvariante MA1G und im Falle von drei bzw. vier Überschneidungsmerkmalen auch durch die Variante MA8G erkennbar.

Im Weiteren werden die durch das Szenario mit vier Überschneidungsmerkmalen erzielten Ergebnisse genauer durchleuchtet, um einen direkten Vergleich mit den folgenden Experimenten 2 und 3 zu ermöglichen, in welchen maximal diese vier Merkmale als Überschneidungsmerkmale auftauchen werden. Das Schaubild 4 zeigt die Verteilung der richtigen Zuordnungen von Unternehmen nach Beschäftigtengrößenklassen.

Bestätigt wird hier die allgemeine Vermutung, dass – unabhängig von der Wahl des Anonymisierungsverfahrens – das

Schaubild 4



Reidentifikationsrisiko mit der Größe des betrachteten Unternehmens steigt; allerdings bietet die Anonymisierungsvariante MA1G auch hier noch einen beachtlichen Schutz.

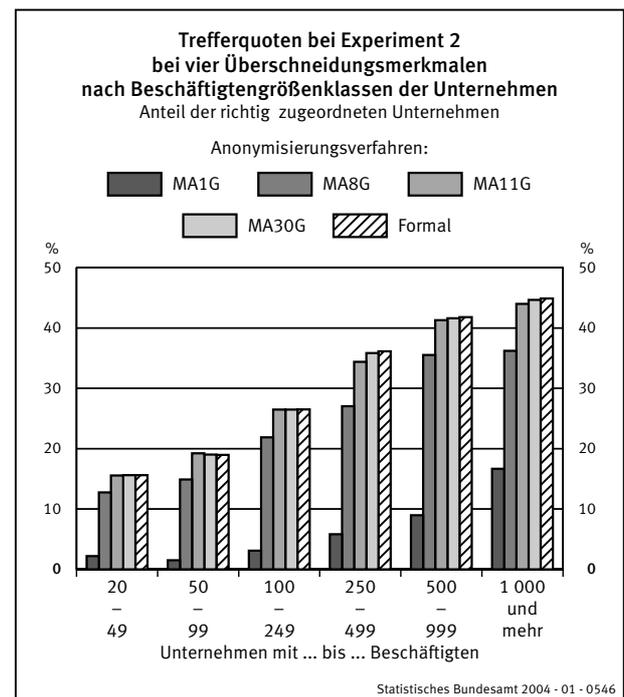
**Experiment 2**

Im zweiten Experiment wird ein Datenangriff unter realitätsnäheren Bedingungen simuliert. Unterstellt wird, dass dem Datenangreifer die kommerziell verfügbare MARKUS-Datenbank zur Verfügung steht.

Die Aussagekraft dieses Experiments ist als sehr hoch einzustufen, da die etwa 9 400 überprüfbar Unternehmen (nicht für alle Unternehmen in der Kostenstrukturerhebung konnten im Vorfeld der Untersuchung die zugehörigen MARKUS-Unternehmen ermittelt werden) die Größenklassenstruktur der Kostenstrukturerhebung repräsentativ abbilden.

Aus dem gewählten Zusatzwissen stehen dem Datenangreifer als Überschneidungsmerkmale die vier Merkmale *Wirtschaftszweigzuordnung*, *Regionalkennung*, *Umsatz* und *Zahl der Beschäftigten* zur Verfügung. Das Schaubild 5 enthält die Anteile der richtig zugeordneten Unternehmen untergliedert nach Beschäftigtengrößenklassen.

Schaubild 5



Auffallend ist zunächst der deutliche Rückgang des Reidentifikationsrisikos gegenüber dem vorherigen Experiment. Sogar bei der schwächsten Form der Anonymisierung (formale Anonymisierung) konnten selbst in der gefährdetsten Größenklasse von Unternehmen mit 1 000 Beschäftigten und mehr weit weniger als die Hälfte der Unternehmen richtig zugeordnet werden. Interessiert sich der Datenangreifer für kleine und mittlere Unternehmen (bis zu 249 Beschäftigte), so muss er mit einer Wahrscheinlichkeit von über 70% mit einer Falschzuordnung rechnen.

### Experiment 3

Im dritten Experiment mit der Kostenstrukturerhebung als Zieldatei werden die Daten der Umsatzsteuerstatistik als mögliches Zusatzwissen für den Datenangreifer verwendet. Dies ist kein wirklich realistisches Szenario, denn die Umsatzsteuerstatistik steht einem Datenangreifer nicht zur Verfügung. Es handelt sich dennoch um eine reale Datei des Zusatzwissens. Aus dem Datenmaterial der Umsatzsteuerstatistik waren etwa 37 000 Unternehmen mit entsprechender Wirtschaftsweangabe und Beschäftigtenzahl für das Experiment relevant, von denen etwa 9 300 Unternehmen (also eine mit Experiment 2 vergleichbare Größenordnung) auf die Richtigkeit der Zuordnung überprüft werden konnten. Dieses Experiment zeigt, dass auch beim Versuch der Verknüpfung zweier amtlicher Erhebungen bei Unkenntnis direkter Identifikatoren – wie zum Beispiel Unternehmensregisternummer – Probleme auftreten können. Verglichen mit Experiment 2 müsste der Datenangreifer hier auf das Überschneidungsmerkmal *Zahl der Beschäftigten* verzichten, da dieses in der Umsatzsteuerstatistik nicht enthalten

ist. Das Schaubild 6 zeigt entsprechend zu den beiden vorherigen Experimenten die Verteilung der richtigen Zuordnungen von Unternehmen nach Beschäftigtengrößenklassen.

Anders als in den beiden vorausgegangenen Experimenten sind hier die Trefferquoten bei den mittleren Unternehmen (50 bis 249 Beschäftigte) sehr ähnlich. Insgesamt ergeben sich bei den stärkeren Anonymisierungsvarianten MA1G und MA8G bei diesem Szenario geringere Trefferquoten als bei dem in Experiment 2 untersuchten realistischen Szenario. Bei den schwächeren Anonymisierungsvarianten MA11G, MA30G und formale Anonymisierung ergeben sich bei den kleinen Unternehmen (20 bis 49 Beschäftigte) höhere Trefferquoten als beim vorherigen Experiment. Auf der anderen Seite waren Unternehmen mit mehr als 249 Beschäftigten etwas schwieriger zuzuordnen (allerdings ist zu beachten, dass bei diesem Experiment ein Überschneidungsmerkmal weniger zur Verfügung steht).

#### 1.3.2 Zuordnungsversuche bei der Umsatzsteuerstatistik

Die folgenden drei Experimente beschäftigen sich mit Zuordnungsversuchen bei der Umsatzsteuerstatistik. Dabei handelt es sich bei den Experimenten 4 und 5 – wie bei den ersten drei Experimenten bei der Kostenstrukturerhebung – um Massenfischzüge, während das Experiment 6 Einzelanriffe simuliert.

#### Experiment 4

In Experiment 4 wurden als Zusatzwissen die Originaldaten der Kostenstrukturerhebung gewählt und versucht, Unternehmen aus dieser Datei Einheiten der Umsatzsteuerstatistik als Zieldatei zuzuordnen. Das Zusatzwissen für überprüfbare Zuordnungen bestand aus knapp 9 300 Unternehmen des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden mit 20 Beschäftigten und mehr. Von den etwa 2,9 Mill. Unternehmen der Umsatzsteuerstatistik kamen wie in Experiment 3 etwa 37 000 Unternehmen als Zielunternehmen mit entsprechender Wirtschaftsweangabe und Beschäftigtenzahl in Frage. Diese werden vollständig in der Umsatzsteuerstatistik erfasst, sodass auf Seiten des Datenangreifers von der Teilnahmekennntnis ausgegangen werden konnte (siehe Abschnitt 1.2).

Diese Versuche haben als Überschneidungsmerkmale zwischen Zusatzwissen und Zieldaten die Abteilungen (Zwei-

Schaubild 6

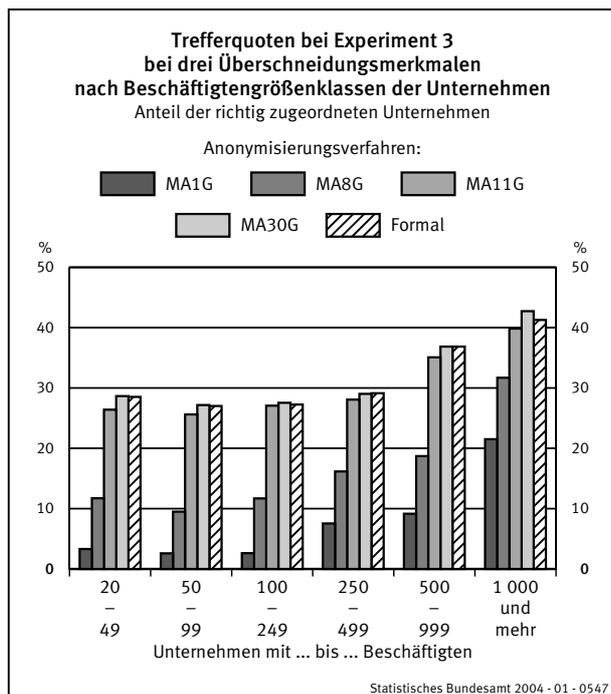


Tabelle 1: Trefferquoten<sup>1)</sup> bei Experiment 4 insgesamt und nach Umsatzgrößenklassen der Unternehmen Prozent

Anonymisierungsverfahren	Gebiets-schlüssel	Insgesamt	Unternehmen mit Umsätzen <sup>2)</sup> von ... bis unter ... EUR				
			16 617 – 1 Mill.	1 Mill. – 10 Mill.	10 Mill. – 100 Mill.	100 Mill. – 1 Mrd.	1 Mrd. und mehr
Formal anonymisiert ...	BBR9	31,7	13,2	27,6	36,0	51,8	54,5
	BBR3	22,6	9,3	18,5	26,9	38,6	43,2
MA21G .....	BBR9	31,7	13,2	27,6	35,8	51,6	54,5
	BBR3	22,6	9,3	18,8	26,7	37,8	47,7
MA1G .....	BBR9	11,2	5,7	6,4	13,8	33,6	59,1
	BBR3	6,0	2,7	3,0	7,7	19,0	47,7
Traditionell .....	BBR9	30,1	12,7	26,5	34,8	45,1	13,6
	BBR3	22,4	9,3	18,7	26,5	38,3	15,9

1) Anteil der richtig zugeordneten Unternehmen. – 2) Lieferungen und Leistungen.

steller) der Klassifikation der Wirtschaftszweige, die Rechtsform<sup>16)</sup>, den Umsatz und einen nichtadministrativen Gebietsschlüssel [siedlungsstruktureller Kreistyp (BBR9) und alternativ der siedlungsstrukturelle Regionsgrundtyp (BBR3)<sup>17)</sup>] verwendet. Dem Datenangreifer steht damit wie bereits in Experiment 3 mit dem Umsatz nur ein metrisches Merkmal zur Verfügung. Er kann dafür mit der Rechtsform ein zusätzliches kategoriales Überschneidungsmerkmal einsetzen, welches in den bisherigen Experimenten keine Rolle spielte. Um die Wirkungen der angewandten Geheimhaltungsmethoden einschätzen zu können, waren die Zieldaten im ersten Schritt nur formal anonymisiert (d. h. Löschung der direkten Identifikatoren wie Name und Adresse). Anschließend wurden die Merkmalsträger der Zieldaten zum einen mit zwei Varianten der Mikroaggregation und zum anderen traditionell anonymisiert. Bei den eingesetzten Varianten der Mikroaggregation wurden im einen Fall sämtliche metrischen Merkmale zusammen gruppiert (MA1G), im anderen Fall wurde jedes der 21 metrischen Merkmale separat mikroaggregiert (MA21G). Damit wurden für die Umsatzsteuerstatistik die stärkste und die schwächste Variante der Mikroaggregation eingesetzt. Bei der traditionellen Anonymisierung wurden u. a. die Verfahren "Topcoding", „Kürzung der Klassifizierung“ und „Umkodierung“ verwendet.<sup>18)</sup>

Tabelle 1 enthält die Trefferquoten für die ausgewählten acht Simulationen innerhalb dieses Experimentes. Die Trefferquoten weisen eine ähnliche Struktur auf wie bei den zuvor behandelten Zuordnungsexperimenten. Je größer ein Unternehmen, desto höher ist die Wahrscheinlichkeit, dieses Unternehmen richtig zuzuordnen. Zumindest auf Basis des BBR9 müssen Unternehmen mit mehr als 100 Mill. Euro Umsatz als stark gefährdet angesehen werden, während die kleineren Unternehmen bereits bei einer formalen Anonymisierung sehr gut geschützt sind.

Wie bereits erwähnt kann ein Datenangreifer die Korrektheit seiner Zuordnung nicht überprüfen. Wird in der Tabelle eine Trefferquote von 20% angegeben, heißt dies nicht nur, dass ein Datenangreifer von fünf Unternehmen lediglich eines richtig zuordnen konnte. Er weiß darüber hinaus nicht, welches der Unternehmen das richtig zugeordnete ist. Je größer ein Unternehmen, desto höher die Wahrscheinlichkeit seiner richtigen Zuordnung. Aber auch dies hat – wie Tabelle 1 zeigt – Grenzen.

Neben der hohen Besetzungsdichte in der Klasse der Unternehmen mit 20 bis 49 Beschäftigten waren die in Abschnitt 1.1 besprochenen Dateninkompatibilitäten zwischen Zusatzwissen und Zieldaten der Hauptgrund für fehlerhafte Zuordnungen.

Während die Variante MA1G über alle Umsatzgrößenklassen die Zuordnungshäufigkeit gegenüber den formal anonymisierten Daten senkt, reduziert die traditionelle Anonymisierung, deren Hauptziel der Schutz der großen Unternehmen ist, diese – wie zuvor erwartet – nur bei den großen Unternehmen. Die Variante MA21G liefert keinen zusätzlichen Schutz vor richtigen Zuordnungen.

**Experiment 5**

In diesem Szenario wird wie in Experiment 2 mit der MARKUS-Datenbank eine kostenpflichtige, aber ansonsten frei zugängliche Datenbank als Zusatzwissen verwendet. Damit rückt auch dieses Szenario stärker an die Realität. Als Zieldatei dient die Probeanonymisierung MA21G der Umsatzsteuerstatistik unter Verwendung der Regionalkennung BBR3. Das Zusatzwissen besteht in diesem Fall aus rund 6 300 Unternehmen. Teilnahmekennntnis kann wie beim ersten Experiment vorausgesetzt werden (s. o.). Eine Übersicht über die Trefferquoten findet sich in Tabelle 2.

Auffallend sind die geringen Trefferquoten, die mit diesem Zusatzwissen erzielt werden konnten. So ist die Gesamttrefferquote bei der Konstellation MA21G/BBR3 von 22,6% im vierten Experiment auf 5% gefallen. Selbst bei den großen Unternehmen bleiben die Trefferquoten deutlich hinter den Erwartungen zurück. Der Grund dafür ist in den wesentlich stärkeren Abweichungen zwischen Zieldaten und Zusatzwissen zu suchen (siehe Abschnitt 1.1).

**Experiment 6**

Unternehmen, die in ihrer Branche eine führende Rolle spielen, werden oftmals im Hinblick auf Reidentifikationsversuche als besonders gefährdet angesehen. Um dies näher zu analysieren, wurde in diesem Alternativszenario die Angriffstechnik geändert: Die bisherige Technik des Massenfischzuges wurde durch den Einzelangriff ersetzt. Der Angriffsversuch konzentrierte sich dabei auf fünf Unternehmen, die in vier unterschiedlichen Branchen die marktfüh-

Tabelle 2: Richtig zugeordnete Unternehmen bei Experiment 5 insgesamt und nach Größenklassen

Anonymisierungsverfahren	Gebiets-schlüssel	Insgesamt	Unternehmen mit Umsätzen <sup>1)</sup> von ... bis unter ... EUR				
			16 617 – 1 Mill.	1 Mill. – 10 Mill.	10 Mill. – 100 Mill.	100 Mill. – 1 Mrd.	1 Mrd. und mehr
MA21G .....	BBR3	319	Anzahl				
			2	55	176	77	9
			%				
MA21G .....	BBR3	5,0	5,4	1,7	4,7	17,4	31,0

1) Lieferungen und Leistungen.

16) Die Rechtsform hatte die Ausprägungen „Personengesellschaft“, „Kapitalgesellschaft“, „Erwerbs- und Wirtschaftsgenossenschaften und Betriebe gewerblicher Art öffentlichen Rechts“ und „sonstige Rechtsformen“.  
 17) Bei der Umsatzsteuerstatistik ist im Projekt noch keine Entscheidung zwischen dem BBR9 und BBR3 gefallen. Daher werden im Folgenden zur Analyse beide nichtadministrativen Gebiets-schlüssel verwendet.  
 18) Zur Probeanonymisierung der Umsatzsteuer siehe Fußnote 14, sowie Lenz, R./Vorgirmler, D., Fußnote 8.

rende Position innehaben (in einer Branche wurden die beiden größten Unternehmen als zu suchende Merkmalsträger gewählt). Als Zusatzwissen standen die Rechtsform, die Wirtschaftszweignummer als Zweisteller und die Regionalerkennung zur Verfügung. Dabei wurden sowohl der Gebietschlüssel BBR9 als auch der Gebietschlüssel BBR3 verwendet, wobei sich die Ergebnisse nicht unterschieden, sodass im Weiteren auf die Differenzierung verzichtet werden kann. Das Wissen, die gemessen am Umsatz größten Unternehmen zu suchen, konnte als zusätzliches Überschneidungsmerkmal bei der Suche eingesetzt werden.

Jedes der fünf gesuchten Unternehmen konnte sowohl bei den formal anonymisierten Daten als auch bei den mit der Mikroaggregation 21G anonymisierten Unternehmen richtig zugeordnet werden. Bei den mit der Mikroaggregation 1G anonymisierten Merkmalsträgern waren noch zwei Unternehmen reidentifizierbar. Die traditionelle Anonymisierung erwies sich in diesem Experiment am effektivsten, nur ein Unternehmen konnte korrekt dem Zusatzwissen zugeordnet werden.

## 2 Unsicherheit über die Qualität von richtig zugeordneten Informationen: Ebene der Einzelmerkmale

Eine erfolgreiche Zuordnung führt nicht unbedingt zu einem Nutzen des Datenangreifers und damit zu einer Verletzung der faktischen Anonymität. Vielmehr ergibt sich der Nutzen erst aus den „brauchbaren“ Informationen, die ein Datenangreifer bei einer erfolgreichen Reidentifikation gewinnen kann. Im Gegensatz zu Wissenschaftlern, für deren Analyse nicht unbedingt die Richtigkeit der Einzelwerte einer Datei, sondern die aus den Einzelwerten errechneten Größen entscheidend sind, hängt der Nutzen für Datenangreifer von der Qualität konkreter Einzelwerte ab.

Im Folgenden werden Unsicherheiten untersucht, die auf der Ebene der einzelnen Merkmale einer Datei bestehen. Wird durch einen Datenangriff ein Datensatz erfolgreich zugeordnet (identifiziert), so gewinnt der Datenangreifer auf der Merkmalebene Informationen über alle weiteren Zielmerkmale im Datensatz. Daneben liefert eine Reidentifikation auch neue Informationen über die Ausprägungen der Überschneidungsmerkmale, die vor allem bei metrischen Größen von den jeweiligen Ausprägungen im Zusatzwissen abweichen können. Der Datenangreifer muss für diese Merkmale selbst entscheiden, welche Werte er für die zuverlässigeren hält. Wird im Folgenden daher von richtig zugeordneten Informationen gesprochen, so sind alle Merkmale im jeweils reidentifizierten Datensatz gemeint.

Die Brauchbarkeit einer Information ist im Falle einer richtigen Zuordnung nicht von vornherein gesichert, da die Merkmalsausprägungen beim Einsatz datenverändernder Verfahren vom Originalwert abweichen können. Dies spielt für die Wahrung der faktischen Anonymität eine wichtige Rolle. Brauchbar ist eine Information nur dann, wenn der gefun-

dene Wert dem „wahren“ Wert<sup>19)</sup> entspricht oder diesem in einem bestimmten Maß ähnelt. Ab einer gewissen Abweichung wird ein Datenangreifer keinen Nutzen mehr aus einer richtig zugeordneten Information erzielen können. Die Abweichungsgrenze, ab welcher ein Wert als unbrauchbar gilt, wird im Folgenden „Nützlichkeitsschwelle“ genannt. In den statistischen Ämtern müssen vernünftige Werte für diese Schwelle festgelegt werden, mit deren Hilfe dann das Risiko eines Datenangreifers bestimmt werden kann. Im Rahmen der Anonymisierung von Tabellen wurden vergleichbare Schwellen bereits festgelegt.

Auch an dieser Stelle ist es wieder wichtig zu beachten: Ob eine korrekte Reidentifikation brauchbare Informationen liefert – also die Abweichung des gefundenen Wertes zu dem entsprechenden Originalwert innerhalb bestimmter Grenzen liegt – kann von einem Datenangreifer nicht geprüft werden. Er kann allenfalls eine Wahrscheinlichkeit für die Brauchbarkeit der gefundenen Informationen abschätzen. Liegt diese Wahrscheinlichkeit unterhalb einer gewissen Schwelle, wird das Risiko, eine unbrauchbare Information mit einer brauchbaren Information zu verwechseln, als zu hoch angesehen werden. Daraus folgt, dass für einen Datenangreifer eine Information wertlos werden kann, unabhängig davon, ob die gefundene Information innerhalb oder außerhalb bestimmter Abweichungsgrenzen liegt. Selbst wenn durch Reidentifikation brauchbare Informationen zugeordnet werden können, muss damit nicht zwingend eine Verletzung der faktischen Anonymität im oben geschilderten Sinne einhergehen.

### 2.1 Qualität von richtig zugeordneten Informationen in der Kostenstrukturerhebung

#### Experiment 1: Abweichungsanalyse auf Basis der Zuordnung zwischen Originaldaten und probe-anonymisierten Daten der Kostenstrukturerhebung

Tabelle 3 enthält die Anteile an nützlichen Informationen über die gefundenen Unternehmen insgesamt und nach Größenklassen der Beschäftigten verteilt bei einer 5%-Nützlichkeitsschwelle. Das bedeutet, dass für die richtig zugeordneten Merkmalsträger der Anteil aller durch den Daten-

Tabelle 3: Anteil nützlicher Informationen bei Experiment 1 insgesamt und nach Beschäftigtengrößenklassen der Unternehmen  
Prozent

Anonymisierungsverfahren	Insgesamt	Unternehmen mit ... bis ... Beschäftigten					
		20 – 49	50 – 99	100 – 249	250 – 499	500 – 999	1 000 und mehr
Formal anonymisiert	100	100	100	100	100	100	100
MA30G	100	100	100	100	99	99	99
MA11G	82	87	86	82	77	74	65
MA8G	58	62	61	59	56	54	50
MA1G	32	36	35	31	29	27	23

19) Als „wahrer“ Wert wird die Originalausprägung in der amtlichen Erhebung definiert.

angriff gefundenen Informationen (Einzelwerte) bestimmt wird, welche um weniger als 5% von ihrem zugehörigen Originalwert abweichen.

Wie zu erwarten, liegen nahezu alle in der Datei MA30G gefundenen Informationen im Bereich des verwendeten Nützlichkeitskorridors, da das zugrunde liegende Anonymisierungsverfahren die Originaldaten nur geringfügig ändert. Anders sieht dies bei den Varianten MA11G, MA8G und MA1G aus. Man erkennt leicht, welche Wirkung diese Anonymisierungsvarianten auf den Nutzen einer gefundenen Information für einen Datenangreifer haben. Bei den kleinen Unternehmen mit 20 bis 49 Beschäftigten liegt der Anteil nützlicher Informationen zwischen 36% (MA1G) und 87% (MA11G), bei den sehr großen Unternehmen zwischen 23% (MA1G) und 65% (MA11G). Zu erkennen ist, dass die Mikroaggregationsverfahren bei den großen Unternehmen stärker wirken, da deren Merkmalsausprägungen stärker voneinander abweichen. Daher ist gerade in solchen Bereichen eine spätere Analyse der Nützlichkeit der zugeordneten Werte für einen Datenangreifer von großer Bedeutung, um die Wirkung(sweise) der Verfahren angemessen bewerten zu können.

**Experiment 2: Abweichungsanalyse auf Basis der Zuordnung zwischen MARKUS-Datenbank und probe-anonymisierten Daten der Kostenstrukturerhebung**

Analog zum ersten Experiment enthält die Tabelle 4 die Anteile an nützlichen Informationen über die gefundenen Unternehmen insgesamt und nach Größenklassen der Beschäftigten verteilt bei einer 5%-Nützlichkeitschwelle für das zweite Experiment.

Die Ergebnisse sind im Wesentlichen mit denen des ersten Experiments vergleichbar. Bei den Varianten MA1G und MA8G ist der Anteil nützlicher Informationen etwas geringer als im ersten Experiment.

Tabelle 4: Anteil nützlicher Informationen bei Experiment 2 insgesamt und nach Beschäftigtengrößenklassen der Unternehmen  
Prozent

Anonymisierungsverfahren	Ins-gesamt	Unternehmen mit ... bis ... Beschäftigten					
		20 – 49	50 – 99	100 – 249	250 – 499	500 – 999	1 000 und mehr
Formal anonymisiert .	100	100	100	100	100	100	100
MA30G .....	99	99	99	99	99	99	98
MA11G .....	82	88	84	81	76	73	65
MA8G .....	55	60	60	57	54	52	45
MA1G .....	28	30	29	26	25	22	15

**Experiment 3: Abweichungsanalyse auf Basis der Zuordnung zwischen Originaldaten der Umsatzsteuerstatistik und probeanonymisierten Daten der Kostenstrukturerhebung**

Ähnlich wie bei den vorhergehenden Experimenten fallen die Ergebnisse der Abweichungsanalyse beim Experiment 3 aus, wie Tabelle 5 zeigt. Mit wachsender Anzahl an Re-

identifikationen nähert sich die relative Häufigkeit der 5%-Abweichungen zwischen den anonymisierten Werten und den Originalwerten der gefundenen Unternehmen der relativen Gesamtabweichung zwischen Originaldaten und anonymisierten Daten an. Dies ist ein Hinweis darauf, dass ein potenzieller Datenangreifer sich möglicherweise eine erste Näherungsgröße für den Nutzen der anonymisierten (vertraulichen) Zieldaten verschaffen kann, indem er sein Zusatzwissen anonymisiert und eine Abweichungsanalyse zwischen anonymisiertem und originalelem Zusatzwissen durchführt. Bei einer stärkeren Anonymisierung (und damit verbundener geringerer Anzahl an Reidentifikationen) ist diese Näherung schlechter, was bei der Betrachtung der Variante MA1G deutlich wird. Hier liegt in Experiment 3 der Anteil gefundener nützlicher Informationen systematisch höher als in den Experimenten 1 und 2.

Tabelle 5: Anteil nützlicher Informationen bei Experiment 3 insgesamt und nach Beschäftigtengrößenklassen der Unternehmen  
Prozent

Anonymisierungsverfahren	Ins-gesamt	Unternehmen mit ... bis ... Beschäftigten					
		20 – 49	50 – 99	100 – 249	250 – 499	500 – 999	1 000 und mehr
Formal anonymisiert .	100	100	100	100	100	100	100
MA30G .....	99	99	99	99	98	98	98
MA11G .....	81	86	83	80	78	72	69
MA8G .....	57	61	59	57	53	50	31
MA1G .....	39	45	42	38	36	34	29

**2.2 Qualität von richtig zugeordneten Informationen in der Umsatzsteuerstatistik**

**Experiment 4: Abweichungsanalyse auf Basis der Zuordnung zwischen Zusatzwissen und probe-anonymisierten Daten der Umsatzsteuerstatistik**

Tabelle 6 enthält die Anteile an nützlichen Informationen über die gefundenen Unternehmen insgesamt und nach Größenklassen des Umsatzes aufgeteilt (bei einer 5%- bzw. 10%-Nützlichkeitschwelle). Leicht einsichtig ist, dass bei einer rein formalen Anonymisierung sämtliche Informationen als nützlich erhalten bleiben und dass demgegenüber bei der Anonymisierungsvariante MA1G nur noch ein geringer Anteil der enthaltenen Informationen als für einen Datenangreifer nützlich deklariert werden kann. Interessanter sind die Fälle MA21G und traditionelle Anonymisierung. Diese lassen einerseits die Anteile nützlicher Informationen bei den kleinen und mittleren Unternehmen praktisch unverändert. Andererseits reduzieren sie die für einen Datenangreifer nützlichen Informationen bei den stärker gefährdeten großen Unternehmen zum Teil erheblich, eine Eigenschaft, die dem Gedanken, dass die höher gefährdeten Merkmalsträger stärker geschützt werden müssen, entgegenkommt. Des Weiteren ist ersichtlich, dass in der Regel der Detaillierungsgrad des Regionalmerkmals (BBR3 oder BBR9) kaum Einfluss auf die Nützlichkeitsanalyse hat.

Tabelle 6: Anteil nützlicher Informationen bei Experiment 4 insgesamt und nach Umsatzgrößenklassen der Unternehmen  
Prozent

Anonymisierungsverfahren	Gebiets-schlüssel	Insgesamt	Unternehmen mit Umsätzen <sup>1)</sup> von ... bis unter ... EUR				
			16 617 – 1 Mill.	1 Mill. – 10 Mill.	10 Mill. – 100 Mill.	100 Mill. – 1 Mrd.	1 Mrd. und mehr
Nützlichkeitschwelle 5%							
Formal anonymisiert ..	BBR9	100	100	100	100	100	100
	BBR3	100	100	100	100	100	100
MA21G .....	BBR9	99,3	98,3	100	99,9	96,2	68,7
	BBR3	99,3	94,0	100	100	96,6	70,0
MA1G .....	BBR9	19,7	16,5	21,8	21,3	14,8	11,1
	BBR3	20,1	17,6	23,9	22,1	14,1	6,5
Traditionell anonymisiert .....	BBR9	99,7	100	100	100	97,2	42,4
	BBR3	99,5	100	100	100	96,1	24,5
Nützlichkeitschwelle 10%							
Formal anonymisiert ..	BBR9	100	100	100	100	100	100
	BBR3	100	100	100	100	100	100
MA21G .....	BBR9	99,6	98,3	100	100	98,0	77,2
	BBR3	99,6	94,0	100	100	98,4	78,0
MA1G .....	BBR9	36,3	31,5	39,4	38,4	29,8	18,6
	BBR3	36,9	33,3	40,6	39,9	29,4	13,8
Traditionell anonymisiert .....	BBR9	99,7	100	100	100	97,5	45,7
	BBR3	99,5	100	100	100	96,5	25,5

1) Lieferungen und Leistungen.

### Experiment 5: Abweichungsanalyse auf Basis der Zuordnung zwischen MARKUS-Datenbank und probe-anonymisierten Daten der Umsatzsteuerstatistik

Wie die Tabelle 7 zeigt, unterscheiden sich die Anteile an nützlichen Informationen für die gefundenen Unternehmen nicht wesentlich gegenüber der analogen Simulation im vierten Experiment (MA21G/BBR3). Insgesamt ist der Anteil an nützlichen Informationen mit über 97% sehr hoch. Lediglich bei den sehr großen Unternehmen mit mehr als 1 Mrd. Euro Umsatz ist ein Rückgang des Anteils an nützlichen Informationen zu beobachten.

### Experiment 6: Abweichungsanalyse auf Basis ausgewählter Einzelangriffe

Die Zielunternehmen des Einzelangriffes waren als Branchenmarktführer ausnahmslos sehr große Unternehmen, bei denen sich die Werteausprägungen aufgrund der Anonymisierung sehr stark verändern. Daher ist die Nützlichkeit der enthüllten Informationen bei den anonymisierten Dateien sehr eingeschränkt. Dies gilt besonders bei MA1G, aber auch bei der MA21G und der traditionellen Anonymisierung. Ein Datenangreifer geht demnach bei Reidentifikationsversuchen bei sehr großen Marktführern ein erhöhtes Risiko ein, unbrauchbare Werte zu enthüllen.

Tabelle 8: Trefferquoten und Anteil nützlicher Informationen beim Einzelangriff (Experiment 6)  
Prozent

Anonymisierungsverfahren	Trefferquoten	Anteil nützlicher Informationen	
		Nützlichkeitschwelle	
		5%	10%
Formal anonymisiert .....	100	100	100
MA21G .....	100	50,5	63,6
MA1G .....	40	2,5	5,0
Traditionell anonymisiert ..	20	64,7	70,6

## 3 Zusammenführung zu einem Maß für die faktische Anonymität: Enthüllungsrisiko

In den beiden vorangegangenen Kapiteln wurden Zuordnungsexperimente für zwei Statistiken ausgewertet. Es wurde zum einen der Anteil richtiger Zuordnungen und zum anderen der Anteil nützlicher Informationen innerhalb der richtigen Zuordnungen berechnet. Die Ergebnisse haben gezeigt, dass auch eine richtige Zuordnung eines Merkmalsträgers  $m$  einen für den potenziellen Datenangreifer

Tabelle 7: Anteil nützlicher Informationen bei Experiment 5 insgesamt und nach Umsatzgrößenklassen der Unternehmen  
Prozent

Anonymisierungsverfahren	Gebiets-schlüssel	Insgesamt	Unternehmen mit Umsätzen <sup>1)</sup> von ... bis unter ... EUR				
			16 617 – 1 Mill.	1 Mill. – 10 Mill.	10 Mill. – 100 Mill.	100 Mill. – 1 Mrd.	1 Mrd. und mehr
Nützlichkeitschwelle 5%							
MA21G .....	BBR3	97,6	100	100	99,9	93,9	64,3
Nützlichkeitschwelle 10%							
MA21G .....	BBR3	98,6	100	100	100	96,8	73,9

1) Lieferungen und Leistungen.

erfolgslosen Angriffsversuch darstellen kann, nämlich wenn der ihn interessierende Einzelwert (bzw. die ihn interessierende Information) um einen festgesetzten Schwellenwert  $\gamma$  (in den Experimenten  $\gamma=0,05$  und  $\gamma=0,1$ ) von dem tatsächlichen Originalwert abweicht. Nun werden die Elemente des Schutzes zu einem Gesamtmaß zusammengeführt: Die Wahrscheinlichkeit, dass ein Datenangreifer einen Einzelwert  $w'$ , welcher um weniger als  $\gamma$  von seinem Originalwert  $w$  relativ abweicht, richtig zuordnet, werde mit  $P_\gamma(w \text{ enthüllt})$  bezeichnet und im Folgenden *Enthüllungsrisiko* genannt. In dem von Höhne/Sturm/Vorgrimler vorgestellten Schutzwirkungskonzept ergibt sich der Nutzen einer erfolgreichen Zuordnung aus den brauchbaren Informationen, die ein Datenangreifer bei einer erfolgreichen Reidentifikation enthüllen kann. Auf Merkmalsebene wird untersucht, in welchem Umfang ein durchgeführter Datenangriff auf einen Einzelwert (bzw. eine Information) als erfolgreich einzustufen ist, also die beiden Ereignisse „erfolgreiche Zuordnung der Merkmalsträger“ und „Zuordnung einer brauchbaren Information“ gleichzeitig auftreten. Als Formalisierung dieser Abschätzung erhalten wir daher die Wahrscheinlichkeit einer Enthüllung:

$$P_\gamma(w \text{ enthüllt}) = P(m \text{ erfolgreich zugeordnet}) * P_\gamma(w' \text{ weicht weniger als } \gamma \text{ von } w \text{ ab} \mid m \text{ erfolgreich zugeordnet}),$$

wobei der erste Faktor durch den Anteil der erfolgreichen Zuordnungen geschätzt werden kann, wie dies in den Abschnitten 1.3.1 und 1.3.2 für die beiden betrachteten Statistiken in den Experimenten 1 bis 6 geschehen ist. Der zweite Faktor lässt sich als bedingte Wahrscheinlichkeit, eine brauchbare Information (auf Merkmalsebene) zu finden, bei einer gegebenen erfolgreichen Zuordnung (auf Merkmalsträgerebene). In Kapitel 2 wurde diese Wahrscheinlichkeit durch den Anteil brauchbarer Informationen innerhalb der erfolgreich zugeordneten Merkmalsträger bei den Experimenten 1 bis 6 geschätzt. Insgesamt erhält man demnach einen Schätzer  $\hat{P}_\gamma(w \text{ enthüllt})$  für das Enthüllungsrisiko  $P_\gamma(w \text{ enthüllt})$ . Ein risikoaverser Datenangreifer wird ab einer bestimmten Wahrscheinlichkeit, aus einem Enthüllungsversuch für ihn unbrauchbare Werte zu erhalten, von einem Datenangriff absehen. Da ihm die Informationen fehlen, um reidentifizierte von nicht reidentifizierten Merkmalsträgern und brauchbare von unbrauchbaren Werten zu unterscheiden, bedeutet das, dass eine Datei als faktisch anonym eingestuft werden kann, auch wenn Merkmalsträger dieser Datei richtig zuordenbar sind *und* dabei brauchbare Einzelwerte enthüllt werden können. Die formale Bedingung für faktische Anonymität lautet demnach, dass eine vorgegebene obere Schwelle  $\tau$  für das Enthüllungsrisiko unterschritten werden muss, das heißt es muss

$$\hat{P}_\gamma(w \text{ enthüllt}) < \tau$$

gelten. Wird diese von den Datenhaltern gewissenhaft von vornherein festgelegte obere Schwelle nicht überschritten, so kann die zugrunde liegende Datei als faktisch anonym eingestuft werden. Je nach betrachteter Statistik kann diese Schwelle mit unterschiedlichen Werten angesetzt werden. Im Abschnitt 3.1 wird das Maß mit einem Vorschlag für eine geeignete Schwelle für die Kostenstrukturerhebung auf Basis der drei durchgeführten Experimente vor-

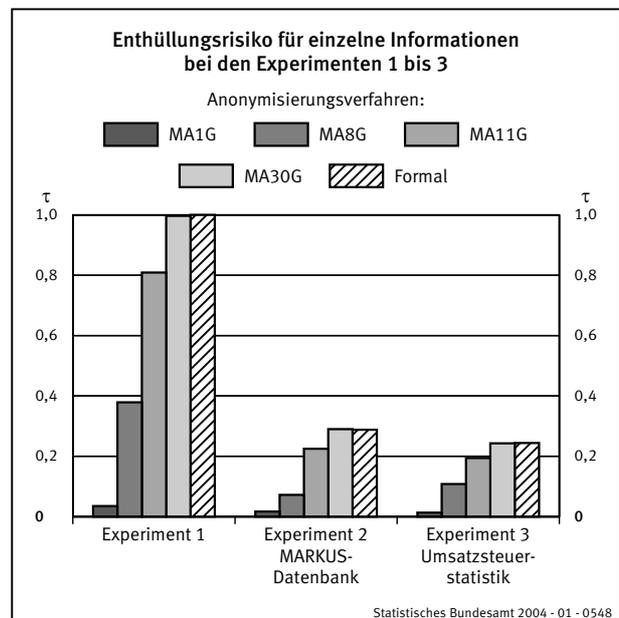
gestellt. Dagegen behandelt Abschnitt 3.2 die drei Experimente für die Umsatzsteuerstatistik und schlägt für diese eine Schwelle vor.

### 3.1 Enthüllungsrisiko in der Kostenstrukturerhebung

#### Zu den Experimenten 1 bis 3:

Wir berücksichtigen nun wie beschrieben die Kombination aus den Zuordnungsexperimenten der Merkmalsträger (Unternehmen/Betriebe) und der Nutzenuntersuchung der durch den Datenangreifer gefundenen Einzelinformationen. Das Schaubild 7 zeigt die berechneten Risiken der Enthüllung einer Einzelinformation für die Experimente 1 bis 3 für jeweils fünf Varianten der Probeanonymisierung.

Schaubild 7

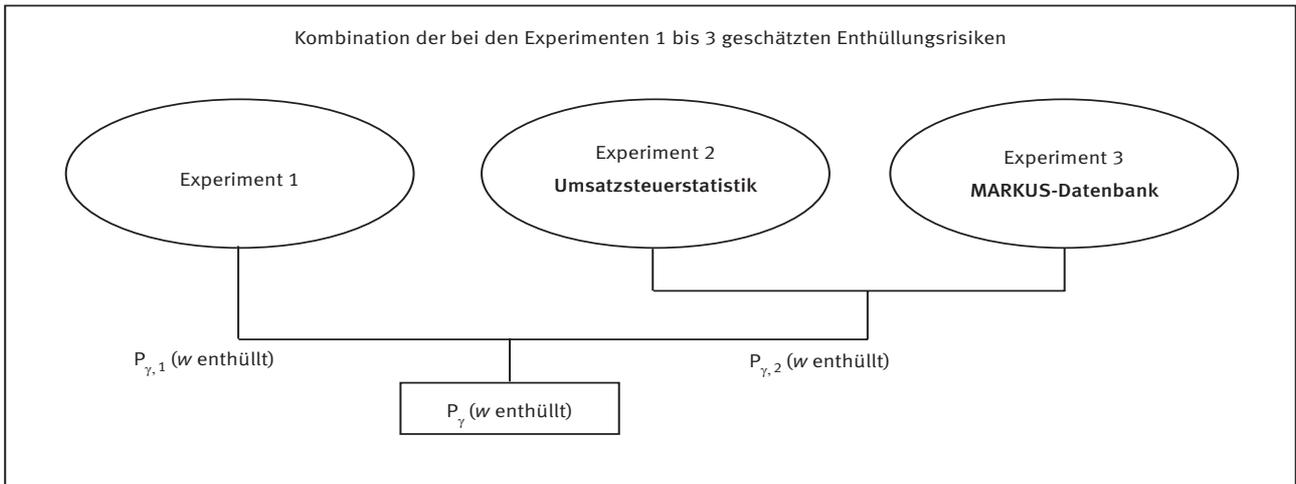


Zum Beispiel beträgt das Risiko der Enthüllung einer Einzelinformation bei Experiment 2 und Verwendung des Anonymisierungsverfahrens MA11G etwa 22,5%. Das Diagramm macht ferner den Unterschied zwischen realistischen und weniger realitätsnahen Szenarien des Angriffswissens deutlich.

Hinsichtlich der Experimente 2 und 3 haben sich interessanterweise, trotz ungleicher Voraussetzungen in den Überschneidungsmerkmalen und natürlichen Abweichungen in den Erhebungen, nach Durchführung der Nutzenanalyse ähnliche Enthüllungsrisiken ergeben. Dies, obwohl die zuvor berechneten Zuordnungsquoten deutlich voneinander verschieden waren (siehe Schaubilder 5 und 6). Das heißt die Schutzwirkung der Anonymisierung auf Basis der Experimente 2 oder 3 ist kaum unterscheidbar.

Obwohl Experiment 1 mit der Verwendung der Originaldaten als Zusatzwissen nicht die Realität abbildet und lediglich *eine* (nicht die kleinste) obere Schranke für das Enthül-

Schaubild 8



lungsrisko bestimmt, kann es angemessen sein, ein solches Szenario in die Abschätzung faktischer Anonymität einfließen zu lassen, und zwar verstärkt in Fällen, in denen wenig Kenntnis über das relevante Zusatzwissen vorhanden ist. Allerdings ist kaum ein Fall denkbar, in dem die faktische Anonymität ausschließlich anhand dieses Szenarios abgeschätzt werden könnte. Stattdessen könnte das Enthüllungsrisiko aus Szenario 1 anteilig in ein Gesamtmaß für das Enthüllungsrisiko einfließen. Wir bezeichnen mit  $\hat{P}_{\gamma,1}(w \text{ enthüllt})$  das beim Experiment 1 geschätzte Risiko der Enthüllung eines Einzelwertes  $w$  und mit  $\hat{P}_{\gamma,2}(w \text{ enthüllt})$  das bei den Experimenten 2 und 3 geschätzte Risiko. Beide Risiken können nun kombiniert werden, wie Schaubild 8 skizziert.

Wenn, wie in den Experimenten 2 und 3 angenommen, geeignetes realitätsnahes Zusatzwissen vorhanden ist, sollte ein Enthüllungsrisiko wie aus Experiment 1 nur mit einem geringen Einfluss in den Schätzer für das zugrunde zu legende gesamte Enthüllungsrisiko eingehen. Im Allgemeinen kann eine Gewichtung der einzelnen Experimente mittels einer Konvexkombination erfolgen:

$$\hat{P}_{\gamma}(w \text{ enthüllt}) = \lambda \hat{P}_{\gamma,1}(o^{(0)} \text{ enthüllt}) + (1 - \lambda) \hat{P}_{\gamma,1}(w \text{ enthüllt})$$

für ein geeignetes  $\lambda \in [0,1]$ .

Um nun eine Aussage über das Vorliegen von faktischer Anonymität zu machen, muss man an dieser Stelle die Schwelle für das Enthüllungsrisiko festlegen. Faktische Anonymität soll wie oben erläutert für diejenigen Probeanonymisierungen vorliegen, für die das durch obige Formel geschätzte Enthüllungsrisiko unterhalb der Risikoschwelle  $\tau$  liegt, das heißt

$$\hat{P}_{\gamma,1}(w \text{ enthüllt}) < \tau$$

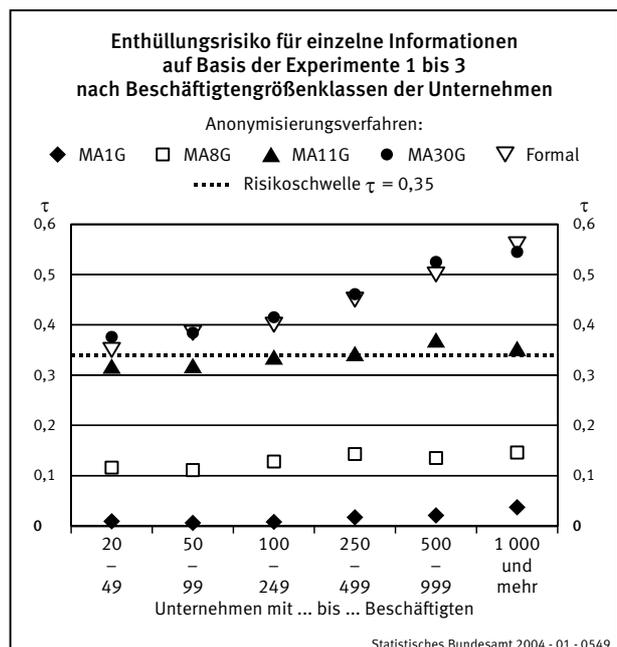
wobei für  $\tau$  ein angemessener Wert gewählt werden sollte.

Unabhängig von der Wahl der Risikoschwelle  $\tau$  sollte allgemein bei qualitativ hochwertigem Zusatzwissen ein  $\lambda$ -Wert nahe bei 0 gewählt werden. Bei einem  $\lambda$ -Wert nahe bei (oder gleich) 1 würde eine absolut anonyme Datei produziert und damit das Wissenschaftsprivileg missach-

tet. Von theoretischer Seite sind Experimente mit realistischem Zusatzwissen immer als beispielhaft einzustufen. Die hundertprozentige Gewissheit, dass einem potenziellen Datenangreifer kein „besseres“ als das in Simulationen verwendete Zusatzwissen verfügbar ist, gibt es seitens des Datenanbieters nicht. Für den Fall der Kostenstrukturerhebung stellen die Autoren nach reiflicher Überlegung einen Wert  $\lambda = 0,2$  zur Diskussion.

Es gibt kritische Teilbereiche in den Daten der Kostenstrukturerhebung, in welchen von vornherein ein hohes Enthüllungsrisiko zu erwarten ist. Aus diesem Grunde sollte das Enthüllungsrisiko im Gegensatz zu Schaubild 7 separat für diese besonders gefährdeten Bereiche untersucht werden. Beispielhaft wurde dies für Beschäftigtengrößenklassen durchgeführt. Als Nutzen- und Risikoschwelle wurden  $\gamma = 0,05$  und  $\tau = 0,35$  gewählt.

Schaubild 9



Wie erhofft wirken die Anonymisierungsmaßnahmen besonders bei größeren Unternehmen. Während in den Schaubildern 4 bis 6 noch ein erhöhtes Risiko für größere Unternehmen festgestellt werden konnte, hat sich nach Berücksichtigung des Nutzenaspektes bei den Anonymisierungsvarianten MA8G und MA11G die Schutzwirkung (bezogen auf Beschäftigtengrößenklassen) deutlich stabilisiert. Bei der Variante MA11G ist sogar ein leichter Rückgang des Enthüllungsrisikos bei den großen Unternehmen zu verzeichnen. Unter der im Schaubild 9 markierten oberen Risikoschwelle  $\tau = 0,35$  würden die mit den Varianten MA1G und MA8G anonymisierten Daten als faktisch anonym eingestuft. Bei der Variante MA11G läge allein die Klasse der Unternehmen mit 500 bis 999 Beschäftigten über der Risikoschwelle  $\tau$ . Es zeigt sich, dass bei den formal anonymisierten Daten die kleinen Unternehmen mit 20 bis 49 Beschäftigten bei dieser Risikoschwelle bereits durch die natürlichen Abweichungen zwischen den betrachteten Erhebungen ausreichend geschützt wären.

### Lehren aus den Experimenten 1 bis 3 zur Kostenstrukturhebung

Die Erstellung eines Scientific Use Files für die Wissenschaft aus der Kostenstrukturhebung scheint auf Grundlage des im Beitrag „Geheimhaltungsmethoden auf dem Prüfstand – eine Analyse anhand der Umsatzsteuerstatistik<sup>420)</sup> in diesem Heft dargestellten Zuordnungsverfahrens und des in Kapitel 3 vorgestellten Ansatzes zur Messung der Schutzwirkung von Anonymisierungsmaßnahmen in naher Zukunft möglich. Kleine und mittlere Unternehmen genießen bereits durch natürliche Abweichungen zwischen Zusatzwissen und Zieldaten einen hohen Schutz. Bei den größeren Unternehmen wurden in den betrachteten Experimenten bei einigen

Anonymisierungsvarianten die wachsenden Trefferquoten durch den sinkenden Nutzen der gewonnenen Informationen für den Datenangreifer kompensiert (siehe Schaubild 7). Bei den Experimenten 2 und 3 wurde das im Projekt zugängliche realitätsnahe Zusatzwissen eingesetzt. Neben den oberen Beschäftigtengrößenklassen stellen auch dünn besetzte Wirtschaftszweige kritische Bereiche mit hohem Enthüllungsrisiko dar. Im Weiteren muss geprüft werden, wie groß die Enthüllungsrisiken für bestimmte Kombinationen von Beschäftigtengrößenklassen und Wirtschaftszweigen sind. Gegebenenfalls muss eine stärkere Behandlung dieser und weiterer kategorialer Merkmale erfolgen (z. B. Verringerung der Tiefe der wirtschaftlichen Gliederung oder Streichung der Regionalkennung aus den Daten). Allerdings darf bei den Untersuchungen zum Schutz von Anonymisierungsmaßnahmen nicht das gleichrangige Ziel des bestmöglichen Erhalts des Analysepotenzials außer Acht gelassen werden.

### 3.2 Enthüllungsrisiko in der Umsatzsteuerstatistik

#### Experiment 4: Enthüllungsrisiko auf Basis der Zuordnung zwischen Zusatzwissen und probe-anonymisierten Daten der Umsatzsteuerstatistik

Auch für die Umsatzsteuerstatistik wird im Folgenden gemäß dem Schutzwirkungskonzept der Indikator für das Enthüllungsrisiko aus dem Produkt des Anteils richtig zugeordneter Unternehmen (Trefferquote oder Zuordnungsrisiko) und dem Anteil nützlicher Informationen errechnet. In der Tabelle 9 sind diese Werte für angenommene Nützlichkeitschwellen von 5 bzw. 10% wiedergegeben. Bei formaler Anonymisierung entspricht das Enthüllungsrisiko dem

Tabelle 9: Enthüllungsrisiko bei Experiment 4 insgesamt und nach Umsatzgrößenklassen der Unternehmen  
Prozent

Anonymisierungsverfahren	Gebiets-schlüssel	Insgesamt	Unternehmen mit Umsätzen <sup>1)</sup> von ... bis unter ... EUR				
			16 617 – 1 Mill.	1 Mill. – 10 Mill.	10 Mill. – 100 Mill.	100 Mill. – 1 Mrd.	1 Mrd. und mehr
Nützlichkeitschwelle 5%							
Formal anonymisiert ..	BBR9	31,7	13,2	27,6	36,0	51,8	54,5
	BBR3	22,6	9,3	18,5	26,9	38,6	43,2
MA21G .....	BBR9	31,4	13,0	27,6	35,8	49,6	37,4
	BBR3	22,4	8,7	18,8	26,7	36,5	33,4
MA1G .....	BBR9	2,2	0,9	1,4	2,9	5,0	6,6
	BBR3	1,2	0,5	0,7	1,7	2,7	3,1
Traditionell anonymisiert .....	BBR9	30,0	12,7	26,5	34,8	43,8	5,8
	BBR3	22,3	9,3	18,7	26,5	36,8	3,9
Nützlichkeitschwelle 10%							
Formal anonymisiert ..	BBR9	31,7	13,2	27,6	36,0	51,8	54,5
	BBR3	22,6	9,3	18,5	26,9	38,6	43,2
MA21G .....	BBR9	31,5	13,0	27,6	35,8	50,6	42,1
	BBR3	22,5	8,7	18,8	26,7	37,2	37,2
MA1G .....	BBR9	4,1	1,8	2,5	5,3	10,0	11,0
	BBR3	2,2	0,9	1,2	3,1	5,6	6,6
Traditionell anonymisiert .....	BBR9	30,0	12,7	26,5	34,8	44,0	6,2
	BBR3	22,3	9,3	18,7	26,5	37,0	4,0

1) Lieferungen und Leistungen.

20) Siehe Fußnote 8.

Zuordnungsrisiko (Trefferquoten), da die Merkmalsausprägungen unverändert bleiben. Dies gilt auch für die kleinen und mittleren Unternehmen bei der vorliegenden traditionellen Anonymisierung und der Anonymisierung nach MA21G. Hier weichen bei diesen Merkmalsträgern die anonymisierten Werte von den wahren Werten nur minimal ab (siehe Abschnitt 1.4.2). Dagegen liegt das Enthüllungsrisiko bei den großen Unternehmen bei diesen Verfahren zum Teil deutlich unter dem Zuordnungsrisiko. Während bei der traditionellen Anonymisierung keine allgemeinen Aussagen gemacht werden können, weil diese in verschiedenen Fällen ganz unterschiedlich aussehen kann<sup>21)</sup>, scheint sich das gefundene Ergebnis bei einer Mikroaggregation, die jedes Merkmal einzeln anonymisiert (wie bei MA21G), bei weiteren Untersuchungen zu erhärten. Beim Verfahren MA1G liegt das Enthüllungsrisiko deutlich unter dem bereits geringen Zuordnungsrisiko, sodass die Chancen eines Angreifers, einen nützlichen Wert zu enthüllen, bei dieser Anonymisierungsmethode als sehr gering einzustufen sind.<sup>22)</sup> Es zeigt sich wie bei der Kostenstrukturerhebung in Abschnitt 3.1, dass unter der Berücksichtigung des Nutzenaspekts besonders das Risiko bei den großen Unternehmen sinkt.

Die nun zu beantwortende Frage ist, welches Ausmaß der Anonymisierung bei der Umsatzsteuerstatistik ausreichend Schutz bietet. Anders formuliert, bei welcher Dosierung der Anonymisierung wird das Risiko für einen Datenangreifer, einen für ihn unbrauchbaren Wert zu enthüllen und fälschlicherweise als brauchbaren anzunehmen, so groß, dass eine Reidentifikation als nutzlos eingestuft werden kann, selbst wenn es theoretisch möglich wäre, brauchbare Werte zu enthüllen. Als Vorschlag wird hier zur Diskussion gestellt, dass solche Daten zwei Kriterien erfüllen sollten: Als erstes Kriterium darf das Enthüllungsrisiko für den gesamten Datenbestand nicht höher als 25% sein. Dies würde bedeuten, dass ein Angreifer bei vier enthüllten Werten drei unbrauchbare enthüllt, ohne zu wissen, welche Werte brauchbar sind. Einem risikoaversen Angreifer dürfte dieses Risiko zu hoch sein. Allerdings kann ein Datenangreifer anhand der Größe der zugeordneten Unternehmen genauere Annahmen über das Risiko treffen, das innerhalb des gesamten Datenbestandes auf denjenigen Datensätzen liegt, die ihn interessieren. Ihm ist die Regel als bekannt zu unterstellen, dass das Enthüllungsrisiko mit der Unternehmensgröße steigt und damit sein Risiko sinkt. Daher wird hier vorgeschlagen, eine zweite Schwelle des akzeptablen Enthüllungsrisikos für Teilmengen des Gesamtdatenbestandes festzulegen. Die Teilmengen werden so geschnitten, dass schutzbedürftige Bereiche, die auch ein Datenangreifer als solche identifizieren kann, auf ihre Gefährdung hin analysiert werden können.

Für die Umsatzsteuerstatistik wird hier zur Diskussion gestellt, dass als zweites Kriterium in keiner der gewählten

Umsatzgrößenklassen das Enthüllungsrisiko höher als 40% sein sollte. Damit wird gleichzeitig verhindert, dass ein Teilbereich völlig ungeschützt ist und dies im Schutzmaß der gesamten Datei durch einen vollständigen Schutz in einem anderen Bereich ausgeglichen wird. Zum Beispiel könnte das Enthüllungsrisiko bei den großen Unternehmen bei 100% liegen und bei den kleinen bei 0%. Da die Anzahl der kleineren Unternehmen deutlich höher ist als die der großen, würde das gesamte Enthüllungsrisiko wahrscheinlich deutlich unter 25% liegen und die Daten würden die Bedingung der faktischen Anonymität erfüllen, obwohl es ein leichtes wäre, Großunternehmen zu reidentifizieren.

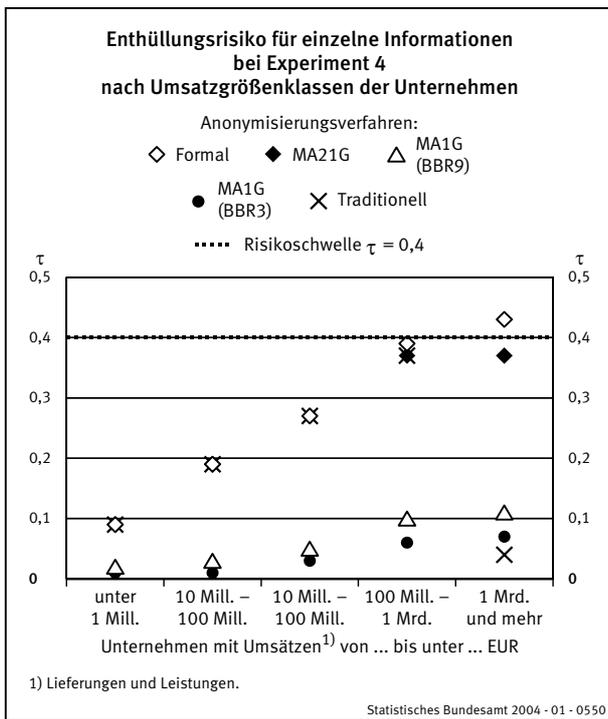
Dass der Schwellenwert des zweiten Kriteriums höher liegt als der des ersten, kann auf zwei Wegen begründet werden. Zum einen ist das erste Kriterium nur dann logisch relevant, wenn das zweite schwächer ist. Da das Enthüllungsrisiko in den Umsatzgrößenklassen unterschiedlich ist, wird mindestens immer eine Größenklasse dabei sein, in der das spezifische Enthüllungsrisiko über dem Gesamtrisiko liegen wird, sodass der Wert des ersten Kriteriums deutlich unter 25% liegen müsste, wollte man in den Umsatzgrößenklassen den gleichen Wert nicht überschreiten. Neben dieser rein technischen Begründung spielen zum anderen auch inhaltliche Aspekte eine Rolle. Der Datenangreifer müsste erstens die Struktur der Enthüllungsrisiken (bzw. Chancen aus Sicht des Angreifers) genau kennen, um diese Information für seinen Angriff auszunutzen. In der Realität kann er diese Struktur aber nur abschätzen, was ein weiteres Risiko für ihn bedeutet. Zweitens müsste er auch wissen, zu welchen Größenklassen die Unternehmen tatsächlich gehören. Auch das ist nicht immer fehlerfrei möglich. Diese zwei Argumente scheinen eine schwächere Bedingung für das Enthüllungsrisiko beim zweiten Kriterium zu rechtfertigen.

Die Bedingung, dass das Enthüllungsrisiko für den gesamten Datenbestand unter 25% liegen soll, wird von allen Daten erfüllt, die als Regionalkennung den Gebietsschlüssel BBR3 aufweisen (siehe Tabelle 9). Ist in den Daten der Gebietsschlüssel BBR9 enthalten, so wird diese Bedingung nur bei der Mikroaggregation 1G erfüllt (gilt für beide Nützlichkeitschwellen). Das Schaubild 10 zeigt – unter der Annahme einer Nützlichkeitschwelle von 10% –, welche Daten, die die Bedingung, dass das Enthüllungsrisiko für den gesamten Datenbestand unter 25% liegen soll, erfüllen, auch die Bedingung erfüllen, dass in keiner Umsatzgrößenklasse das Enthüllungsrisiko über 40% liegen soll. Liegt bei Teilbereichen einer Probeanonymisierung ein Punkt oberhalb der Trennlinie, dann gilt die Datei *insgesamt* als nicht faktisch anonym. Nach diesem Kriterium muss die lediglich formal anonymisierte Datei als nicht faktisch anonym eingestuft werden. Die weiteren Alternativen erfüllen das Kriterium. Welche der vier „faktisch anonymen“ Dateien allerdings für ein Scientific Use File geeignet erscheint,

21) So könnte z. B. innerhalb einer traditionellen Anonymisierung das Verfahren „Bottomcoding“ zur Anwendung kommen. Hierbei wird eine Mindestgrenze eingeführt. Alle Werte, die unter der Grenze liegen, werden auf diese Grenze gesetzt. Dadurch werden die nützlichen Informationen bei kleinen Unternehmen reduziert und damit liegt auch das Enthüllungsrisiko bei kleinen Unternehmen unter dem Zuordnungsrisiko.

22) Zu beachten ist allerdings, dass hiermit MA1G nur eines von zwei Kriterien eines Scientific Use Files erfüllt, nämlich die Datensicherheit. Inwieweit durch eine solche Anonymisierung auch das Analysepotenzial als zweite Bedingung erhalten bleibt, erscheint nach ersten Analysen sehr zweifelhaft. Siehe Rosemann, M.: „Erste Ergebnisse von vergleichenden Untersuchungen mit anonymisierten und nicht anonymisierten Einzeldaten am Beispiel der Kostenstrukturerhebung und der Umsatzsteuerstatistik“, Band 42 der Schriftenreihe „Forum der Bundesstatistik“, Wiesbaden 2003, S. 154 ff.

Schaubild 10



hängt von der Beurteilung des Analysepotenzials ab. Unter der Annahme, dass bei steigendem Schutz das Analysepotenzial sinkt, sollte die Datei gewählt werden, welche die Bedingungen der faktischen Anonymität am geringsten erfüllen. Dies wären MA21G und eventuell die traditionelle Anonymisierung.

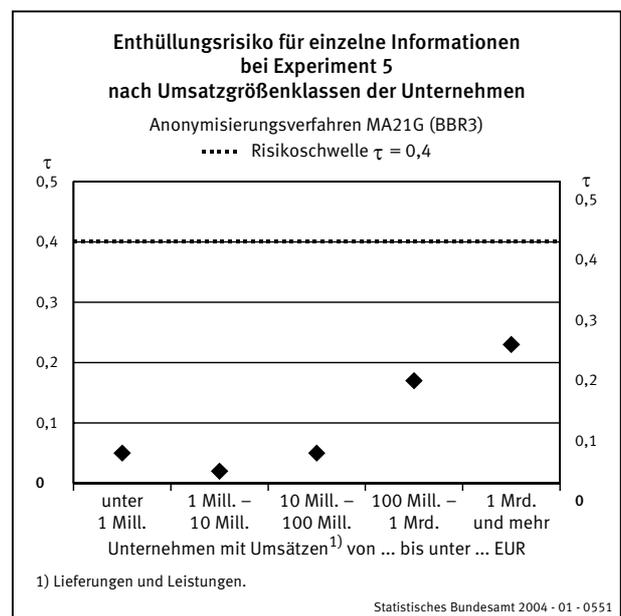
**Experiment 5: Enthüllungsrisiko auf Basis der Zuordnung zwischen MARKUS-Datenbank und probe-anonymisierten Daten der Umsatzsteuerstatistik**

Da sich die Mikroaggregation MA21G in Kombination mit dem BBR3 beim vierten Experiment als möglicher Kandidat für ein Scientific Use File herauskristallisiert hat, kann dieses fünfte Experiment abschließend als Test des Ergebnisses des vierten Experiments angesehen werden.

In der Tabelle 10 sind neben dem Enthüllungsrisiko noch einmal die Ergebnisse der „Zwischenschritte“ (Trefferquote

und Anteil nützlicher Informationen) aufgelistet. Ersichtlich wird, dass sich zwar der Anteil an brauchbaren Informationen gegenüber dem vierten Experiment nicht unterscheidet, allerdings in diesem Experiment deutlich geringere Trefferquoten erzielt wurden. Aufgrund dieser geringen Trefferquoten ist auch das Enthüllungsrisiko deutlich niedriger als im vierten Experiment. Selbst bei den größten Unternehmen liegt das Enthüllungsrisiko nur um die 20% und damit deutlich unter den im vierten Experiment geforderten Grenzen. Die faktische Anonymität dieser Daten wäre nach den vorgeschlagenen Schwellenwerten daher durch einen Datenangriff mittels Zuordnung mit Hilfe der MARKUS-Datenbank nicht in Frage gestellt. Dies verdeutlicht auch Schaubild 11.

Schaubild 11



**Experiment 6: Enthüllungsrisiko auf Basis ausgewählter Einzelangriffe**

In der Tabelle 11 sind die Ergebnisse der Enthüllungsversuche, die im Rahmen des Experimentes 6 erzielt wurden, zusammengefasst. Der besseren Übersicht wegen wurden auch hier die Ergebnisse der Zwischenschritte in die Tabelle aufgenommen.

Tabelle 10: Richtig zugeordnete Unternehmen, Anteil nützlicher Informationen und Enthüllungsrisiko bei Experiment 5 insgesamt und nach Umsatzgrößenklassen der Unternehmen

Gegenstand der Nachweisung	Insgesamt	Unternehmen mit Umsätzen <sup>1)</sup> von ... bis unter ... EUR				
		16 617 – 1 Mill.	1 Mill. – 10 Mill.	10 Mill. – 100 Mill.	100 Mill. – 1 Mrd.	über 1 Mrd.
Richtig zugeordnete Unternehmen						
Anzahl .....	319	2	55	176	77	9
% .....	5,0	5,4	1,7	4,7	17,4	31,0
Anteil nützlicher Informationen						
Nützlichkeitsschwelle 5% .....	97,6	100	100	99,9	93,9	64,3
Nützlichkeitsschwelle 10% .....	98,6	100	100	100	96,8	73,9
Enthüllungsrisiko						
Nützlichkeitsschwelle 5% .....	4,9	5,4	1,7	4,7	16,3	19,9
Nützlichkeitsschwelle 10% .....	4,9	5,4	1,7	4,7	16,8	22,9

1) Lieferungen und Leistungen.

Tabelle 11: Trefferquoten, Anteil nützlicher Informationen und Enthüllungsrisiko beim simulierten Einzelangriff (Experiment 6)  
Prozent

Anonymisierungs- angaben	Treffer- quoten	Anteil nützlicher Informationen		Enthüllungsrisiko	
		$\gamma = 5\%$	$\gamma = 10\%$	bei $\gamma = 5\%$	bei $\gamma = 10\%$
Formal anonymisiert ...	100	100	100	100	100
MA21G .....	100	50,5	63,6	50,5	63,6
MA1G .....	40	2,5	5,0	1,0	2,0
Traditionell anonymisiert .....	20	64,7	70,6	12,9	14,1

$\gamma$  = Nützlichkeitschwelle

Da es nicht nur problemlos war, die Unternehmen bei einer rein formalen Anonymisierung zu reidentifizieren, sondern auch die enthüllten Werte den wahren Werten entsprechen, ergibt sich in diesem Fallbeispiel für die formale Anonymisierung ein Enthüllungsrisiko von 100%. Wie beim Massenfischzug kann durch eine MA21G Anonymisierung die Anzahl der richtigen Zuordnungen nicht reduziert werden. Allerdings wirkt sich die Veränderung der einzelnen Merkmalsausprägungen erheblich auf die Nützlichkeit der Information für einen Datenangreifer aus. Ist für ihn eine Information nur brauchbar, wenn sie um weniger als 5% von der wahren Information abweicht, dann ist für ihn jeder zweite Wert nutzlos (dabei kann er wertvollen und wertlosen Wert nicht unterscheiden). Bei den Verfahren MA1G und der traditionellen Anonymisierung werden sowohl die Trefferquoten als auch die Nützlichkeit der Informationen derartig gesenkt, dass die Einzelangriffe als gescheitert angesehen werden können. Anders formuliert, diese Daten können bei den vorgeschlagenen Schwellenwerten ohne Einschränkung auch bei den Branchenmarktführern als faktisch anonym eingestuft werden.

## Lehren aus den Experimenten 4 bis 6 zur Umsatzsteuerstatistik

Auf Basis des in dieser Zeitschrift bereits vorgestellten Konzepts zur Schutzwirkung faktischer Anonymisierung<sup>23)</sup> kann festgehalten werden, dass Scientific Use Files für die Umsatzsteuerstatistik unter bestimmten Bedingungen möglich sind. Für den Großteil der kleinen und mittleren Unternehmen ist hierzu eine verhältnismäßig schwache Anonymisierung notwendig, eventuell kann hier auf datenverändernde Verfahren ganz verzichtet werden. Dagegen erscheinen datenverändernde Verfahren bei den Großunternehmen notwendig. Diese unterschiedliche Behandlung kann einerseits dadurch erreicht werden, dass die Unternehmensgrößen getrennt anonymisiert werden (wie bei der traditionellen Anonymisierung geschehen). Andererseits werden mit Hilfe einer Mikroaggregation die Unternehmensgrößenklassen automatisch unterschiedlich behandelt. Wie gesehen führt eine Mikroaggregation dazu, dass die Ausprägungen der kleinen und mittleren Unternehmen nur auf vernachlässigbare Weise verändert werden, die der großen jedoch zum Teil erheblich. Gibt es nach der Anwendung die-

ser Anonymisierung kleine Risikogruppen – wie zum Beispiel die größten Unternehmen einer Branche – dann sollte zusätzlich überlegt werden, diese wenigen Unternehmen separat zu behandeln.

## 4 Fazit und Ausblick

Gemäß dem im Projekt entwickelten Konzept zur Schutzwirkung von Anonymisierungen werden Gefährdungspotenziale für Einzeldaten an den Möglichkeiten gemessen, als Datenangreifer durch Deanonymisierungen nutzbringende Informationen zu gewinnen. In diesem Beitrag wird anhand echter Einzeldaten der amtlichen Statistik und verschiedener Arten des Zusatzwissens gezeigt, wie das Schutzwirkungskonzept angewendet wird. Damit ist den statistischen Ämtern ein Hilfsmittel an die Hand gegeben, mit dem es möglich wird, die faktische Anonymität von Merkmalsträgern zu prüfen. Das Schutzwirkungskonzept ist explizit auf die faktische und nicht auf die absolute Anonymität ausgerichtet und bietet damit einen Weg, die in § 16 Absatz 6 BStatG formulierten Bedingungen zur Weitergabe von Einzeldaten an die Wissenschaft zu operationalisieren. Eine wesentliche Leistung des Konzeptes liegt darin, dass nicht nur die Kostenseite des Datenangreifers, sondern auch die Nutzenseite in die Abschätzung einbezogen wird. Da der Nutzen eines unbekanntem Datenangreifers aber nicht als Wert, zum Beispiel als Geldäquivalent, zu ermitteln ist, werden die Unsicherheiten betrachtet, mit denen ein Datenangreifer konfrontiert ist, wenn er eine Enthüllung versucht. Gerade unter der Berücksichtigung des Nutzens eines Datenangriffs und der Risiken eines Datenangreifers wird die Möglichkeit, Merkmalsträger wirtschaftsstatistischer Einzeldaten so zu anonymisieren, dass die Bedingung der faktischen Anonymität erfüllt wird, wesentlich verbessert. Die Untersuchung der Unsicherheiten für einen Datenangreifer erlaubt es, das abstrakte Unverhältnismäßigkeitskriterium des § 16 Absatz 6 BStatG in ein messbares Entscheidungskriterium für die faktische Anonymität zu überführen. Die vorgestellten Analysen zur Kostenstrukturherhebung im Verarbeitenden Gewerbe und zur Umsatzsteuerstatistik belegen die Anwendbarkeit des Konzeptes.

Die im Beitrag und auf dem genannten Workshop gezogenen Schlussfolgerungen sind abhängig von den gewählten Schwellenwerten für die Risikomaße. Die Schwellenwerte wurden in diesem Beitrag für zwei Erhebungen beispielhaft gesetzt. Die Entscheidung über die faktische Anonymität einer Datei wird – wie bei Haushalts- und Personendaten bereits seit langem praktiziert – durch Abstimmungsprozesse in den statistischen Ämtern und unter maßgeblicher Mitarbeit der für die jeweilige Erhebung zuständigen Fachstatistiker getroffen. Wesentliche Vorarbeit für die Entscheidung über die angemessenen Schwellenwerte ist die Untersuchung des für einen Datenangriff verfügbaren Zusatzwissens.

In diesem Beitrag wurde auch diskutiert, dass die Festlegung der Schwellenwerte statistikspezifisch erfolgen kann.

23) Siehe Fußnote 3.

Darüber hinaus kann es angemessen sein, die Schwellen für eine spezielle Statistik auch weiter zu differenzieren, zum Beispiel in Abhängigkeit von der Sensibilität verschiedener Merkmale, aus der sich eine unterschiedlich große Motivation zum Datenangriff ergeben kann. Um die Gefährdungsbereiche innerhalb der Gesamtdatensätze zu erkennen, ist die Berechnung der Risikomaße zum Beispiel abhängig von der Unternehmensgröße oder der Tiefe der wirtschaftlichen Gliederung angezeigt. Mit diesem Wissen kann dann problemadäquat anonymisiert werden, das heißt dass wirklich nur das durch Anonymisierungsmaßnahmen zusätzlich geschützt wird, was des Schutzes bedarf. Auf diese Weise wird das Analysepotenzial der Daten am besten erhalten.

Es liegt nun an den statistischen Ämtern, sinnvolle Schwellen für ihre jeweiligen Statistiken zu entwickeln und damit den Weg frei zu machen, um der Wissenschaft faktisch anonymisierte wirtschaftsstatistische Einzeldaten anbieten zu können. [u](#)

Dr. Rainer Lenz, Dr. Daniel Vorgrimler

# Geheimhaltungsmethoden auf dem Prüfstand – eine Analyse anhand der Umsatzsteuerstatistik

*In der wissenschaftlichen Diskussion zur Geheimhaltungsmethoden von Einzeldaten ist bisher relativ wenig über die Möglichkeit bekannt, unter realitätsnahen Bedingungen anonymisierten Merkmalsträgern mit Hilfe von so genanntem Zusatzwissen direkte Identifikatoren korrekt zuzuordnen. Ebenso ist unzureichend geklärt, in welchem Umfang Geheimhaltungsmethoden unter realen Bedingungen in der Lage sind, richtige Zuordnungen zu verhindern. Realitätsnah bedeutet dabei, Einzeldaten einer externen Datei als Zusatzwissen zu verwenden und dieses mit einem verfügbaren Zuordnungsalgorithmus einer anonymisierten Zieldatei zuzuordnen.*

*Der vorliegende Beitrag nimmt sich dieser beiden Fragestellungen an und will mit seinen Resultaten die Diskussion zu den Fragen weiter voranbringen. Hierzu wird im Rahmen von Simulationsexperimenten eine externe Datei als Zusatzwissen einem Teilbereich der Umsatzsteuerstatistik 2000 zugeordnet. Die verschiedenen Ergebnisse zeigen, welche Chancen bestehen, Merkmalsträger zu „reidentifizieren“ bzw. ob die Daten durch einen „natürlichen Schutz“ bereits so sicher sind, dass die statistische Geheimhaltung bereits durch „formale“ Anonymisierung gewahrt bleibt. Verwendet wird dazu ein Zuordnungsalgorithmus, der im Rahmen von CASC – einem europäischen Forschungsprojekt zur Geheimhaltung – entwickelt wurde. Dieser soll noch in diesem Jahr allgemein zugänglich gemacht werden. Die wiederholte Durchführung der Experimente mit anonymisierten Dateien zeigt anschließend die Effektivität verschiedener Geheimhaltungsmethoden bei der Verhinderung von rich-*

*tigen Zuordnungen auf. Neben traditionellen Anonymisierungsmaßnahmen steht dabei die Anonymisierung durch Mikroaggregation im Mittelpunkt des Interesses.*

## Vorbemerkung

Empirische Wirtschaftsforscher klagen zunehmend über einen mangelnden Zugang zu wirtschaftsstatistischen Einzeldaten der amtlichen Statistik. Durch diesen sei das Potenzial an möglichen Analysen, die mit amtlichen Erhebungen denkbar sind, nur ungenügend ausgeschöpft. Dabei wurde bereits 1987 im Bundesstatistikgesetz<sup>1)</sup> mit dem § 16 Abs. 6 der Wissenschaft ein privilegierter Zugang zu Mikrodaten der amtlichen Statistik eingeräumt. Dieser Paragraph erlaubt die Übermittlung von Einzeldaten an die Wissenschaft, sofern diese nur mit unverhältnismäßig hohem Aufwand reidentifiziert werden können. „Unverhältnismäßig“ bedeutet, dass die Kosten einer Reidentifikation deren Nutzen übersteigen (faktische Anonymität).

Während die Bereitstellung von faktisch anonymisierten Personendaten bereits seit längerem bewährte Praxis ist, ergeben sich für wirtschaftsstatistische Einzeldaten besondere Probleme, die bisher eine faktische Anonymisierung unmöglich erscheinen ließen. Verglichen mit Personenerhebungen liegen bei Unternehmenserhebungen wesentlich kleinere Grundgesamtheiten zugrunde, sodass die Besetzungszahlen einzelner Gruppen häufig kleiner sind. Die Verteilungen der quantitativen Merkmale sind wesentlich heterogener und es treten mit größerer Wahrscheinlichkeit

1) Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) vom 22. Januar 1987 (BGBl. I S. 462, 565), zuletzt geändert durch Artikel 16 des Gesetzes vom 21. August 2002 (BGBl. I S. 3322).

dominierende Fälle auf. Des Weiteren sind die Stichprobenauswahlsätze bei Unternehmenserhebungen wesentlich höher als bei Personenerhebungen. Publizitätspflichten der Unternehmen einerseits und kommerzielle Datenbanken andererseits bewirken, dass einem Externen, der Einzeldaten ihrem Träger zuordnen will, bei Wirtschaftsdaten besseres Zusatzwissen zur Verfügung steht als bei Personendaten. Schließlich wird der Nutzen aus der Kenntnis von Unternehmens- und Betriebsdaten wesentlich höher eingestuft als von Daten aus Personen- und Haushaltserhebungen. Ein rationaler Datenangreifer wird deshalb auch höhere Kosten für eine Deanonymisierung akzeptieren.

Diese Besonderheiten wirtschaftsstatistischer Einzeldaten stellen besondere Ansprüche an die verwendeten Anonymisierungsmethoden. Um diesen Ansprüchen gerecht zu werden, entwickelte die Anonymisierungsforschung neue Verfahren, welche die Vertraulichkeit der Daten sichern sollen. Im folgenden Beitrag wird mit der Mikroaggregation ein Verfahren auf den Prüfstand gestellt, welches in der internationalen Literatur größere Bedeutung erlangt hat.<sup>2)</sup> Verglichen wird dabei die Wirkung des Verfahrens sowohl in Kombination mit verschiedenen traditionellen Anonymisierungsmaßnahmen als auch mit einer rein traditionellen Anonymisierung. Darüber hinaus wird untersucht, ob eine formale Anonymisierung (d. h. Löschung der direkten Identifikatoren) bereits genügend Schutz bietet.

Diese Fragen werden mit Hilfe von Simulationen analysiert. Es wird einerseits versucht, über so genannte Massenfischzüge<sup>3)</sup> möglichst viele Merkmalsträger (in diesem Falle Unternehmen) und andererseits über Einzelangriffe bestimmte Merkmalsträger zu reidentifizieren. Das grundsätzliche Vorgehen und die Beschreibung des hierzu verwendeten Algorithmus ist Thema der folgenden Kapitel 1 und 2. Die Ergebnisse der eigentlichen Simulationen stehen in Kapitel 3 im Mittelpunkt. Hier werden zunächst die Zieldaten und das so genannte Zusatzwissen beschrieben sowie die Höhe des „natürlichen Schutzes“ der Daten gegen Reidentifikationsversuche bestimmt. Anschließend werden in Abschnitt 3.2 die Ergebnisse der Simulationen vorgestellt. Ein Fazit in Kapitel 4 schließt den Beitrag ab.

## 1 Die Grundvoraussetzungen eines Datenangriffes

Um Merkmalsträger einer vertraulichen Datei erfolgreich reidentifizieren zu können, sind folgende Grundannahmen für einen Datenangreifer nötig<sup>4)</sup>:

- Zusatzwissen über die gesuchten Merkmalsträger (etwa in Form einer externen Unternehmensdatenbank)
- Kenntnis über die Teilnahme des gesuchten Merkmalsträgers an der Erhebung (Zieldaten)

- Merkmale, welche sowohl in externen als auch in Zieldaten enthalten sind (Überschneidungsmerkmale).

Darüber hinaus muss der Datenangreifer persönlich von der Richtigkeit der Zuordnung überzeugt sein, was bei einem Massenfischzug nahezu unmöglich erscheint.

Im Folgenden seien  $A = \{a_1, \dots, a_m\}$  und  $B = \{b_1, \dots, b_n\}$  Mengen von Merkmalsträgern der externen Daten und der Zieldaten, welche eine nichtleere Menge von Überschneidungsmerkmalen teilen. Die Menge der Überschneidungsmerkmale werde mit  $\{v_1, \dots, v_k\}$  bezeichnet. Wir zerlegen die Überschneidungsmerkmale in zwei Klassen, *kategoriale* und *metrische* Merkmale. Bei kategorialen Merkmalen wird zwischen *nominalen* Merkmalen (es gibt keine lineare Ordnung zwischen den Kategorien) und *ordinalen* Merkmalen (die Kategorien besitzen eine lineare Ordnung, wobei Differenzen zwischen Kategorien keinen Sinn ergeben) unterschieden. Unter metrischen Merkmalen versteht man Merkmale, bei denen eine Differenzbildung von Ausprägungen Bedeutung hat, wie zum Beispiel „Größe“ und „Gewicht“ einer Person oder in unserem Falle „Umsatz“ und „Anzahl der Beschäftigten“ eines Unternehmens. Offenbar kann jedes metrische Merkmal über Klassenbildung des Wertebereiches in ein kategoriales Merkmal umgewandelt werden (man betrachte etwa Umsatz- oder Beschäftigtengrößenklassen).

Bei der Simulation eines Datenangriffes sei nun  $(a, b) \in A \times B$  ein Kandidatenpaar für eine mögliche Zuordnung. Wir fordern, dass beide Merkmalsträger  $a$  und  $b$  in zuvor festgelegten Überschneidungsmerkmalen übereinstimmen. Solche Merkmale werden *Blockmerkmale* genannt, da sie den gesamten Datenbestand in überschneidungsfreie (disjunkte) Blöcke unterteilen. Es erscheint vernünftig, solche Merkmale als Blockmerkmale zu bestimmen, welche einerseits die Merkmalsträger gut charakterisieren und andererseits wenig fehlerbehaftet sind. Letzteres gilt bei unseren Betrachtungen insbesondere für einige kategoriale Merkmale, die nicht durch das auf die Zieldaten angewendete Anonymisierungsverfahren verändert wurden. In der Praxis können allerdings auch in solchen Merkmalen signifikante Abweichungen zwischen verschiedenen Erhebungen auftreten (siehe Abschnitt 3.1.1).

## 2 Der Zuordnungsalgorithmus

Ein potenzieller Datenangreifer steht vor dem Entscheidungsproblem, ob ein Paar  $(a, b) \in A \times B$  von Merkmalsträgern zu demselben zugrunde liegenden Individuum (oder Unternehmen) gehört. Hierzu ist ein vernünftiger Ähnlichkeitsbegriff notwendig. Grob gesprochen tritt die größtmögliche für den Datenangreifer messbare Ähnlichkeit zwischen zwei Merkmalsträgern dann auf, wenn sie in sämtlichen

2) Stellvertretend siehe Domingo-Ferrer, J./Mateo-Sanz, J. M.: "Practical data-oriented microaggregation for statistical disclosure control", IEEE Transactions on Knowledge and Data Engineering, Vol. 14(1), 2002, S. 189 ff.

3) Bei einem Massenfischzug versucht ein Datenangreifer mit Hilfe einer externen Datenbank als Zusatzwissen soviel Merkmalsträger der Zieldaten wie möglich zu identifizieren. Im Gegensatz zum Einzelangriff ist er dabei nicht an einem speziellen Merkmalsträger interessiert.

4) Siehe Brand, R./Bender, S./Kohaut, S.: "Possibilities for the creation of a scientific-use-file for the IAB-establishment-panel", Proceedings of the Joint Eurostat/UN-ECE Work Session on Statistical Data Protection, Thessaloniki, 1999, S. 57 ff.

Überschneidungsmerkmalen übereinstimmen. Da im Falle geringer Abweichungen in den Überschneidungsmerkmalen die beiden betrachteten Merkmalsträger als stark verwandt empfunden werden, hängt die Qualität der Zuordnung offenbar im Wesentlichen von der Wahl des Ähnlichkeitsmaßes ab. Wir führen hierzu folgende Distanztypen ein:

Sei  $k$  die Anzahl der Überschneidungsmerkmale. Für das metrische Merkmal  $v_i$  und das Paar  $(a, b)$  von Merkmalsträgern wird die  $i$ -te Komponentendistanzfunktion definiert durch  $d_i(a, b) = (a^{(i)} - b^{(i)})^2$ , wobei  $a = (a^{(1)}, \dots, a^{(k)})$  und  $b = (b^{(1)}, \dots, b^{(k)})$  die auf die  $k$  Überschneidungsmerkmale reduzierten Merkmalsträger sind. Weiterhin definieren wir mit

$$d_i(a, b) := \begin{cases} 0, & \text{wenn } a^{(i)} = b^{(i)} \\ 1 & \text{andernfalls} \end{cases}$$

eine Komponentendistanzfunktion für nominale Merkmale  $v_i$  und mit

$$d_i(a, b) := \frac{|\{c_j \mid \min(a^{(i)}, b^{(i)}) \leq c_j < \max(a^{(i)}, b^{(i)})\}|}{r}$$

eine Komponentendistanzfunktion für ordinale Merkmale  $v_i$ , ausgenommen Blockmerkmale, wobei  $c_1 < c_2 < \dots < c_r$  der geordnete Wertebereich des Merkmals  $v_i$  ist. Um Skalierungsprobleme zu vermeiden, werden die berechneten Komponentendistanzen normiert mittels der bewährten max-min-Standardisierung

$$\tilde{d}_i(a, b) := \frac{d_i(a, b) - \min_{(\alpha, \beta) \in A \times B} d_i(\alpha, \beta)}{\max_{(\alpha, \beta) \in A \times B} d_i(\alpha, \beta) - \min_{(\alpha, \beta) \in A \times B} d_i(\alpha, \beta)}$$

Es sei im Weiteren  $NV$  die Indexmenge der metrischen und  $CV$  die Indexmenge der kategorialen Merkmale. Die Gesamtdistanz ergibt sich als gewichtete Summe über alle Komponentendistanzen:

$$\begin{aligned} d(a, b) &:= \sum_{i \in NV} \lambda_i \tilde{d}_i(a, b) + \sum_{i \in CV} \lambda_i \tilde{d}_i(a, b) \\ &= \sum_{i \in NV} \lambda_i \tilde{d}_i(a, b) + \sum_{i \in C_{ord}} \lambda_i \tilde{d}_i(a, b) + \sum_{i \in C_{nom}} \lambda_i \tilde{d}_i(a, b), \end{aligned}$$

wobei der Distanzbeitrag der kategorialen Merkmale gemäß der vorherigen Überlegungen in einen ordinalen ( $C_{ord}$ ) und einen nominalen ( $C_{nom}$ ) Anteil zerfällt. Es ist sinnvoll, die Blockung der Daten ebenfalls in die Distanzberechnung einzubeziehen<sup>5)</sup>, da das in Komplexitätsuntersuchungen gewöhnlich nicht berücksichtigte blockweise Einlesen der Daten sehr zeitaufwändig sein kann.

Wir sind nun in der Lage, die einzelnen Schritte des Zuordnungsalgorithmus zu skizzieren:

1. Input:  $T = (\tau_1, \dots, \tau_k)$  Vektor der Merkmalstypen,

$\Lambda = (\lambda_1, \dots, \lambda_k)$  Vektor der Merkmalsgewichte,

$A = \{a_1, \dots, a_m\}$  und  $B = \{b_1, \dots, b_n\}$  wie in Kapitel 1 beschrieben.

Dabei ist  $T$  ein  $k$ -Tupel ganzer Zahlen mit den folgenden Eigenschaften:

$\tau_i = 0$  genau dann, wenn  $v_i$  das Identifikatormerkmal (z. B. Unternehmensregisternummer),

$\tau_i = 1$  genau dann, wenn  $v_i$  ein Blockmerkmal,

$\tau_i = 2$  genau dann, wenn  $v_i$  ein ordinales Merkmal und

$\tau_i = 3$  genau dann, wenn  $v_i$  ein nominales Merkmal ist.

Andernfalls wird  $v_i$  als metrisches Merkmal angenommen.

Mit  $\Lambda$  wird ein mit  $T$  verträgliches  $k$ -Tupel reeller Zahlen bezeichnet. Das heißt,  $\tau_i = 0$  impliziert  $\lambda_i = 0$  und  $\tau_i = 1$  impliziert  $\lambda_i = 1$ . Die verbleibenden Gewichte werden geeignet standardisiert, sodass die Gleichung  $\sum_{i=1}^k \lambda_i = 1$  erfüllt ist.

2. Berechnung der Gesamtdistanzen  $d_{ij} := d(a_i, b_j)$  für  $i = 1, \dots, m$  und  $j = 1, \dots, n$ .

3. Aufsteigende Sortierung der Gesamtdistanzen in eine Liste  $L$ .

4. Solange  $L$  nicht leer ist:

Betrachte das erste Element  $d_{ij}$  in  $L$  und ordne  $(a_i, b_j)$  zu.

Entferne alle Elemente  $d_{rs}$ , für welche  $r = i$  oder  $s = j$  gilt.

5. Output: Relative Häufigkeit der korrekten Zuordnungen.

### 3 Anwendung auf reale Daten

Dieses Kapitel enthält die Ergebnisse der Anwendung des oben beschriebenen Algorithmus auf reale Daten. Die Massenfischzüge werden sowohl mit formal anonymisierten Daten (Weglassen direkter Identifikatoren wie Name und Adresse) als auch mit Daten, die über unterschiedliche Methoden probeanonymisiert wurden, durchgeführt.

#### 3.1 Die Zieldaten und das Zusatzwissen

Als Zieldaten für den Massenfischzug wurden die Daten der Umsatzsteuerstatistik 2000 verwendet. Die Umsatzsteuerstatistik dient der Beurteilung der Struktur und Wirkungsweise der Umsatzsteuer und ihrer wirtschaftlichen Bedeutung. Aus der Beobachtung der Umsätze ergeben sich wertvolle Informationen für die Haushaltsplanungen und Steuerschätzungen des Bundes, der Länder und der Gemeinden. Die Umsatzsteuerstatistik ist nicht allein ein Instrument der Fiskal- und Steuerpolitik; sie dient darüber hinaus auch der allgemeinen Wirtschaftsbeobachtung. Mit ihren Angaben über die Entwicklung der Umsätze in allen Bereichen der Volkswirtschaft liefert sie Informationen, die in dieser Vollständigkeit in keiner anderen Bundesstatistik enthalten sind. Die Ergebnisse der Umsatzsteuerstatistik sind eine wichtige Datenbasis für die Erstellung der volkswirtschaftlichen Gesamtrechnungen. Aufgrund ihrer tiefen

5) Siehe Lenz, R.: "Disclosure of confidential information by means of multi objective optimisation", Comparative analysis of enterprise (micro) data conference, London 2003.

wirtschaftssystematischen Gliederung lassen sich mit Hilfe der Umsatzsteuerstatistik auch branchenspezifische Analysen sowie Konzentrationsuntersuchungen durchführen.

Die Umsatzsteuerstatistik ist eine Sekundärstatistik, die auf die Daten zurückgreift, die bei der Finanzverwaltung anfallen. Erfasst werden alle Unternehmen, die Umsatzsteuer-Voranmeldungen abgeben, mit einem Jahresumsatz (ohne Umsatzsteuer) von über 32 500 DM (16 617 Euro). In der Umsatzsteuerstatistik für das Jahr 2000 sind rund 2,9 Mill. Unternehmen erfasst.

Für die Durchführung der Reidentifikationsexperimente wurden die Merkmale Wirtschaftsweignummer, Umsatz, Rechtsform und Regionalkennung<sup>6)</sup> als Schlüssel- bzw. Überschneidungsmerkmale verwendet.

Das verwendete Zusatzwissen enthielt knapp 9 300 Unternehmen mit 20 oder mehr Beschäftigten aus den Abteilungen 10 bis 37 (Abschnitt C „Bergbau und Gewinnung von Steinen und Erden“ sowie D „Verarbeitendes Gewerbe“) der Klassifikation der Wirtschaftszweige, Ausgabe 1993 (WZ 93). Rund 37 000 Unternehmen wiesen seitens der Umsatzsteuerstatistik diese Charakteristika auf, sodass diese prinzipiell als gesuchte Unternehmen in Frage kamen. Daraus ergab sich das Ziel, die 9 300 Unternehmen, für die Zusatzwissen vorlag, innerhalb der 37 000 Unternehmen der Zieldaten zu reidentifizieren (siehe die Übersicht „Prinzip des Massenfischzuges“).

Prinzip des Massenfischzuges

Zusatzwissen (9 300 Unternehmen)		
Identifikatoren (Name, Anschrift)	Umsatz, Rechtsform, Regionalkennung, Wirtschaftszweignummer	Zielmerkmale (z.B. Umsatzwachstum 1999 bis 2000)
	Umsatz, Rechtsform, Regionalkennung, Wirtschaftszweignummer	
Zieldaten (37 000 Unternehmen)		

In Anlehnung an Höhne, J./Sturm, R./Vorgrimler, D.: „Konzept zur Beurteilung der Schutzwirkung von faktischer Anonymisierung“ in WiSta 4/2003, S. 288.

Für die Massenfischzüge wurden die kategorialen Merkmale unterschiedlich anonymisiert. So wurden Massenfischzüge auf der Ebene des WZ-Vier- bis Einstellers<sup>7)</sup> durchgeführt und es wurden Massenfischzüge simuliert, bei denen ganz auf eine Untergliederung der Abschnitte C und D verzichtet wurde. Die Rechtsform wurde einmal mit acht Ausprägungen und einmal mit vier Ausprägungen verwendet. Der Regionalschlüssel wies neun, sieben oder drei Ausprägungen auf.

Die verwendeten Umsatzsteuerstatistikdaten wurden darüber hinaus nach vier unterschiedlichen Verfahren anonymisiert:

- Das erste Verfahren war die formale Anonymisierung als schwächste Anonymisierungsform. Dabei werden lediglich die direkten Identifikatoren (Name, Anschrift) gelöscht.
- Die zweite Anonymisierung beschränkte sich auf traditionelle Anonymisierungsmaßnahmen wie zum Beispiel Vergrößerungen von Merkmalen (siehe Abschnitt 3.1.2).
- Das dritte Verfahren war die schwächste Form der Anonymisierung durch Mikroaggregation. Dabei wird jedes stetige Merkmal separat mikroaggregiert (siehe Abschnitt 3.1.3).
- Die vierte Art beinhaltete dagegen die stärkste Form der Mikroaggregation, bei der sämtliche stetigen Merkmale gemeinsam behandelt werden (siehe Abschnitt 3.1.3).

### 3.1.1 Natürlicher Schutz der Zieldaten

In diesem Beitrag werden zwei Arten des Schutzes analysiert: zum einen der durch die Anonymisierungsmaßnahme bewirkte Schutz und zum anderen der Schutz, der dadurch entsteht, dass die Merkmalsausprägungen bereits zwischen formal anonymisierten Zieldaten und Zusatzwissen für einen Merkmalsträger erheblich voneinander abweichen können. Dies kann man als natürlichen Schutz der Daten gegen Reidentifikation bezeichnen. Tabelle 1 zeigt den Grad der Übereinstimmung zwischen beiden Erhebungen für die Merkmale Umsatz und Wirtschaftsklassifikation. Es wird beispielsweise ersichtlich, dass bei fast 30% (20%) aller

Tabelle 1: Übereinstimmungen in den Merkmalsausprägungen zwischen Zieldaten und Zusatzwissen

Gegenstand der Nachweisung	Unternehmen	
	Anzahl	%
mit Abweichungen in der Ausprägung des Merkmals Umsatz		
Abweichung geringer als ...		
1% .....	3 546	38,2
5% .....	6 541	70,5
10% .....	7 539	81,2
25% .....	8 395	90,4
50% .....	8 706	93,8
Insgesamt ...	9 283	100
mit identischer WZ 93 <sup>1)</sup> -Klassifizierung auf der Ebene der		
4-Steller .....	5 206	56,1
3-Steller .....	5 917	63,7
2-Steller .....	7 007	74,5
1-Steller .....	7 823	84,3
Insgesamt ...	9 283	100

1) Klassifikation der Wirtschaftszweige, Ausgabe 1993. Als 1-Steller wird hier die Zehnerstelle des Codes für die Abteilungen (Zweisteller) bezeichnet.

6) Als Regionalkennung wurden die siedlungsstrukturellen Kreistypen verwendet. Diese nichtadministrativen Schlüssel dienen dem intraregionalen Vergleich. Es wird nach „Kernstädten“ und sonstigen Kreisen bzw. Kreisregionen unterschieden. Als Kernstädte werden kreisfreie Städte mit mehr als 100 000 Einwohnern ausgewiesen. Kreisfreie Städte unterhalb dieser Größe werden mit ihrem Umland zu Kreisregionen zusammengefasst. Die Typisierung der Kreise und Kreisregionen erfolgt außerhalb der Kernstädte nach der Bevölkerungsdichte. Um den großräumigen Kontext zu berücksichtigen, wird dann weiter nach der Lage im siedlungsstrukturellen Regionstyp differenziert. Mit dieser Einordnung wird der Überlegung Rechnung getragen, dass die Lebensbedingungen in den Kreisen sowie ihre Entwicklung wesentlich auch von der Entwicklung und der Struktur der jeweiligen Region bzw. des Regionstyps abhängig sind. Insgesamt ergeben sich neun Kreistypen, die im Folgenden unter dem Merkmal BBR9 angegeben werden. Nach einer Zusammenfassung auf regionale Grundtypen ergeben sich drei Ausprägungen, die im Folgenden unter dem Merkmal BBR3 ausgewiesen werden (siehe Bundesamt für Bauordnung und Raumwesen, [www.bbr.bund.de/raumordnung/raumbearbeitung/gebietstypen2.htm](http://www.bbr.bund.de/raumordnung/raumbearbeitung/gebietstypen2.htm)).

7) Als „Einsteller“ wird hier die Zehnerstelle des (zweistelligen) Codes für die Abteilungen in der WZ 93 bezeichnet (mit den Ausprägungen 1, 2 und 3).

Unternehmen der Umsatzwert im Zusatzwissen um mindestens 5% (10%) vom Umsatzwert in der Umsatzsteuerstatistik abweicht. Des Weiteren zeigt die Tabelle, dass nur 56% aller Unternehmen die gleiche vierstellige Wirtschaftszweignummer in Zusatzwissen und Zieldaten aufweisen.

Darüber hinaus ist noch von Bedeutung, dass etwa 2% der Unternehmen Unterschiede in der Regionalkennung aufweisen. Aufgrund dieser zum Teil mangelnden Übereinstimmung ist es von vornherein unmöglich, alle im Zusatzwissen vorhandenen Unternehmen richtig zuzuordnen bzw. zu reidentifizieren. Die aus diesem Grund nicht auffindbaren Unternehmen sind durch den natürlichen Schutz bereits ausreichend geschützt. Tabelle 2 enthält die Anzahl der geschützten und ungeschützten Unternehmen in Abhängigkeit von der verwendeten Gliederungstiefe der Wirtschaftszweigklassifikation. Es zeigt sich, dass beispielsweise bei einem Reidentifikationsversuch, der auf der Ebene der Vierstelliger durchgeführt wird, rund 45% aller Unternehmen von vornherein geschützt sind (37% bei Verwendung der Dreistelliger der Wirtschaftszweigklassifikation usw.).

Tabelle 2: Durch Abweichungen geschützte und ungeschützte Unternehmen

Auf Ebene der WZ 93 <sup>1)</sup>	Ungeschützte Unternehmen		Geschützte Unternehmen	
	Anzahl	%	Anzahl	%
4-Steller .....	5 120	55,2	4 163	44,8
3-Steller .....	5 812	62,6	3 471	37,4
2-Steller .....	6 878	74,1	2 405	25,9
1-Steller .....	7 673	82,7	1 610	17,3
0-Steller .....	9 097	98,0	186	2,0
Insgesamt <sup>2)</sup> ...	9 283	100	0	0,0

1) Klassifikation der Wirtschaftszweige, Ausgabe 1993. Als 1-Steller wird hier die Zehnerstelle des (zweistelligen) Codes für die Abteilungen bezeichnet; der 0-Steller steht für die einbezogenen Abschnitte C „Bergbau und Gewinnung von Steinen und Erden“ und D „Verarbeitendes Gewerbe“ ohne weitere Untergliederung. – 2) Ohne Abweichungen im Regionalschlüssel.

### 3.1.2 Traditionelle Anonymisierung

Bei einer der betrachteten Probeanonymisierungen der Zieldaten wurden die Merkmalsträger durch traditionelle Methoden anonymisiert. Dabei wurden folgende Maßnahmen durchgeführt:

- Kürzung der Wirtschaftszweignummer auf zwei Stellen (Abteilungen),
- jede Abteilung musste mit mindestens 3 500 Unternehmen besetzt sein, weshalb kleinere Abteilungen zusammengefasst wurden,
- Vergrößerung der Rechtsform auf vier Ausprägungen (anstelle von acht),
- Topcoding des Umsatzes unter Berücksichtigung zweier Abschneidegrenzen (500 Mill. Euro und 1 Mrd. Euro Umsatz). Die Umsätze der Unternehmen oberhalb der Abschneidegrenze wurden durch den Durchschnitt der Umsätze aller Unternehmen oberhalb der Abschneidegrenze ersetzt (Replacement),

- Runden des Umsatzes bei Unternehmen mit bis zu 500 Mill. Euro Umsatz.

### 3.1.3 Anonymisierung durch Mikroaggregation

Erfahrungen früherer Untersuchungen anhand der Kostenstrukturerhebung zeigen, dass es mit Hilfe der (mehrdimensionalen) Mikroaggregation möglich erscheint, anonymisierte Daten zu erstellen, die genügend Analysepotenzial beinhalten<sup>8)</sup>. Die Mikroaggregation unterteilt zunächst die metrischen Merkmale in Gruppen, wobei anzuraten ist, hoch korrelierte Merkmale zusammen zu gruppieren. Innerhalb der Gruppen werden die Merkmale standardisiert und für jeden Merkmalsträger aufsummiert, sodass die Merkmalsträger nach diesen so genannten Z-scores geordnet werden können. Im nächsten Schritt werden für eine vorgegebene ganze Zahl  $k$  (hier  $k=3$ ) die Merkmalsträger mit dem größten bzw. kleinsten Z-score mit ihren  $k-1$  nächsten Nachbarn (bei mehrdimensionaler Mikroaggregation bezüglich der euklidischen Norm) zusammengelegt und bei den so entstandenen  $k$ -Tupeln (hier: Tripeln) merkmalsweise die Werte gemittelt.

Um eine Ober- und eine Untergrenze für das mit der Mikroaggregation verbundene Reidentifikationsrisiko bestimmen zu können, werden im nächsten Abschnitt die schwächste und die stärkste Variante der Mikroaggregation auf die Zieldaten angewendet. Während bei der schwächsten Variante (MA21G) jedes metrische Merkmal seine eigene Gruppe definiert, werden bei der stärksten Variante (MA1G) alle metrischen Merkmale zusammen gruppiert, sodass Tripel von Merkmalsträgern erzeugt werden, welche in allen metrischen Merkmalen übereinstimmen und sich lediglich in den kategorialen Merkmalen (wie z. B. Wirtschaftszweiguordnung oder Regionalkennung) unterscheiden können. In der Tat kann man bei der Variante MA1G große Abweichungen der anonymisierten von den Originaldaten beobachten. Mehr als 60% der veränderten Werte weichen um mehr als 10% von ihrem Originalwert ab. Bei der Variante MA21G hingegen werden die Originaldaten nur sehr geringfügig modifiziert. Hier weichen mehr als 99,9% der veränderten Werte um weniger als 5% von ihren Originalwerten ab. Dies gilt sogar für beinahe 90% der Unternehmen mit mehr als 500 Beschäftigten, welche bekanntermaßen besonders reidentifikationsgefährdet sind.

## 3.2 Der Massenfischzug

Bei den nun folgenden Ergebnissen der Massenfischzüge werden zwei Arten von Trefferquoten unterschieden: Bei der ersten Art werden die Treffer in Beziehung zu allen gesuchten Unternehmen gesetzt. Bei der zweiten Art dagegen werden die Treffer lediglich im Verhältnis zu den Unternehmen betrachtet, die nicht aufgrund des natürlichen Schutzes per se sicher sind (siehe Tabelle 2). Diese Trefferquote wird im Folgenden als „korrigierte Trefferquote“ bezeichnet. Als Blockmerkmale wurden die kategorialen Merkmale „Rechtsform“, „Wirtschaftszweiguordnung“ und „Regionalken-

8) Siehe Rosemann, M./Vorgrimler, D./Lenz, R.: „Erste Ergebnisse faktischer Anonymisierung wirtschaftsstatistischer Einzeldaten“ in Allgemeines Statistisches Archiv, Heft 1, 2004, S. 73 ff.

nung“ verwendet. Als metrisches Überschneidungsmerkmal stand das Merkmal „Umsatz“ zur Verfügung.

3.2.1 Ergebnisse in Abhängigkeit von der Tiefe der wirtschaftlichen Gliederung

Tabelle 3 zeigt die Trefferquoten in Abhängigkeit von der Tiefe der wirtschaftlichen Gliederung bis hin zum Verzicht auf eine Untergliederung der einbezogenen Abschnitte der WZ 93 (hier als „0-Steller“ bezeichnet). Letzteres kommt der Annahme gleich, dass ein Datenangreifer keinerlei Wissen über die ökonomischen Aktivitäten des gesuchten Unternehmens hat, außer dass es sich um ein Unternehmen des Bergbaus und der Gewinnung von Steinen und Erden oder des Verarbeitenden Gewerbes handelt.

Es wird offensichtlich, dass die schwächste Form der Mikroaggregation (MA21G) die Trefferquoten nicht reduzieren kann.<sup>9)</sup> Die Unterschiede zwischen den Erhebungen erschweren einen Datenangriff bereits so stark, dass eine zusätzliche minimale Veränderung der Werte durch MA21G bei einem Massenfischzug nicht ins Gewicht fällt. Betrachtet man die Ergebnisse des Matches auf Basis der 0-Steller, so erkennt man sogar einen enthüllenden Effekt der Mikroaggregation. Es wurden mehr Unternehmen nach der Anonymisierung durch MA21G gefunden als bei lediglich formaler Anonymisierung (1 270 gegenüber 1 259). MA1G dagegen generiert nahezu sichere Daten. Allerdings sei an dieser Stelle darauf hingewiesen, dass diese Form der Mikroaggregation das Analysepotenzial stark einschränkt.<sup>10)</sup> Die Trefferquoten bei Verwendung der vierstelligen und der dreistelligen Wirtschaftszweignummer ähneln sich bei den formal und den durch MA21G anonymisierten Daten sehr. Beim

Übergang von vier auf drei Stellen werden zwar die Blöcke, innerhalb derer versucht wird die Unternehmen zu reidentifizieren, größer (und damit auch das Risiko der Falschzuordnung, siehe Tabelle 5), andererseits nimmt die Anzahl derjenigen Unternehmen ab, die durch den natürlichen Schutz geschützt sind. Diese beiden gegenläufigen Effekte heben sich in diesem Fallbeispiel auf. Wird die Wirtschaftszweignummer aber weiter gekürzt, so kann ein Anstieg der Schutzwirkung beobachtet werden. Bei der traditionellen Anonymisierung zeigt sich, dass die großen Unternehmen hierbei besser geschützt werden als bei den anderen Verfahren. Dies liegt darin begründet, dass die Anonymisierungsmethoden in diesem Fall besonders auf die großen Unternehmen ausgerichtet waren (z. B. Topcodingverfahren beim Umsatz).

Tabelle 4 zeigt die korrigierten Trefferquoten. Diese Raten zeigen die reine Schutzwirkung aufgrund der getroffenen Anonymisierungsmaßnahmen für die metrischen Merkmale. Es zeigt sich einerseits wiederum, dass durch die Variante MA21G die Trefferquote nicht wesentlich reduziert wird. Andererseits wird in der Tabelle die Schutzwirkung einer Vergrößerung der Gliederung nach Wirtschaftszweigen deutlich, da dieser Effekt nicht mehr durch einen Verlust an natürlichem Schutz konterkariert wird. Daher ist erkennbar, dass die Trefferquoten beim Übergang vom Vier- auf den Dreisteller der WZ 93 zurückgehen. Die erste Zeile der Tabelle zeigt die Effektivität des verwendeten Matchingverfahrens. Ohne Inkompatibilitäten bei den Blockmerkmalen und ohne zusätzliche Anonymisierung kann der Algorithmus drei von vier Unternehmen richtig zuordnen. Bei den größten Unternehmen wurden sogar sämtliche Unternehmen richtig reidentifiziert.

Tabelle 3: Korrekt zugeordnete Unternehmen

Zieldaten	WZ 93 <sup>1)</sup>	Insgesamt		Beschäftigtengrößenklasse <sup>2)</sup>											
				1		2		3		4		5		6	
		Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Formal anonymisiert	4-Steller	3726	40,1	188	35,3	1764	35,7	1583	45,7	169	54,9	15	57,7	7	70
	3-Steller	3720	40,1	189	35,5	1781	36,1	1565	45,1	162	52,6	17	65,4	6	60
	2-Steller	3287	35,4	168	31,6	1557	31,5	1371	39,5	167	54,2	16	61,5	8	80
	1-Steller	1951	21,0	95	17,9	912	18,5	800	23,1	131	42,5	9	34,6	4	40
	0-Steller	1259	13,6	61	11,5	586	11,9	508	14,7	89	28,9	11	42,3	4	40
MA21G <sup>3)</sup>	4-Steller	3723	40,1	188	35,3	1763	35,7	1580	45,6	170	55,2	15	57,7	7	70
	3-Steller	3709	39,9	192	36,1	1785	36,1	1545	44,6	164	53,3	15	57,7	8	80
	2-Steller	3282	35,4	168	31,6	1558	31,5	1362	39,3	170	55,2	16	61,5	8	80
	1-Steller	1934	20,8	94	17,7	906	18,3	790	22,8	130	42,2	9	34,6	5	50
	0-Steller	1270	13,7	60	11,3	593	12,0	505	14,5	100	32,5	8	30,8	4	40
MA1G <sup>4)</sup>	4-Steller	2593	27,9	114	21,4	1076	21,8	1214	35,0	166	53,9	17	65,4	6	60
	3-Steller	2169	23,4	84	15,8	853	17,3	1043	30,1	162	52,6	19	73,1	8	80
	2-Steller	1332	14,4	37	6,9	471	9,5	656	18,9	144	46,8	18	69,2	6	60
	1-Steller	497	5,4	11	2,1	165	3,3	244	7,0	65	21,2	9	34,6	3	30
	0-Steller	241	2,6	5	0,9	73	1,5	117	3,4	36	11,7	7	26,9	3	30
Traditionell anonymisiert .....		2792	30,0	142	26,7	1335	27,0	1181	34,0	127	41,2	5	19,2	2	20

1) Klassifikation der Wirtschaftszweige, Ausgabe 1993. Als 1-Steller wird hier die Zehnerstelle des (zweistelligen) Codes für die Abteilungen bezeichnet; der 0-Steller steht für die einbezogenen Abschnitte C „Bergbau und Gewinnung von Steinen und Erden“ und D „Verarbeitendes Gewerbe“ ohne weitere Untergliederung. – 2) 1 = weniger als 25 Beschäftigte, 2 = 25 bis unter 100 Beschäftigte, 3 = 100 bis unter 1 000 Beschäftigte, 4 = 1 000 bis unter 5 000 Beschäftigte, 5 = 5 000 bis unter 15 000 Beschäftigte, 6 = 15 000 und mehr Beschäftigte. – 3) Schwächste Variante der Mikroaggregation. – 4) Stärkste Variante der Mikroaggregation.

9) Die Reduzierung der Trefferquoten ist aber nur einer der beiden in Höhe u. a. beschriebenen Aspekte der Schutzwirkung faktischer Anonymisierung. Eine Analyse, die beide Aspekte der Schutzwirkung berücksichtigt, beinhaltet der Beitrag von Lenz, R./Sturm, R./Vorgriemer, D.: „Maße für die faktische Anonymität von Mikrodaten“ auf S. 621 ff. in diesem Heft.  
 10) Zu den Auswirkungen der Anonymisierung auf das Analysepotenzial siehe Rosemann, M.: „Erste Ergebnisse von vergleichenden Untersuchungen mit anonymisierten und nicht anonymisierten Einzeldaten am Beispiel der Kostenstrukturerhebung und der Umsatzsteuerstatistik“ in Gnoss, R./Ronning, G. (Hrsg.): „Anonymisierung wirtschaftsstatistischer Einzeldaten“, Band 42 der Schriftenreihe „Forum der Bundesstatistik“, Wiesbaden 2003, S. 154 ff.

Tabelle 4: Korrigierte Trefferquoten  
Prozent

Zieldaten	WZ 93 <sup>1)</sup>	Insgesamt	Beschäftigtengrößenklasse <sup>2)</sup>					
			1	2	3	4	5	6
Formal anonymisiert	4-Steller .....	72,8	70,7	70,0	75,1	84,9	83,3	100
	3-Steller .....	64,0	62,0	61,0	66,7	76,4	81,0	75,0
	2-Steller .....	47,8	46,3	44,4	50,2	68,4	72,7	100
	1-Steller .....	25,4	22,8	23,0	26,7	49,1	39,1	50,0
	0-Steller .....	13,8	11,7	12,1	15,0	29,7	42,3	44,4
MA21G <sup>3)</sup>	4-Steller .....	72,7	70,7	69,9	75,0	85,4	83,3	100
	3-Steller .....	63,8	62,9	61,2	65,8	77,4	71,4	100
	2-Steller .....	47,7	46,3	44,4	49,9	69,7	72,7	100
	1-Steller .....	25,2	22,5	22,9	26,4	48,7	36,1	62,5
	0-Steller .....	14,0	11,5	12,2	14,9	33,3	30,8	44,4
MA1G <sup>4)</sup>	4-Steller .....	50,6	42,9	42,7	57,6	83,4	94,4	85,7
	3-Steller .....	37,3	27,5	29,2	44,4	76,4	90,5	100
	2-Steller .....	19,4	10,2	13,4	24,0	59,0	81,8	75,0
	1-Steller .....	6,5	2,6	4,2	8,2	24,4	39,1	37,5
	0-Steller .....	2,7	1,0	1,5	3,5	12,0	26,9	33,3
Traditionell anonymisiert .....		40,6	39,1	38,0	43,2	52,1	22,7	25,0

1) Klassifikation der Wirtschaftszweige, Ausgabe 1993. Als 1-Steller wird hier die Zehnerstelle des (zweistelligen) Codes für die Abteilungen bezeichnet; der 0-Steller steht für die einbezogenen Abschnitte C „Bergbau und Gewinnung von Steinen und Erden“ und D „Verarbeitendes Gewerbe“ ohne weitere Untergliederung. – 2) 1 = weniger als 25 Beschäftigte, 2 = 25 bis unter 100 Beschäftigte, 3 = 100 bis unter 1 000 Beschäftigte, 4 = 1 000 bis unter 5 000 Beschäftigte, 5 = 5 000 bis unter 15 000 Beschäftigte, 6 = 15 000 und mehr Beschäftigte. – 3) Schwächste Variante der Mikroaggregation. – 4) Stärkste Variante der Mikroaggregation.

Tabelle 5 zeigt, dass je größer einzelne Datenblöcke sind (d.h. je mehr Unternehmen eine bestimmte Kombination aus Rechtsform, Regionalkennung und Wirtschaftszweiguordnung erfüllen), desto geringer ist die Wahrscheinlichkeit, ein Unternehmen richtig zuzuordnen zu können. Dies spricht dafür, eine Mindestanzahl von Unternehmen je Gliederungsebene der Wirtschaftszweigklassifikation nicht zu unterschreiten.

Tabelle 5: Korrelation zwischen Trefferquoten und Besetzungszahlen der Wirtschaftszweige

Zieldaten	WZ 93 <sup>1)</sup>		
	4-Steller	3-Steller	2-Steller
Formal anonymisiert .....	-0,504	-0,683	-0,777
MA21G <sup>2)</sup> .....	-0,499	-0,665	-0,779
MA1G <sup>3)</sup> .....	-0,474	-0,644	-0,661
Traditionell anonymisiert ....			-0,413

1) Klassifikation der Wirtschaftszweige, Ausgabe 1993. – 2) Schwächste Variante der Mikroaggregation. – 3) Stärkste Variante der Mikroaggregation.

Tabelle 6 enthält ausgewählte beschreibende Statistiken der richtig zugeordneten Unternehmen. Zum Vergleich zeigt die letzte Zeile die Statistik aller Unternehmen des Zusatzwissens. Man erkennt, dass mit steigendem Anonymisierungsgrad die durchschnittliche Zahl der Beschäftigten der jeweils richtig zugeordneten Unternehmen steigt. Eine Ausnahme stellt die traditionelle Anonymisierung dar. Es zeigt sich daher auch bei dieser Betrachtung, dass mit Ausnahme der traditionellen Anonymisierung die Anonymisierungsmaßnahmen stärker bei den kleinen Unternehmen als bei den großen wirken. Betrachtet man allerdings die Größe des jeweils kleinsten gefundenen Unternehmens, dann erkennt man, dass auch die kleinsten Unternehmen richtig zugeordnet werden können.

Tabelle 6: Beschreibende Statistik der reidentifizierten Unternehmen (Zahl der Beschäftigten)

Zieldaten	WZ 93 <sup>1)</sup>	Identifikation des größten Unternehmens	Durchschnittliche Größe	Kleinstes Unternehmen	Standardabweichung	Unternehmen Anzahl
			Beschäftigte	Beschäftigte		
Formal anonymisiert	4-Steller	ja	431,3	20	4 326,3	3 726
	3-Steller	ja	421,8	20	4 279,4	3 720
	2-Steller	ja	488,6	20	4 673,4	3 287
	1-Steller	nein	445,5	20	5 469,0	1 951
	0-Steller	nein	626,0	20	6 897,7	1 259
MA21G <sup>2)</sup>	4-Steller	ja	433,2	20	4 328,9	3 723
	3-Steller	ja	443,3	20	4 396,3	3 709
	2-Steller	ja	491,5	20	4 677,9	3 282
	1-Steller	ja	532,4	20	8 924,5	1 934
	0-Steller	ja	644,5	20	5 883,1	1 270
MA1G <sup>3)</sup>	4-Steller	ja	561,3	20	5 117,5	2 593
	3-Steller	ja	688,4	20	5 743,1	2 169
	2-Steller	ja	936,6	20	7 126,3	1 332
	1-Steller	ja	1 200,6	21	8 924,5	497
	0-Steller	ja	1 946,6	20	12 750,1	241
Traditionell anonymisiert .....		nein	305,6	20	2 279,2	2 792
Alle Unternehmen des Zusatzwissens .....		-	295,8	20	2 888,7	9 283

1) Klassifikation der Wirtschaftszweige, Ausgabe 1993. Als 1-Steller wird hier die Zehnerstelle des (zweistelligen) Codes für die Abteilungen bezeichnet; der 0-Steller steht für die einbezogenen Abschnitte C „Bergbau und Gewinnung von Steinen und Erden“ und D „Verarbeitendes Gewerbe“ ohne weitere Untergliederung. – 2) Schwächste Variante der Mikroaggregation. – 3) Stärkste Variante der Mikroaggregation.

### 3.2.2 Effekte durch die Vergrößerung der Rechtsform

Im nächsten Schritt wurde die Rechtsform wie bei der traditionellen Anonymisierung auf vier Kategorien vergrößert und die Massenfischzüge bei den ansonsten formal anonymisierten und bei den mit der Mikroaggregation anonymisierten Daten wiederholt.<sup>11)</sup> Im Folgenden wird die Schutzwirkung dieser Vergrößerung beschrieben.

11) Eine Wiederholung der Massenfischzüge bei den mit traditionellen Maßnahmen anonymisierten Merkmalsträgern erübrigt sich, da dort diese Maßnahme von vornherein angewandt wurde. Daher enthält die Tabelle 7 auch keine Ergebnisse für die traditionelle Anonymisierung.

Tabelle 7: Korrekt zugeordnete Unternehmen absolut nach Vergrößerung der Rechtsform auf vier Kategorien

Zieldaten	WZ 93 <sup>1)</sup>	Insgesamt		Beschäftigtengrößenklasse <sup>2)</sup>											
				1		2		3		4		5		6	
		Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Formal anonymisiert	4-Steller	3 527	37,9	175	32,9	1 644	33,3	1 526	44,0	163	52,9	13	50,0	6	60
	3-Steller	3 501	37,7	175	32,9	1 656	33,5	1 492	43,0	156	50,7	15	70,0	7	70
	2-Steller	2 945	31,7	146	27,4	1 389	28,1	1 244	35,9	151	49,0	12	46,1	3	30
	1-Steller	1 587	17,1	75	14,4	718	14,5	669	19,3	115	37,3	8	30,8	2	20
	0-Steller	1 081	11,4	51	9,6	488	9,9	441	12,7	92	12,7	7	26,9	2	20
MA21G <sup>3)</sup>	4-Steller	3 527	37,9	175	32,9	1 645	33,3	1 525	44,0	163	52,9	13	50,0	6	60
	3-Steller	3 499	37,7	175	32,9	1 656	33,5	1 491	43,0	155	50,3	15	57,7	7	70
	2-Steller	2 941	31,7	146	27,4	1 393	28,2	1 236	35,6	151	49,0	11	42,3	4	40
	1-Steller	1 569	16,9	76	14,3	712	14,4	660	19,0	108	35,1	8	30,8	5	50
	0-Steller	1 055	11,4	52	9,8	484	9,8	426	12,3	81	26,3	7	26,9	5	50
MA1G <sup>4)</sup>	4-Steller	2 340	25,2	100	18,8	919	18,6	1 143	32,9	158	51,3	15	57,7	5	50
	3-Steller	1 881	20,3	69	13,0	690	14,0	957	27,6	141	45,8	17	65,4	7	70
	2-Steller	1 037	11,2	23	4,3	355	15,1	525	15,1	116	37,7	11	42,3	7	70
	1-Steller	317	3,4	7	1,3	89	1,8	160	4,6	49	15,9	6	23,1	6	60
	0-Steller	199	2,1	7	1,3	72	1,5	85	2,5	28	9,1	3	11,5	4	40

1) Klassifikation der Wirtschaftszweige, Ausgabe 1993. Als 1-Steller wird hier die Zehnerstelle des (zweistelligen) Codes für die Abteilungen bezeichnet; der 0-Steller steht für die einbezogenen Abschnitte C „Bergbau und Gewinnung von Steinen und Erden“ und D „Verarbeitendes Gewerbe“ ohne weitere Untergliederung. – 2) 1 = weniger als 25 Beschäftigte, 2 = 25 bis unter 100 Beschäftigte, 3 = 100 bis unter 1 000 Beschäftigte, 4 = 1 000 bis unter 5 000 Beschäftigte, 5 = 5 000 bis unter 15 000 Beschäftigte, 6 = 15 000 und mehr Beschäftigte. – 3) Schwächste Variante der Mikroaggregation. – 4) Stärkste Variante der Mikroaggregation.

Der Schutzeffekt, der mit einer Vergrößerung der Rechtsform einhergeht, ist deutlich schwächer ausgefallen als erwartet. Das kommt unerwartet, da die Rechtsform in beiden Erhebungen für die gemeinsamen Merkmalsträger identisch ausgewiesen war. Die Vergrößerung konnte daher ihre volle Schutzwirkung entfalten und wurde nicht – wie bei der Vergrößerung der wirtschaftlichen Gliederung – durch einen geringeren natürlichen Schutz konterkariert. Der Vergleich der Tabellen 3 und 7 verdeutlicht darüber hinaus die Beobachtung, dass eine Verringerung der Tiefe der wirtschaftlichen Gliederung eine höhere Schutzwirkung als die vorgeschlagene Vergrößerung der Rechtsform bewirkt.

Die durch die Vergrößerung der Rechtsform einhergehende leichte Verstärkung der Anonymisierung hat wiederum zur Folge, dass die durchschnittliche Größe der richtig zugeordneten Unternehmen ansteigt. Auch hier werden demnach besonders die kleineren Unternehmen geschützt.

### 3.2.3 Effekte durch die Vergrößerung der Regionalkennung

Neben der Vergrößerung der Rechtsform stellt die Vergrößerung der Regionalkennung eine weitere Anonymisierungsmaßnahme dar. Als Regionalkennung wurde wie erwähnt der siedlungsstrukturelle Kreistyp verwendet. Dieser besteht aus einer Hauptstufe mit drei Ausprägungen und einer Unterstufe, durch die der BBR9<sup>12)</sup> seine insgesamt neun Ausprägungen erhält. Durch den Verzicht auf die Unterstufe reduziert sich die Anzahl der Ausprägungen auf drei. Im Folgenden wird daher vom BBR3 gesprochen, wenn auf die Unterstufe verzichtet wird. Alternativ könnte auch der regionsstrukturelle Kreistyp verwendet werden. Dieser besitzt sieben Ausprägungen, also nur zwei weniger als der BBR9. Die Schutzwirkung, die durch eine solche „Vergrößerung“ entsteht, ist allerdings vernachlässigbar, sodass auf

eine weitere Diskussion auf Basis des BBR7 verzichtet werden kann.

Im Gegensatz zur Vergrößerung der Rechtsform wird die Vergrößerung auf den BBR3 auch auf die traditionelle Anonymisierung angewendet. Tabelle 8 enthält die Ergebnisse der Massenfischzüge.

Mit der Vergrößerung der Regionalkennung wird eine größere Anonymisierungswirkung erzielt als mit der vorhergehenden Vergrößerung der Rechtsform. Dies verdeutlicht die Bedeutung von Regionalkennungen bei einer versuchten Deanonymisierung. Ein Vergleich von Tabelle 3 und Tabelle 8 zeigt, dass eine Anonymisierung durch Reduzierung der Tiefe der wirtschaftlichen Gliederung von Vierstellern auf Zweisteller der Klassifikation der Wirtschaftszweige sowie durch Vergrößerung der Rechtsform und der Regionalkennung zu einem Rückgang der Trefferquoten von teilweise mehr als 40% führt, was zeigt, dass traditionelle Maßnahmen bei der Anonymisierung wirtschaftsstatistischer Einzeldaten eine wichtige Rolle spielen können. Dies deuteten bereits die Ergebnisse an, die bei den Massenfischzügen mit den traditionell anonymisierten Daten erzielt wurden. Allerdings sei darauf hingewiesen, dass die Vergrößerung der Regionalkennung bei der traditionellen Anonymisierung bei den großen Unternehmen sogar eine enthüllende Wirkung hatte, das heißt es konnten nach der Vergrößerung mehr Unternehmen reidentifiziert werden. Demnach kann bei solchen Simulationen auch der Zufall eine Rolle spielen.

### 3.3 Der Einzelangriff

Das Ziel eines Einzelangriffs ist die Gewinnung von Informationen über einen spezifischen Merkmalsträger. Dabei sammelt ein Datenangreifer aus verschiedenen externen

12) Zur Definition siehe Fußnote 6.

Tabelle 8: Korrekt zugeordnete Unternehmen nach Vergrößerung der Rechtsform und der Regionalkennung

Zieldaten	WZ 93 <sup>1)</sup>	Insgesamt		Beschäftigtengrößenklasse <sup>2)</sup>											
				1		2		3		4		5		6	
		Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Formal anonymisiert	4-Steller	3114	33,6	150	28,2	1425	28,9	1365	39,4	153	49,7	14	53,9	7	70,0
	3-Steller	2900	31,2	145	27,3	1334	27,0	1253	36,1	150	48,7	12	46,2	6	60,0
	2-Steller	2100	22,6	95	17,9	980	19,8	896	25,8	117	37,9	10	38,5	2	20,0
	0-Steller	889	9,6	44	8,3	386	7,8	376	10,8	74	24,0	7	26,9	2	20,0
MA21G <sup>3)</sup>	4-Steller	534	5,8	30	5,6	239	4,8	214	6,2	45	14,6	5	19,2	1	10,0
	3-Steller	3112	33,5	150	28,2	1425	28,9	1361	39,2	155	50,3	14	53,9	7	70,0
	2-Steller	2904	31,3	146	27,4	1339	27,1	1251	36,1	150	48,7	12	46,2	6	60,0
	0-Steller	2105	22,7	96	18,1	991	20,1	890	25,7	113	36,7	10	38,5	5	50,0
MA1G <sup>4)</sup>	4-Steller	897	9,7	42	7,9	396	8,0	374	10,8	70	22,7	9	34,6	6	60,0
	3-Steller	511	5,5	30	5,6	232	4,7	199	5,7	41	13,3	6	23,1	3	30,0
	2-Steller	1706	18,4	63	11,8	599	12,1	882	25,4	142	46,1	14	53,9	6	60,0
	0-Steller	1250	13,5	43	8,1	400	8,1	655	18,9	129	41,9	15	57,7	8	80,0
Traditionell anonymisiert .....	4-Steller	557	6,0	13	2,4	171	3,5	282	8,1	73	23,7	11	42,3	7	70,0
	3-Steller	139	1,5	2	0,4	40	0,8	63	1,8	24	7,8	4	15,4	6	60,0
	2-Steller	69	0,7	3	0,6	19	0,4	28	0,8	11	3,6	5	19,2	3	30,0
	0-Steller	2081	22,2	94	16,7	982	19,9	883	25,5	111	36,0	6	23,1	5	50,0

1) Klassifikation der Wirtschaftszweige, Ausgabe 1993. Als 1-Steller wird hier die Zehnerstelle des (zweistelligen) Codes für die Abteilungen bezeichnet; der 0-Steller steht für die einbezogenen Abschnitte C „Bergbau und Gewinnung von Steinen und Erden“ und D „Verarbeitendes Gewerbe“ ohne weitere Untergliederung. – 2) 1 = weniger als 25 Beschäftigte, 2 = 25 bis unter 100 Beschäftigte, 3 = 100 bis unter 1 000 Beschäftigte, 4 = 1 000 bis unter 5 000 Beschäftigte, 5 = 5 000 bis unter 15 000 Beschäftigte, 6 = 15 000 und mehr Beschäftigte. – 3) Schwächste Variante der Mikroaggregation. – 4) Stärkste Variante der Mikroaggregation.

Quellen Informationen über das gesuchte Individuum bzw. Unternehmen und versucht anschließend, mit diesem Wissen den gesuchten Merkmalsträger in den Zieldaten zu reidentifizieren. Ein solcher „Angriff“ wurde anhand der Umsatzsteuerstatistik simuliert. Dabei wurde versucht 15 Unternehmen zu reidentifizieren, die in der Erhebung lediglich formal anonymisiert enthalten sind. Die Überschneidungsmerkmale waren die Regionalkennung, die Wirtschaftszweiguordnung, die Rechtsform und die Umsätze der Jahre 2000 und 1999 (die Überschneidungsmerkmale standen aber nicht in jedem Fall zur Verfügung). Mit Hilfe dieser Überschneidungsmerkmale gelang es 6 der 15 Unternehmen eindeutig und richtig zu identifizieren.

Wie beim Massenfischzug besteht die Hauptfehlerquelle darin, dass die Ausprägungen des Zusatzwissens sehr deutlich von den Ausprägungen der Umsatzsteuerstatistik abweichen können, die Merkmalsträger daher bereits natürlich geschützt sind. Darüber hinaus kann man sogar feststellen, dass im Gegensatz zu anderen Statistiken<sup>13)</sup> die Umsatzsteuerstatistik dem Datenangreifer im Rahmen eines Einzelangriffs keine zusätzlichen Überschneidungsmerkmale gegenüber einem Massenfischzug bietet. Aus diesem Grund ist das Risiko der Reidentifikation für ein spezifisches Unternehmen der Umsatzsteuerstatistik in einem Einzelangriffsszenario nicht höher zu bewerten als im Szenario des Massenfischzugs.

## 4 Fazit und Ausblick

Im vorliegenden Beitrag wurden verschiedene Methoden der Anonymisierung wirtschaftsstatistischer Einzeldaten

dahingehend getestet, inwieweit sie geeignet erscheinen, Reidentifikationen zu verhindern. Hierzu wurde ein Algorithmus gewählt, der im Rahmen des EU-Projektes CASC (Computational Aspects of Statistical Confidentiality) entwickelt wurde und in näherer Zukunft allgemein zugänglich gemacht werden soll.<sup>14)</sup> Die Leistungsfähigkeit des gewählten Algorithmus zeigte sich bei Zuordnungsversuchen, bei denen auf Anonymisierungen verzichtet wurde und bei denen diejenigen Unternehmen nicht betrachtet wurden, die aufgrund von Dateninkompatibilitäten bereits einen natürlichen Schutz genießen. Dass dieser natürliche Schutz bereits einen wesentlichen Beitrag zur Erreichung faktischer Anonymität darstellt, konnte ebenfalls gezeigt werden.

Eine Anonymisierung der Merkmalsträger durch Vergrößerung der diskreten Merkmale wird teilweise durch die Aufhebung der Abweichungen zwischen Zusatzwissen und Zieldaten – wodurch der natürliche Schutz der Daten sinkt – konterkariert. So erweist sich eine Verringerung der Gliederungstiefe von vier auf drei Stellen der Klassifikation der Wirtschaftszweige als nicht sonderlich hilfreich. Erst die Reduzierung auf zwei Stellen führt zu einem signifikanten Rückgang der Zahl der richtig zugeordneten Unternehmen. Trotzdem zeigt sich, dass die Verwendung traditioneller Verfahren einen großen Beitrag zur Erreichung der faktischen Anonymität auch bei wirtschaftsstatistischen Einzeldaten leisten kann. Allerdings scheint es notwendig, diese Maßnahmen mit datenverändernden Verfahren zumindest bei den großen Unternehmen zu flankieren.

Auch wenn nur ein kleiner Ausschnitt der Umsatzsteuerstatistik auf die Vertraulichkeit der Daten hin getestet wurde, erscheinen die Ergebnisse geeignet für eine Verallgemei-

13) Siehe z. B. die Ergebnisse zur Kostenstrukturerhebung im Verarbeitenden Gewerbe; siehe Vorgrimler, D.: „Reidentifikationsmöglichkeiten am Beispiel eines konkreten Angriffsszenarios“ in Gnoss, R./Ronning, G. (Hrsg.): „Anonymisierung wirtschaftsstatistischer Einzeldaten“, Band 42 der Schriftenreihe „Forum der Bundesstatistik“, Wiesbaden 2003, S. 40 ff.

14) Siehe Lenz, R.: „A graph theoretical approach to record linkage“, Joint UN-ECE/Eurostat work session on statistical data confidentiality, Luxemburg 2003.

nerung auf die gesamte Erhebung zu sein. Dabei spielen zwei Gründe eine wesentliche Rolle. Zunächst einmal sind fast alle sehr großen Unternehmen in dem betrachteten Ausschnitt enthalten und die durchschnittliche Größe der betrachteten Unternehmen liegt signifikant höher als die der nicht betrachteten (durchschnittlich 4,3 Mill. gegenüber 1 Mill. Euro Umsatz). Da sich die Annahme bestätigt hat, dass die größeren Unternehmen stärker gefährdet sind als die kleineren, ist davon auszugehen, dass die Trefferquoten sinken, sobald es gelingt, sämtliche Unternehmen in die Betrachtung mit einzubeziehen. Ebenfalls hat sich gezeigt, dass Unternehmen in geringer besetzten Wirtschaftszweigen einem erhöhten Risiko ausgesetzt sind. In dem betrachteten Ausschnitt sind nicht nur sieben der zehn am dünnsten besetzten Abteilungen enthalten, die durchschnittliche Besetzung ist mit 12 000 Unternehmen auch deutlich geringer als bei den übrigen Abteilungen, die mit durchschnittlich 94 000 Unternehmen besetzt sind. Dies wird ebenfalls dazu führen, dass die Trefferquoten bei einer Betrachtung der gesamten Umsatzsteuerstatistik geringer sein werden, als dies bei dem betrachteten Ausschnitt der Fall war. Für die Vertraulichkeit der Daten heißt das, dass man bei einer Verallgemeinerung der ermittelten Trefferquote das Risiko eher über- als unterschätzt.

Aufgrund dieser möglichen Verallgemeinerung liefern die Ergebnisse einen wichtigen Schritt hin zu einem Scientific Use File für die Umsatzsteuerstatistik. Allerdings sind die Ergebnisse für die Sicherheitsanalyse noch unvollständig, da bisher nur die Möglichkeit einer Zuordnung betrachtet wurde. Der Nutzen, den ein Angreifer durch eine Zuordnung erzielen kann, wurde bisher ausgeblendet, spielt aber zur Erreichung der faktischen Anonymität eine wesentliche Rolle. Im Beitrag „Maße für die faktische Anonymität von Mikrodaten“ in diesem Heft auf S. 621 ff. wird auf diesen Aspekt näher eingegangen. Es sind allerdings nicht nur die Analysen zur Sicherheit unvollständig; ein Scientific Use File muss darüber hinaus noch ausreichend Analysepotenzial für die Datennutzer aufweisen. Nur wenn dies gegeben ist, macht es für die statistischen Ämter Sinn, der Wissenschaft faktisch anonyme Daten als Scientific Use Files anzubieten. [UU](#)

Dipl.-Soziologe Joachim Fritz, Dr. Josef Hartmann, Matthias Sacher, M.A.<sup>1)</sup>

# Arbeitsmarkt in Deutschland – Analysen zu Ausfallprozessen in der Piloterhebung zum ILO-Erwerbsstatus

*In der seit April 2003 vom Statistischen Bundesamt in Zusammenarbeit mit TNS Infratest Sozialforschung unter dem Namen „Arbeitsmarkt in Deutschland“ durchgeführten Piloterhebung zum Erwerbsstatus der Bevölkerung sind – wie in anderen sozialwissenschaftlichen Erhebungen ohne Auskunftspflicht – Ausfallprozesse zu beobachten. Dabei spielen insbesondere soziodemographische Merkmale der Befragten, wie Geschlecht, Alter, Staatsangehörigkeit und Bildung, sowie die Größe des Wohnortes eine Rolle. Zur Qualitätssicherung der Erhebung ist es unerlässlich, die Ausfallprozesse mit statistischen Methoden zu untersuchen und im Zeitverlauf zu kontrollieren. Die Erkenntnisse aus der Analyse der Ausfallprozesse fließen in das Gewichtung- und Hochrechnungsverfahren ein, das sowohl die Präzision der Ergebnisse erhöhen als auch Verzerrungen der Stichprobe nach Möglichkeit korrigieren soll.*

*Der vorliegende Beitrag widmet sich zum einen der Analyse der Ausfallprozesse, insbesondere unter Nutzung des Panelcharakters der Erhebung, zum anderen skizziert er das Hochrechnungs- und Gewichtungsverfahren, welches für diese Erhebung entwickelt wurde. Für die Analyse der Ausfallprozesse werden sowohl bivariate als auch multivariate Verfahren eingesetzt.*

## Vorbemerkung

Seit April 2003 wird vom Statistischen Bundesamt in Zusammenarbeit mit TNS Infratest Sozialforschung unter dem

Namen „Arbeitsmarkt in Deutschland“ eine Piloterhebung zum Erwerbsstatus der Bevölkerung nach den Definitionen der Internationalen Arbeitsorganisation (ILO)<sup>2)</sup> durchgeführt. Ziel der Piloterhebung ist es, methodische Erfahrungen mit der Erhebung des ILO-Erwerbsstatus mittels einer ausschließlich telefonischen Befragung ohne Auskunftspflicht zu sammeln. Seit April 2003 werden monatlich 10 000 Personen telefonisch zum ILO-Erwerbsstatus und zu weiteren Merkmalen ihrer Erwerbsbeteiligung befragt.

Um die monatliche Veränderung des Erwerbsstatus hinreichend genau messen zu können, ist die Erhebung als rotierendes Panel mit sechsmaliger Befragung und einer monatlichen Nettoüberlappung von mindestens 70% angelegt. Die zeitliche Belastung der Befragten konnte durch ein zwischen Erst- und Wiederholungsbefragung unterscheidendes Fragebogenkonzept minimiert werden.<sup>3)</sup>

Aus den bisher vorliegenden Daten wurden die Daten der ersten zwölf Erhebungsmonate – das heißt die Monate April 2003 bis einschließlich März 2004 – ausgewählt, um statistische Analysen zu den Ausfallprozessen durchzuführen.

Der Aufsatz gliedert sich folgendermaßen: Zunächst wird mit Hilfe bivariater statistischer Verfahren der Frage nachgegangen, ob bei dem gewählten Stichprobenverfahren alle Personen der interessierenden Grundgesamtheit eine Auswahlchance haben. Der Vergleich der Ergebnisse mit externen Referenzstatistiken liefert erste Hinweise auf die Ausfallprozesse der Erhebung.

1) Joachim Fritz und Matthias Sacher sind wissenschaftliche Mitarbeiter beim Statistischen Bundesamt, Dr. Josef Hartmann ist Projektleiter bei TNS Infratest Sozialforschung, München.

2) International Labour Organization, Genf.

3) Zur Konzeption der Erhebung siehe Riede, T./Sacher, M.: „Arbeitsmarkt in Deutschland – erster Baustein der neuen ILO-Statistik“ in WiSta 2/2004, S. 148 ff.

Daran anschließend wird unter Verwendung von bivariaten und multivariaten Verfahren der Panelcharakter der Befragung genutzt, um die Ausfallprozesse noch besser analysieren und verstehen zu können.

Abschließend wird kurz auf das Gewichtung- und Hochrechnungsverfahren der Erhebung eingegangen, das die vorher beschriebenen Erkenntnisse aufnimmt und u. a. dazu dient, die vorgefundenen Ausfallprozesse nach Möglichkeit zu korrigieren.

## 1 Die Stichprobe der Untersuchung

### 1.1 Abdeckung der Grundgesamtheit durch die Stichprobe

Die Grundgesamtheit der Erhebung ist die in Deutschland in Privathaushalten lebende Bevölkerung im Alter zwischen 15 und 74 Jahren. Da die aus der Stichprobe gewonnenen Ergebnisse für die Grundgesamtheit aussagekräftig sein sollen, ist es wichtig, dass durch das gewählte Stichprobenverfahren alle Personen der Grundgesamtheit eine angebbare Auswahlchance besitzen.

Die Befragung findet ausschließlich telefonisch statt und es besteht keine Auskunftspflicht, das heißt die Teilnahme an der Befragung ist freiwillig. Die Ziehung der Stichprobe erfolgt anhand des am Zentrum für Umfragen, Methoden und Analysen (ZUMA) in Mannheim entwickelten „Gabler-Häder-Verfahrens“, das gewährleistet, dass sowohl eingetragene als auch nicht eingetragene Festnetztelefonnummern eine Auswahlchance erhalten.<sup>4)</sup> Durch die Stichprobenziehung haben prinzipiell alle Personen eine Auswahlchance, die in einem Haushalt mit mindestens einem Festnetzanschluss leben.

Trotz der wachsenden Verbreitung von Mobiltelefonen haben heute fast 95% der Haushalte in Deutschland einen Festnetzanschluss, wobei zu berücksichtigen ist, dass der Ausstattungsgrad mit Festnetzanschlüssen nach Haushaltstypen variiert. Nach der Einkommens- und Verbrauchsstichprobe (EVS) 2003 ist der Ausstattungsgrad bei Paaren mit zwei Kindern mit 98,4% am größten. Während 93,6% der allein lebenden Frauen einen Festnetzanschluss haben, ist dies bei allein lebenden Männern nur bei 84,1% der Fall. Bei Haushalten mit einem Haupteinkommensbezieher zwischen 65 und 70 Jahren liegt der Ausstattungsgrad bei 97,3%, bei Haushalten von unter 25-Jährigen hingegen

nur bei 74%.<sup>5)</sup> Die Befunde zeigen, dass vor allem jüngere heute öfter auf einen Festnetzanschluss verzichten.

Neben Personen ohne Festnetzanschluss haben allerdings einige spezifische Bevölkerungsgruppen systematisch keine oder nur eine sehr geringe Auswahlchance – insbesondere, weil sie schwer zu erreichen sind. Hierzu zählen „Personen mit ungewöhnlichen Arbeits-, Lebens- und Wohnsituationen“<sup>6)</sup>, die aber auch auf dem Arbeitsmarkt aktiv werden können, so zum Beispiel Obdachlose und Nichtsesshafte oder auch das Personal von Binnenschiffen, Personen aus dem Schaustellergewerbe sowie Personen, die werktags auf Baustellen oder in Bauzügen wohnen. Schließlich kann es aber auch vorkommen, dass Personen in Anstaltshaushalten in die Stichprobe geraten, sofern sie über einen eigenen Festnetztelefonanschluss verfügen.

### 1.2 Ausschöpfungsquote und Nettoüberlappung

Die Bereitschaft, an Befragungen der Sozialforschung teilzunehmen, ist in Deutschland seit Jahren rückläufig.<sup>7)</sup> Eine immer schlechter werdende Erreichbarkeit insbesondere jüngerer, allein lebender Personen in größeren Gemeinden sowie eine abnehmende Kooperationsbereitschaft, möglicherweise auch zunehmende Viktimisierungängste, werden als hauptsächliche Gründe rückläufiger Ausschöpfungsquoten genannt.<sup>8)</sup>

Die Piloterhebung liegt mit einer Ausschöpfungsquote der erstmalig Befragten zwischen 41% und 49% (siehe Tabelle 1) im Bereich dessen, was heute bei sozialwissenschaftlichen Erhebungen ohne Auskunftspflicht in Deutschland üblicherweise zu erreichen ist.<sup>9)</sup> Bei den Wiederholungsbefragten lag die Antwortquote zwischen 83% und 91%. Die große Mehrheit der Befragten, die sich einmal dazu bereit erklärt hat, an der Befragung teilzunehmen, ist auch bereit, an den Folgebefragungen teilzunehmen. Im Durchschnitt konnte bei mehr als 90% der Befragten eine Zustimmung zu einem Wiederanruf in den folgenden Monaten erreicht werden. Umgekehrt bedeutet dies, dass weniger als 10% der Befragten die Teilnahme über mehrere Erhebungswellen prinzipiell abgelehnt haben. Dies spricht für die Akzeptanz der Erhebung.

Die kurze Interviewdauer von durchschnittlich sieben Minuten bei den Erstbefragten und nur zwei Minuten bei den Wiederholungsbefragten trägt sicherlich zur hohen Akzeptanz der Erhebung bei.

4) Siehe hierzu Fußnote 3, S. 152 ff.; Gabler, S./Häder, S.: „Erfahrungen beim Aufbau eines Auswahlrahmens für Telefonstichproben in Deutschland“, ZUMA-Nachrichten 44, 1999, S. 45 ff., sowie Gabler, S./Häder, S. (Hrsg.): „Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland“, Münster 2002.

5) Siehe hierzu Statistisches Bundesamt (Hrsg.): „Ausstattung und Wohnsituation privater Haushalte. Einkommens- und Verbrauchsstichprobe 2003“, Presseexemplar, Wiesbaden 2003; Deckl, S./Krebs, T.: „Ausstattung mit Gebrauchsgütern und Wohnsituation privater Haushalte“ in WiSta 2/2004, S. 209 ff.

6) Siehe Schnell, R.: „Wer ist das Volk? Zur faktischen Grundgesamtheit bei „allgemeinen Bevölkerungsumfragen““ in Kölner Zeitschrift für Soziologie und Sozialpsychologie 43, 1991, Ausgabe 1, S. 106 ff., hier S. 125 f.

7) Siehe hierzu beispielsweise Schneekloth, U./Leven, I.: „Woran bemisst sich eine ‚gute‘ allgemeine Bevölkerungsumfrage? Analysen zu Ausmaß, Bedeutung und zu den Hintergründen von Nonresponse in zufallsbasierten Stichprobenerhebungen am Beispiel des ALLBUS“, ZUMA-Nachrichten 53, 2003, S. 16 ff.

8) Siehe hierzu auch die umfangreiche Monografie von Schnell, R.: „Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen“, Opladen 1997. Freilich zeigt Däubler für die Stichprobe F des SOEP, dass auf die Vermutung, „dass auf Grund von Viktimisierungängsten die Tür nicht geöffnet wurde“ keine Hinweise feststellbar sind (siehe Däubler, T.: „Nonresponseanalysen der Stichprobe F des SOEP“, Berlin 2002, DIW-Materialien 15, S. 42).

9) Beim ALLBUS (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften) 2000 lag die Ausschöpfungsquote beispielsweise bei 46,9% im früheren Bundesgebiet ohne Berlin und 53,7% in den neuen Ländern mit Berlin. Beim ALLBUS werden allerdings persönliche Interviews durchgeführt und die Feldzeiten sind wesentlich länger. Hierdurch können i. d. R. höhere Antwortquoten erreicht werden. Zur Methodik des ALLBUS siehe Koch, A./Wasmer, M./Harkness, J./Scholz, E.: „Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2000“, ZUMA-Methodenbericht 2001/05, Mannheim 2001.

In den Ferienmonaten im Sommer und im Dezember waren die Befragten erwartungsgemäß schwieriger zu erreichen, sodass die Ausschöpfungsquoten in diesen Monaten etwas niedriger lagen.

Für die Messung der monatlichen Veränderung der Erwerbslosen- und Erwerbstätigenzahlen ist eine möglichst hohe Nettoüberlappung im Panel wichtig, da sich auf diese Weise der Standardfehler der Messung verringern lässt. Angestrebt wird eine monatliche Nettoüberlappung von mindestens 70%, das heißt wenigstens 70% der Befragten eines Monats sollen auch im Folgemonat befragt werden. Diese Vorgabe konnte bislang in allen Befragungsmonaten eingehalten und in den meisten Monaten sogar deutlich überschritten werden (siehe Tabelle 1).

Tabelle 1: Ausschöpfungsquoten<sup>1)</sup> der Erst- und Wiederholungsbefragten und Nettoüberlappung<sup>2)</sup> des Panels  
Ungewichtete Ergebnisse in Prozent

Jahr Monat	Ausschöpfungs- quote <sup>1)</sup> Erstbefragte	Ausschöpfungs- quote <sup>1)</sup> Wiederholungs- befragte	Nettoüber- lappung <sup>2)</sup> des Panels
2003			
April .....	47,2	X	X
Mai .....	41,4	82,7	82,6
Juni .....	44,2	84,6	74,2
Juli .....	43,4	85,1	71,6
August .....	42,6	84,8	76,0
September .....	44,8	88,8	75,9
Oktober .....	46,4	90,5	76,8
November .....	46,7	90,2	80,9
Dezember .....	42,3	89,1	76,3
2004			
Januar .....	46,6	90,9	77,5
Februar .....	48,7	89,7	78,9
März .....	46,6	89,7	76,3
Durchschnitt April 2003 bis März 2004 .....	45,1	87,8	77,0

1) Die Ausschöpfungsquote der Erst- und der Wiederholungsbefragten ist das Verhältnis der Zahl der durchgeführten Interviews (bei Wiederholungsbefragung: ohne so genannte fortgeschriebene Fälle, d.h. nicht erwerbstätige Personen im Alter von 65 Jahren und älter ohne Interesse an einer Erwerbstätigkeit, deren Angaben aus der Erstbefragung in die Folgezellen übernommen werden) zur Summe von durchgeführten Interviews und Ausfällen. – 2) Die Nettoüberlappung berechnet sich als das Verhältnis der Zahl der durchgeführten Wiederholungsbefragten einschließlich fortgeschriebener Fälle und ohne aus dem Panel herausrotierte Personen zur Anzahl der Befragten eines Monats insgesamt.

## 2 Vergleich der Ergebnisse mit anderen Statistiken

Niedrige Ausschöpfungsquoten müssen nicht bedeuten, dass sich damit auch die Qualität der Ergebnisse verschlechtert. Schneekloth und Leven legten kürzlich in einer Studie mit den Daten des ALLBUS schlüssig dar, dass sinkende Ausschöpfungsquoten nicht notwendig mit einer „zunehmenden systematischen Verzerrung einhergehen.“<sup>10)</sup> Dennoch ist es in jeder Erhebung für die Qualitätskontrolle und die Entwicklung eines angemessenen Gewichtung- und

Hochrechnungsverfahrens unerlässlich, die Ausfallprozesse zu analysieren.

Eine erste Möglichkeit, die Strukturen einer Stichprobenerhebung zu untersuchen, besteht im Vergleich der Ergebnisse mit den Randverteilungen anderer Statistiken.<sup>11)</sup> Die Gestaltung der Erhebung als rotierendes Panel bietet darüber hinaus die Möglichkeit, Ausfallprozesse im Längsschnitt zu betrachten und somit Informationen über die Zusammensetzung der Stichprobe als Ganzes zu erhalten.

Im Folgenden werden zunächst die frei hochgerechneten Ergebnisse der Piloterhebung mit den Randverteilungen anderer Statistiken verglichen. Für die Untersuchung der Ausfallprozesse der Erhebung sind frei hochgerechnete Ergebnisse besser geeignet als gebunden hochgerechnete Ergebnisse, weil bei gebundener Hochrechnung bereits bestimmte Strukturveränderungen vorgenommen werden.<sup>12)</sup>

### 2.1 Vergleich mit anderen Statistiken, in denen soziodemographische Merkmale erhoben werden

Für den Vergleich der frei hochgerechneten Ergebnisse der Piloterhebung mit den Ergebnissen anderer Statistiken, in denen soziodemographische Merkmale erhoben werden, werden die Ergebnisse der Piloterhebung im April 2003 genutzt, da in diesem Monat ausschließlich Erstbefragungen durchgeführt wurden und somit ein Vergleich der Ausfallmechanismen beim Erstkontakt mit denjenigen im Panel möglich ist.

Tabelle 2: Vergleich der frei hochgerechneten Ergebnisse der Piloterhebung mit denen der Bevölkerungsfortschreibung<sup>1)</sup>  
1 000

Gegenstand der Nachweisung	Piloterhebung April 2003	Bevölkerungs- fortschreibung 31.12.2002	Differenz
nach dem Geschlecht			
Männer .....	28 987	32 028	- 3 041
Frauen .....	34 913	31 872	+ 3 041
nach Altersgruppen			
von ... bis ... Jahren			
15 – 24 .....	8 334	9 514	- 1 180
25 – 34 .....	9 473	10 751	- 1 278
35 – 44 .....	15 719	14 012	+ 1 707
45 – 54 .....	13 617	11 301	+ 2 316
55 – 64 .....	10 508	10 103	+ 405
65 – 74 .....	6 248	8 217	- 1 969
nach der Staatsangehörigkeit			
Deutsche .....	61 334	57 901	+ 3 434
Nichtdeutsche .....	2 565	5 999	- 3 434

1) Die frei hochgerechneten Ergebnisse der Piloterhebung wurden dem Eckwert der Bevölkerungsfortschreibung 2002 angepasst (zur Bevölkerungsfortschreibung siehe Fachserie 1 „Bevölkerung und Erwerbstätigkeit“, Reihe 1 „Gebiet und Bevölkerung“, 1999, S. 14 f.). Die Befragten, die keine Angabe zu ihrem Alter oder ihrer Staatsangehörigkeit machten, wurden im entsprechenden Verhältnis auf die einzelnen Klassen verteilt.

10) Siehe Schneekloth, U./Leven, I., Fußnote 7, S. 49.

11) Siehe Schnell, R., 1997, Fußnote 6, S. 133 ff.

12) Die „freie Hochrechnung“ beinhaltet in diesem Fall eine so genannte Designgewichtung, welche die unterschiedlichen Auswahlwahrscheinlichkeiten einzelner Personen aufgrund der unterschiedlichen Anzahl von Zielpersonen und Telefonanschlüssen in Haushalten berücksichtigt. Bei der „gebundenen Hochrechnung“ erfolgt eine Anpassung an die Eckwerte von ausgewählten Referenzstatistiken.

Die frei hochgerechneten Ergebnisse der Piloterhebung weichen von den Strukturen der Bevölkerungsfortschreibung 2002 ab. Männer, ausländische Staatsbürger sowie jüngere (15 bis 34 Jahre) und auch ältere Personen (65 bis 74 Jahre) sind in der Piloterhebung seltener vertreten als in der Bevölkerungsfortschreibung. Besonders die Altersgruppe der 45- bis 54-Jährigen wurde in der Piloterhebung weitaus häufiger befragt, als man es aufgrund der Strukturen der Bevölkerungsfortschreibung erwarten würde. Es ist aus vielen Bevölkerungsumfragen bekannt, dass sowohl jüngere als auch ältere Personen seltener an Befragungen teilnehmen als Personen der mittleren Altersgruppe. Das Gleiche gilt – insbesondere bei Telefonbefragungen – für Personen mit ausländischer Staatsangehörigkeit im Vergleich mit Deutschen, wobei die Gründe hierfür sehr unterschiedlich sein können. Während Jüngere im Allgemeinen mobiler und daher schwieriger über einen Festnetzanschluss zu erreichen sind, sind ältere Menschen oft misstrauisch gegenüber Befragungen am Telefon. Hinzu kommt die bereits angesprochene Problematik, dass Jüngere häufiger als andere Altersgruppen nur über das Handy zu erreichen sind. Auch das Interesse am Thema der Befragung könnte sowohl bei älteren als auch bei jüngeren Personen geringer sein. Ausländische Mitbürger haben hingegen am Telefon öfter Verständigungsschwierigkeiten als Deutsche.

Ein weiteres wichtiges soziodemographisches Merkmal für die Untersuchung von Ausfallprozessen ist der Bildungsgrad der Befragten. Der Mikrozensus stellt in Deutschland die verlässlichste Quelle für Angaben über die Bildungsstruktur der Bevölkerung dar. Im Mikrozensus 2003 machten allerdings ungefähr 5% der Befragten keine Angabe zu ihrer Schulbildung. Da im Mikrozensus die Beantwortung der entsprechenden Fragen für Personen im Alter von 51 Jahren und älter freiwillig ist, lag der Anteil der Befragten, die ihren Bildungsabschluss nicht angaben, in dieser Altersgruppe besonders hoch.

In der Piloterhebung sind weniger Personen mit niedriger Bildung – Personen, die keinen Schulabschluss oder höchstens einen Haupt- oder Volksschulabschluss haben – vertreten als im Mikrozensus. Personen mit einem mittleren [mit Abschluss einer Polytechnischen Oberschule (POS) oder mit Realschulabschluss (Mittlere Reife)] oder hohen Bildungsabschluss [mit Fachhochschulreife oder mit allgemeiner oder fachgebundener Hochschulreife (Abitur)] sind hingegen in der Piloterhebung häufiger vertreten als im Mikrozensus. Würde man die Fälle ohne Angabe zur Schulbildung unter Berücksichtigung der Tatsache, dass Personen im Alter von 51 Jahren und älter häufiger einen niedrigeren Bildungsabschluss haben als jüngere Personen, nach den vorhandenen Bildungsstrukturen aufteilen, so würde der Unterschied zwischen Piloterhebung und Mikrozensus bei Personen mit niedriger Bildung noch stärker ausfallen, die absolute Differenz zwischen Piloterhebung und Mikrozensus würde in der mittleren und hohen Bildungsklasse hingegen abnehmen.

Auch aus anderen Telefonbefragungen ist bekannt, dass Personen mit niedriger Bildung seltener an Befragungen

teilnehmen als Personen mit höherer Bildung.<sup>13)</sup> Größere Kommunikationsschwierigkeiten – insbesondere am Telefon – und möglicherweise Viktimisierungsängste können hierfür Ursachen sein.

Die Anzahl der Personen in Ausbildung (Schüler/-innen, Studierende und Auszubildende) lag in der Piloterhebung im April 2003 nur geringfügig über den Ergebnissen des Mikrozensus. Die etwas höhere Zahl in der Piloterhebung geht hauptsächlich auf einen höheren Anteil von Studierenden unter den Befragten zurück. Der Anteil von Schülerinnen und Schülern sowie Auszubildenden unter den Befragten entspricht sehr gut den Strukturen des Mikrozensus.

Tabelle 3: Vergleich der frei hochgerechneten Ergebnisse der Piloterhebung mit denen des Mikrozensus hinsichtlich der Bildungsstruktur<sup>1)</sup>  
1 000

Gegenstand der Nachweisung	Piloterhebung April 2003	Mikrozensus 2003	Differenz
höchster Schulabschluss			
Kein Schulabschluss oder Haupt-/Volksschulabschluss .....	18 475	25 882	-7 406
Abschluss Polytechnische Oberschule (POS) oder Realschulabschluss (Mittlere Reife) .....	21 207	15 617	+5 590
Fachhochschulreife, allgemeine oder fachgebundene Hochschulreife (Abitur) .....	16 094	11 392	+4 702
in Ausbildung			
Schüler/-innen, Studierende, Auszubildende .....	7 785	7 457	+328
Keine Angabe .....	176	3 286	-3 110

1) Bevölkerung in Privathaushalten im Alter von 15 bis 74 Jahren.

## 2.2 Vergleich mit der Arbeitslosenstatistik der Bundesagentur für Arbeit

Die bisherigen Ergebnisse zeigen, dass sich die Ausfallprozesse in der Piloterhebung kaum von denen in anderen Bevölkerungsumfragen ohne Auskunftspflicht unterscheiden. Eine wichtige Frage ist, ob die Ausfallprozesse nicht nur von soziodemographischen Merkmalen der Befragten, sondern auch von deren Erwerbsstatus abhängen. Wenn das Interesse am Thema der Befragung das Teilnahmeverhalten beeinflusst, würde man erwarten, dass Erwerbstätige und Personen auf Arbeitsuche häufiger an der Erhebung teilnehmen als Nichterwerbspersonen.

Um einen Hinweis über die Zusammensetzung der Stichprobe der Piloterhebung bezüglich des Erwerbsstatus zu bekommen, bietet es sich an, die Ergebnisse der Piloterhebung mit den Ergebnissen des Mikrozensus zu vergleichen. Im Mikrozensus wird der Erwerbsstatus ebenfalls nach dem Labour-Force-Konzept, das heißt nach der internationalen Definition der ILO, erhoben.

Weitergehende Analysen haben jedoch gezeigt, dass die methodischen Unterschiede in der Erhebung des Erwerbs-

13) Siehe Schnell, R., 1997, Fußnote 6, S. 202 f.

status zwischen Mikrozensus und Piloterhebung zu groß sind, als dass ein Vergleich der Ergebnisse dazu geeignet wäre, Ausfallprozesse in der Piloterhebung in Bezug auf den Erwerbsstatus zu beschreiben: Zu nennen sind zum Beispiel die unterschiedlichen Frageformulierungen zur Erwerbstätigkeit und Erwerbslosigkeit, die Methodik der Haushaltsbefragung mit der Möglichkeit von Proxy-Interviews beim Mikrozensus im Vergleich zur Personenbefragung in der Piloterhebung oder das persönliche bzw. schriftliche Interview im Vergleich mit einem Telefoninterview.<sup>14)</sup>

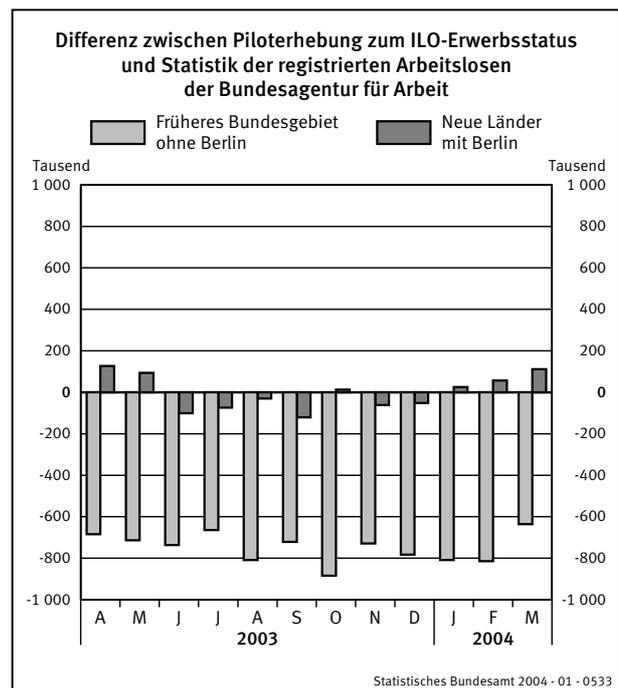
An dieser Stelle soll deshalb nur ein Vergleich der Ergebnisse der Piloterhebung mit der Statistik der registrierten Arbeitslosen der Bundesagentur für Arbeit vorgenommen werden. Die Befragten der Piloterhebung werden nicht nur nach ihrer Erwerbslosigkeit gemäß den internationalen Kriterien der ILO gefragt, sondern auch danach, ob sie beim Arbeitsamt als arbeitslos oder arbeitsuchend gemeldet sind. Zwischen beiden Merkmalen besteht eine hohe Korrelation.

Die Statistik der registrierten Arbeitslosen beruht jedoch auf den komplexen und vielen Befragten in den einzelnen Detailspekten wenig bekannten sozialrechtlichen Definitionen. In einer mündlich-telefonischen Befragung ist die Erhebung eines solchen komplexen Sachverhalts unvermeidlich mit Messfehlern behaftet, da nicht allen Befragten bewusst ist, ob sie in der Statistik der Bundesagentur für Arbeit als Arbeitslose oder als nicht arbeitslose Arbeitssuchende oder als keines von beiden geführt werden.

Der Unterschied zwischen der Zahl der Arbeitslosen, die sich in der Befragung als solche einstufen, und der Anzahl der laut den Statistiken der Bundesagentur für Arbeit registrierten Arbeitslosen differiert zwischen dem früheren Bundesgebiet ohne Berlin und den neuen Ländern mit Berlin (siehe das Schaubild). Die frei hochgerechneten Ergebnisse der Piloterhebung entsprechen im Osten in etwa den Zahlen der Bundesagentur für Arbeit, während sie im Westen seit Beginn der Erhebung unter den Zahlen der Bundesagentur für Arbeit liegen.

Die Differenz zwischen der Zahl der Arbeitslosen laut Piloterhebung und der Zahl der registrierten Arbeitslosen der Bundesagentur für Arbeit vom April 2003 kann als Maßstab für die Teilnahmebereitschaft der Arbeitslosen beim erstmaligen Kontakt betrachtet werden. Demnach nehmen Arbeitslose im früheren Bundesgebiet ohne Berlin seltener an der Erhebung teil, als man es nach den Zahlen der Bundesagentur für Arbeit erwarten würde, während in den neuen Ländern mit Berlin die Zahl der befragten Arbeitslosen geringfügig über dem erwarteten Wert liegt.

Es ist jedoch möglich, dass es sich hierbei nicht um einen reinen Stichprobeneffekt handelt, sondern auch Messprobleme eine Rolle spielen. Denkbar wäre beispielsweise die



Hypothese, dass der Status der Arbeitslosigkeit im Westen mit einer stärkeren Stigmatisierung einhergeht als im Osten und deshalb westdeutsche Befragte eine Arbeitslosmeldung nicht immer zugeben wollen.

### 3 Ausfallanalysen im Panel

#### 3.1 Bivariate Ausfallanalysen

Das grundlegende Problem bei der Analyse von Ausfallprozessen ist, dass man über diejenigen Personen, die an der Befragung nicht teilgenommen haben, keine oder nur sehr wenige Informationen besitzt, anhand derer man feststellen könnte, ob die Ausfallwahrscheinlichkeiten mit dem zu untersuchenden Merkmal korrelieren. Neben dem Vergleich mit externen Referenzstatistiken bietet sich bei Panelerhebungen eine weitere Möglichkeit an, Ausfallprozesse zu untersuchen: Für die bereits befragten Personen liegt eine Vielzahl an Informationen vor, unter anderem zum zentralen Merkmal der Untersuchung, dem Erwerbsstatus. Diese Informationen können für Analysen des Ausfallgeschehens herangezogen werden.

Dazu wurden für verschiedene Gruppen die Ausfallrisiken über alle Wiederholungsbefragungen von April/Mai 2003 bis Februar/März 2004 hinweg berechnet und miteinander verglichen (siehe Tabelle 4).<sup>15)</sup> Um die Vergleichbarkeit mit den noch folgenden multivariaten Ausfallanalysen zu gewährleisten, wurden die bivariaten Odds Ratios

14) Das Fragenprogramm der Piloterhebung wurde ausgehend vom Fragenprogramm der Arbeitskräfteerhebung entwickelt. Allerdings mussten die im Mikrozensus bzw. der Arbeitskräfteerhebung angewandten und für die persönlich-mündliche bzw. schriftliche Befragung entwickelten Fragen für den Einsatz im Rahmen einer ausschließlich telefonischen Befragung an einigen Stellen sprachlich modifiziert werden (siehe Fußnote 3, S. 153).

15) Der für die Analysen verwendete Datensatz umfasst alle potenziellen Wiederholungsbefragten, jedoch ohne die fortgeschriebenen Fälle und ohne die beim Aufbau des Panels angefallenen Rotationsfälle (zu den fortgeschriebenen Fällen und deren Begründung siehe Fußnote 3, S. 154).

ermittelt:<sup>16)</sup> Das Odds Ratio einer Merkmalsausprägung berechnet sich als Verhältnis aus der Ausfallwahrscheinlichkeit (p) zur komplementären Wahrscheinlichkeit, wieder teilzunehmen (1 – p), für die Personen mit dieser Merkmalsausprägung und der Ausfallwahrscheinlichkeit relativ zur komplementären Wahrscheinlichkeit in der Referenzgruppe. Ein Odds Ratio zwischen null und eins weist auf ein geringeres Ausfallrisiko als in der Referenzgruppe hin, ein Wert über eins entsprechend auf ein höheres Ausfallrisiko.<sup>17)</sup>

Tabelle 4: Bivariate Odds Ratios der Panelfälle Mai 2003 bis März 2004 bezüglich ausgewählter Merkmale, ungewichtete Ergebnisse

Merkmale	Bivariate Odds Ratios
nach dem Geschlecht	
Männer .....	Referenzgruppe
Frauen .....	0,85
nach dem Alter	
von ... bis ... Jahren	
15 – 24 .....	Referenzgruppe
25 – 34 .....	0,76
35 – 44 .....	0,60
45 – 54 .....	0,60
55 – 64 .....	0,72
65 – 74 .....	0,79
Keine Angabe .....	4,29
nach der Staatsangehörigkeit	
Deutsch .....	Referenzgruppe
Nicht deutsch, keine Angabe .....	1,74
nach der Schulbildung	
Kein Schulabschluss oder Haupt-/ Volksschulabschluss .....	Referenzgruppe
Abschluss Polytechnische Oberschule (POS) oder Realschulabschluss (Mittlere Reife) ..	0,84
Fachhochschulreife, allgemeine oder fachgebundene Hochschulreife (Abitur) ...	0,91
Keine Angabe .....	2,57
in Ausbildung	
Nicht in Ausbildung, keine Angabe .....	Referenzgruppe
Schüler/-innen .....	1,27
Studierende .....	1,36
Auszubildende .....	1,42
Arbeitslos gemeldet	
Nicht arbeitslos gemeldet .....	Referenzgruppe
Arbeitslos gemeldet, West .....	1,34
Arbeitslos gemeldet, Ost .....	1,12
nach dem Gebietsstand	
Früheres Bundesgebiet ohne Berlin .....	Referenzgruppe
Neue Länder mit Berlin .....	1,01

Die Ausfallrisiken im Panel entsprechen weitestgehend den Erkenntnissen aus dem Vergleich der April-Ergebnisse der Piloterhebung mit externen Referenzstatistiken. Diejenigen Merkmalsgruppen, deren frei hochgerechnete Ergebnisse aus der Piloterhebung unter den Ergebnissen vergleichbarer Referenzstatistiken lagen, haben auch im Panel in der Regel ein höheres Ausfallrisiko. Entsprechend zeigen die meisten Merkmalsgruppen, deren frei hochgerechnete Ergebnisse aus der Piloterhebung über den Ergebnissen entsprechender externer Statistiken lagen, im Panel ein niedrigeres Ausfallrisiko.

Die bivariaten Ausfallanalysen ergeben für die Panelfälle folgende Ergebnisse:

- Frauen haben ein geringeres Ausfallrisiko als Männer.
- Die jüngsten Teilnehmer an der Erhebung (15 bis 24 Jahre) weisen – abgesehen von den Personen ohne Angabe zu ihrem Alter – das höchste Ausfallrisiko im Panel auf. Personen zwischen 35 und 54 Jahren nehmen wiederum am häufigsten an Wiederholungsbefragungen teil.
- Die Erkenntnis aus dem Vergleich mit der Bevölkerungsfortschreibung, dass Personen ohne deutsche Staatsangehörigkeit ein höheres Ausfallrisiko besitzen als Deutsche, wird bestätigt.
- Bei der Bildung sind es die Personen ohne Schulabschluss oder nur mit Volks- oder Hauptschulabschluss, die seltener an Wiederholungsbefragungen teilnehmen als Personen mit einer höheren Schulbildung.
- Im Gegensatz zu den Erkenntnissen aus dem Vergleich mit externen Statistiken haben Schüler/-innen, Studierende und Auszubildende höhere Ausfallrisiken als Personen, die sich nicht in Ausbildung befinden. Möglicherweise hängt dieser Befund mit dem Alter der Befragten zusammen, da es sich bei Personen in Ausbildung fast ausschließlich um jüngere Personen handelt, die entsprechend hohe Ausfallrisiken im Panel aufweisen.
- Während es zwischen dem früheren Bundesgebiet ohne Berlin und den neuen Ländern mit Berlin insgesamt kaum Unterschiede gibt, bestätigt sich das Ergebnis, dass Arbeitslose im Westen ein höheres Ausfallrisiko besitzen als Arbeitslose im Osten. Sowohl Arbeitslose aus dem früheren Bundesgebiet ohne Berlin als auch Arbeitslose aus den neuen Ländern mit Berlin haben zudem ein höheres Ausfallrisiko im Panel als Personen, die nicht arbeitslos gemeldet sind.

Schließlich interessiert noch das Teilnahmeverhalten der Befragten in Bezug auf ihren Erwerbsstatus. Bei den Nichterwerbspersonen wird nochmals zwischen Personen unterschieden, die kein Interesse an der Aufnahme einer Erwerbstätigkeit haben, und solchen, die grundsätzlich an der Aufnahme einer Tätigkeit interessiert sind. Es ist zu erwarten, dass Nichterwerbspersonen mit Interesse an der Aufnahme einer Tätigkeit geringere Ausfallrisiken haben als Nichterwerbspersonen ohne ein solches Interesse.

Die Tabelle 5 zeigt, dass Erwerbstätige am häufigsten an den Wiederholungsbefragungen teilnehmen. Erwerbslose und Nichterwerbspersonen haben im Panel höhere Ausfallrisiken als Erwerbstätige. Die Annahme, dass Nichterwerbspersonen ohne Interesse an der Aufnahme einer Erwerbstätigkeit seltener an Wiederholungsbefragungen teilnehmen als Nichterwerbspersonen mit einem solchen Interesse, wird durch die Ergebnisse bestätigt. Der Erwerbsstatus scheint demnach einen gewissen Einfluss auf das Teilnah-

16) Zu Odds Ratios und deren Interpretation siehe Andreß, H.-J./Hagenaars, J. A./Kühnel, S.: „Analyse von Tabellen und kategorialen Daten. Log-lineare Modelle, latente Klassenanalyse, logistische Regression und GSK-Ansatz“, Berlin 1997, S. 23 ff.

17) Sowohl die bivariaten als auch die noch folgenden multivariaten Ausfallanalysen wurden ungewichtet gerechnet, da die Ausfallrisiken der Befragten nach bestimmten Merkmalsausprägungen unabhängig von ihrer Wahrscheinlichkeit, in die Stichprobe zu gelangen, interessieren.

meverhalten der Befragten zu haben. Inwieweit dieser Effekt bestehen bleibt, wenn andere Merkmale kontrolliert werden, werden die multivariaten Analysen im weiteren Verlauf des Aufsatzes zeigen.

Tabelle 5: Bivariate Odds Ratios der Panelfälle Mai 2003 bis März 2004 in Bezug auf den Erwerbsstatus, ungewichtete Ergebnisse

Erwerbsstatus	Bivariate Odds Ratios
Erwerbstätige .....	Referenzgruppe
Erwerbslose .....	1,30
Nichterwerbspersonen mit Interesse an der Aufnahme einer Tätigkeit .....	1,22
Nichterwerbspersonen ohne Interesse an der Aufnahme einer Tätigkeit .....	1,29

Die Ergebnisse machen aber bereits deutlich, dass ein geeignetes Gewichtungsverfahren die unterschiedlichen Ausfallrisiken von Erwerbstätigen, Erwerbslosen und Nichterwerbspersonen berücksichtigen sollte.

Neben den bisher analysierten Merkmalen hat die Dauer der Teilnahme am Panel großen Einfluss auf das Teilnahmeverhalten bei den Wiederholungsbefragungen (siehe Tabelle 6). Die Wahrscheinlichkeit, an der nächsten Befragung nicht teilzunehmen, ist umso geringer, je öfter eine befragte Person schon an der Erhebung teilgenommen hat.

Tabelle 6: Bivariate Odds Ratios der Panelfälle Mai 2003 bis März 2004 in Bezug auf die Dauer der Teilnahme am Panel, ungewichtete Ergebnisse

Person wurde bereits ...-mal befragt	Bivariate Odds Ratios
1 .....	Referenzgruppe
2 .....	0,42
3 .....	0,28
4 .....	0,25
5 .....	0,20

Der Übergang von der Erstbefragung zur ersten Wiederholungsbefragung stellt die Hürde mit dem höchsten Ausfallrisiko dar. Personen, die bereits zweimal befragt wurden, zeigen ein deutlich niedrigeres Ausfallrisiko als Personen, die erst einmal befragt wurden. Die Abnahme des Ausfallrisikos setzt sich in entsprechender Weise fort. Am häufigsten können Personen, die bereits fünfmal an der Befragung teilgenommen haben, für eine weitere Befragung gewonnen werden.

### 3.2 Multivariate Analysen

Die bisherigen Befunde haben auf bivariate Beziehungen zwischen verschiedenen Merkmalen und dem Teilnahmeverhalten hingewiesen. In diesem Abschnitt soll nun in einem multivariaten Modell der Einfluss einer Reihe von Merkmalen auf das Ausfallrisiko überprüft werden. Multivariate Modelle haben den Vorteil, dass der Effekt der einzelnen Merkmale unter Kontrolle der jeweils anderen, in dem Modell ebenfalls berücksichtigten Merkmale geschätzt

wird. Wie auch schon bei den bivariaten Ausfallanalysen bieten die Längsschnittdaten die Möglichkeit, das Teilnahmeverhalten im Zeitverlauf zu analysieren. Geeignet sind beispielsweise ereignisanalytische Modelle mit diskreter Zeit. Als Zeiteinheiten werden die Erhebungsmonate betrachtet; der zu erklärende Sachverhalt ist, ob die Person an der jeweiligen Befragung wieder teilnimmt oder nicht. Da es sich bei dem zu erklärenden Merkmal, der Wiederteilnahme, um ein dichotomes Merkmal handelt, stellt das binomiale Logit-Modell ein geeignetes multivariates Analyseverfahren dar.

Die Art der Modellbildung beim binomialen Logit-Modell zeigt die folgende Gleichung:

$$\ln \left[ \frac{P_{(Y_i=1|X_{ij})}}{1 - P_{(Y_i=1|X_{ij})}} \right] = \sum_j \beta_j X_{ij} \quad \text{bzw.} \quad \frac{P_{(Y_i=1|X_{ij})}}{1 - P_{(Y_i=1|X_{ij})}} = e^{\sum_j \beta_j X_{ij}} = \prod_j e^{\beta_j X_{ij}}$$

Die zu erklärende Variable wird in der Gleichung mit Y, die erklärenden Variablen werden mit X bezeichnet, i verweist auf das Individuum, j auf die einzelnen erklärenden Variablen. Y hat die Ausprägungen 0 oder 1: In der hier vorgenommenen Modellierung zeigt 0 an, dass die Person wieder teilnahm, und 1, dass sie nicht mehr befragt werden konnte. Man erkennt in der Gleichung, dass der Logarithmus des Odds Ratios, also des Verhältnisses der Wahrscheinlichkeit von  $Y_i=1$  zur komplementären Wahrscheinlichkeit, als linear abhängig von den erklärenden Variablen  $X_{ij}$  modelliert wird.<sup>18)</sup>

In den unten dargestellten Ergebnistabellen werden die Werte zum Einfluss der erklärenden Variablen auf das Odds Ratio, also die Werte für  $\exp(\beta_j)$ , berichtet. Bei einem Wert des Koeffizienten der logistischen Regression  $\beta_j$  größer als Null ist  $\exp(\beta_j)$  größer als 1, das heißt beim Anstieg einer erklärenden Variablen um den Wert eins ist das Odds Ratio beim höheren Wert  $\exp(\beta_j)$ -mal so hoch wie beim niedrigeren Wert. Vereinfacht, aber auch verkürzt ausgedrückt: Die entsprechende Wahrscheinlichkeit für  $Y_i=1$ , also für einen Ausfall, nimmt zu. Analog ist ein Wert des Koeffizienten der logistischen Regression kleiner als Null mit einer Verringerung der Wahrscheinlichkeit für einen Ausfall bei steigendem Wert der erklärenden Variable gleichbedeutend.

Die in dem Modell berücksichtigten erklärenden Merkmale auf Personenebene sind in der Übersicht 1 dargestellt. Dabei handelt es sich zunächst um soziodemographische Merkmale, für die bekannt ist, dass sie mit der Teilnahmebereitschaft an Umfragen zusammenhängen und die auch bereits in den bivariaten Ausfallanalysen betrachtet wurden.

Daneben werden weitere Merkmale kontrolliert, die mit der Untersuchungsthematik zusammenhängen: Hierbei handelt es sich um die Tatsache, ob jemand in Ausbildung ist, um den Erwerbsstatus sowie um die Tatsache, dass jemand arbeitslos gemeldet ist. Hintergrund ist die Erwartung, dass Personen, für die die inhaltliche Fragestellung der Untersuchung eine größere persönliche Relevanz besitzt, eher an der Befragung teilnehmen. Hierbei könnten – wie im Kon-

18) Eine Konstante wird in dem Modell ebenfalls berücksichtigt. In der hier dargestellten Formel ergibt sie sich, wenn eine konstante erklärende Variable mit dem Wert 1 eingeführt wird.

Übersicht 1: Im Modell verwendete personenbezogene Merkmale

Geschlecht:	männlich weiblich
Altersgruppen:	15 bis 24 Jahre 25 bis 34 Jahre 35 bis 44 Jahre 45 bis 54 Jahre 55 bis 64 Jahre 65 bis 74 Jahre keine Angabe
Staatsangehörigkeit:	deutsch nicht deutsch, keine Angabe
Bildungsabschluss:	niedrig (kein Abschluss, Volks-, Hauptschule), keine Angabe mittel (Realschule, Polytechnische Oberschule der ehemaligen DDR) hohes Bildungsniveau (Abitur, Fachhochschulreife)
Aktueller Ausbildungsstatus:	nicht in Schule oder Ausbildung, keine Angabe Schüler/-innen Studierende Auszubildende
ILO-Erwerbsstatus:	erwerbstätig erwerbslos Nichterwerbsperson mit Interesse an Tätigkeitsaufnahme Nichterwerbsperson ohne Interesse an Tätigkeitsaufnahme
Arbeitslos gemeldet:	nein ja
Bruttoerwerbseinkommen:	nicht erfragt bis 400 Euro über 400 bis unter 1 200 Euro 1 200 bis unter 2 000 Euro 2 000 bis unter 3 000 Euro 3 000 bis unter 4 000 Euro 4 000 Euro oder mehr keine Angabe
Region:	Früheres Bundesgebiet ohne Berlin Neue Länder mit Berlin
Gemeindegröße:	Kleine Gemeinde (bis 5 000 Einwohner) Mittlere Gemeinde (über 5 000 bis 100 000 Einwohner) Große Gemeinde (über 100 000 Einwohner)

text der bivariaten Analysen bereits vermutet wurde – auch Stigmatisierungseffekte einen Einfluss auf das Teilnahmeverhalten ausüben.

Als weiteres Merkmal wurde in den Modellen berücksichtigt, ob in der jeweiligen Befragung das Einkommen aus der Tätigkeit erhoben wurde und, wenn ja, wie hoch es war.<sup>19)</sup> Die Fragen nach dem Bruttoeinkommen aus der jeweiligen Erwerbstätigkeit wurden erstmals im Oktober 2003 in das Erhebungsinstrument aufgenommen. Sie werden an Erstbefragte gestellt, die erwerbstätig sind, oder an Panelfälle, die im aktuellen Monat erwerbstätig sind und bei denen sich diesbezüglich im Vergleich zum Vormonat eine Veränderung ergeben hat. Es ist anzunehmen, dass die Befragten die Information über die Höhe ihres Einkommens als vertraulich ansehen, daher wird die Frage danach als problematisch eingestuft und es wird erwartet, dass sie zu einer Erhöhung des Ausfallrisikos führt. Durch die gleichzeitige Kontrolle der Höhe des Einkommens in den Analysen kann zudem überprüft werden, ob verstärkt Personen

in bestimmten Einkommenskategorien ausfallen und, wenn ja, in welchem Ausmaß.

Daneben wurden in den Modellen die Gemeindegröße und die Region kontrolliert. Es ist zu erwarten, dass die Teilnahmebereitschaft mit zunehmender Gemeindegröße sinkt. Die weiter oben berichteten bivariaten Untersuchungen haben zu dem Ergebnis geführt, dass sich bei einigen Kovariaten deutliche Unterschiede zwischen dem früheren Bundesgebiet ohne Berlin und den neuen Ländern mit Berlin zeigen. Daher wurde diese Kovariate auch in das multivariate Modell eingeführt.

Übersicht 2: Im Modell verwendete erhebungstechnische Merkmale

Vorher bereits ...-mal befragt:	1 2 3 4 5
Erhebungsmonat:	Mai 2003 Juni 2003 Juli 2003 August 2003 September 2003 Oktober 2003 November 2003 Dezember 2003 Januar 2004 Februar 2004 März 2004
Interviewdauer:	Minuten

Neben den die befragten Personen beschreibenden Variablen wurden Merkmale, die mit der Untersuchung und deren Ablauf in Zusammenhang stehen, in das Modell eingeführt. Die Befragten sollen an sechs aufeinander folgenden Monaten an der Untersuchung teilnehmen. Auf der Ebene des einzelnen Panelbefragten variiert daher, um welche Erhebungswelle es sich jeweils für sie handelt. Daher wurde für jede Person kontrolliert, wie oft sie bereits teilgenommen hat. Aufgrund der Erfahrungen mit anderen Panelerhebungen sowie der Befunde aus den bivariaten Analysen ist zu erwarten, dass der Übergang von der ersten zur zweiten Befragung am problematischsten ist. Der Kalendermonat stellt die zweite zeitveränderliche Kovariate dar. Zu vermuten ist, dass die Wiederteilnahme insbesondere in Ferien- und Urlaubszeiten zurückgeht. Schließlich wurde die Dauer der Interviews in den Modellen als erklärendes Merkmal berücksichtigt. Es ist zu erwarten, dass das Ausfallrisiko mit steigender Interviewdauer zunimmt.

Die Resultate der multivariaten Analysen mit den angesprochenen Kovariaten zeigt Tabelle 7.<sup>20)</sup>

Das Modell 1 berücksichtigt die gerade vorgestellten Merkmale und geht von der Annahme aus, dass deren Einfluss über die verschiedenen Wellen hinweg konstant ist. Nun ist es aber auch denkbar, dass sich die Effekte dieser Merkmale in Abhängigkeit davon unterscheiden, in welcher

19) Wenn die befragte Person mehreren Erwerbstätigkeiten nachging, wurde in den Analysen nur das Einkommen aus der ersten Tätigkeit berücksichtigt.

20) Die neben den Modellen für Deutschland für das frühere Bundesgebiet ohne Berlin und die neuen Länder mit Berlin getrennt durchgeführten Analysen weisen auf keine grundsätzlichen Unterschiede bezüglich der Ausfallmechanismen hin. Daher werden hier nur die Ergebnisse für Deutschland berichtet.

Tabelle 7: Multivariate Modelle zum Ausfallrisiko: alle Übergänge zusammen (Modell 1) gegenüber getrennter Betrachtung der Übergänge (Modell 2), ungewichtete Ergebnisse

Merkmale	Modell 1		Modell 2									
	eß	S.E.	W1 → W2		W2 → W3		W3 → W4		W4 → W5		W5 → W6	
			eß	S.E.								
nach dem Geschlecht												
Männlich .....	Ref.											
Weiblich .....	0,85 ***	0,02	0,88 ***	0,03	0,76 ***	0,05	0,93	0,07	0,93	0,07	0,80 **	0,09
nach dem Alter												
15 – 24 .....	Ref.											
25 – 34 .....	0,71 ***	0,05	0,70 ***	0,07	0,83 *	0,11	0,68 ***	0,14	0,65 ***	0,15	0,82	0,19
35 – 44 .....	0,58 ***	0,05	0,60 ***	0,07	0,64 ***	0,11	0,53 ***	0,14	0,53 ***	0,15	0,57 ***	0,20
45 – 54 .....	0,57 ***	0,05	0,60 ***	0,07	0,66 ***	0,11	0,50 ***	0,15	0,46 ***	0,16	0,59 **	0,21
55 – 64 .....	0,62 ***	0,05	0,65 ***	0,07	0,63 ***	0,12	0,61 ***	0,15	0,53 ***	0,16	0,69 *	0,21
65 – 74 .....	0,67 ***	0,09	0,73 ***	0,12	0,68 *	0,20	0,50 **	0,29	0,56 *	0,30	0,69	0,37
Keine Angabe .....	3,07 ***	0,13	4,38 ***	0,18	1,87 *	0,35	1,43	0,51	1,98	0,53	0,65	1,06
nach der Staatsangehörigkeit												
Deutsch .....	Ref.											
Nicht deutsch, keine Angabe .....	1,52 ***	0,05	1,50 ***	0,07	1,68 ***	0,10	1,55 ***	0,14	1,43 **	0,16	1,06	0,23
nach der Schulbildung												
Niedrige Bildung, keine Angabe .....	Ref.											
Mittlere Bildung .....	0,87 ***	0,03	0,87 ***	0,04	0,95	0,06	0,88	0,08	0,76 ***	0,09	0,95	0,12
Hohe Bildung .....	0,91 ***	0,03	0,85 ***	0,04	0,97	0,06	0,91	0,09	0,91	0,10	1,35 **	0,12
in Ausbildung												
Nicht in Schule, Ausbildung keine Angabe .....	Ref.											
Schüler/-innen .....	0,76 ***	0,07	0,76 ***	0,09	0,67 ***	0,15	0,91	0,19	0,67 *	0,22	0,96	0,27
Studierende .....	1,03	0,06	1,04	0,08	1,03	0,13	0,93	0,18	1,23	0,18	0,87	0,24
Auszubildende .....	1,00	0,07	0,98	0,10	1,12	0,16	0,91	0,22	0,93	0,24	1,22	0,29
nach dem Erwerbsstatus												
Erwerbstätig .....	Ref.											
Erwerbslos .....	1,02	0,06	0,90	0,08	0,84	0,13	0,93	0,17	1,07	0,19	1,40	0,23
Nichterwerbsperson, Interesse .....	1,11 **	0,05	0,97	0,07	1,10	0,11	1,10	0,15	1,03	0,17	1,64 **	0,20
Nichterwerbsperson, kein Interesse .....	1,24 ***	0,04	1,42 ***	0,05	1,06	0,08	0,98	0,11	0,88	0,12	1,12	0,15
arbeitslos gemeldet												
Nicht arbeitslos gemeldet ..	Ref.											
Arbeitslos gemeldet .....	1,30 ***	0,05	1,33 ***	0,07	1,13	0,12	1,12	0,16	1,10	0,18	1,06	0,22
Bruttoerwerbseinkommen												
Nicht erfragt .....	Ref.											
von ... bis unter ... EUR												
unter 400 .....	0,92 *	0,05	0,92	0,07	0,77 **	0,11	0,71 **	0,14	0,87	0,15	1,11	0,18
400 – 1 200 .....	1,13 *	0,07	1,02	0,10	1,09	0,15	1,12	0,20	0,81	0,24	1,16	0,28
1 200 – 2 000 .....	1,12 *	0,06	1,18 *	0,09	0,88	0,16	0,80	0,22	0,65 *	0,26	1,36	0,27
2 000 – 3 000 .....	1,03	0,06	1,06	0,08	0,77 *	0,15	0,77	0,20	0,96	0,20	0,84	0,27
3 000 – 4 000 .....	0,95	0,07	0,99	0,09	0,90	0,16	0,83	0,23	0,41 ***	0,33	0,68	0,36
4 000 und mehr .....	1,26 ***	0,07	1,32 ***	0,10	1,05	0,17	1,20	0,24	0,49 *	0,38	1,02	0,39
Keine Angabe .....	2,21 ***	0,08	2,50 ***	0,11	1,65 ***	0,19	0,98	0,33	1,62	0,31	1,72	0,39
nach der Gemeindegröße												
von ... bis unter ... Einwohner												
unter 5 000 .....	Ref.											
5 000 – 100 000 .....	1,04	0,03	1,11 **	0,04	0,93	0,07	1,07	0,09	0,94	0,10	0,95	0,13
100 000 und mehr .....	1,16 ***	0,03	1,20 ***	0,04	1,04	0,07	1,34 ***	0,10	1,04	0,11	1,17	0,13
nach dem Gebietsstand												
Früheres Bundesgebiet ohne Berlin .....	Ref.											
Neue Länder mit Berlin .....	1,04	0,03	1,03	0,04	1,00	0,06	1,17 **	0,08	1,00	0,09	0,98	0,11
Interviewdauer .....	1,01	0,01	0,98 ***	0,01	1,17 ***	0,02	1,14 ***	0,03	1,13 ***	0,03	1,12 ***	0,04

eß = Ausfallrisiko; S.E. = Standard Error (Standardfehler); Ref. = Referenzgruppe; Signifikanzniveau\* p &lt; 0,1, \*\* p &lt; 0,05, \*\*\* p &lt; 0,01

noch Tabelle 7: Multivariate Modelle zum Ausfallrisiko: alle Übergänge zusammen (Modell 1) gegenüber getrennter Betrachtung der Übergänge (Modell 2), ungewichtete Ergebnisse

Merkmale	Modell 1		Modell 2									
	eß	S.E.	W1 → W2		W2 → W3		W3 → W4		W4 → W5		W5 → W6	
			eß	S.E.								
nach dem Erhebungsmonat												
Mai 2003 .....	Ref.											
Juni 2003 .....	1,45 ***	0,05	1,40 ***	0,06								
Juli 2003 .....	1,50 ***	0,05	1,43 ***	0,06	1,08	0,10						
August 2003 .....	1,67 ***	0,04	1,40 ***	0,06	1,07	0,09	1,51 ***	0,11				
September 2003 .....	1,56 ***	0,05	1,38 ***	0,06	0,95	0,09	0,87	0,11	0,79 *	0,12		
Oktober 2003 .....	1,06	0,05	1,10	0,06	0,73 ***	0,10	0,68 ***	0,13	0,52 ***	0,13	0,37 ***	0,16
November 2003 .....	1,04	0,05	1,06	0,07	0,80 **	0,09	0,72 **	0,13	0,40 ***	0,14	0,33 ***	0,15
Dezember 2003 .....	1,19 ***	0,05	1,31 ***	0,07	0,92	0,11	0,78 **	0,12	0,46 ***	0,14	0,38 ***	0,15
Januar 2004 .....	0,94	0,05	1,00	0,07	0,65 ***	0,13	0,64 ***	0,15	0,43 ***	0,13	0,33 ***	0,16
Februar 2004 .....	1,05	0,05	1,18 **	0,07	0,74 **	0,12	0,68 **	0,16	0,56 ***	0,16	0,38 ***	0,14
März 2004 .....	1,08	0,06	1,14 *	0,07	0,83	0,11	0,74 **	0,14	0,62 ***	0,16	0,30 ***	0,18
nach der Befragungswelle												
Übergang von 1. auf 2. Welle	Ref.											
Übergang von 2. auf 3. Welle	0,39 ***	0,04										
Übergang von 3. auf 4. Welle	0,27 ***	0,05										
Übergang von 4. auf 5. Welle	0,28 ***	0,05										
Übergang von 5. auf 6. Welle	0,26 ***	0,06										
LN (Konstante) .....	-1,04 ***	0,08	-0,93 ***	0,09	-1,82 ***	0,14	-2,17 ***	0,18	-1,54 ***	0,19	-1,82 ***	0,25
Fallzahl (ohne Rotationsfälle und ohne fortgeschriebene Fälle) .....	84 677		28 958		19 660		15 050		12 016		8 993	
-2LL Null-Modell (nur Konstante) .....	65 789		30 636		13 310		7 974		6 371		4 302	
Modellverbesserung .....	4 287 ***		635 ***		273 ***		206 ***		234 ***		207 ***	
Freiheitsgrade .....	42		38		37		36		35		34	
Pseudo-R <sup>2</sup> (Mc Fadden) ....	0,07		0,02		0,02		0,03		0,04		0,05	

eß = Ausfallrisiko; S.E. = Standard Error (Standardfehler); Ref. = Referenzgruppe; Signifikanzniveau\* p < 0,1, \*\* p < 0,05, \*\*\* p < 0,01

Erhebungswelle die Befragten sich befinden. Dies wurde in Modell 2 überprüft.<sup>21)</sup>

Erwartungsgemäß ist zunächst für Frauen ein geringeres Ausfallrisiko als für Männer festzustellen (Modell 1). Dieser Effekt tritt am stärksten beim Übergang von Welle 2 auf Welle 3 auf, gilt aber – zumindest tendenziell – auch für alle anderen Übergänge (Modell 2). Auch im Hinblick auf das Alter der befragten Personen bestätigen sich die Vermutungen und die Ergebnisse der bivariaten Ausfallanalysen: Im Vergleich zur jüngsten Altersgruppe der 15- bis 24-jährigen nimmt das Ausfallrisiko zunächst ab, ist bei den Personen zwischen 35 und 54 Jahren am niedrigsten und steigt dann wieder etwas an. Zwischen den Erhebungswellen sind keine bedeutenden Unterschiede festzustellen. Bemerkenswert ist darüber hinaus das sehr hohe Ausfallrisiko von Personen, die die Angabe zu ihrem Alter verweigern. Besonders hoch ist es beim Übergang von der ersten auf die zweite Befragung, allerdings ist das Odds Ratio auch bei den folgenden Übergängen – den letzten ausgenommen – noch größer als eins.

Das signifikant höhere Ausfallrisiko von Personen ohne deutsche Staatsangehörigkeit bestätigt ebenfalls die Ergebnisse der bivariaten Analysen. Dies verändert sich im Verlauf der Erhebung nur wenig. In Bezug auf die Bildung entspricht das Muster der Vermutung, dass Personen mit niedriger Bildung mit größerer Wahrscheinlichkeit nicht mehr teilnehmen. Zwischen Personen mit mittlerer und höherer Bildung zeigen sich nur geringe Unterschiede. Mit Ausnahme des letzten Übergangs – hier ist für Hochgebildete ein erhöhtes Ausfallrisiko festzustellen – bleibt dies im Zeitverlauf zumindest in der Tendenz konstant. Im Vergleich zu Personen, die nicht in Ausbildung sind, ist das Ausfallrisiko von Schülern und Schülerinnen signifikant niedriger, Studierende und Auszubildende dagegen unterscheiden sich nicht von der Referenzgruppe.<sup>22)</sup> Das Muster ist im Zeitverlauf in der Tendenz stabil.

In Bezug auf den Erwerbsstatus ist zu erkennen, dass Nichterwerbspersonen im Vergleich zu Personen, die erwerbstätig sind, ein höheres Ausfallrisiko haben. Dabei sind über die Erhebungswellen hinweg verschiedene Entwicklungs-

21) Ein alternativ zu dem hier gewählten Vorgehen geschätztes Modell mit Interaktionstermen aus den Merkmalen und der jeweiligen Erhebungswelle erlaubte es zu prüfen, ob die Annahme von Ausfallmechanismen, die sich im Zeitverlauf ändern, zu einer Modellverbesserung führt. Tatsächlich war dies nicht der Fall.

22) Der den bivariaten Ergebnissen für Schüler/-innen im multivariaten Modell entgegengesetzte Effekt ist, wie hier nicht dargestellte Analysen zeigen, auf die Kontrolle des Alters in den Logit-Modellen zurückzuführen.

muster für die beiden Teilgruppen der Nichterwerbstätigen festzustellen: Nichterwerbspersonen mit Interesse an der Aufnahme einer Erwerbstätigkeit weisen nur beim Übergang zur sechsten Welle ein signifikant erhöhtes Ausfallrisiko auf. Bei Nichterwerbspersonen ohne Interesse an der Aufnahme einer Erwerbstätigkeit dagegen ist das erhöhte Ausfallrisiko beim Übergang vom ersten auf den zweiten Befragungsmonat feststellbar, in den Folgewellen geht es zurück und ist nicht mehr signifikant. Ein signifikant höheres Ausfallrisiko von Erwerbslosen im Vergleich zu Erwerbstätigen konnte durch die multivariaten Ausfallanalysen nicht bestätigt werden. Die Ergebnisse der bivariaten Ausfallanalysen werden an dieser Stelle demnach relativiert.

Personen, die beim Arbeitsamt als arbeitslos gemeldet sind, zeigen hingegen ein erhöhtes Ausfallrisiko. Dies ist vor allem für den Übergang von der Erst- zur Zweitbefragung festzustellen, anschließend ist das Ausfallrisiko nicht mehr signifikant höher als das der nicht arbeitslos Gemeldeten.

Für Personen, denen die Frage nach dem Verdienst aus ihrer Erwerbstätigkeit gestellt wurde, ist in der Gesamtbetrachtung nicht generell ein höheres Ausfallrisiko feststellbar. Vielmehr hängt das Ausfallrisiko von der Höhe des Einkommens ab: Nur Personen mit einem Erwerbseinkommen von 400 bis unter 2000 Euro und ab 4000 Euro nehmen seltener an den Wiederholungsbefragungen teil. Ebenso wie beim Alter ist allerdings das höchste Ausfallrisiko für Personen erkennbar, die die Angabe zum Einkommen verweigern. Ähnlich wie bei den Nichterwerbspersonen ohne Interesse an der Aufnahme einer Erwerbstätigkeit konzentriert sich diesbezüglich das Hauptausfallgeschehen auf den Übergang zwischen der ersten und zweiten Welle. Anschließend zeigen sich teilweise gegenläufige Tendenzen. Anfängliche Verzerrungen verstärken sich im Verlauf der Erhebung also nicht unbedingt, sondern werden teilweise sogar wieder korrigiert.

In Bezug auf die Wohnortgröße ist festzustellen, dass in größeren Orten, insbesondere in Städten ab 100 000 Einwohnern, die Bereitschaft, wieder teilzunehmen, signifikant geringer ist als in kleinen Gemeinden. In der Tendenz gilt dies für die großen Städte für alle Übergänge. Schließlich ist noch zu erkennen, dass zwischen den neuen und den alten Bundesländern in der Gesamtbetrachtung keine Unterschiede bezüglich des Ausfallrisikos bestehen, wenn gleich eine leichte Tendenz festzustellen ist, dass Befragte in den neuen Ländern mit Berlin etwas seltener wieder an einer Folgebefragung teilnehmen.

Als erhebungstechnisches Merkmal wurde die Dauer des Interviews in das Modell eingeführt. Mit Ausnahme des ersten Übergangs ist zu erkennen, dass das Ausfallrisiko mit steigender Interviewdauer zunimmt. Ab dem Übergang von der zweiten auf die dritte Erhebungswelle nimmt das Odds Ratio Werte deutlich über 1 an. Dies macht deutlich, dass insbesondere bei späteren Befragungen eine kurze Inter-

viewdauer für die Sicherstellung der weiteren Teilnahme besonders wichtig ist.

Saisonale Effekte bezüglich des Teilnahmeverhaltens spiegeln sich in den Variablen für die Kalendermonate wider. Hier ist zu erkennen, dass das Ausfallrisiko von Juni bis September 2003 und im Dezember 2003 höher war als im Mai 2003.<sup>23)</sup> Dieser Befund und die Tatsache, dass es in den Sommermonaten noch wesentlich stärker ausgeprägt ist als im Dezember, spricht für die These, dass die Wiederteilnahme abhängig von Urlaubs- und Ferienzeiten ist.

Die Betrachtung der „Panelalterseffekte“, das heißt wie lange die Befragten bereits an der Erhebung teilnehmen, bestätigt ebenfalls das Ergebnis der bivariaten Analysen, dass der Übergang von der ersten zur zweiten Befragung der schwierigste ist. Relativ dazu ist in allen Folgewellen die Wahrscheinlichkeit, nicht mehr an der Befragung teilzunehmen, signifikant niedriger. Zudem verringert sich das Ausfallrisiko im Zeitverlauf deutlich.

Somit lässt sich festhalten: Sowohl erhebungstechnische als auch persönliche Merkmale haben einen Einfluss auf die Ausfallprozesse. Bei den meisten Merkmalen verändern sich die Ausfallmechanismen im Verlauf der Panelbefragung nicht, die entsprechende Modellierung (hier nicht dargestellt) führte zu keiner Modellverbesserung.

Das Ausfallrisiko erhöht sich in der Ferienzeit und ab der zweiten Befragung mit zunehmender Länge des Interviews. Personen, die die Auskunft bei bestimmten Fragen verweigern – konkret handelte es sich um das Alter und das Bruttoerwerbseinkommen –, nehmen in Folgewellen mit einer höheren Wahrscheinlichkeit nicht mehr teil. Möglicherweise stellt der Item-Nonresponse ein erstes Signal in Bezug auf den Unit-Nonresponse dar.

Dem Problem selektiver Ausfallprozesse in Bezug auf inhaltliche Merkmale wird mit einem eigens entwickelten Gewichtungsverfahren Rechnung getragen. Hierauf wird im nächsten Abschnitt eingegangen.

## 4 Zusammenfassung und Gewichtungsverfahren

### 4.1 Zusammenfassung

Die Analysen zu den Ausfallprozessen haben gezeigt, dass in der Piloterhebung zum ILO-Erwerbsstatus ähnliche Ausfallprozesse auftreten wie in den meisten Bevölkerungsumfragen ohne Auskunftspflicht, die heute in Deutschland durchgeführt werden. Insbesondere jüngere Personen im Alter zwischen 15 und 34 Jahren sowie ältere Menschen im Alter zwischen 65 und 74 Jahren nehmen seltener an der Befragung teil als 35- bis 64-Jährige. Frauen zeigen eine höhere Teilnahmebereitschaft als Männer. Ausländische Mitbürger nehmen – möglicherweise aufgrund von Sprach-

23) Bei diesem Merkmal ist zu beachten, dass der Referenzmonat sich in Modell 2 in Abhängigkeit vom betrachteten Übergang ändert. Beim Übergang von Welle 5 auf Welle 6 beispielsweise ist es der September 2003, da dies der erste Monat ist, in dem Personen zum sechsten Mal befragt wurden. Da im September das Ausfallrisiko so hoch ist, erklärt dies teilweise, warum die Odds Ratios in den Folgemonaten so niedrig sind.

barrieren – seltener an der Befragung teil als Deutsche. Das Teilnahmeverhalten wird außerdem vom Bildungsniveau der kontaktierten Personen beeinflusst. Personen mit einem mittleren oder einem höheren Bildungsabschluss beteiligen sich öfter an der Befragung als Personen ohne Schulabschluss oder mit einem Volks- oder Hauptschulabschluss. Sowohl der Vergleich mit externen Referenzstatistiken als auch die bivariaten und multivariaten Ausfallanalysen im Panel ergaben in Bezug auf diese soziodemographischen Merkmale einheitliche Ergebnisse. Die Ausfallrisiken von

Tabelle 8: Ausfallrisiken der Panelfälle Mai 2003 bis März 2004 bezüglich ausgewählter Merkmale: Vergleich der Ergebnisse der bivariaten und der multivariaten Ausfallanalysen, ungewichtete Ergebnisse

Merkmale der Panelfälle	Bivariate Odds Ratios	Multivariate Odds Ratios (Modell 1)
nach dem Geschlecht		
Männer .....	Referenzgruppe	Referenzgruppe
Frauen .....	0,85	0,85 ***
nach dem Alter		
von ... bis ... Jahren	Referenzgruppe	Referenzgruppe
15 – 24 .....	0,76	0,71 ***
25 – 34 .....	0,60	0,58 ***
35 – 44 .....	0,60	0,57 ***
45 – 54 .....	0,72	0,62 ***
55 – 64 .....	0,79	0,67 ***
65 – 74 .....	4,29	3,07 ***
Keine Angabe .....		
nach der Staatsangehörigkeit		
Deutsch .....	Referenzgruppe	Referenzgruppe
Nicht deutsch, keine Angabe .....	1,74	1,52 ***
nach der Schulbildung		
Kein Schulabschluss oder Haupt-/ Volksschulabschluss (bzw. keine Angabe) .....	Referenzgruppe	Referenzgruppe
Abschluss Polytechnische Oberschule (POS) oder Realschulabschluss (Mittlere Reife) .....	0,84	0,87 ***
Fachhochschulreife, allgemeine oder fachgebundene Hochschulreife (Abitur) .....	0,91	0,91 ***
in Ausbildung		
Nicht in Ausbildung, keine Angabe ...	Referenzgruppe	Referenzgruppe
Schüler/-innen .....	1,27	0,76 ***
Studierende .....	1,36	1,03
Auszubildende .....	1,42	1,00
nach dem Gebietsstand		
Früheres Bundesgebiet ohne Berlin ...	Referenzgruppe	Referenzgruppe
Neue Länder mit Berlin .....	1,01	1,04
arbeitslos gemeldet		
Nicht arbeitslos gemeldet .....	Referenzgruppe	Referenzgruppe
Arbeitslos gemeldet .....	1,32	1,30 ***
nach dem Erwerbsstatus		
Erwerbstätige .....	Referenzgruppe	Referenzgruppe
Erwerbslose .....	1,30	1,02
Nichterwerbspersonen mit Interesse an der Aufnahme einer Tätigkeit ...	1,22	1,11 **
Nichterwerbspersonen ohne Interesse an der Aufnahme einer Tätigkeit ...	1,29	1,24 ***
nach der Dauer der Teilnahme am Panel		
Person wurde bereits ... -mal befragt	Referenzgruppe	Referenzgruppe
1 .....	0,42	0,39 ***
2 .....	0,28	0,27 ***
3 .....	0,25	0,28 ***
4 .....	0,20	0,26 ***
5 .....		

Signifikanzniveau \* p < 0,10, \*\* p < 0,05, \*\*\* p < 0,01

Personen in Ausbildung unterscheiden sich in den bivariaten Analysen von denjenigen in den multivariaten Analysen, was auf den Zusammenhang dieses Merkmals mit dem Alter zurückzuführen ist (siehe Tabelle 8).

Die Dauer der Teilnahme am Panel (Panelalter) beeinflusst ebenfalls das Teilnahmeverhalten der Befragten. Je häufiger eine Person bereits am Panel teilgenommen hat, desto geringer ist die Wahrscheinlichkeit, dass sie in der nächsten Welle nicht mehr an der Befragung teilnimmt. Die Analysen haben gezeigt, dass das Ausfallrisiko beim Übergang von der ersten auf die zweite Befragung überdurchschnittlich hoch ist. Darüber hinaus beeinflussen saisonale Besonderheiten das Ausfallgeschehen: In den Ferienmonaten von Juni bis September 2003 und im Dezember 2003 war das Ausfallrisiko im Panel signifikant höher als in den übrigen Monaten.

Die multivariaten Ausfallanalysen ergaben weiterhin, dass Personen in bestimmten Einkommensklassen – insbesondere beim Übergang von der ersten auf die zweite Befragung – ein erhöhtes Ausfallrisiko aufweisen. Hierbei handelt es sich vor allem um Personen mit einem Einkommen zwischen 400 und 2 000 Euro und über 4 000 Euro. Das Ausfallgeschehen in den Folgewellen korrigiert diesen Effekt aber teilweise wieder.

Die logistischen Regressionsmodelle legen den Schluss nahe, dass Nichterwerbspersonen seltener an der Befragung teilnehmen als Erwerbstätige, insbesondere wenn sie kein Interesse daran haben, zukünftig eine Erwerbstätigkeit aufzunehmen. Ein signifikant höheres Ausfallrisiko von Erwerbslosen im Vergleich zu Erwerbstätigen konnte durch die multivariaten Analysen nicht nachgewiesen werden. Arbeitslos gemeldete Personen haben hingegen ein höheres Ausfallrisiko als nicht arbeitslos gemeldete Befragte.

## 4.2 Gewichtungungsverfahren

Trotz verschiedener Maßnahmen im Zuge der Stichprobenziehung und eines breit gefächerten Informationsangebots für die Befragten lassen sich bei der Piloterhebung – wie bei jeder auf Freiwilligkeit basierenden Befragung – Nonresponse-bedingte Selektivitäten nicht vollständig vermeiden. Zur Korrektur der dadurch hervorgerufenen Abbildungsprobleme bedürfte es einer Theorie des Ausfallgeschehens und daraus abgeleiteter Korrekturprozesse. Eine solche Theorie, aus der Gewichtungsverfahren ableitbar sind, die den Ausfallprozess nachvollziehen und korrigieren, gibt es derzeit nicht. Ausfallgewichtungen bei Querschnittsstichproben können sich daher nur an den Strukturen der realisierten Stichprobe und deren Vergleich mit Strukturen der Grundgesamtheit orientieren, das heißt die Stichprobenstrukturen werden denen der Grundgesamtheit mittels Gewichtungsfaktoren angepasst. Voraussetzung hierfür ist die Verfügbarkeit von Referenzstatistiken, die – durch Konvention – als Abbildung der Grundgesamtheit gelten und die interessierenden Strukturmerkmale in vergleichbarer Weise abbilden.

Die Erkenntnisse aus den Analysen der Ausfallprozesse fließen in das Gewichtungungsverfahren der Piloterhebung ein. Durch eine geeignete Anpassung der Strukturen der Pilot-

erhebung an für die Grundgesamtheit als repräsentativ betrachtete Eckwerte externer Statistiken kann zum einen die Präzision der Messung erhöht werden, zum anderen sollen die Auswirkungen der beschriebenen Ausfallprozesse korrigiert werden.<sup>24)</sup> Es empfiehlt sich dabei, Merkmale und deren Kombinationen zu verwenden, die möglichst hoch mit den Zielmerkmalen der Untersuchung korreliert sind. Ausfälle in der Folgebefragung einer Panelerhebung sind anhand vorliegender Merkmale aus der vorangegangenen Panelwelle beschreibbar. Damit können Effekte der Panelmortalität auf die Struktur der Stichprobe festgestellt und durch faktorielle Gewichtung ausgeglichen werden.

Aus den genannten Gründen wurde für die Piloterhebung ein mehrstufiges Gewichtungsverfahren eingeführt, das in Übersicht 3 dargestellt ist.

Übersicht 3: Das mehrstufige Gewichtungsverfahren in der Piloterhebung

1. Stufe: Designgewichtung auf Haushaltsbasis – Anzahl der Telefonanschlüsse im Haushalt
2. Stufe: Umwandlung in eine Personstichprobe – Anzahl der Zielpersonen im Haushalt
3. Stufe: Personengewichtung: Panelgewichtung und Anpassung an externe Eckwerte – Panelgewichtung: Erwerbsstatus Vorwelle – Regionale Merkmale: Früheres Bundesgebiet ohne Berlin sowie neue Länder mit Berlin bzw. Bundesländer (zu acht Gruppen zusammengefasst, so genannte Nielsen-Gebiete) • Schleswig-Holstein, Hamburg, Niedersachsen, Bremen • Nordrhein-Westfalen • Hessen, Rheinland-Pfalz, Saarland • Baden-Württemberg • Bayern • Berlin • Brandenburg, Mecklenburg-Vorpommern, Sachsen-Anhalt • Sachsen, Thüringen – Geschlecht – Altersklassen von ... bis ... Jahren 15 – 24 25 – 40 41 – 50 51 – 64 65 – 74 – Staatsangehörigkeit • Deutsch • Nicht deutsch – Schulbildung • Schüler/-innen, Studierende, Auszubildende • Kein Schulabschluss, Hauptschulabschluss • Abschluss Polytechnische Oberschule, Realschulabschluss • Fachhochschulreife, Abitur • Keine Angabe – Berufliche Ausbildung • Schüler/-innen, Studierende, Auszubildende • Kein Berufsabschluss, Anlernausbildung • Lehrausbildung, Facharbeiterqualifikation, Berufsfachschulabschluss • Meister-, Techniker-, Fachschulabschluss, Abschluss einer Fachhochschule oder Universität • Keine Angabe – Arbeitslosenzahlen der Bundesagentur für Arbeit

Die ersten beiden Stufen des Gewichtungsverfahrens gleichen die Tatsache aus, dass nicht alle Personen der Grundgesamtheit die gleiche Wahrscheinlichkeit haben, in die

Stichprobe zu gelangen. Je mehr Telefonanschlüsse in einem Haushalt vorhanden sind und je weniger Zielpersonen im Alter von 15 bis 74 Jahren in dem angetroffenen Haushalt wohnen, desto größer ist diese Wahrscheinlichkeit.

Die Anpassung an die einzelnen Eckwerte externer Statistiken erfolgt jeweils in unterschiedlicher demographischer Untergliederung und Kombination der Merkmale. Durch die Anpassung der Ergebnisse der Piloterhebung an die Eckwerte der Bevölkerungsfortschreibung des Statistischen Bundesamtes nach Geschlecht, Alter und Nationalität wird der Bedeutsamkeit dieser Merkmale für das Teilnahmeverhalten an der Erhebung Rechnung getragen. Schulbildung und berufliche Ausbildung haben Einfluss auf die Arbeitsmarktchancen von Personen und damit auf ihren Erwerbsstatus. Durch eine Anpassung der gemessenen Bildungsstrukturen der Piloterhebung an die Strukturen einer geeigneten Merkmalskombination aus Schulbildung und beruflicher Ausbildung des Mikrozensus, welche Unterschiede im Risiko der Erwerbslosigkeit erklären kann, wird einerseits die Präzision der Messung des Erwerbsstatus erhöht, andererseits das höhere Ausfallrisiko von weniger gebildeten Personen ausgeglichen.

Das erhöhte Ausfallrisiko von Nichterwerbspersonen im Panelverlauf im Vergleich mit Erwerbstätigen wurde durch eine separate Gewichtung der Panelfälle ausgeglichen. Hierbei wird die Struktur des Erwerbsstatus der Vorwelle bei allen Wiederholungsbefragten konstant gehalten, unabhängig davon, welchen Erwerbsstatus diese in der aktuellen Welle tatsächlich hatten.

Die Analysen haben gezeigt, dass Erwerbslose zwar kein signifikant höheres Ausfallrisiko haben als Erwerbstätige, beim Arbeitsamt als arbeitslos gemeldete Personen im Vergleich zu nicht arbeitslos Gemeldeten dagegen schon. Durch das Gewichtungsverfahren der Piloterhebung wird deshalb eine Anpassung der Zahl der Personen, die in der Piloterhebung angaben, bei der Bundesagentur für Arbeit arbeitslos gemeldet zu sein, an die Arbeitslosenzahl der Bundesagentur für Arbeit vorgenommen. Das Teilnahmeverhalten der Arbeitslosen im Osten ist anders als das der Arbeitslosen im Westen. Aus diesem Grund erfolgt die Anpassung an die Zahlen der registrierten Arbeitslosen getrennt für die neuen Länder mit Berlin sowie das frühere Bundesgebiet ohne Berlin. Da die Merkmale Arbeitslosigkeit und ILO-Erwerbslosigkeit miteinander korreliert sind, hat diese Anpassung auch einen positiven Effekt auf die Präzision der Messung des Erwerbsstatus.

Das gewählte Gewichtungsverfahren bietet einen Ansatz für die Korrektur der beschriebenen Ausfallprozesse. Ob dies gelingt, hängt von Annahmen über die Strukturen der nicht befragten Personen im Vergleich mit den befragten Personen in Bezug auf das Zielmerkmal der Erhebung ab.

Über das hier geschilderte Gewichtungs- und Hochrechnungsverfahren hinaus sind Weiterentwicklungen des Schätzverfahrens zur Ermittlung von Ergebnissen für die

24) Voraussetzung für die Korrektur der Auswirkungen der Ausfallprozesse ist immer die Annahme, dass sich diejenigen Personen, die nicht an der Erhebung teilgenommen haben, in der Struktur der interessierenden Merkmale innerhalb der Anpassungsklassen nicht von den erfolgreich befragten Personen unterscheiden.

angestrebte Grundgesamtheit der erwerbsfähigen Bevölkerung nach ihrem Erwerbsstatus denkbar. Hier ist insbesondere darüber nachzudenken, im Rahmen der Erhebung auf Basis der kürzlich verabschiedeten Erwerbsstatistikverordnung (ErwerbStatV)<sup>25)</sup> mit 30 000 Befragten ab August 2004 die Panelgewichtung nicht nur mit dem Merkmal des Erwerbsstatus durchzuführen, sondern dieses mit anderen Merkmalen, die in den berichteten Ausfallanalysen einen signifikanten Einfluss gezeigt haben, zu ergänzen. [u](#)

---

<sup>25)</sup> Verordnung über statistische Erhebungen zum Erwerbsstatus der Bevölkerung (Erwerbsstatistikverordnung – ErwerbStatV) vom 10. Mai 2004 (BGBl. I S. 870).

Dipl.-Volkswirt Holger Breiholz

# Ergebnisse des Mikrozensus 2003

*Im Mai 2003 gab es in Deutschland 40,2 Mill. Erwerbspersonen, das waren 170 000 mehr als im April 2002. Zurückzuführen ist dieser Anstieg allein auf die Entwicklung der Zahl weiblicher Erwerbspersonen im früheren Bundesgebiet, die gegenüber dem Vorjahr um 180 000 auf 14,3 Mill. zulegte. Damit trugen die Frauen im früheren Bundesgebiet maßgeblich zu einer weiteren Annäherung der Frauenerwerbsquoten im Westen und Osten Deutschlands bei.*

*Teilzeitarbeit ist am stärksten in kleineren Betrieben verbreitet. So gingen im Mai 2003 50% der abhängig Beschäftigten (ohne Auszubildende) in Arbeitsstätten mit ein oder zwei Mitarbeitern einer Teilzeittätigkeit nach, in Arbeitsstätten mit 11 bis 19 (20 bis 49) tätigen Personen waren es 29% (23%). Mit 16% fiel die Teilzeitquote für Betriebe mit 50 und mehr Mitarbeitern am niedrigsten aus.*

*58% der männlichen und 46% der weiblichen Erwerbslosen im früheren Bundesgebiet gaben im Frühjahr 2003 als Begründung ihrer Arbeitsuche an, den früheren Arbeitsplatz durch Entlassung verloren zu haben. In den neuen Ländern und Berlin-Ost suchten sogar rund drei Viertel der erwerbslosen Männer (77%) und Frauen (75%) eine Stelle nach einer Kündigung durch den Arbeitgeber.*

*1,3 Mill. Nichterwerbspersonen suchten im Mai 2003 trotz Erwerbwunsch keinen Arbeitsplatz. Charakteristisch für diesen als Arbeitskräftereserve zu betrachtenden Personenkreis ist, dass die Frauen mit einem Anteil von fast zwei Dritteln (64%) deutlich überrepräsentiert waren. Der größte Teil (45%) von ihnen gab an, dass persönliche oder familiäre Verpflichtungen einer Arbeitsuche entgegenstanden.*

*Staatliche Transferleistungen gewinnen für den Lebensunterhalt an Bedeutung. So nannten im Frühjahr 2003 23%*

*der Bevölkerung Einkommen aus Renten oder Pensionen als wichtigste Unterhaltsquelle; im April 1991 war dieser Anteil um gut vier Prozentpunkte niedriger ausgefallen. 4% der Bevölkerung lebten vorwiegend von Arbeitslosengeld oder -hilfe; 12 Jahre zuvor waren es lediglich 2%.*

## Vorbemerkung

Der Mikrozensus ist eine amtliche Repräsentativstatistik über die Bevölkerung und den Arbeitsmarkt. Als 1%-Haushaltsstichprobe liefert er bereits seit 1957 in jährlichem Rhythmus statistische Informationen in tiefer fachlicher und regionaler Gliederung zu einer Vielzahl unterschiedlicher Themen, insbesondere zur sozio-ökonomischen Grundstruktur der Bevölkerung, zum Erwerbsleben, zur Altersvorsorge und zur Ausbildung der Erwerbsbevölkerung. Hinzu kommen mehrjährige Zusatzprogramme, die beispielsweise die Wohnsituation und – im Rahmen des Mikrozensus 2003 – die Gesundheit der Bevölkerung (u. a. das Vorliegen einer Krankheit oder Unfallverletzung, Rauchgewohnheiten und Körpermaße) betreffen. Dabei ist ein großer Vorzug des Mikrozensus, dass Daten nicht nur auf der Ebene von Personen, sondern auch im Haushalts- und Familienzusammenhang zur Verfügung gestellt werden können.

Mit der Einführung des Mikrozensus in den neuen Ländern und Berlin-Ost im Jahr 1991 wurde eine wichtige gemeinsame Datenbasis für einen Vergleich der Bevölkerungs- und Arbeitsmarktstrukturen in allen 16 Bundesländern geschaffen. Vergleichbare Daten für alle Mitgliedstaaten der Europäischen Union (EU) stellt die Arbeitskräfteerhebung der EU bereit, die bereits seit 1968 als Unterstichprobe in den Mikrozensus integriert ist. Mit dem derzeit geltenden Mikro-

zensusgesetz<sup>1)</sup> erfolgte eine weitgehende Harmonisierung der Erhebungsmerkmale von Mikrozensus und Arbeitskräfteerhebung hinsichtlich der verwendeten Definitionen und Abgrenzungen sowie der Periodizität und des Auswahlsatzes. Auf diese Weise wurde vor allem der Belastung der Befragten, dem organisatorischen Aufwand und den Kosten der Erhebung Rechnung getragen.

Mit der Realisierung des europäischen Binnenmarktes nimmt die Bedeutung vergleichbarer Daten über den Arbeitsmarkt in den Mitgliedstaaten der EU ständig zu. Erkennbar ist dies daran, dass das jährliche Frageprogramm der Arbeitskräfteerhebung mit neuen Variablen und Antwortkategorien ab 1992<sup>2)</sup> und 1998<sup>3)</sup> deutlich erweitert wurde. Darüber hinaus sind seit 1998 neben dem Grundprogramm jährlich wechselnde Ad-hoc-Module zu speziellen Fragestellungen vorgesehen.<sup>4)</sup> Im gemeinsamen Erhebungsprogramm der Arbeitskräfteerhebung und des Mikrozensus 2003 gehen bereits mehr als zwei Drittel (68%) aller Merkmale auf Datenanforderungen der EU zurück. Dabei ist bemerkenswert, dass – als Folge der erwähnten Harmonisierungsanstrengungen – über die Hälfte (58%) aller Merkmale beiden Erhebungen zuzuordnen sind.

Im Rahmen des Ad-hoc-Moduls der Arbeitskräfteerhebung 2003 wurde die aktuelle Thematik „Lebenslanges Lernen“ aufgegriffen. Mit diesem Zusatzprogramm konnten für Deutschland wichtige EU-weit vergleichbare Daten nicht nur zum Bereich der formellen Bildung, sondern erstmals auch zu Aktivitäten der informellen allgemeinen oder beruflichen Weiterbildung bereitgestellt werden.

Der vorliegende Beitrag kann nur einen kleinen Ausschnitt aus dem riesigen Informationspotenzial des Mikrozensus 2003 präsentieren. Dabei liegt das Schwergewicht auf dem Stand und der Entwicklung des Arbeitsmarktes. Eine detaillierte Darstellung der Ergebnisse des Mikrozensus findet sich in den vom Statistischen Bundesamt herausgegebenen Reihen 4.1.1 „Stand und Entwicklung der Erwerbstätigkeit“, 4.1.2 „Beruf, Ausbildung und Arbeitsbedingungen der Erwerbstätigen“ und 3 „Haushalte und Familien“ der Fachserie 1 „Bevölkerung und Erwerbstätigkeit“. Eine Fülle von Informationen enthält auch die zum siebten Mal erschienene Broschüre „Leben und Arbeiten in Deutschland“. Diese Publikation ist über das Internetangebot des Statistischen Bundesamtes unter <http://www.destatis.de> (Pfad: Presse/ Presseveranstaltungen) kostenlos abrufbar.

## Rückgang der Bevölkerung im erwerbsfähigen Alter

Im Mai 2003 lebten in Deutschland 82,5 Mill. Menschen, das waren knapp 50000 mehr als im April 2002 (siehe

Tabelle 1). Im früheren Bundesgebiet legte die Bevölkerungszahl gegenüber dem Vorjahr um rund 160000 auf fast 67,7 Mill. zu. In den neuen Ländern und Berlin-Ost fiel sie im gleichen Zeitraum um etwa 120000 auf knapp 14,9 Mill. In diesen gegenläufigen Entwicklungen spiegelt sich wider, dass die Ost-West-Wanderungen in Deutschland nach wie vor wesentlich zum Bevölkerungswachstum im Westen Deutschlands beitragen.

Tabelle 1: Bevölkerung<sup>1)</sup> nach Geschlecht und Staatsangehörigkeit  
Ergebnisse des Mikrozensus  
1 000

Geschlecht Staatsangehörigkeit	Insgesamt		Darunter: 15- bis 64-Jährige	
	Mai 2003	April 2002	Mai 2003	April 2002
Deutschland				
Bevölkerung .....	82 502	82 455	55 058	55 230
Männer .....	40 330	40 282	27 824	27 927
Frauen .....	42 172	42 173	27 234	27 303
Deutsche .....	75 341	75 327	49 571	49 804
Männer .....	36 585	36 532	24 968	25 084
Frauen .....	38 756	38 795	24 603	24 721
Ausländer/-innen .....	7 161	7 128	5 487	5 426
Männer .....	3 745	3 750	2 856	2 843
Frauen .....	3 416	3 378	2 631	2 583
Früheres Bundesgebiet				
Bevölkerung .....	67 651	67 489	44 780	44 853
Männer .....	33 049	32 954	22 556	22 606
Frauen .....	34 602	34 534	22 225	22 245
Neue Länder und Berlin-Ost				
Bevölkerung .....	14 850	14 968	10 278	10 378
Männer .....	7 280	7 328	5 269	5 321
Frauen .....	7 570	7 640	5 009	5 058

1) Bevölkerung am Ort der alleinigen bzw. der Hauptwohnung.

Auswirkungen auf den Arbeitsmarkt sind von der Entwicklung der Bevölkerung im erwerbsfähigen Alter (15 bis 64 Jahre) zu erwarten, die für das gesamtwirtschaftliche Arbeitskräftepotenzial eine obere Grenze vorgibt. Von April 2002 bis Mai 2003 ist die Zahl der Personen in der Altersgruppe der 15- bis 64-Jährigen in Deutschland um etwa 170000 auf knapp 55,1 Mill. gesunken. Dabei darf nicht übersehen werden, dass dieser Rückgang ohne den Zuwachs bei den Ausländerinnen und Ausländern um rund 60000 Personen noch höher ausgefallen wäre. Im früheren Bundesgebiet verringerte sich die Zahl der Personen im erwerbsfähigen Alter gegenüber April 2002 um gut 70000 auf knapp 44,8 Mill. Infolge von Wanderungsverlusten und demographischer Alterung lebten in den neuen Ländern und Berlin-Ost im Mai 2003 nur noch knapp 10,3 Mill. Menschen im Erwerbsalter, das waren 100000 weniger als ein Jahr zuvor.

1) Mikrozensusgesetz und Gesetz zur Änderung des Bundesstatistikgesetzes vom 17. Januar 1996 (BGBl. I S. 34), Artikel 1: Gesetz zur Durchführung einer Repräsentativstatistik über die Bevölkerung und den Arbeitsmarkt sowie die Wohnsituation der Haushalte (Mikrozensusgesetz).

2) Siehe Verordnung (EWG) Nr. 3711/91 des Rates vom 16. Dezember 1991 zur Durchführung einer jährlichen Stichprobenerhebung über Arbeitskräfte in der Gemeinschaft (Amtsbl. der EG Nr. L 351, S. 1).

3) Siehe Verordnung (EG) Nr. 577/98 des Rates vom 9. März 1998 zur Durchführung einer jährlichen Stichprobenerhebung über Arbeitskräfte in der Gemeinschaft (Amtsbl. der EG Nr. L 77, S. 3).

4) Im Hinblick auf die bereits erreichte Belastungsgrenze der Befragten im Mikrozensus wurden die Ad-hoc-Module in Deutschland bislang nicht oder nur zum Teil umgesetzt. Letzteres ist ab 2001 durch die Verordnung zur Aussetzung einzelner Merkmale des Mikrozensusgesetzes vom 3. April 2000 (BGBl. I S. 442) ermöglicht worden.

## Erwerbsverhalten der Frauen: trotz Annäherung immer noch deutliche Unterschiede zwischen West und Ost

Auf der Angebotsseite des Arbeitsmarktes standen im Mai 2003 fast 40,2 Mill. Erwerbspersonen [d. h. Erwerbstätige<sup>5)</sup> und Erwerbslose<sup>6)</sup>] zur Verfügung. Damit erhöhte sich die Zahl der Erwerbspersonen gegenüber dem Vorjahr um gut 170 000. Dieser Zuwachs geht allein auf die Entwicklung der Zahl weiblicher Erwerbspersonen im früheren Bundesgebiet zurück, die gegenüber April 2002 um 180 000 auf fast 14,3 Mill. angestiegen ist. Demgegenüber blieb die Zahl der im Westen lebenden männlichen Erwerbspersonen mit rund 18,2 Mill. nahezu konstant. Für die neuen Länder und Berlin-Ost lassen sich keine markanten Veränderungen feststellen. Sowohl bei den Männern (4,2 Mill.) als auch bei den Frauen (3,6 Mill.) lagen die Erwerbspersonenzahlen im Mai 2003 nur geringfügig unter dem Vorjahresniveau (siehe Tabelle 2).

Zuwächse bei der Zahl der Erwerbspersonen trotz sinkender Bevölkerungszahl bei den 15- bis 64-Jährigen deuten darauf hin, dass das Arbeitsangebot nicht nur von demographischen Veränderungen, sondern auch vom Erwerbsverhalten der Menschen beeinflusst wird. Der Analyse des Erwerbsverhaltens der Bevölkerung dienen insbesondere alters- und geschlechtsspezifische Erwerbsquoten, die als Anteil der Erwerbspersonen an allen Personen eines bestimmten Alters und Geschlechts definiert sind und als Maß für die Stärke der Erwerbsorientierung interpretiert werden können.

Bezogen auf die Altersgruppe der 15- bis 64-Jährigen lag die Erwerbsquote der Männer im Mai 2003 bei 79,2%. Dabei bestanden nur geringe Unterschiede zwischen West und Ost: Mit 79,4% fiel die Erwerbsquote der Männer im früheren Bundesgebiet um etwa einen Prozentpunkt höher aus als in den neuen Ländern und Berlin-Ost (78,3%). Große Übereinstimmung zeigt sich auch bei den Erwerbsquotenkurven, die nennenswerte Ost-West-Differenzen nur im oberen Altersbereich und in den unteren Altersgruppen erkennen lassen. Letztere sind vor dem Hintergrund zu sehen, dass die Erwerbsquoten der Jugendlichen unter 25 Jahren im Osten seit der Vereinigung Deutschlands drastisch gesunken sind. So lässt sich gegenüber 1991 ein Rückgang der Erwerbsquote bei den 15- bis 19-Jährigen von 58,4 auf 37,5% und bei den 20- bis 24-Jährigen von 86,7 auf 79,3% feststellen (siehe Tabelle 3). Im Westen entwickelten sich die Vergleichsquoten – ausgehend von einem niedrigeren Niveau – moderater: Unter den 15- bis 19-Jährigen (20- bis 24-Jährigen) zählten hier im April 1991 40,0% (77,0%), im Mai 2003 32,0% (73,4%) zu den Erwerbspersonen. Aus diesen Veränderungen ist ersichtlich, dass sich bei der jüngeren Generation in den neuen Bundesländern eine Anpassung an das Bildungsverhalten der Altersgenossen im Westen (d. h. längere Verweildauer im Schulsystem und späterer Berufseinstieg) vollzogen hat.

Im Gegensatz zu den Männern unterscheiden sich die Frauen im Westen und Osten hinsichtlich ihrer Erwerbsneigung deutlich voneinander. Mit 71,4% im Mai 2003 fiel die Erwerbsquote der Frauen im erwerbsfähigen Alter im Osten knapp acht Prozentpunkte höher aus als im Westen

Tabelle 2: Bevölkerung<sup>1)</sup> nach Beteiligung am Erwerbsleben  
Ergebnisse des Mikrozensus

Beteiligung am Erwerbsleben	Mai 2003			April 2002			Veränderung Mai 2003 gegenüber April 2002					
	insgesamt	männlich	weiblich	insgesamt	männlich	weiblich	insgesamt	männlich	weiblich	insgesamt	männlich	weiblich
	1 000						%	1 000	%	1 000	%	
Deutschland												
Bevölkerung .....	82 502	40 330	42 172	82 455	40 282	42 173	+47	+0,1	+48	+0,1	-1	+0,0
Erwerbspersonen .....	40 195	22 312	17 883	40 022	22 318	17 704	+173	+0,4	-6	+0,0	+179	+1,0
Erwerbstätige .....	36 172	19 996	16 176	36 536	20 336	16 200	-364	-1,0	-340	-1,7	-24	-0,1
Erwerbslose <sup>2)</sup> .....	4 022	2 316	1 707	3 486	1 982	1 504	+536	+15,4	+334	+16,6	+203	+13,5
Nichterwerbspersonen <sup>3)</sup> .	42 307	18 018	24 289	42 433	17 964	24 469	-126	-0,3	+54	+0,3	-180	-0,7
Früheres Bundesgebiet												
Bevölkerung .....	67 651	33 049	34 602	67 489	32 954	34 534	+162	+0,2	+95	+0,3	+68	+0,2
Erwerbspersonen .....	32 457	18 163	14 294	32 268	18 158	14 110	+189	+0,6	+5	+0,0	+184	+1,3
Erwerbstätige .....	29 847	16 601	13 246	30 132	16 880	13 252	-285	-0,9	-279	-1,7	-6	+0,0
Erwerbslose <sup>2)</sup> .....	2 609	1 561	1 048	2 136	1 278	858	+473	+22,1	+283	+22,1	+190	+22,1
Nichterwerbspersonen <sup>3)</sup>	35 195	14 887	20 308	35 220	14 796	20 424	-25	-0,1	+91	+0,6	-116	-0,6
Neue Länder und Berlin-Ost												
Bevölkerung .....	14 850	7 280	7 570	14 968	7 328	7 640	-118	-0,8	-48	-0,7	-70	-0,9
Erwerbspersonen .....	7 739	4 150	3 589	7 754	4 160	3 594	-15	-0,2	-10	-0,2	-5	-0,1
Erwerbstätige .....	6 325	3 395	2 930	6 404	3 456	2 948	-79	-1,2	-61	-1,8	-18	-0,6
Erwerbslose <sup>2)</sup> .....	1 414	755	659	1 350	704	646	+64	+4,7	+51	+7,2	+13	+2,0
Nichterwerbspersonen <sup>3)</sup>	7 111	3 130	3 981	7 214	3 186	4 046	-103	-1,4	-56	-1,8	-65	-1,6

1) Bevölkerung am Ort der alleinigen bzw. der Hauptwohnung. – 2) Erwerbslose sind gemäß ILO-Konzept für den Arbeitsmarkt sofort verfügbar. – 3) Einschließlich für den Arbeitsmarkt nicht sofort verfügbarer, Arbeit suchender Nichterwerbstätiger.

5) Als Erwerbstätige gelten Personen im Alter von 15 Jahren oder älter, die in der Berichtswoche in einem Arbeitsverhältnis standen bzw. als Selbstständige oder mithelfende Familienangehörige tätig waren. Dabei spielt es keine Rolle, ob die Tätigkeit regelmäßig oder nur gelegentlich ausgeübt wurde.

6) Bei der Abgrenzung der Erwerbslosen folgt der Mikrozensus den international geltenden Standards der Internationalen Arbeitsorganisation (ILO). Als Erwerbslose gelten danach nicht erwerbstätige Personen über 14 Jahre, die sich in den letzten drei Wochen vor der Berichtswoche bzw. in der Berichtswoche aktiv um eine Erwerbsarbeit bemüht haben und für eine (neue) Tätigkeit sofort (d. h. innerhalb von zwei Wochen) verfügbar sind.

(63,6%). Dabei ist jedoch zu berücksichtigen, dass die entsprechenden Frauenerwerbsquoten im April 1991 noch fast 18 Prozentpunkte auseinander lagen. Somit hat in Bezug auf den Umfang der Frauenerwerbsbeteiligung in den beiden Teilen Deutschlands im Laufe der Jahre eine Annäherung, jedoch noch keine Anpassung stattgefunden. Dazu haben in erster Linie die Frauen im früheren Bundesgebiet beigetragen, deren Erwerbsquote in der Altersgruppe der 15- bis 64-jährigen gegenüber April 1991 um über sechs Prozentpunkte angestiegen ist. Im gleichen Zeitraum hat die Vergleichsquote in den neuen Ländern und Berlin-Ost um gut drei Prozentpunkte abgenommen. Die nach wie vor stärkere Erwerbsorientierung der Frauen in den neuen Ländern und Berlin-Ost dürfte einerseits auf wirtschaftliche Notwendigkeiten, andererseits auf das noch von den sozialen Normen und ökonomischen Zwängen in der ehemaligen DDR geprägte Lebenskonzept zurückzuführen sein.

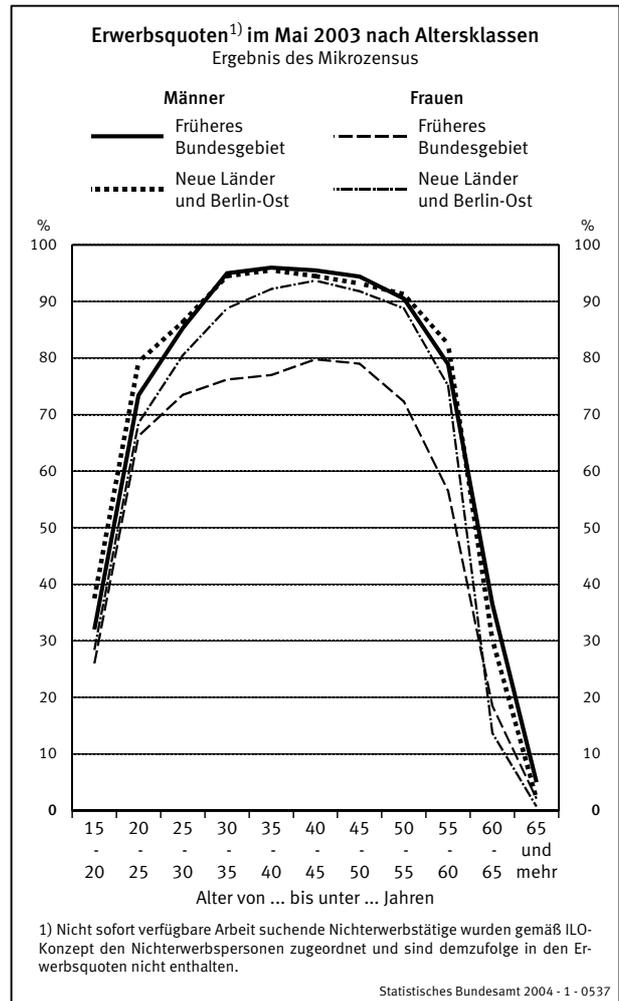
Die Erwerbsquotenkurven dokumentieren, dass die Frauen in den neuen Ländern und Berlin-Ost im Mai 2003 in allen Altersgruppen unter 60 Jahren relativ häufiger am Erwerbsleben teilnahmen als im früheren Bundesgebiet. Bei den 60- bis 64-Jährigen machen sich offensichtlich immer noch – ebenso wie bei den Männern – die nach der deutschen Vereinigung eingeführten Vorruhestandsregelungen in den neuen Bundesländern bemerkbar. Bezüglich der Erwerbsquoten der 15- bis 24-jährigen Frauen gelten die Ausführungen zur Erwerbsbeteiligung der Männer in diesem Alter entsprechend. Erwartungsgemäß groß sind die Unterschiede in den Altersgruppen, in denen familiäre Pflichten wie Kinderbetreuung die größte Rolle spielen. So waren in den neuen Ländern und Berlin-Ost 88,8% der 30- bis 34-jährigen und 92,2% der 35- bis 39-jährigen Frauen erwerbstätig

Tabelle 3: Erwerbsquoten nach Altersklassen<sup>1)</sup>  
Ergebnisse des Mikrozensus  
Prozent

Alter von ... bis unter ... Jahren	Männer		Frauen	
	Mai 2003	April 1991	Mai 2003	April 1991
Früheres Bundesgebiet				
15 – 20 .....	32,0	40,0	26,0	33,8
20 – 25 .....	73,4	77,0	66,2	72,0
25 – 30 .....	85,3	85,6	73,5	69,1
30 – 35 .....	95,0	94,6	76,2	64,9
35 – 40 .....	96,0	96,3	77,0	67,6
40 – 45 .....	95,5	96,3	79,8	69,5
45 – 50 .....	94,4	94,9	79,0	66,1
50 – 55 .....	90,5	91,9	72,3	57,6
55 – 60 .....	78,9	79,4	56,5	43,0
60 – 65 .....	36,5	34,2	18,5	11,9
15 – 65 .....	79,4	81,1	63,6	57,2
Neue Länder und Berlin-Ost				
15 – 20 .....	37,5	58,4	28,4	44,6
20 – 25 .....	79,3	86,7	68,5	83,2
25 – 30 .....	86,4	94,4	80,5	92,5
30 – 35 .....	94,5	96,9	88,8	93,7
35 – 40 .....	95,5	97,9	92,2	95,8
40 – 45 .....	94,5	97,7	93,7	94,8
45 – 50 .....	93,2	97,3	91,8	93,9
50 – 55 .....	91,3	94,1	88,8	89,0
55 – 60 .....	82,5	71,2	75,1	34,6
60 – 65 .....	30,2	25,8	13,7	4,6
15 – 65 .....	78,3	84,5	71,4	74,8

1) Nicht sofort verfügbare, Arbeit suchende Nichterwerbstätige wurden gemäß ILO-Konzept den Nichterwerbspersonen zugeordnet und sind demzufolge in den Erwerbsquoten nicht enthalten.

Schaubild 1



tig oder erwerbslos; im früheren Bundesgebiet lagen diese Anteile lediglich bei 76,2 bzw. 77,0%. Dennoch ist für die Frauen im früheren Bundesgebiet zu konstatieren, dass sich ihre ehemals M-förmige – das heißt durch einen ausgeprägten „Knick“ im Familiengründungsalter gekennzeichnete – Erwerbsquotenkurve mittlerweile nicht nur niveau-, sondern auch gestaltmäßig jener der Frauen in den neuen Ländern und Berlin-Ost angenähert hat. So ist ihre Erwerbsquote in der Altersgruppe der 30- bis 34-Jährigen gegenüber April 1991 um gut 11, in der folgenden Altersgruppe der 35- bis 39-Jährigen um gut 9 Prozentpunkte angewachsen. Dies deutet darauf hin, dass die Frauen im Westen immer seltener bzw. immer kürzere Familienpausen in Anspruch nehmen.

## Negative Arbeitsmarktentwicklung in den neuen und alten Bundesländern

Für die Beurteilung der Arbeitsmarktsituation ist eine Differenzierung der Erwerbspersonen in Erwerbslose und Erwerbstätige erforderlich. Sie lässt erkennen, dass die Zahl der Erwerbstätigen in Deutschland von April 2002 bis Mai

2003 um gut 360 000 auf knapp 36,2 Mill. gesunken, die Zahl der Erwerbslosen hingegen im gleichen Zeitraum um fast 540 000 auf 4,0 Mill. gestiegen ist (siehe Tabelle 2). Die verschlechterte Arbeitsmarktlage trifft für beide Teile Deutschlands zu: Im früheren Bundesgebiet gab es im Mai 2003 rund 29,8 Mill. Erwerbstätige, das waren 285 000 weniger als im Vorjahr. Die Zahl der Erwerbslosen legte seit April 2002 kräftig um gut 470 000 auf rund 2,6 Mill. zu. Die neuen Länder und Berlin-Ost verzeichneten gegenüber April 2002 einen Rückgang der Erwerbstätigkeit um knapp 80 000 und einen Anstieg der Erwerbslosigkeit um gut 60 000 Personen. Somit waren dem Arbeitsmarkt im Osten im Frühjahr 2003 rund 6,3 Mill. Erwerbstätige und 1,4 Mill. Erwerbslose zuzurechnen.

Besonders drastische Veränderungen vollzogen sich bei den Männern im früheren Bundesgebiet: Ihre gegenüber dem Vorjahr beinahe unveränderte Erwerbspersonenzahl beruht auf einem Plus bei den Erwerbslosen (+ 283 000) und einem Minus bei den Erwerbstätigen (- 279 000) in etwa gleicher Höhe. Die gestiegene Erwerbspersonenzahl der Frauen im früheren Bundesgebiet rührt daher, dass seit April 2002 netto rund 190 000 Frauen in den Erwerbslosenstatus übergewechselt sind. Die Zahl der erwerbstätigen Frauen blieb hingegen geringfügig (- 6 000) unter dem Vergleichswert vom April 2002. Kennzeichnend für die geschlechtsspezifische Arbeitsmarktsituation in den neuen Ländern und Berlin-Ost ist, dass hier die Männer (- 1,8%) und Frauen (- 0,6%) vom Rückgang der Erwerbstätigkeit relativ stärker betroffen waren als im Westen (- 1,7 bzw. 0,0%). Dagegen fiel der Anstieg der Erwerbslosenzahl mit + 7,2% bei den Männern und + 2,0% bei den Frauen deutlich moderater aus als im früheren Bundesgebiet (beide Geschlechter: + 22,1%).

## Teilzeitbeschäftigung in kleinen Betrieben am stärksten verbreitet

Vom Rückgang der Erwerbstätigkeit war die Teilzeitbeschäftigung nicht betroffen: Unter den 32,0 Mill. abhängig Beschäftigten stufen sich im Mai 2003 knapp 7,2 Mill. als Teilzeitbeschäftigte ein<sup>7)</sup>; im April 2002 waren es lediglich 6,9 Mill. Damit stieg die Teilzeitquote der abhängig Beschäftigten – der Anteil der Teilzeitbeschäftigten an allen abhängig Beschäftigten – gegenüber dem Vorjahr um einen Prozentpunkt. Der Vergleichswert von 1991 – dem Jahr der ersten Durchführung des Mikrozensus in den neuen Ländern und Berlin-Ost – wurde sogar um über acht Prozentpunkte übertroffen; dies entspricht einem Zuwachs der Zahl der abhängig Teilzeitbeschäftigten um 2,4 Mill. bzw. 51%.

Im früheren Bundesgebiet ist die abhängige Teilzeitbeschäftigung nach wie vor deutlich stärker verbreitet als in den neuen Ländern und Berlin-Ost. Zwar stieg die Teilzeitquote im Osten von April 1991 bis Mai 2003 um etwa sieben Prozentpunkte auf 15,6%, der Abstand zu der – im

Tabelle 4: Abhängig Teilzeitbeschäftigte  
Ergebnisse des Mikrozensus

Gegenstand der Nachweisung	Mai 2003		April 1991	
	Abhängig Teilzeitbeschäftigte	Teilzeitquote <sup>1)</sup>	Abhängig Teilzeitbeschäftigte	Teilzeitquote <sup>1)</sup>
	1 000	%	1 000	%
Deutschland .....	7 168	22,4	4 736	14,0
Männer .....	1 037	6,0	402	2,1
Frauen .....	6 131	41,4	4 334	30,2
Früheres Bundesgebiet ...	6 276	23,8	4 075	15,4
Männer .....	868	6,1	355	2,3
Frauen .....	5 408	44,7	3 720	34,3
Neue Länder und				
Berlin-Ost .....	892	15,6	661	8,9
Männer .....	169	5,7	47	1,2
Frauen .....	724	26,6	614	17,5

1) Anteil der abhängig Teilzeitbeschäftigten an allen abhängig Erwerbstätigen.

gleichen Zeitraum um über acht Prozentpunkte auf 23,8% gestiegenen – Teilzeitquote im Westen ist jedoch noch größer geworden (siehe Tabelle 4). Der Trend zu mehr Teilzeitarbeit zeigt sich insbesondere bei den Frauen, die bei einer Teilzeitquote von 41,4% (1991: 30,2%) im Mai 2003 in Deutschland rund 86% aller Teilzeittätigen in abhängiger Beschäftigung stellten. Im früheren Bundesgebiet ist die Zahl weiblicher Teilzeitkräfte seit April 1991 um 45% auf 5,4 Mill., in den neuen Ländern und Berlin-Ost um 18% auf rund 720 000 angewachsen. Parallel dazu legte die Teilzeitquote der Frauen im Westen von 34,3 auf 44,7% und im Osten von 17,5 auf 26,6% zu.

Bei den Männern gewinnt die Teilzeitarbeit allmählich an Bedeutung: Seit April 1991 stieg ihre Teilzeitquote um vier Prozentpunkte auf 6% an; im gleichen Zeitraum erhöhte sich ihr Anteil an allen Teilzeitbeschäftigten von 8 auf 14%. Somit gab es im Mai 2003 rund 1 Mill. abhängig erwerbstätiger Männer, deren Beschäftigungsumfang nach eigenen Angaben nicht dem Vollzeitstandard entsprach.

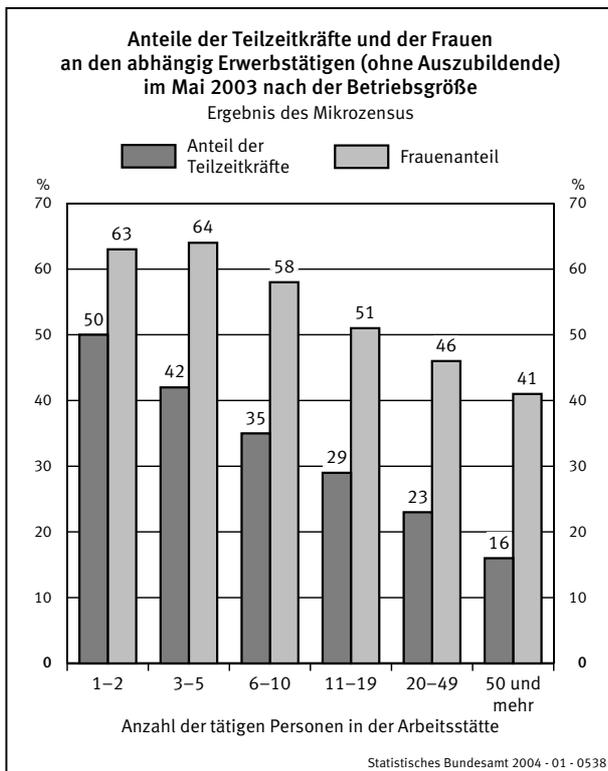
Teilzeitarbeit ist am stärksten in kleineren Betrieben<sup>8)</sup> verbreitet. So gingen im Mai 2003 rund 50% der abhängig Beschäftigten (ohne Auszubildende) in Arbeitsstätten mit ein bis zwei Mitarbeitern einer Teilzeittätigkeit nach; in Arbeitsstätten mit drei bis fünf Mitarbeitern waren es 42% (siehe Schaubild 2). Mehr als jede(r) dritte (35%) abhängig Beschäftigte (ohne Auszubildende) hatte einen Teilzeitarbeitsplatz in einem Betrieb inne, dem sechs bis zehn Mitarbeiter angehörten. Für Arbeitsstätten mit 11 bis 19 (20 bis 49) tätigen Personen beläuft sich der Vergleichsanteil auf etwa 29% (23%). Mit 16% fiel die Teilzeitquote für Arbeitsstätten mit 50 und mehr Mitarbeitern am niedrigsten aus.

Der in Abhängigkeit von der Größe des Betriebes dargestellte Verbreitungsgrad der abhängigen Teilzeitbeschäftigung ist vor dem Hintergrund zu betrachten, dass der Frauenanteil an den in der Arbeitsstätte abhängig beschäf-

7) Für die Ermittlung des Umfangs der Teilzeitarbeit bietet der Mikrozensus alternative Erhebungstatbestände: Zum einen können sich die Erwerbstätigen selbst als Vollzeit- oder Teilzeittätige einstufen, zum anderen wird nach der normalerweise geleisteten Wochenarbeitszeit gefragt. In diesem Beitrag beruhen die Ergebnisse zur Teilzeitbeschäftigung auf der Selbsteinstufung der Befragten.

8) Die Ergebnisse zur Betriebsgröße beruhen auf den Angaben der Befragten.

Schaubild 2



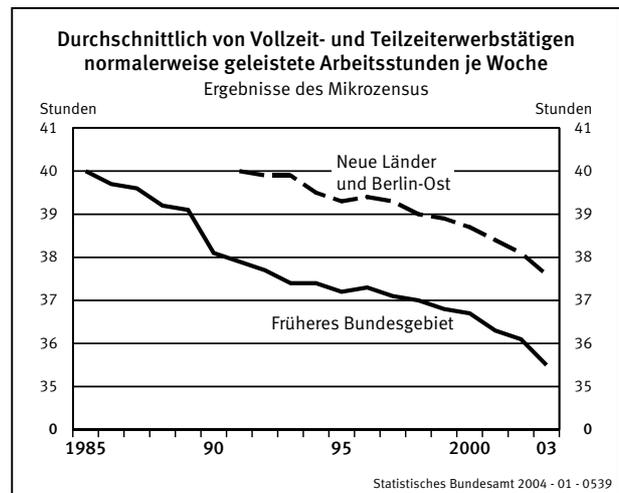
tigten Personen mit wachsender Betriebsgröße abnimmt. So waren im Mai 2003 63% der abhängig Beschäftigten (ohne Auszubildende) in Arbeitsstätten mit ein oder zwei Mitarbeitern weiblichen Geschlechts. Betriebe mit 20 bis 49 Beschäftigten wiesen lediglich einen Frauenanteil von 46% auf. Den mit 41% niedrigsten Wert hatte der Frauenanteil in Betrieben mit 50 und mehr Beschäftigten. Der Anteil der Frauen an allen Erwerbstätigen in Deutschland betrug im Mai 2003 45%.

### Durchschnittliche Wochenarbeitszeit immer kürzer

Tiefere Einblicke in den zeitlichen Umfang der Erwerbstätigkeit gestatten die Ergebnisse des Mikrozensus zur normalerweise geleisteten Arbeitszeit. Aus diesen lässt sich für die Vergangenheit eine deutliche Arbeitszeitverkürzung in Deutschland ablesen. So ist im früheren Bundesgebiet die durchschnittlich normalerweise geleistete Wochenarbeitszeit der Erwerbstätigen von April 1985 bis Mai 2003 von 40,0 auf 35,5 Stunden gesunken. Für die neuen Länder und Berlin-Ost ist seit April 1991 ein Arbeitszeitrückgang von 40,0 auf 37,6 Stunden feststellbar (siehe Schaubild 3). Damit lag die normalerweise geleistete Arbeitszeit je Woche im Mai 2003 im Osten auch 13 Jahre nach der Vereinigung immer noch deutlich höher als im Westen Deutschlands. Hinsichtlich der genannten Stundenzahlen ist zu beachten, dass Arbeitszeiten aus Nebenerwerbstätigkeiten nicht eingerechnet wurden.

Bei der Interpretation der zunehmend kürzeren Wochenarbeitszeit je Erwerbstätigen muss berücksichtigt werden,

Schaubild 3



dass – wie bereits dargelegt – die Teilzeitarbeit insbesondere bei den weiblichen Beschäftigten enorm an Bedeutung gewonnen hat. Dies hatte zur Folge, dass die durchschnittliche Wochenarbeitszeit der erwerbstätigen Frauen in Deutschland von April 1991 bis Mai 2003 von 34,4 auf 30,8 Stunden gesunken ist (früheres Bundesgebiet: von 33,2 auf 29,8 Stunden, neue Länder und Berlin-Ost: von 38,3 auf 35,0 Stunden). Bei den Männern fiel der Rückgang der durchschnittlichen Wochenarbeitszeit im gleichen Zeitraum (von 41,2 auf 40,0 Stunden) merklich schwächer aus (siehe Tabelle 5). Wegen der relativ niedrigen durchschnittlichen Stundenzahl lag der Anteil der Frauen am gesamten Arbeitsvolumen (d. h. an den von allen Erwerbstätigen normalerweise geleisteten Arbeitsstunden) im Mai 2003 nur bei 38%, das heißt sieben Prozentpunkte niedriger als der bereits genannte Anteil der Frauen an den Erwerbstätigen.

Tabelle 5: Durchschnittlich von Erwerbstätigen normalerweise geleistete Arbeitsstunden je Woche  
Ergebnisse des Mikrozensus

Gegenstand der Nachweisung	Deutschland	Früheres Bundesgebiet	Neue Länder und Berlin-Ost
Mai 2003			
Insgesamt .....	35,8	35,5	37,6
Männer .....	40,0	40,0	39,8
Frauen .....	30,8	29,8	35,0
April 1991			
Insgesamt .....	38,4	37,9	40,0
Männer .....	41,2	41,1	41,4
Frauen .....	34,4	33,2	38,3

### Entlassungen häufigster Grund für die Arbeitsuche

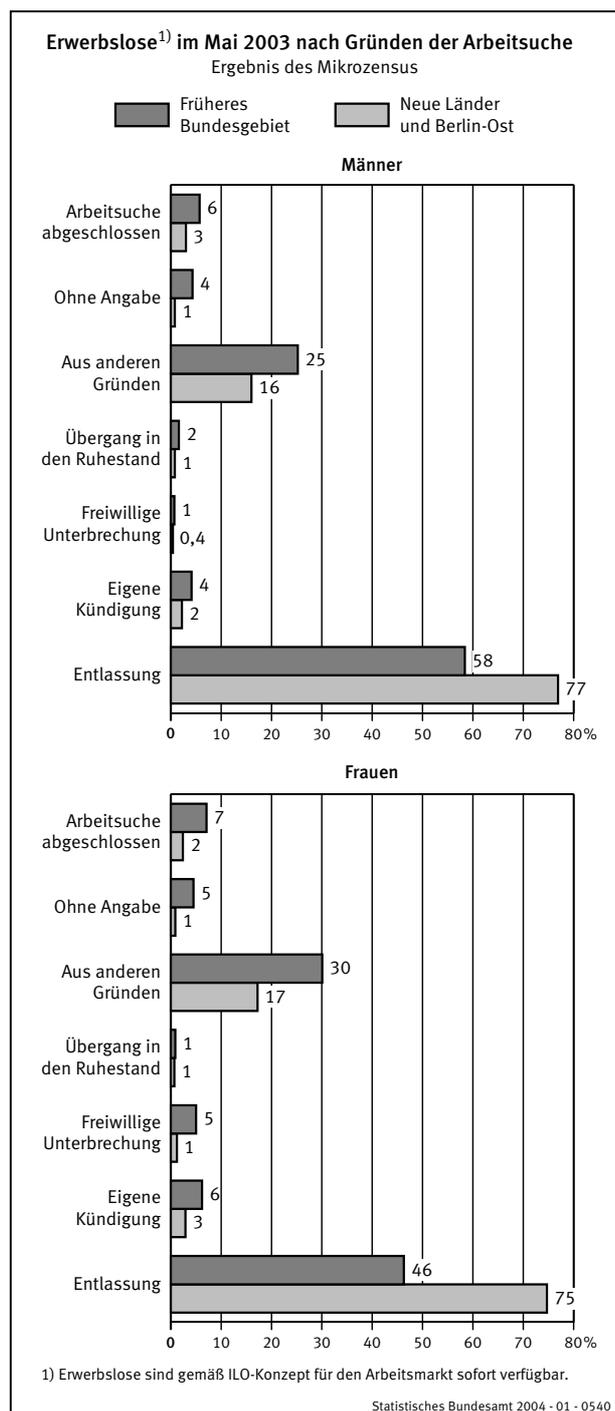
Der hohe Stand der Erwerbslosigkeit stellt für die Arbeitsmarktpolitik gegenwärtig die größte Herausforderung dar. Der Mikrozensus liefert eine Fülle von aufschlussreichen Informationen zur Struktur der Erwerbslosen, u. a. zu den Gründen und zur Dauer ihrer Arbeitsuche.

Entlassungen stellten im Mai 2003 bei den Erwerbslosen den wichtigsten Grund für die Suche nach einer Erwerbstätigkeit dar.

tigkeit dar: 58% der männlichen und 46% der weiblichen Erwerbslosen im früheren Bundesgebiet gaben als Begründung ihrer Arbeitsuche an, den früheren Arbeitsplatz durch Entlassung verloren zu haben. Aufgrund des massiven Stellenabbaus als Folge des wirtschaftlichen Umbruchs und des nur langsamen Aufbauprozesses suchten in den neuen Ländern und Berlin-Ost sogar rund drei Viertel der erwerbslosen Männer (77%) und Frauen (75%) einen Arbeitsplatz nach einer Kündigung durch den Arbeitgeber (siehe Schaubild 4). Alle weiteren Gründe spielten demgegenüber im Osten nur eine untergeordnete Rolle. Eine eigene Kündigung hat

ten im Westen 5%, im Osten nur knapp 3% der Erwerbslosen als Grund angegeben. 5% der als erwerbslos einzustufenden Frauen im früheren Bundesgebiet waren nach einer freiwilligen Unterbrechung erwerbslos geworden; diesen Grund nannten nur rund 1% der erwerbslosen Frauen in den neuen Ländern und Berlin-Ost. Überproportional häufig begründeten freiwillige Unterbrechungen die Erwerbslosigkeit verheirateter Frauen (7%) im früheren Bundesgebiet. In den neuen Ländern und Berlin-Ost stellten nur 1% der erwerbslosen verheirateten Frauen ihre berufliche Tätigkeit aus eigenem Antrieb vorübergehend ein. Dies ist ein Hinweis darauf, dass Familienpflichten die Frauenerwerbstätigkeit in West und Ost immer noch in unterschiedlicher Weise beeinflussen.

Schaubild 4



## Jeder zweite Erwerbslose seit einem Jahr oder länger auf Arbeitsuche

Etwa jeder zweite Erwerbslose (51%) in Deutschland mit Angabe zur Dauer der Arbeitsuche bemühte sich im Frühjahr 2003 seit einem Jahr oder länger um einen Arbeitsplatz; 1991 zählte nur jeder dritte Erwerbslose (32%) zu diesem als langzeitarbeitslos eingestuftem Personenkreis (siehe Tabelle 6). Der Anteil der Erwerbslosen, die schon zwei Jahre und länger auf Arbeitsuche waren, erhöhte sich ebenfalls, und zwar von 18% im April 1991 auf 31% im Mai 2003. Fast jeder dritte Erwerbslose (30%) gab im Frühjahr 2003 an, seit weniger als sechs Monaten Ausschau nach einer Erwerbstätigkeit zu halten (1991: 44%), weniger als drei Monate nannten 15% der Erwerbslosen (1991: 23%).

Tabelle 6: Erwerbslose nach Dauer der Arbeitsuche<sup>1)</sup>  
Ergebnisse des Mikrozensus  
1 000

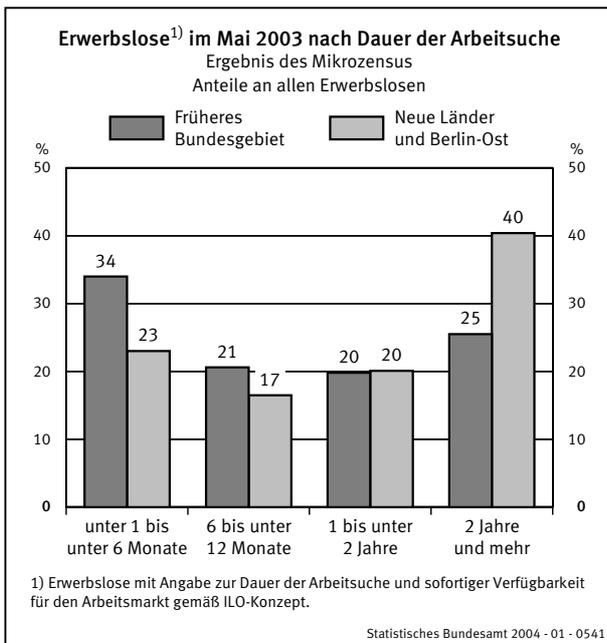
Gegenstand der Nachweisung	Mai 2003	April 1991
Insgesamt .....	4 022	1 931
dar.: mit Angabe zur Dauer der Arbeitsuche .....	3 652	1 883
unter 1 Monat .....	181	142
1 bis unter 3 Monate .....	371	300
3 bis unter 6 Monate .....	543	392
6 bis unter 12 Monate .....	697	454
1 bis unter 1 ½ Jahre .....	406	170
1 ½ bis unter 2 Jahre .....	321	91
2 Jahre und mehr .....	1 131	334

1) Erwerbslose sind gemäß ILO-Konzept für den Arbeitsmarkt sofort verfügbar.

In den neuen Ländern und Berlin-Ost war das Problem der Langzeitarbeitslosigkeit im Frühjahr 2003 deutlich größer als im früheren Bundesgebiet (siehe Schaubild 5). Der Anteil der seit einem Jahr oder länger Arbeit suchenden Erwerbslosen betrug im Westen – gegenüber April 1991 unverändert – 45%. Für den Osten lässt sich ein drastischer Anstieg des Anteils der Langzeitarbeitslosen von 10% im April 1991 auf 60% im Mai 2003 verzeichnen. Dabei ist allerdings zu berücksichtigen, dass Ende der 1980er-Jahre in der ehemaligen DDR zumindest offiziell keine Arbeitslosigkeit existierte. Besonders groß waren die Ost-West-Unterschiede hinsichtlich der Langzeitarbeitslosen, die schon seit zwei Jahren oder länger erfolglos nach einer Arbeit Ausschau hielten. Dieser Personenkreis, für den die Gefahr der Ver-

festigung der Erwerbslosigkeit besonders groß sein dürfte, stellte in den neuen Ländern und Berlin-Ost 40% (Männer: 33%, Frauen: 48%), im früheren Bundesgebiet nur 25% (Männer: 27%, Frauen: 23%) der Erwerbslosen. Weniger als sechs Monate befand sich jeder Dritte (34%) im Westen lebende Erwerbslose auf Arbeitsuche; eine so kurze Dauer nannte im Osten nur knapp jeder vierte Erwerbslose (23%).

Schaubild 5



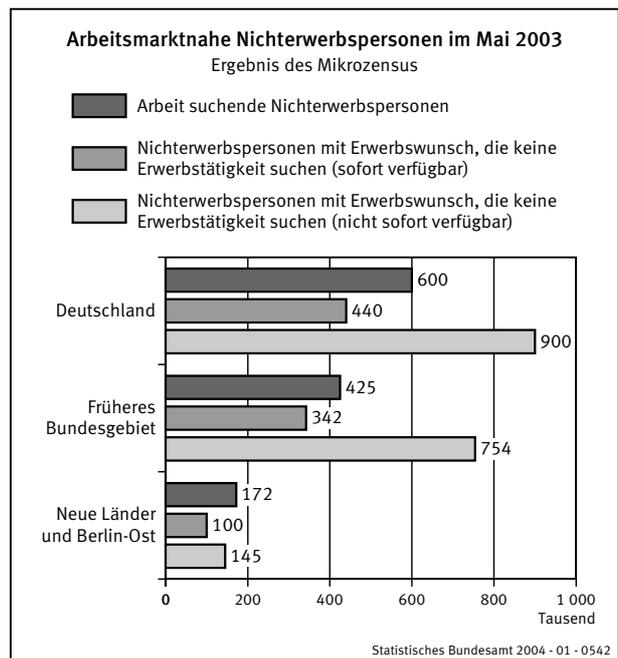
## Über 1 Mill. Nichterwerbspersonen suchen trotz Erwerbwunsch keinen Arbeitsplatz

Den verbleibenden Teil der Bevölkerung nach Abzug der Erwerbstätigen und Erwerbslosen bilden die Nichterwerbspersonen. Quantitativ gesehen sind die Nichterwerbspersonen jedoch längst keine Restgröße mehr, denn mit 42,3 Mill. Menschen stellen sie im Mai 2003 über die Hälfte (51%) der Gesamtbevölkerung in Deutschland. Mehr als ein Viertel (29%) davon waren Kinder unter 15 Jahren, die aufgrund ihres Alters definitionsgemäß – das heißt im begrifflichen Rahmen des dem Mikrozensus zu Grunde liegenden ILO-Konzepts – zu den Nichterwerbspersonen gehören.

Nichterwerbspersonen mit hinreichend großer Arbeitsmarktnähe bilden als verdecktes Erwerbpersonspotenzial eine zusätzliche Arbeitskräftereserve. Demzufolge dürfen sie aus der Betrachtung des Arbeitsmarktes nicht ausgeschlossen werden. Nach den Ergebnissen des Mikrozensus waren im Mai 2003 in Deutschland gut 1,9 Mill. Nichterwerbspersonen als arbeitsmarktnah zu bewerten. Darunter gab es knapp 600 000 Arbeit Suchende, die dem Arbeitsmarkt nicht sofort zur Verfügung standen und daher gemäß ILO-Konzept nicht der Gruppe der Erwerbslosen zuzuordnen sind. Die übrigen gut 1,3 Mill. Nichterwerbspersonen suchten zwar nicht nach einem Arbeitsplatz, wünschten sich aber trotzdem eine Erwerbsarbeit. Von diesen

erklärten rund 440 000 Personen, für die Aufnahme einer Erwerbstätigkeit sofort zur Verfügung zu stehen; die restlichen 900 000 stufen sich als kurzfristig nicht verfügbar ein (siehe Schaubild 6). Unter Einbeziehung der Erwerbslosen gaben somit in Deutschland im Frühjahr 2003 insgesamt knapp 6,0 Mill. nichterwerbstätige Personen an, nach einer Erwerbsarbeit zu suchen oder zumindest den Wunsch nach einer Erwerbsarbeit zu haben. Dieses Ergebnis verdeutlicht, dass die bestehende Beschäftigungslücke auf dem deutschen Arbeitsmarkt deutlich über die Zahl der Erwerbslosen hinausgeht.

Schaubild 6



Das frühere Bundesgebiet verzeichnete im Mai 2003 gut 1,5 Mill. arbeitsmarktnahe Nichterwerbspersonen, von denen 425 000 auf Arbeitsuche waren und knapp 1,1 Mill. eine Erwerbstätigkeit wünschten, ohne sich auf aktive Arbeitssuche begeben zu haben. In den neuen Ländern und Berlin-Ost lebten rund 170 000 Arbeit suchende Nichterwerbspersonen; die Zahl der nicht nach einer Erwerbsarbeit suchenden Nichterwerbspersonen mit Erwerbwunsch lag bei 245 000.

Charakteristisch für die über den Erwerbwunsch ermittelte Arbeitskräftereserve ist, dass die Frauen im Mai 2003 mit einem Anteil von fast zwei Dritteln (64%) deutlich überrepräsentiert waren. Von ihnen gaben 45% an, aufgrund von persönlichen oder familiären Verpflichtungen keine Suchaktivitäten zu entfalten, obwohl der Wunsch nach einer Erwerbstätigkeit besteht. In dieser Situation befanden sich lediglich gut 2% der männlichen Nichterwerbspersonen mit Erwerbwunsch.

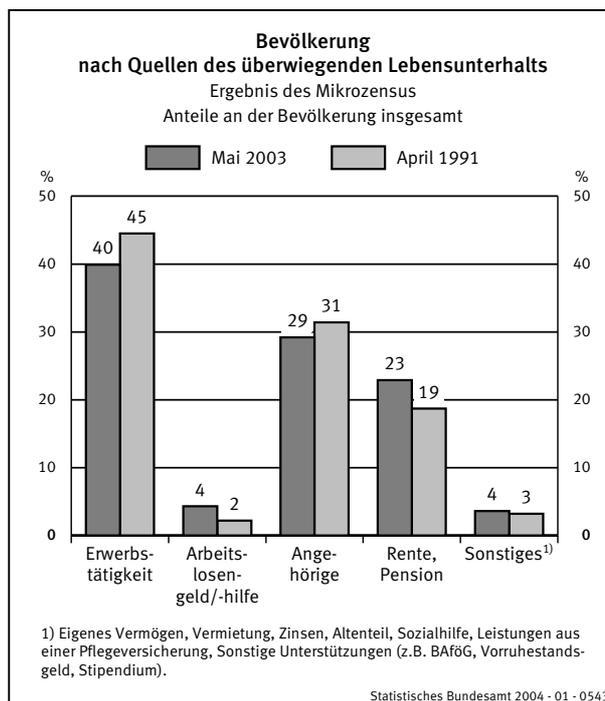
Mehr als jede vierte Nichterwerbsperson (27%) mit Wunsch nach einer Erwerbstätigkeit ergriff im Mai 2003 wegen einer schulischen oder beruflichen Ausbildung keine Maßnahmen, um einen Arbeitsplatz zu finden (siehe Tabelle 7). Dieser Grund spielte bei den Männern mit einem Anteil von

Tabelle 7: Nichterwerbspersonen mit Erwerbwunsch, die keine Erwerbstätigkeit suchen, nach Gründen der Arbeitsuche  
Ergebnis des Mikrozensus 2003  
Prozent

Gegenstand der Nachweisung	Insgesamt	Männer	Frauen
Ausbildung .....	27,1	37,8	21,3
Ruhestand .....	11,2	17,4	7,7
Krankheit oder (vorübergehende) Arbeitsunfähigkeit .....	9,0	12,3	7,2
Keine Beschäftigungsmöglichkeiten .....	4,4	4,5	4,3
Persönliche oder familiäre Verpflichtungen .....	29,7	2,4	44,8
Frühinvalidität oder sonstige Behinderung .....	7,5	12,5	4,7
Sonstiges .....	11,1	13,1	10,0

38% die größte Rolle (Frauen: 21%). Gut jede zehnte Nichterwerbsperson (11%), die gerne einer Erwerbsarbeit nachgehen würde, unternahm wegen ihres Ruhestandes keine Schritte der Arbeitsuche (Männer: 17%, Frauen: 8%). Weitere 4% begaben sich aufgrund fehlender Beschäftigungsmöglichkeiten nicht auf die Suche nach einer bezahlten Tätigkeit. Krankheit oder (vorübergehende) Arbeitsunfähigkeit nannten 9% der erwerbsorientierten Nichterwerbspersonen als Grund ihrer Nichtarbeitsuche.

Schaubild 7



## Staatliche Transfereinkommen gewinnen für den Lebensunterhalt an Bedeutung

Der Mikrozensus ist nicht nur in der Lage, alle im oder außerhalb des Erwerbslebens stehenden Personengruppen (Erwerbstätige, Erwerbslose, arbeitsmarktnahe und

sonstige Nichterwerbspersonen) umfassend abzubilden. Durch die jährlich gestellte Frage nach dem überwiegenden Lebensunterhalt ermöglicht er auch eine Kategorisierung der Bevölkerung nach der wichtigsten Quelle des Lebensunterhalts (Unterhaltskonzept). Durch die Kombination dieser Variablen mit demographischen und haushalts- bzw. familienbezogenen Merkmalen sowie in Verbindung mit den Angaben zur Höhe des Einkommens und zur Erwerbs-

Tabelle 8: Bevölkerung nach Beteiligung am Erwerbsleben und überwiegendem Lebensunterhalt  
Ergebnisse des Mikrozensus

Personengruppe	Insgesamt		Überwiegender Lebensunterhalt durch									
			Erwerbstätigkeit		Arbeitslosengeld/-hilfe		Angehörige		Renten und Pensionen		Sonstiges <sup>3)</sup>	
	Mai 2003	April 1991	Mai 2003	April 1991	Mai 2003	April 1991	Mai 2003	April 1991	Mai 2003	April 1991	Mai 2003	April 1991
	1 000		%									
Bevölkerung insgesamt .....	82 502	79 829	39,9	44,5	4,3	2,2	29,2	31,4	22,9	18,7	3,6	3,2
Männlich .....	40 330	38 548	47,4	55,5	5,4	2,3	22,4	23,7	21,4	15,4	3,4	3,1
Weiblich .....	42 172	41 281	32,7	34,3	3,3	2,0	35,8	38,6	24,4	21,7	3,9	3,3
dar.: verheiratet .....	19 784	19 817	39,5	42,8	3,7	2,7	33,9	38,6	20,5	13,6	2,4	2,3
Erwerbstätige zusammen ...	36 172	37 445	90,9	94,9	0,6	0,1	5,9	3,7	1,5	0,8	1,0	0,6
Männer .....	19 996	21 875	95,5	97,7	0,5	0,0	1,7	1,2	1,5	0,6	0,7	0,4
Frauen .....	16 176	15 570	85,1	90,8	0,7	0,1	11,1	7,1	1,6	0,9	1,4	1,1
dar.: verheiratet .....	9 519	9 498	82,1	89,3	0,6	0,1	15,5	9,3	1,0	0,4	0,8	0,9
Erwerbslose <sup>1)</sup> zusammen ...	4 022	1 931	X	X	73,9	67,8	15,5	20,3	1,0	1,8	9,5	10,1
Männer .....	2 316	930	X	X	79,3	71,9	10,3	13,5	0,7	1,9	9,6	12,7
Frauen .....	1 707	1 001	X	X	66,7	64,1	22,6	26,5	1,4	1,7	9,4	7,7
dar.: verheiratet .....	927	655	X	X	62,8	62,6	30,6	32,3	/	/	6,3	4,4
Nichterwerbspersonen <sup>2)</sup>												
zusammen .....	42 307	40 453	X	X	0,9	0,9	50,5	57,7	43,3	36,1	5,3	5,3
Männer .....	18 018	15 743	X	X	1,3	1,2	46,9	55,6	46,2	36,8	5,6	6,4
Frauen .....	24 289	24 710	X	X	0,6	0,8	53,1	59,0	41,1	35,6	5,1	4,6
dar.: verheiratet .....	9 339	9 663	X	X	0,9	1,3	53,0	67,8	42,4	27,4	3,6	3,5

1) Erwerbslose sind gemäß ILO-Konzept für den Arbeitsmarkt sofort verfügbar. – 2) Einschl. für den Arbeitsmarkt nicht sofort verfügbarer, Arbeit suchender Nichterwerbstätiger. – 3) Eigenes Vermögen, Zinsen, Altenteil, Sozialhilfe oder sonstige Unterstützung; 2003 zusätzlich Leistungen aus einer Pflegeversicherung.

beteiligung erlaubt es der Mikrozensus, ein differenziertes Bild über die wirtschaftliche Situation der Bevölkerung zu zeichnen. Die nachfolgenden Ausführungen beschränken sich auf eine Betrachtung der Bevölkerung nach ihrer Unterhaltsstruktur.

Im Mai 2003 gaben in Deutschland rund 40% der Menschen – knapp 33% der Frauen und gut 47% der Männer – das eigene Erwerbseinkommen als wichtigste Unterhaltsquelle an. 1991 hatte der Anteil der vornehmlich von Erwerbseinkommen lebenden Personen in Deutschland noch nahezu 45% betragen (siehe Schaubild 7). Zu dieser Entwicklung hat zum einen der sinkende Anteil der Erwerbstätigen an der Gesamtbevölkerung (von 47% im April 1991 auf 44% im Mai 2003), zum anderen die rückläufige Bedeutung der Erwerbstätigkeit für den überwiegenden Lebensunterhalt der Erwerbstätigen beigetragen. Während im April 1991 noch 95% der Erwerbstätigen das eigene Erwerbseinkommen für den größten Teil der Lebenshaltungskosten einsetzten, waren es im Mai 2003 nur 91% (siehe Tabelle 8). Darin dürfte sich in erster Linie die Zunahme der Teilzeitarbeit widerspiegeln, die in vielen Fällen lediglich der Aufbesserung des Familieneinkommens dient.

Bedingt durch die demographische Alterung der Gesellschaft und die steigende Arbeitslosigkeit stellen Transferleistungen des Staates für einen steigenden Bevölkerungsanteil die wichtigste Einkunftsquelle dar. So gaben im Frühjahr 2003 rund 23% aller Bundesbürgerinnen und Bundesbürger Einkommen aus Renten bzw. Pensionen als wichtigste Unterhaltsquelle an; 12 Jahre zuvor war dieser Anteil um gut vier Prozentpunkte niedriger ausgefallen. In den neuen Ländern und Berlin-Ost lebten im Mai 2003 sogar rund 28% der Menschen überwiegend von Altersruhegeldern, im früheren Bundesgebiet waren es 22% (siehe Tabelle 9). Zwei Einflussfaktoren für diesen deutlichen West-Ost-Unterschied wurden bereits genannt: der relativ starke Rückgang der Bevölkerung im erwerbsfähigen Alter und die vergleichsweise niedrigen Erwerbsquoten älterer Menschen im Osten.

Leistungen durch Angehörige stellten im Westen für etwa 32% der Bevölkerung die überwiegende Unterhaltsquelle dar; im Osten lag dieser Anteil nur bei gut 18%. Bundesweit lebten im Mai 2003 29% der Bevölkerung hauptsächlich vom Unterhalt durch Angehörige. Auffallend sind dabei die deutlichen Unterschiede zwischen Männern und Frauen: Für fast 36% der weiblichen, jedoch für nur gut ein Fünftel (22%) der männlichen Bevölkerung stellten Angehörige den wichtigsten Unterhaltsträger dar. [\[U\]](#)

Tabelle 9: Bevölkerung nach der Quelle des überwiegenden Lebensunterhalts  
Ergebnisse des Mikrozensus  
Prozent

Gegenstand der Nachweisung	Früheres Bundesgebiet		Neue Länder und Berlin-Ost	
	Mai 2003	April 1991	Mai 2003	April 1991
Erwerbstätigkeit .....	39,8	43,7	40,0	47,6
Rente, Pension .....	21,8	18,7	27,9	18,5
Arbeitslosengeld/-hilfe	3,2	1,5	9,5	4,9
Unterstützung durch Angehörige .....	31,6	33,2	18,4	24,4
Sonstiges .....	3,6	2,9	4,2	4,6

Für gut 4% der Bevölkerung Deutschlands war im Mai 2003 das Arbeitslosengeld oder die Arbeitslosenhilfe die vorherrschende Einnahmequelle (1991: rund 2%). Dabei lag dieser Anteil in den – von Arbeitslosigkeit besonders stark betroffenen – neuen Ländern und Berlin-Ost (knapp 10%) wesentlich höher als im früheren Bundesgebiet (gut 3%).

Dipl.-Volkswirt Uwe Reim

# Unternehmen der Binnenschifffahrt 2002

Mitte 2002 waren insgesamt 1232 in Deutschland ansässige Unternehmen in der Binnenschifffahrt tätig, das waren 77 weniger als im Vorjahr (-5,9%). Vor allem in der Trockengüterschifffahrt gab es einen Rückgang um 43 auf 708 Unternehmen (-5,7%). Aber auch Personenschifffahrt übten 27 Unternehmen weniger als im Vorjahr aus (-7,7%).

Die Unternehmen verfügten über 2868 Binnenschiffe zur Güter- und Personenbeförderung (-3,8%), davon 1868 Güterschiffe, 236 Schub- und Schleppboote und 764 Personenschiffe. Der Rückgang der Zahl der verfügbaren Binnenschiffe ist hauptsächlich auf Abnahmen der Güterschiffe um 55 Einheiten (-2,9%) und der Personenschiffe um 54 Einheiten (-6,6%) zurückzuführen.

In der Binnenschifffahrt waren Mitte 2002 in deutschen Unternehmen 7689 Personen beschäftigt, 60 weniger als im Vorjahr (-0,8%). Zum fahrenden Personal zählten 6162 Personen (-2,4%), zum Landpersonal 1527 Personen (+6,3%). 2002 erzielten die deutschen Unternehmen einen Beförderungsumsatz aus Binnenschifffahrt von 724 Mill. Euro (-7,6%) und einen Umsatz aus Befrachtung von 511 Mill. Euro (+8,4%).

## Vorbemerkung

Die Binnenschifffahrt nimmt in Deutschland nach Straße, Schiene und Seeschifffahrt im Güterverkehr den vierten Platz ein. Im Jahr 2002 wurden 231,7 Mill. Tonnen (t) Güter mit Binnenschiffen befördert. Dies waren 4,4 Mill. t oder 1,8% weniger als im Vorjahr. Die Beförderungsleistung nahm im gleichen Zeitraum nur um 1,0% auf 64,2 Mrd. Tonnenkilometer (tkm) ab, was darauf zurückzuführen ist,

dass sich der durchschnittliche Transportweg etwas verlängert hat. Insgesamt setzte sich mit dieser Entwicklung die Tendenz des Vorjahres fort, das auch schon durch konjunkturelle Schwächen geprägt war. Im Vergleich zur Eisenbahn, dem wichtigsten Konkurrenten der Binnenschifffahrt, wurde dabei etwas an Boden verloren. Dort nahm im Jahr 2002 die Beförderungsmenge um 1,0% und damit weniger als in der Binnenschifffahrt ab, die Beförderungsleistung der Eisenbahn sank dagegen mit 1,2% etwas stärker als die der Binnenschifffahrt. Die Beförderungsmenge der Eisenbahn lag damit um 23,2% über der der Binnenschifffahrt, im Jahr 2001 hatte dieser Abstand noch 22,2% betragen. Bei der Beförderungsleistung blieb der Abstand dagegen fast gleich. Hier erbrachte die Eisenbahn im Jahr 2002 17,5% mehr Tonnenkilometer als die Binnenschifffahrt (2001: 17,8%).

Inländische Binnenschifffahrtsunternehmen stehen nicht nur im Wettbewerb mit Unternehmen anderer Verkehrszweige, sie konkurrieren auch mit ausländischen Binnenschifffahrtsunternehmen um Transportaufträge: Im Jahr 2002 beförderten die unter deutscher Flagge fahrenden Schiffe mit 84,9 Mill. t 3,8% weniger Güter als 2001, ihr Anteil an der Güterbeförderung reduzierte sich dadurch von 37,3% auf 36,6%. Die Beförderungsleistung der deutschen Flagge nahm um 2,5% auf 22,0 Mrd. tkm ab, ihr Anteil an der gesamten Beförderungsleistung sank daher von 34,8% auf 34,3%. Damit hat sich der seit Jahren beobachtete Trend der abnehmenden Beteiligung deutscher Schiffe am Gütertransport weiter fortgesetzt. Auch daher ist die wirtschaftliche Lage der deutschen Binnenschifffahrtsunternehmen gesondert von der Entwicklung der insgesamt innerhalb Deutschlands auf Binnenwasserstraßen beförderten Gütermenge zu betrachten.

Neben der Güterschifffahrt betreiben deutsche Binnenschifffahrtsunternehmen in nennenswertem Umfang Personenbeförderung mit Fahrgast- und Fahrgastkabinenschiffen. Eine allein auf die Güterbinnenschifffahrt gerichtete Darstellung würde daher nur ein unvollständiges Abbild dieses Verkehrsträgers liefern. In dieser Zeitschrift wird daher regelmäßig über den Stand und die Entwicklung aller inländischen Unternehmen der Fahrgast- und Güterbinnenschifffahrt informiert. Der Beitrag gliedert sich dabei in einen einleitenden Abschnitt zur Methodik der Statistik der Unternehmen der Binnenschifffahrt, gefolgt von der Darstellung ihrer wichtigsten quantitativen Ergebnisse insbesondere hinsichtlich der Unternehmens- und Schiffszahl sowie der Beschäftigten- und Umsatzentwicklung. Der Beitrag schließt mit einem Ausblick auf das Jahr 2003.

### Methodik

Zum Berichtskreis der Statistik der Unternehmen der Binnenschifffahrt zählen Unternehmen mit Sitz im Inland, die Personen- und Güterbeförderung mit Binnenschiffen<sup>1)</sup> gewerblich (als Haupt- oder Nebentätigkeit) oder im Werkverkehr (als Hilfstätigkeit) durchführen. Binnenschifffahrt als Haupttätigkeit liegt vor, wenn Unternehmen ausschließlich oder überwiegend Personen- oder Güterbeförderungen mit Binnenschiffen betreiben, das heißt wenn ihr wirtschaftlicher Schwerpunkt in der Binnenschifffahrt liegt. Liegt der Schwerpunkt der wirtschaftlichen Tätigkeit außerhalb des Bereichs Binnenschifffahrt, also zum Beispiel im Bereich Handel, wird die Binnenschifffahrt aber gewerblich betrieben, so handelt es sich um eine Nebentätigkeit. Werkverkehr bedeutet, dass Unternehmen ihre Binnenschiffe nur für eigene betriebliche Zwecke einsetzen und deren Transportkapazität somit nicht Dritten anbieten.

Einbezogen sind somit alle auf diesem Markt tätigen inländischen Unternehmen. Diese Unternehmen werden ausschließlich über ihren fachlichen Unternehmensteil „Binnenschifffahrt“ zu ihren Schiffen, den in der Binnenschifffahrt Beschäftigten und den Umsätzen aus Binnenschifffahrt befragt. Diese funktionale Sicht stellt somit die Transporte erbringenden Einheiten von Unternehmen in den Mittelpunkt der Betrachtung. Dabei ist es von untergeordneter Bedeutung, ob der Transport von einem Unternehmen erbracht wird, das ausschließlich oder mit wirtschaftlichem Schwerpunkt oder nur als Nebenerwerb Verkehrsleistungen anbietet bzw. für eigene Zwecke im Werkverkehr erstellt.

In diesem Aufsatz erfolgt daher die statistische Abbildung der Unternehmen der Binnenschifffahrt in funktionaler Abgrenzung. Waren Unternehmen außer in der Binnenschifffahrt auch in anderen Wirtschaftsbereichen tätig, gehen bei dieser Art der Betrachtung Umsätze und Beschäf-

tigte in den binnenschifffahrtsfremden Tätigkeiten nicht in die statistische Auswertung ein. Die Anzahl der Unternehmen, Binnenschiffe und Beschäftigten wird zum Stichtag 30. Juni eines Berichtsjahres erhoben, der Umsatz ist für das Berichtsjahr anzugeben.

Bis zur Verabschiedung des Verkehrsstatistikgesetzes im Jahr 1999 wurden zudem mit der Statistik der Unternehmen der Binnenschifffahrt einige Angaben in institutioneller Abgrenzung erhoben. Mit der Einführung einer Dienstleistungsstatistik war dies nicht mehr erforderlich, da durch die Dienstleistungsstatistik vergleichbare institutionelle Strukturdaten für alle Dienstleistungsbereiche bereitgestellt werden. Um indes Doppelbefragungen zu vermeiden, wurden in Abstimmung mit dem Bundesministerium für Verkehr, Bau- und Wohnungswesen die bisher erfragten institutionellen Merkmale gestrichen und die Statistik der Unternehmen der Binnenschifffahrt rein funktional ausgerichtet. Durch das Verkehrsstatistikgesetz entfielen somit 1999 in dieser Erhebung die bis dahin erfragten institutionellen Merkmale „Gesamtzahl der Beschäftigten“ und „Gesamtumsatz“ der Unternehmen mit Schwerpunkt in der Binnenschifffahrt.

Mit dem Verkehrsstatistikgesetz waren weitere Änderungen verbunden. Zum einen werden seit dem Berichtsjahr 1999 sämtliche Merkmale für die gleiche Berichtsperiode erfragt.<sup>2)</sup> Zum anderen wurde eine Regelung in das Verkehrsstatistikgesetz aufgenommen, mit der verschiedene Institutionen verpflichtet wurden, dem Statistischen Bundesamt Namen und Anschrift der Binnenschifffahrt betreibenden Unternehmen zu übermitteln. Obwohl damit keine inhaltliche Änderung des Berichtskreises verbunden war, zeigte sich bei der erstmaligen Anforderung der Anschriften, dass bis 1998 gewisse Untererfassungen vorgekommen waren, da insbesondere kleinere Unternehmen und Neugründungen dem Statistischen Bundesamt nicht bekannt gewesen und somit auch nicht in die Erhebungen einbezogen worden waren. Durch das Verkehrsstatistikgesetz und dessen Umsetzung konnte somit der Berichtskreis vollständig erfasst werden. Aus den beiden genannten Gründen ist allerdings die Vergleichbarkeit von Ergebnissen des aktuellen Berichtssystems mit solchen der Berichtsjahre vor 1999 nur eingeschränkt möglich.

### Gesamtüberblick

Einen zusammenfassenden Überblick über die Entwicklung der deutschen Binnenschifffahrtsunternehmen seit 1999 gibt Tabelle 1. Am Erhebungsstichtag, dem 30. Juni 2002, waren 1 232 Unternehmen in der Binnenschifffahrt tätig. Sie verfügten über 2 868 Binnenschiffe zur Güter- und Personenbeförderung, beschäftigten 7 689 Personen in ihren Unternehmensteilen „Binnenschifffahrt“ und erzielten einen Umsatz aus Binnenschifffahrt von 1 257 Mill. Euro. Im

1) Von der Befragung ausgenommen sind Unternehmen mit wirtschaftlicher Tätigkeit im Fluss-, See- und Kanalfährverkehr oder in der Hafenschifffahrt, die daneben keine weiteren Binnenschifffahrtstätigkeiten ausüben. Nicht zum Berichtskreis dieser Statistik gehören Unternehmen der Küstenschifffahrt, Fluss- und Seefischerei sowie Unternehmen, die Bunkerboote, Bilgenorgansschiffe und Proviantschiffe betreiben, die Schiffe als Lager-, Messe- und Ausstellungsschiffe nutzen oder die festliegende Schiffe ausschließlich als Restaurants, Hotels, Supermärkte, Büros und für ähnliche Zwecke einsetzen.

2) Davor erfolgte die Erhebung von Unternehmenszahl, Schiffsbestand und Beschäftigten zu einem Stichtag im jeweils laufenden Kalender- bzw. Geschäftsjahr, die Erhebung des Umsatzes jedoch für das Vorjahr. Damit war die Verknüpfbarkeit der auf unterschiedliche Zeitpunkte bzw. Zeiträume bezogenen Merkmale und damit auch die Aussagefähigkeit von Kennziffern wie Umsatz je Unternehmen oder je Beschäftigten stark eingeschränkt.

Tabelle 1: Strukturdaten der Binnenschifffahrt

Jahr	Unternehmen	Verfügbare Binnenschiffe	Beschäftigte	Umsatz
	Anzahl			Mill. EUR
1999	1400	3513	8380	1036,2
2000	1370	3375	8057	1231,1
2001	1309	2980	7749 r	1274,6
2002	1232	2868	7689	1256,6
Veränderung gegenüber dem Vorjahr in %				
2000	-2,1	-3,9	-3,9	+18,8
2001	-4,5	-11,7	-3,8	+3,5
2002	-5,9	-3,8	-0,8	-1,4

Vergleich der Jahre untereinander ist dabei die Zahl der Unternehmen, der Schiffe und der Beschäftigten entsprechend dem langfristigen Trend weiter gesunken. Aber auch der Umsatz, der in der Vergangenheit häufig eine positive Tendenz aufwies, nahm im Jahr 2002 gegenüber dem Vorjahr ab.

## Zahl der Unternehmen weiter abnehmend

Von den deutschen Binnenschifffahrtsunternehmen betrieben – wie im Vorjahr – 97,9% Binnenschifffahrt als Gewerbe und die übrigen 2,1% ausschließlich als Werkverkehr (siehe Tabelle 2). Die Gesamtzahl der Unternehmen sank im Vergleich zum Vorjahr per saldo um 77 Unternehmen (-5,9%). Dabei standen 67 Neuzugängen 144 Einstellungen der Binnenschifffahrtstätigkeit gegenüber. Der weitaus größte Teil der Unternehmen (83,0%) betreibt ausschließlich Binnenschifffahrt, 11,1% der Unternehmen üben neben der Binnenschifffahrt als Hauptaktivität noch weitere Tätigkeiten aus, für 3,8% der Unternehmen stellt die Binnenschifffahrt eine gewerbliche Nebentätigkeit dar. Im Jahr 2002 hat die Zahl der ausschließlich Binnenschifffahrt betreibenden Unternehmen um 57 (-5,3%) abgenommen. Diese Entwicklung entspricht der bisherigen Tendenz, dass vor allem Unternehmen mit ausschließlicher Binnenschifffahrtstätigkeit aus dem Markt ausscheiden. Entgegen der Entwicklung in den Vorjahren hat sich im Jahr 2002 aber auch die Zahl der

Tabelle 2: Unternehmen, verfügbare Schiffe, Beschäftigte und Umsatz der Binnenschifffahrt 2002

Gegenstand der Nachweisung	Einheit Veränderung gegenüber dem Vorjahr	Unternehmen					
		insgesamt	mit gewerblicher Binnenschifffahrt				mit Binnenschifffahrt nur im Werkverkehr
			zusammen	davon			
				ausschließlich in der Binnenschifffahrt tätig	überwiegend in der Binnenschifffahrt tätig	mit neben-gewerblicher Binnenschifffahrtstätigkeit	
Unternehmen am 30. Juni 2002 .....	Anzahl	1232	1206	1022	137	47	26
	%	-5,9	-5,9	-5,3	-6,8	-14,5	-7,1
Verfügbare Schiffe am 30. Juni 2002 .....	Anzahl	2868	2754	2153	513	88	114
	%	-3,8	-3,6	-3,0	-3,0	-18,5	-7,3
Güterschiffe .....	Anzahl	1868	1787	1542	215	30	81
	%	-2,9	-2,6	-2,6	-4,0	+7,1	-8,0
Schubboote, Schub-Schleppboote und Schleppboote .....	Anzahl	236	204	194	8	2	32
	%	-1,3	-0,5	+14,8	-72,4	-71,4	-5,9
Fahrgast- und Fahrgastkabinenschiffe .....	Anzahl	764	763	417	290	56	1
	%	-6,6	-6,6	-10,9	+5,1	-23,3	-
Beschäftigte am 30. Juni 2002 .....	Anzahl	7689	7527	5088	2174	265	162
	%	-0,8	-0,6	-5,1	+17,1	-24,5	-9,5
Fahrendes Personal .....	Anzahl	6162	6034	4307	1502	225	128
	%	-2,4	-2,2	-7,1	+18,8	-15,7	-11,7
dar.: Schiffseigner und unbezahlt mithelfende Familienangehörige .....	Anzahl	1216	1214	1074	119	21	2
	%	-7,9	-7,8	-5,7	-20,1	-25,0	-50,0
Landpersonal .....	Anzahl	1527	1493	781	672	40	34
	%	+6,3	+6,4	+7,4	+13,5	-52,4	-
Umsatz <sup>1)</sup> aus Binnenschifffahrtstätigkeit 2002 .....	Mill. EUR	X	1256,6	660,8	506,3	89,5	X
	%	X	-1,4	-1,4	+3,8	-23,3	X
aus Schub- und Schleppleistungen .....	Mill. EUR	X	21,7	19,3	2,1	0,2	X
	%	X	+9,7	+22,1	-11,9	-87,1	X
aus Beförderung in der Güterschifffahrt ..	Mill. EUR	X	528,3	388,4	130,7	9,1	X
	%	X	-9,4	-8,0	-22,7	-59,3	X
in der Trockengüterschifffahrt .....	Mill. EUR	X	287,2	213,3	69,4	4,4	X
	%	X	-4,6	-3,7	+17,0	-77,9	X
in der Tankschifffahrt .....	Mill. EUR	X	241,1	175,1	61,3	4,7	X
	%	X	-14,5	+3,0	-44,2	+96,3	X
aus Beförderung in der Personenschifffahrt .....	Mill. EUR	X	196,1	99,8	89,9	6,4	X
	%	X	-2,3	-15,5	+25,2	-40,1	X
aus Befrachtung .....	Mill. EUR	X	510,6	153,3	283,6	73,7	X
	%	X	+8,4	+5,9	+16,0	-10,0	X
dar.: an Unterfrachtführer weitergeleitet ..	Mill. EUR	X	441,7	140,7	267,5	33,6	X
	%	X	+14,5	+19,4	+19,5	-24,0	X

1) Ohne Umsatzsteuer.

überwiegend Binnenschifffahrt betreibenden Unternehmen um zehn Unternehmen (-6,8%) verringert. Die insgesamt schwierige Situation der deutschen Binnenschifffahrtsunternehmen führte somit dazu, dass auch Unternehmen, die neben der Binnenschifffahrt ein zweites Standbein hatten, ihre Binnenschifffahrtstätigkeit einstellten. Es kam hier auch nicht zu Schwerpunktverlagerungen auf die anderen Geschäftsfelder, denn auch die Zahl der Unternehmen mit nebegewerblicher Binnenschifffahrtstätigkeit sank um acht auf 47 Unternehmen. Weiterhin an Bedeutung verliert der Werkverkehr: Mit nur noch 26 Unternehmen waren hier zwei Unternehmen weniger als im Vorjahr aktiv.

Von den 1232 Unternehmen betrieben 708 Trockengüterschifffahrt, 170 Tankschifffahrt, 55 Schub- und Schleppschifffahrt für andere Unternehmen sowie 325 Personenschifffahrt. Vor allem in der Trockengüterschifffahrt war ein hoher absoluter Rückgang um 43 Unternehmen (-5,7%) zu verzeichnen. Dagegen konnte sich die Tankschifffahrt behaupten: Lediglich ein Unternehmen stellte diese Tätigkeit ein (-0,6%). Die Schub- und Schleppschifffahrt gaben fünf Unternehmen auf, was eine Abnahme um 8,3% bedeutet. Ein relativ hoher prozentualer Rückgang war ebenfalls in der Personenschifffahrt zu registrieren, zurückzuführen auch auf die allgemeine Konsumzurückhaltung der privaten Haushalte: Im Jahr 2002 betrieben 27 Unternehmen weniger Personenschifffahrt als im Vorjahr, dies entspricht einer Abnahme um 7,7%. Bei dieser Gliederung nach Schifffahrtssparten werden alle unternehmerischen Tätigkeiten berücksichtigt. Unternehmen, die mehr als eine dieser Schifffahrtssparten betreiben, werden somit auch mehrfach gezählt, sodass die Summe der Einzelpositionen höher ist als die Gesamtzahl der Unternehmen. Unterscheidet man hier explizit, so führten als Binnenschifffahrtstätigkeit 687 Unternehmen ausschließlich Trockengüterschifffahrt, 163 ausschließlich Tankschifffahrt, 37 ausschließlich Schub- und Schleppschifffahrt für andere Unternehmen und 323 Unternehmen ausschließlich Personenschifffahrt durch; die übrigen 22 der 1232 Unternehmen waren in mehreren Schifffahrtssparten aktiv.

### Schiffsbestand wiederum rückläufig

Der Schiffsbestand der Unternehmen umfasst die Schiffe für den Personen- und Gütertransport auf Binnengewässern. Unberücksichtigt bleiben Fähren, Schuten, Hafenschlepper und Schlepp-Barkassen sowie sonstige nicht primär für den Personen- und Gütertransport auf Binnengewässern konstruierte Schiffe. Nachgewiesen werden die verfügbaren Binnenschiffe, das heißt die eigenen (ohne vermietete) sowie die gemieteten oder geleasteten Schiffe. Eingeschlossen sind dabei auch stillgelegte Binnenschiffe, die jederzeit wieder eingesetzt werden könnten. Der hier nachgewiesene verfügbare Schiffsbestand bildet somit die Produktionsgrundlage der deutschen Binnenschifffahrt<sup>3)</sup>.

Tabelle 3: Verfügbare Binnenschiffe am 30. Juni 2002

Schiffsgattung	Bestand inländischer Unternehmen	Veränderung gegenüber dem 30. Juni 2001
	Anzahl	%
Trockengutmotorschiffe .....	939	-3,7
Tankmotorschiffe .....	298	+16,0
Trockengutschubleichter .....	606	-7,8
Tanschubleichter .....	23	-17,9
Trockengutschleppkähne .....	2	-66,7
Schubboote .....	180	+7,1
Schleppboote .....	26	-36,6
Schub-Schleppboote .....	30	-
Fahrgastkabinenschiffe .....	15	-34,8
Fahrgastschiffe .....	749	-5,8
<b>Insgesamt ...</b>	<b>2868</b>	<b>-3,8</b>
nach der Motorleistung		
	1 000 kW	%
Trockengutmotorschiffe .....	543	-1,3
Tankmotorschiffe .....	218	+15,0
Schubboote .....	107	+31,6
Schleppboote .....	9	-7,3
Schub-Schleppboote .....	11	-26,7
<b>Insgesamt ...</b>	<b>888</b>	<b>+5,1</b>
nach der Ladekapazität		
	1 000 t	%
Trockengutmotorschiffe .....	1 230	-1,3
Tankmotorschiffe .....	448	+19,3
Trockengutschubleichter .....	573	+3,4
Tanschubleichter .....	31	-21,5
Trockengutschleppkähne .....	1	-66,7
<b>Insgesamt ...</b>	<b>2 282</b>	<b>+3,0</b>
nach der Platzkapazität		
	Personenplätze	%
Fahrgastkabinenschiffe .....	1 585	-36,0
Fahrgastschiffe .....	183 072	-2,5
<b>Insgesamt ...</b>	<b>184 657</b>	<b>-2,9</b>

1 868 Güterschiffe, 764 Fahrgastschiffe sowie 236 Schub-, Schub-Schlepp- und Schleppboote bildeten den verfügbaren Schiffsbestand (siehe Tabellen 2 und 3) der Unternehmen von insgesamt 2 868 Schiffen. Der Rückgang gegenüber dem Vorjahr um insgesamt 112 Schiffe resultiert vor allem aus einer Abnahme der Güterschiffe um 55 Einheiten (-2,9%) und der Personenschiffe um 54 Einheiten (-6,6%). Die Schub-, Schub-Schlepp- und Schleppboote verzeichneten dagegen nur leichte Rückgänge: 236 verfügbare Einheiten bedeuten eine Abnahme um 1,3% bzw. drei Einheiten gegenüber dem Vorjahr.

Innerhalb des Güterschiffsbereichs verlief die Entwicklung allerdings sehr unterschiedlich. Verantwortlich für den genannten Rückgang war insbesondere die Abnahme der Zahl der Trockengutmotorschiffe um 36 auf 939 Einheiten (-3,7%) und die der Trockengutschubleichter um 51 auf 606 Einheiten (-7,8%); dagegen erhöhte sich die Zahl der Tankmotorschiffe um 41 auf 298 Einheiten (+16,0%). In detaillierter Untergliederung nach Ladekapazität und Motorleistung ändert sich das Bild nicht wesentlich. Rückgängen bei den Trockengutmotorschiffen standen Zunah-

3) Die Ergebnisse der Unternehmensstatistik zu den verfügbaren Schiffen weichen aus methodischen Gründen zum Teil vom Schiffsbestand in der Binnenschiffsbestandsdatei ab, die von der Wasser- und Schifffahrtsdirektion Südwest auf Basis inländischer Binnenschiffsregister zusammengestellt wird. Die Binnenschiffsbestandsdatei basiert auf dem Eigentümerkonzept, Erfassungskriterium für die Statistik der Binnenschifffahrtsunternehmen ist dagegen das Betreiberkonzept. Zur ausführlichen Erläuterung der Unterschiede beider Konzepte siehe Stede, H.-J.: „Unternehmen der Binnenschifffahrt 1995“ in WiSta 6/1996, S. 366 ff., hier S. 367 f.

men bei den Tankmotorschiffen gegenüber. Insgesamt nahm die Ladekapazität aller Güterschiffe um 3,0% zu, da insbesondere kleinere Einheiten aus dem Verkehr genommen wurden (Trockengutmotorschiffe) bzw. größere Einheiten zusätzlich in Betrieb genommen wurden (Tankmotorschiffe). So stieg die durchschnittliche Ladekapazität der Güterschiffe nach 1 152 t im Vorjahr auf 1 222 t an, darunter die der Trockengutmotorschiffe auf 1 310 t (Vorjahr: 1 279 t) und die der Tankmotorschiffe auf 1 502 t (Vorjahr: 1 459 t).

Der bereits im vorherigen Kapitel beschriebene Rückgang bei den Personenschiffahrt betreibenden Unternehmen spiegelt sich auch in einem Rückgang der Zahl der Personenschiffe wider. Dies weist darauf hin, dass hier insgesamt im Jahr 2002 Kapazitäten (vorübergehend oder endgültig) stillgelegt wurden und nicht im Rahmen von Konzentrationsprozessen von anderen Unternehmen weiterbetrieben werden. Dabei wurden aber vor allem kleinere Schiffe aus dem Markt genommen, denn dem Rückgang der Zahl der Personenschiffe um 6,6% steht ein geringerer Rückgang der Zahl der Personenplätze um 2,9% gegenüber. Die durchschnittliche Zahl der Personenplätze je Schiff lag im Jahr 2002 bei 242 Plätzen und somit um neun Plätze höher als im Jahr 2001.

## Leichter Rückgang der Beschäftigtenzahl

Im Jahr 2002 setzte sich der Rückgang der Zahl der Beschäftigten nur leicht, aber dennoch weiter fort (siehe Tabelle 2). Von Mitte 2001 bis Mitte 2002 sank ihre Zahl um 60 auf 7 689 Personen und damit um 0,8%.

Vom Beschäftigungsrückgang stärker betroffen war in absoluten Zahlen das fahrende Personal: 150 Personen weniger als im Vorjahr und damit nur noch 6 162 Personen (-2,4%) führen bei deutschen Unternehmen. Dabei nahm die Zahl der Eigner und unbezahlt mithelfenden Familienangehörigen besonders stark um 104 auf 1 216 Personen (-7,9%) ab, das abhängig beschäftigte Schiffpersonal verringerte sich dagegen „nur“ um 0,9% auf 4 946 Personen. Zunahmen waren dagegen bei der Zahl der an Land beschäftigten Personen zu verzeichnen: Hier arbeiteten in der deutschen Binnenschiffahrt im Jahr 2002 1 527 Personen, das waren 90 oder 6,3% mehr als im Jahr 2001. Diese gegenläufige Entwicklung könnte zum einen damit zusammenhängen, dass sich im Geschäftsfeld „Befrachtung“ durchaus positive Entwicklungen zeigten (siehe dazu auch das folgende Kapitel zur Umsatzentwicklung), zum anderen könnte eventuell trotz sinkender Beförderungen Personal vorerst an Land weiter beschäftigt worden sein, damit dessen Know-how bei wieder anspringender Binnenschiffahrtskonjunktur für die Unternehmen nicht verloren ist.

Von den 7 689 Beschäftigten entfielen 3 717 Beschäftigte auf solche Unternehmen, die ausschließlich in einer einzigen der oben genannten Schiffahrtssparten der Güterschiffahrt (also entweder in der Trockengüterschiffahrt, der Tankschiffahrt oder in der Schub- und Schleppschiffahrt) aktiv waren. 3 181 Beschäftigte arbeiteten in Unternehmen, die ausschließlich Personenschiffahrt betrieben. Die übr-

gen 791 Beschäftigten entfielen auf Unternehmen, die mehrere Binnenschiffahrtstätigkeiten durchführten.

Hinsichtlich der drei bisher betrachteten Merkmale Unternehmenszahl, verfügbare Schiffe und Beschäftigte sind im aktuellen Berichtsjahr zum Teil sehr heterogene Entwicklungen je nach wirtschaftlichem Schwerpunkt der Unternehmen zu verzeichnen. Bei Unternehmen, die ausschließlich Binnenschiffahrt betreiben, gingen die drei genannten Merkmale sowie der Umsatz zurück. Bei den überwiegend in der Binnenschiffahrt tätigen Unternehmen, also denjenigen Unternehmen, die neben ihrer Binnenschiffahrtstätigkeit noch andere Tätigkeiten ausführen, wurden dagegen zum Teil stärkere Rückgänge (Zahl der Unternehmen und der Schiffe), zum Teil aber auch Zunahmen (Zahl der Beschäftigten und Umsatz) verzeichnet. Auch hierfür dürfte wieder die positive Entwicklung im Segment „Befrachtung“ mit verantwortlich sein.

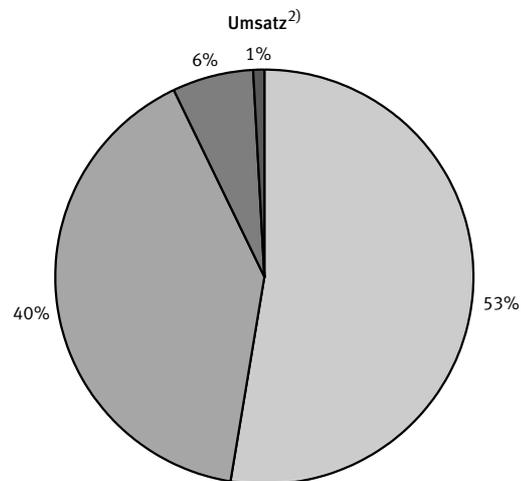
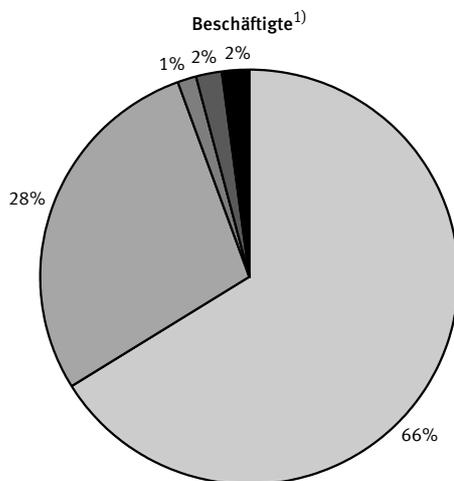
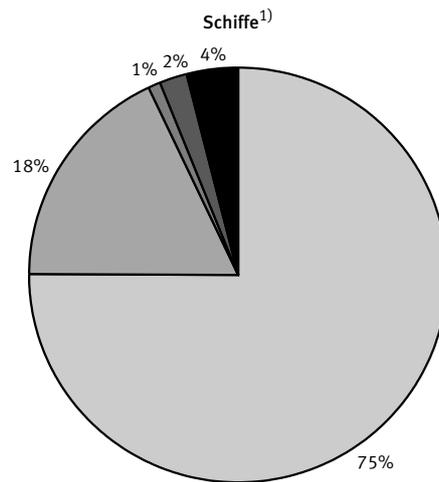
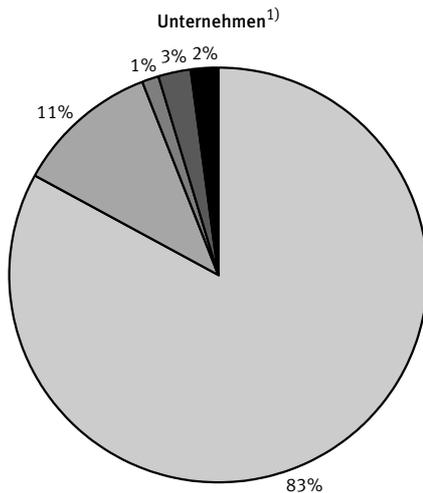
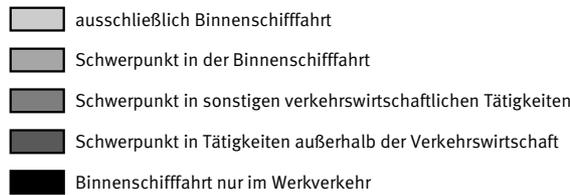
Im Durchschnitt hatte ein ausschließlich in der Binnenschiffahrt tätiges Unternehmen 2,1 Schiffe, beschäftigte 5,0 Personen und erzielte dabei einen Umsatz aus Binnenschiffahrtstätigkeit von 647 000 Euro. Ein überwiegend in der Binnenschiffahrt tätiges Unternehmen hatte dagegen durchschnittlich 3,7 Schiffe, beschäftigte 15,9 Personen und erzielte einen Umsatz von 3,7 Mill. Euro. Immer geringere Bedeutung kommt Unternehmen zu, die Binnenschiffahrt im Werkverkehr betreiben. Hier sind gleichzeitig weniger Unternehmen, weniger Schiffe und weniger Beschäftigte zu verzeichnen. Eine detaillierte Aufteilung der genannten Merkmale sowie des Umsatzes nach dem Schwerpunkt der Tätigkeit der Unternehmen zeigt ergänzend das Schaubild.

## Abnahme beim Umsatz

Der Umsatz aus gewerblicher Binnenschiffahrt sank im Jahr 2002 um 1,4% auf 1 257 Mill. Euro. Da es in der Binnenschiffahrt allerdings durchaus üblich ist, innerhalb des Wirtschaftszweiges Geschäfte der Unternehmen untereinander zu tätigen, umfasst der Gesamtumsatz aus Binnenschiffahrt gewisse Doppelzählungen. Ein genaueres Bild ergibt erst eine nähere Untersuchung der verschiedenen Umsatzarten. Zu unterscheiden sind hier Umsätze aus Beförderung und aus Befrachtung. Bei der Umsatzart „Befrachtung“ handelt es sich um Aufträge für Güterbeförderungen per Binnenschiff, die Binnenschiffahrtsunternehmen – hauptsächlich Reedereien – akquirieren, aber nicht selbst durchführen, sondern an andere Binnenschiffahrtsunternehmen – im Regelfall an Partikuliere – weitergeben bzw. diese mit der Ausführung beauftragen. Vom Gesamtumsatz entfielen 57,6% auf Beförderung (davon 22,9% in der Trockengüterschiffahrt, 19,2% in der Tankschiffahrt und 15,6% in der Personenschiffahrt), 40,6% auf Befrachtung sowie 1,7% auf Schub- und Schleppleistungen. Die Umsatzart Befrachtung hat dabei gegenüber dem Vorjahr in ihrer Bedeutung stark zugenommen, während parallel dazu der Anteil der Umsatzart Beförderung um 3,8 Prozentpunkte abnahm.

Aus Güterbeförderung wurden 2002 um 9,4% niedrigere Umsätze als im Vorjahr erwirtschaftet. Dabei kam es zu

Unternehmen, Schiffe, Beschäftigte und Umsatz 2002 nach wirtschaftlicher Tätigkeit



1) Am 30. Juni. - 2) Ohne Umsatzsteuer.

gleich gerichteten, aber in ihrer Höhe stark differierenden Entwicklungen in den einzelnen Zweigen der Güterschifffahrt. Die Gesamtabnahme resultiert insbesondere aus einer hohen Abnahme der Beförderungsumsätze aus Tankschifffahrt (-14,5%), während die Umsätze aus Trockengüterschifffahrt zwar ebenfalls, aber weniger stark zurückgingen (-4,6%). Hierin spiegeln sich auch die Entwicklung der Beförderungsmengen an trockenem Massengut (-2,0% im Jahr 2002) und flüssigem Massengut (-3,5%) sowie

der Rückgang der Beförderungsmenge der unter deutscher Flagge fahrenden Binnenschiffe wider. Damit musste die Trockengüterschifffahrt hinsichtlich aller relevanten Merkmale (Zahl der Unternehmen, Zahl der Schiffe, Beförderungsmenge, Umsatz) Rückgänge verkräften, während in der Tankschifffahrt bei nahezu konstanter Unternehmenszahl zunehmende Schiffskapazitäten auf sinkende Beförderungsnachfrage trafen<sup>4)</sup>, was die genannten Umsatzabnahmen verursachte.

4) Damit hat sich die Entwicklung im Jahr 2002 gegenüber der des Jahres 2001 genau umgekehrt. Im Jahr 2001 wurden Schiffskapazitäten abgebaut, die auf steigende Nachfrage trafen, was im Jahr 2001 zu stark steigenden Umsätzen in der Tankschifffahrt geführt hatte. Der Umsatz des Jahres 2002 entspricht daher fast dem des Jahres 2000 (243,6 Mill. Euro).

Der Umsatz aus Befrachtung trug im aktuellen Berichtsjahr zur Stabilisierung des deutschen Binnenschiffahrtsgewerbes bei: Mit 511 Mill. Euro lag der Befrachtungsumsatz um 8,4% höher als im Vorjahr. Davon wurden 442 Mill. Euro an Unterfrachtführer weitergeleitet.

Insgesamt kommen bei dieser Entwicklung mehrere Tendenzen zum Tragen, die sich negativ auf die deutschen Güterbeförderungsunternehmen, gegebenenfalls aber positiv auf die Befrachtungsunternehmen auswirken. Das einzige Wachstumssegment in der Binnenschiffahrt war im Jahr 2002 – mit einem Anstieg der Tonnage um 8,3%<sup>5)</sup> – der Transport von Containern, wobei Container insbesondere von und zu den niederländischen Rheinhäfen befördert werden. Gerade am Containerverkehr hat die deutsche Flagge aber einen eher geringen Anteil. Schiffe unter niederländischer Flagge haben im Jahr 2002 fast 70% der TEU<sup>6)</sup> befördert, der deutsche Anteil lag bei unter 20%<sup>7)</sup>.

Die Gesamtlage des deutschen Beförderungsgewerbes stellt sich etwas weniger negativ dar als die bisher genannten Ergebnisse, da die Personenschiffahrt mit nur relativ geringen Rückgängen zu einem gewissen Grad zur Stabilisierung der Umsatzergebnisse beitragen konnte. Allerdings muss auch darauf hingewiesen werden, dass in der Binnenschiffahrt – wie in anderen Verkehrszweigen auch – Personenbeförderung einerseits und Güterbeförderung andererseits Tätigkeitsfelder sind, die unabhängig voneinander sind und sich in ihren wirtschaftlichen Rahmenbedingungen und Einflussgrößen grundlegend unterscheiden. Die hier dargestellte Personenschiffahrt ist dabei weitestgehend den Ausflugs- und Flusskreuzfahrten zuzuordnen, da Personenverkehr mit Fähren nicht erfasst wird, sodass der Personenverkehr mit Binnenschiffen vor allem durch die Urlaubs- und Freizeitgestaltung von Privatpersonen

bestimmt wird. Der Umsatz aus Personenschiffahrt zeigte im Berichtsjahr 2002 den geringsten Rückgang der in der Statistik der Unternehmen der Binnenschiffahrt nachgewiesenen Beförderungskategorien: Mit 196 Mill. Euro lag der Umsatz aus dieser Umsatzart um 4,5 Mill. Euro niedriger als im Vorjahr, das entspricht einem Rückgang um 2,3%. Diese weniger negative Entwicklung als in den übrigen Sparten führte insgesamt dazu, dass der Beförderungsumsatz der deutschen Binnenschiffahrt aus Trockengüterschiffahrt, Tankschiffahrt und Personenschiffahrt „nur“ um 7,6% auf 724 Mill. Euro zurückgegangen ist.

Im Folgenden werden noch einige Eckdaten in der Gliederung nach Größenklassen der Unternehmen vorgestellt, um die unterschiedlichen Strukturen in Abhängigkeit von der Größe der Unternehmen zu dokumentieren. Herangezogen werden hierzu die Ladekapazität der Güterschiffe sowie der Umsatz.

## Gewerbliche Güterbinnenschiffahrt nach Ladekapazität

Einen Überblick über die gewerbliche Güterbinnenschiffahrt und den Werkverkehr nach Ladekapazität gibt Tabelle 4. In die folgende Darstellung sind vor allem jene Unternehmen einbezogen, die gewerbliche Binnenschiffahrt in Form der Güterschiffahrt betreiben, das heißt ihrem Gewerbe durch den Einsatz von mindestens einem Schiff zum Transport von Trocken- oder Flüssiglading nachgehen. Damit sind reine Befrachtungsunternehmen nicht enthalten. Im aktuellen Berichtsjahr zählten 845 Unternehmen zu diesem Binnenschiffahrtszweig; Mitte 2002 fuhren damit 47 Unternehmen weniger als Mitte 2001 in der gewerblichen Güterbinnenschiffahrt (- 5,3%). Der Hauptteil des Rückgangs entfiel

Tabelle 4: Unternehmen, verfügbare Güterschiffe, Beschäftigte und Umsatz der Binnenschiffahrt 2002 nach Ladekapazitätsgrößenklassen

Gegenstand der Nachweisung	Einheit	Unternehmen mit einer Ladekapazität von ... bis unter ... t					
		insgesamt	unter 1 000	1 000 – 3 000	3 000 – 10 000	10 000 – 50 000	50 000 und mehr
<b>Gewerbliche Binnenschiffahrt</b>							
Unternehmen der Binnenschiffahrt am 30. Juni 2002 ...	Anzahl	845	222	559	47	13	4
Verfügbare Güterschiffe am 30. Juni 2002 .....	Anzahl	1 787	227	603	144	163	650
	1 000 t Trgf.	2 211,7	166,9	822,7	211,5	306,6	703,9
Beschäftigte am 30. Juni 2002 .....	Anzahl	3 997	527	1 960	527	456	527
dar.: fahrendes Personal .....	Anzahl	3 336	465	1 812	410	278	371
Umsatz <sup>1)</sup> 2002 .....	Mill. EUR	803,4	41,6	196,0	152,7	259,1	154,1
dar.: aus Beförderung .....	Mill. EUR	523,7	29,7	184,2	104,0	126,8	79,0
in der Trockengüterschiffahrt .....	Mill. EUR	287,2	28,1	114,7	23,5	63,2	57,6
in der Tankschiffahrt .....	Mill. EUR	236,5	1,6	69,5	80,6	63,6	21,3
in der Personenschiffahrt .....	Mill. EUR	0,0	–	0,0	–	–	–
<b>Werkverkehr</b>							
Unternehmen der Binnenschiffahrt am 30. Juni 2002 ...	Anzahl	21	7	10	4	–	–
Verfügbare Güterschiffe am 30. Juni 2002 .....	Anzahl	81	20	32	29	–	–
	1 000 t Trgf.	71,2	4,2	19,0	48,1	–	–
Beschäftigte am 30. Juni 2002 .....	Anzahl	136	15	49	72	–	–
dar.: fahrendes Personal .....	Anzahl	110	11	31	68	–	–

1) Ohne Umsatzsteuer.

5) Siehe Reim, U.: „Kombinierter Verkehr 2002“ in WiSta 1/2004, S. 63 ff.

6) Twenty-foot-Equivalent-Unit; 20-Fuß-Einheit (Container von etwa 6 m Länge). Über die TEU können unterschiedliche Containergrößen vergleichbar gemacht werden.

7) Siehe Winter, H.: „Binnenschiffahrt 2002 – Fortsetzung des konjunkturbedingten Transportrückgangs“ in WiSta 7/2003, S. 623 ff.

auf Unternehmen mit einer Ladekapazität von unter 1 000 t (– 19 Unternehmen) und auf Unternehmen mit einer Ladekapazität von 1 000 bis unter 3 000 t (– 28 Unternehmen). Auch im Jahr 2002 wurde damit insbesondere von kleinen Beförderungsunternehmen mit einem oder zwei Schiffen die Binnenschifffahrtstätigkeit aufgegeben.

In der gewerblichen Güterbinnenschifffahrt wurden 1 787 Güterschiffe eingesetzt und 3 997 Personen beschäftigt. Gegenüber dem Jahr 2001 nahm die Zahl der verfügbaren Güterschiffe um 2,6% und das Personal um 0,9% ab. Die Unternehmen der gewerblichen Güterbinnenschifffahrt erwirtschafteten im Jahr 2002 einen Umsatz von 803 Mill. Euro, darunter 524 Mill. Euro aus Beförderung. Der Binnenschifffahrtsumsatz sank im Vorjahresvergleich um 11,5%. Dabei nahm der Umsatz aus Beförderung um 9,5% und der Umsatz aus Befrachtung um 16,4% ab. Der entgegen der allgemeinen Entwicklung verlaufende und sehr hohe Rückgang der Befrachtungsumsätze dürfte dabei daraus resultieren, dass sich Unternehmen mit in den Vorjahren noch eigenen Kapazitäten aus der Beförderung zurückgezogen haben und somit in der Tabelle nicht mehr mit aufgenommen sind.

Unternehmen mit weniger als 3 000 t Ladekapazität stellten mit 92,4% den überwiegenden Teil der Unternehmen und beschäftigten knapp zwei Drittel des Personals (62,2%), sie verfügten aber lediglich über 46,4% der Güterschiffe und erwirtschafteten nur 29,6% des Umsatzes der gewerblichen Güterbinnenschifffahrt, allerdings 40,8% des Beförderungsumsatzes.

Über mindestens 10 000 t eigene Ladekapazität konnten 17 Unternehmen disponieren; sie hatten damit einen Anteil von 2,0% an der Gesamtzahl der Unternehmen und beschäftigten 24,6% des Personals, verfügten jedoch über 45,5% der Schiffe und erzielten 51,4% des Umsatzes der gewerblichen Güterbinnenschifffahrt.

## Strukturdaten nach Umsatzgrößenklassen

Die weit überwiegende Zahl der Unternehmen der gewerblichen Binnenschifffahrt (988 Unternehmen bzw. 81,9%) erzielte im Jahr 2002 einen Umsatz von weniger als 500 000 Euro (siehe Tabelle 5). Im Durchschnitt erreichten diese vier Fünftel aller Unternehmen einen Umsatz je Unternehmen von 203 000 Euro (Vorjahr: 201 000 Euro) und einen Umsatz je Beschäftigten von 60 500 Euro (Vorjahr: 61 500 Euro). Die kleineren Unternehmen verfügten über 44,2% der Schiffe und beschäftigten 44,1% des Personals. Ihr Anteil am Gesamtumsatz betrug lediglich 16,0%, ihr Anteil am Umsatz aus Beförderung allerdings 27,1%. Mit wachsender Größe der Unternehmen sinkt der Anteil des Umsatzes aus Beförderung am Gesamtumsatz. Wurde bei den kleineren Unternehmen der gesamte Umsatz fast vollständig aus Beförderung erwirtschaftet, betrug dieser Anteil bei den umsatzstärksten Unternehmen nur noch rund ein Drittel.

Im Jahr 2002 sind nicht nur, wie in den Vorjahren, vor allem Unternehmen mit einem Jahresumsatz von unter 500 000 Euro aus dem Markt ausgeschieden, sondern Unternehmen fast aller Größenklassen. Bei den kleineren Unternehmen wurden 42 Unternehmen, bei den Unternehmen mit einem Jahresumsatz zwischen 500 000 und 2,5 Mill. Euro 31 Unternehmen und bei den Unternehmen mit einem Jahresumsatz zwischen 2,5 und 12,5 Mill. Euro zwei Unternehmen weniger als im Vorjahr registriert. Lediglich die Zahl der Großunternehmen blieb konstant.

Mindestens jeweils 12,5 Mill. Euro Umsatz im Jahr 2002 erwirtschafteten 19 Unternehmen. Diese 1,6% der Unternehmen erzielten 54,1% des Gesamtumsatzes aus Binnenschifffahrt mit 30,8% der Schiffe, beschäftigten aber nur 15,1% aller tätigen Personen. Von den Großunternehmen wurde ein durchschnittlicher Umsatz je Unternehmen

Tabelle 5: Unternehmen, verfügbare Schiffe, Beschäftigte und Umsatz der Binnenschifffahrt 2002 nach Umsatzgrößenklassen

Gegenstand der Nachweisung	Einheit	Unternehmen mit einem Umsatz von ... bis unter ... EUR						
		insgesamt	unter 125 000	125 000 – 500 000	500 000 – 2,5 Mill.	2,5 Mill. – 12,5 Mill.	12,5 Mill. und mehr	ohne Angaben <sup>1)</sup>
Unternehmen der gewerblichen Binnenschifffahrt am 30. Juni 2002 ..	Anzahl	1 206	272	716	155	44	19	–
Verfügbare Schiffe am 30. Juni 2002 ...	Anzahl	2 754	315	903	367	322	847	–
Güterschiffe .....	Anzahl	1 787	128	605	193	130	731	–
	1 000 t Trgf.	2 211,7	93,0	752,8	289,4	206,7	869,8	–
Schubboote, Schub-Schleppboote und Schleppboote .....	Anzahl	204	9	42	48	4	101	–
	1 000 kW	113,2	1,3	16,5	23,7	1,7	70,0	–
Fahrgast- und Fahrgastkabinenschiffe	Anzahl	763	178	256	126	188	15	–
	1 000 Pers.-Pl.	184,6	23,5	50,1	39,0	61,9	10,2	–
Beschäftigte am 30. Juni 2002 .....	Anzahl	7 527	717	2 599	1 474	1 600	1 137	–
dar.: fahrendes Personal .....	Anzahl	6 034	653	2 309	1 253	1 197	622	–
Umsatz <sup>2)</sup> 2002 .....	Mill. EUR	1 256,6	20,5	180,1	144,2	232,1	679,8	X
dar.: aus Beförderung und Befrachtung	Mill. EUR	1 235,0	20,2	175,8	133,3	231,9	673,8	X
aus Beförderung .....	Mill. EUR	724,4	20,2	175,8	130,0	166,6	231,9	X
in der Trockengüterschifffahrt .....	Mill. EUR	287,2	11,1	116,3	25,6	39,3	94,8	X
in der Tankschifffahrt .....	Mill. EUR	241,1	0,2	27,3	68,2	22,4	123,0	X
in der Personenschifffahrt .....	Mill. EUR	196,1	8,9	32,2	36,1	104,8	14,1	X
aus Befrachtung .....	Mill. EUR	510,6	–	0	3,3	65,4	441,9	X

1) Unternehmen, die im Berichtsjahr gegründet worden sind. – 2) Ohne Umsatzsteuer.

von 35,8 Mill. Euro (Vorjahr: 34,2 Mill. Euro) und ein durchschnittlicher Umsatz je Beschäftigten von 598 000 Euro (Vorjahr: 588 000 Euro) erreicht.

## Zusammenfassung und Ausblick

Das Jahr 2002 war insgesamt für die deutschen Binnenschiffahrtsunternehmen ein Jahr mit sehr viel mehr Schatten als Licht. Der gesamte Beförderungssektor musste Einbußen verkraften, besonders stark in der Tankschiffahrt (mit Ausnahme der Transportkapazitäten), aber auch in der Trockengüterschiffahrt. Lediglich in der Sparte „Befrachtung“ konnten steigende Umsätze erwirtschaftet werden. Auch im Jahr 2003 dürfte sich die wirtschaftliche Lage der deutschen Binnenschiffahrtsunternehmen nicht wesentlich verbessert haben: Zum einen zeigen die zu den Verkehrsleistungen bereits vorliegenden Daten einen Rückgang der auf deutschen Binnenwasserstraßen beförderten Mengen um 5,1% und der dabei erbrachten Beförderungsleistungen um 9,4%, was insbesondere auf die niedrigen Flusspegel durch die Trockenheit im Sommer und die damit verbundenen Schiffahrtseinschränkungen zurückzuführen ist. Zum anderen war auch im Jahr 2003 ein weiterer Rückgang des Anteils der unter deutscher Flagge transportierten Gütermengen auf 36,3% (2002: 36,6%) zu registrieren. [III](#)

Dr. Stefan Linz, Dipl.-Kaufmann Timm Behrmann, Dipl.-Betriebswirt (FH) Ulf Becker

# Hedonische Preismessung bei EDV-Investitionsgütern

Ab Berichtsmonat Mai 2004 nutzt das Statistische Bundesamt auch für die Preismessung bei EDV-Investitionsgütern die so genannte hedonische Methode. Dies wirkt sich auf die Ergebnisse der Erzeuger-, Einfuhr- und Ausführungspreisindizes sowie auf die Ergebnisse des Großhandelsverkaufspreisindex aus. Bisher wurden für EDV-Investitionsgüter in der Regel so genannte Matched-Model-Indizes berechnet, bei denen nur die Preise derjenigen Produkte gemessen werden, die im aktuellen Monat und im Vormonat in gleicher Qualität abgesetzt wurden. Durch die Einführung der hedonischen Qualitätsbereinigung können nun auch Preiswirkungen gemessen werden, die indirekt durch eine verbesserte Produktqualität bei etwa gleich bleibenden Verkaufspreisen entstehen. Diese Preiswirkungen spielen wegen des technischen Fortschritts bei EDV-Produkten eine besonders wichtige Rolle. Im Folgenden werden das Verfahren der Qualitätsbereinigung sowie die Auswirkungen der Implementierung auf die Ergebnisse der Preisindizes beschrieben.

## 1 Hedonische Preismessung im Statistischen Bundesamt

Das Ziel der amtlichen Preisstatistik besteht darin, die so genannte „reine“ Preisänderung, unbeeinflusst von Ände-

rungen der Verbrauchsgewohnheiten, Güterarten oder Güterqualitäten, zu messen. Das entspricht im Grundsatz dem Laspeyres-Prinzip, bei dem ein Warenkorb mit bestimmten Gütern festgelegt und über einen definierten Zeitraum nach Möglichkeit konstant gehalten wird.<sup>1)</sup>

Insbesondere können die Preise eines Gutes aus zwei Perioden nur dann aussagekräftig miteinander verglichen werden, wenn die Qualität des Gutes konstant bleibt. Ist dies nicht gegeben – zum Beispiel aufgrund des technischen Fortschritts – so wird in der Regel eine Qualitätsbereinigung vorgenommen. Diese zielt darauf ab, den Geldwert der veränderten Güterqualität beim Preisvergleich zu berücksichtigen.<sup>2)</sup>

Bei der hedonischen Qualitätsbereinigung wird mit Hilfe der Regressionsanalyse ein rechnerischer Zusammenhang zwischen der Güterqualität und den Verkaufspreisen hergestellt. Auf diese Weise kann der Geldwert von Qualitätsänderungen geschätzt und bei der Preismessung berücksichtigt werden.<sup>3)</sup>

Die hedonische Qualitätsbereinigung wurde im Statistischen Bundesamt bisher nur in Teilbereichen des Verbraucherpreisindex eingesetzt. Seit Juni 2002 wird dort der Teilindex für Personalcomputer<sup>4)</sup> und seit Mai 2003 der Teil-

1) Siehe hierzu auch Kunz, D.: „Ausgewählte methodische und praktische Probleme des zeitlichen Preisvergleichs“ in Allgemeines Statistisches Archiv, Bd. 55, Heft 1/1971, S. 23 ff.

2) Zu den Qualitätsbereinigungsverfahren siehe Szenzenstein, J.: „Preisindizes für industrielle Güter in der amtlichen Statistik“ in Harhoff, D./Müller, M. (Hrsg.): „Preismessung und technischer Fortschritt“, ZEW-Wirtschaftsanalysen, Baden-Baden 1995.

3) Zur hedonischen Preismessung siehe z. B. Griliches, Z. (Hrsg.): „Introduction: Hedonic Price Indexes Revisited“ in Griliches, Z. (Hrsg.): „Price Indexes and Quality Change, Studies in New Methods of Measurement“, Harvard University Press, Cambridge 1971, oder Harhoff, D.: „Methodik und Einsatz hedonischer Preisindizes – Ein Überblick“ in Harhoff, D./Müller, M. (Hrsg.): „Preismessung und technischer Fortschritt“, ZEW-Wirtschaftsanalysen, Baden-Baden 1995, sowie Brachinger, H.: „Statistical Theory of Hedonic Price Indices“, Working Paper, zu finden im Internet unter <http://www.unifr.ch/stat/en-home.php>.

4) Siehe hierzu Linz, S./Eckert, G.: „Zur Einführung hedonischer Methoden in die Preisstatistik“ in WiSta 10/2002, S. 857 ff.

index für Gebrauchtwagen<sup>5)</sup> nach der hedonischen Methode berechnet. Für den Bereich „Neuwagen“ im Verbraucherpreisindex wurde Anfang 2003 eine Evaluation des bestehenden Index anhand von hedonischen Methoden abgeschlossen. Da sich keine systematischen Abweichungen zwischen dem bisherigen Qualitätsbereinigungsverfahren und dem hedonischen Index zeigten, wurde das bestehende Berechnungsverfahren für Neuwagen beibehalten.<sup>6)</sup> Anfang 2004 wurde die hedonische Preismessung außerdem für die Berechnung eines Häuserpreisindex im Rahmen einer europäischen Pilotstudie eingesetzt.<sup>7)</sup> Gegenwärtig in Bearbeitung befinden sich hedonische Verbraucherpreisindizes für die Bereiche „elektrische Haushaltsgeräte“ und „Unterhaltungselektronik“, die bis Januar 2005 in die laufende Berechnung implementiert werden sollen.

Mit der Berechnung von hedonischen Erzeuger-, Einfuhr-, Ausfuhr- und Großhandelspreisindizes für EDV-Güter wird diese Methode nun auch bei der Preismessung in vorgelagerten Wirtschaftsstufen eingesetzt.

Der Index der Erzeugerpreise gewerblicher Produkte misst die Entwicklung der Preise für die vom Verarbeitenden Gewerbe, von der Energie- und Wasserversorgung sowie vom Bergbau und der Gewinnung von Steinen und Erden in Deutschland hergestellten und abgesetzten Erzeugnisse. Die Bezugsgröße des Gesamtindex ist die Summe aller gewerblichen Inlandsumsätze im Basisjahr, das gegenwärtig das Jahr 2000 ist. Damit werden sowohl der Absatz industrieller Güter an inländische Abnehmer anderer Wirtschaftsstufen als auch die Umsätze zwischen den gewerblichen Unternehmen berücksichtigt. Der Gesamtindex berechnet sich als gewichteter Durchschnitt der bei einer repräsentativen Auswahl gewerblicher Produkte gemessenen Preisänderungen. Als Gewichtungszahlen dienen die Umsatzanteile der ausgewählten Erzeugnisse im Basisjahr.

Die Außenhandelspreisindizes (Einfuhr- und Ausfuhrpreise) messen die Entwicklung der Preise der Waren, die zwischen Deutschland und dem Ausland gehandelt werden. Die Bezugsgröße der Gesamtindizes ist der Gesamtwert der Ein- bzw. Ausfuhren im Basisjahr 2000. Öffentliche Abgaben (Zölle, Abschöpfungen, Währungsausgleichsbeträge, Einfuhrumsatzsteuer bzw. bei der Ausfuhr Umsatzsteuer, Verbrauchsteuern sowie Exporthilfen) sind in den Preisen nicht enthalten. In die Indexberechnungen gehen nur Preise in Euro ein.

Mit dem Index der Großhandelsverkaufspreise wird die Entwicklung der Preise für die von Großhändlern im Inland abgesetzten Waren gemessen, soweit diese Verkäufe der Großhandelsfunktion zuzurechnen sind (z. B. Abschlüsse mit Wiederverkäufern, mit Verarbeitern oder mit anderen Großabnehmern). Dabei kann es sich um im Inland erzeugte oder um eingeführte Waren handeln. Die Bezugsgröße des Gesamtindex ist die Summe aller Umsätze des Großhandels im Basisjahr 2000, vermindert um den Einzelhandelsabsatz

und andere Verkäufe, die nicht der Großhandelsfunktion zuzurechnen sind, sowie um die Ausfuhrwerte. Damit werden also auch die Umsätze zwischen den Großhandelsunternehmen berücksichtigt.

## 2 Datenquellen

### 2.1 Prinzip der Approximation

Da bei der hedonischen Qualitätsbereinigung die beobachteten Produktpreise und -qualitäten in eine Regressionsfunktion eingesetzt werden, ist die Anwendung dieser Methode nur möglich, wenn sich die Qualität der betreffenden Produkte durch eine überschaubare Anzahl von Merkmalen beschreiben lässt. Die Produkte müssen also weitgehend standardisiert sein. Darüber hinaus kann das Verfahren nur angewandt werden, wenn monatlich eine größere Anzahl von Preisbeobachtungen verfügbar ist, die sich auf verschiedene Varianten derselben Produktart beziehen.

Sollen in Deutschland Erzeuger- oder Ausfuhrpreisindizes für EDV-Güter berechnet werden, so besteht zunächst das Problem, dass in dieser Sparte hier nur wenige produzierende Firmen angesiedelt sind. EDV-Güter werden in Deutschland weitgehend importiert. Auch die Daten deutscher Unternehmen, die in Deutschland standardisierte EDV-Investitionsgüter herstellen, reichen bei weitem nicht aus, um hedonische Indizes zu berechnen.

Das Problem wird – wie in der Statistik international üblich – durch eine Approximation der Erzeugerpreise gelöst: Der europäische Markt für standardisierte EDV-Güter ist ein globaler Markt, der preislich sehr homogen ist. Die Preisbildung im EDV-Bereich erfolgt praktisch nur in den wichtigen Herstellerländern wie den Vereinigten Staaten, Japan und Taiwan. Die wenigen deutschen EDV-Hersteller sind bei den standardisierten Produkten daher gezwungen, ihre Preise den globalen Marktpreisen anzupassen, um ihre Erzeugnisse absetzen zu können. Daher kann man davon ausgehen, dass sich die Preise der in Deutschland hergestellten EDV-Güter parallel zu den in Deutschland beobachteten Großhandelspreisen dieser Gütergruppe bewegen. Die Großhandelspreise messen die Preisentwicklung von Gütern, die im Inland auf der Großhandelsstufe verkauft werden, wobei diese Güter auch im Ausland hergestellt sein können.

Im Schaubild 1 sind die Warenströme grob skizziert. Im Bereich der standardisierten EDV-Güter (Server, Desktop-Computer, Notebooks, Drucker, Prozessoren usw.) werden die Produkte meist über den Großhandel importiert und dann an deutsche Unternehmen (Firmen, Zwischenhändler, Einzelhändler) verkauft. Die Einkaufspreise des Großhandels bilden dadurch eine Richtgröße, an der sich auch die deutschen Hersteller dieser Produkte orientieren müssen. Die Verkaufspreise der deutschen Erzeuger (Pfeil 4) können

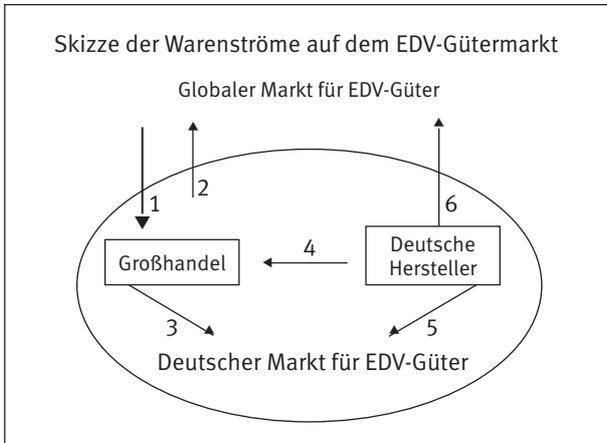
5) Siehe hierzu Linz, S./Dexheimer, V./Kathe, A.: „Hedonische Preismessung bei Gebrauchtwagen“ in WiSta 6/2003, S. 538 ff. Eine aktualisierte Version dieses Artikels findet sich im Internet unter [http://www.destatis.de/themen/d/thm\\_preise.htm](http://www.destatis.de/themen/d/thm_preise.htm).

6) Siehe hierzu die Pressemitteilung des Statistischen Bundesamtes vom 17. Februar 2003 (<http://www.destatis.de/presse/deutsch/pm2003/p0580051.htm>).

7) Siehe hierzu Behrmann, T./Kathe, A.: „Zur Anwendung hedonischer Methoden beim Häuserpreisindex“ in WiSta 5/2004, S. 525 ff.

also durch Großhandelseinkaufspreise (Pfeil 1) für solche Produkte approximiert werden.

Schaubild 1



Weiterhin gibt es Direktverkäufe deutscher Hersteller an Firmen, Zwischenhändler oder Einzelhändler in Deutschland (Pfeil 5). Die Preise der so verkauften Güter lassen sich durch Recherchen im Internet erheben, weil es sich um relativ wenige Produkte handelt, die nur ergänzend einbezogen werden. Daher ist hier keine Approximation notwendig.

Im Bereich der Ausführpreisstatistik besteht im Prinzip dasselbe Problem wie bei den Erzeugerpreisen und auch hier können die Ausführpreise durch Großhandelspreise ersetzt werden. Die Daten haben gezeigt, dass die deutschen Hersteller bei Verkäufen ins Ausland ihre Produkte zu 80% (Umsatzanteil) indirekt über den Großhandel absetzen und zu 20% direkt ins Ausland verkaufen. Die Preise der Produkte deutscher EDV-Hersteller, die über den Großhandel ins Ausland abgesetzt werden, sind hier durch den Großhandelsverkaufspreis approximiert. Die Preise der von deutschen Firmen direkt ins Ausland verkauften Produkte werden durch den Großhandelseinkaufspreis approximiert, weil die deutschen Erzeuger im Ausland an den gleichen Markt verkaufen, an dem auch deutsche Großhändler einkaufen. Hierbei werden wiederum die Daten der Direktvermarkter berücksichtigt.

## 2.2 Preisstatistiken mit hedonischen Indizes für EDV-Investmentsgüter

Im vorigen Abschnitt wurde gezeigt, dass hedonische Großhandelspreisindizes für EDV-Güter für verschiedene Indizes der vorgelagerten Wirtschaftsstufen repräsentativ zu verwenden sind. In der Übersicht 1 sind die Statistiken aufgeführt, in welche die hedonischen Indizes für EDV-Investmentsgüter implementiert wurden.

Die Erzeugerpreise für EDV-Güter werden – wie oben dargestellt – durch Großhandelseinkaufspreise approximiert und ergänzt durch Preise von Erzeugern, die ihre Produkte selbst direkt vermarkten. Die Ausführpreise werden im EDV-Bereich außer durch die Großhandelseinkaufspreise noch durch Großhandelsverkaufspreise und durch Verkaufspreise der Direktvermarkter approximiert. Da der Import von EDV-Investmentsgütern zum größten Teil über den Großhandel erfolgt, kann

Übersicht 1: Preisstatistiken und die eingehenden hedonischen Indizes im EDV-Bereich

Statistik	Approximiert durch:
Erzeugerpreisstatistik	Großhandelseinkaufspreis + Preise von Direktvermarktern
Ausführpreisstatistik	Großhandelseinkaufspreis + Großhandelsverkaufspreis + Preise von Direktvermarktern
Einfuhrpreisstatistik	Großhandelseinkaufspreis
Großhandelsverkaufspreisstatistik	Großhandelsverkaufspreis + Preise von Direktvermarktern

die Einfuhrpreisentwicklung direkt an den Großhandelseinkaufspreisen gemessen werden. Für die Großhandelsverkaufspreisstatistik werden die Großhandelsverkaufspreise durch Verkaufspreise von Direktvermarktern ergänzt.

## 2.3 Verfügbare Datenquellen

Der nächste Schritt besteht darin, auf der Großhandelsebene die benötigten Preise und Qualitätsmerkmale der EDV-Güter zu beschaffen. Im Bereich des Großhandels gibt es einerseits große, europaweit agierende Unternehmen, die in bedeutenden Mengen direkt am Weltmarkt einkaufen und ihre Produkte an deutsche Zwischenhändler und Unternehmen absetzen. Bei den Einkaufspreisen dieser Großhändler handelt es sich um Preise, die sich nahe an der Produktionsstufe befinden und sich somit für eine Approximation von Erzeugerpreisen eignen. Diese Großhändler haben ihren Firmensitz überwiegend nicht in Deutschland. Andererseits gibt es kleinere, in Deutschland angesiedelte Großhändler (Zwischenhändler), die ihre Produkte ihrerseits von den großen Großhändlern beziehen. Mit deren Einkaufspreisen können die Erzeugerpreise weniger gut approximiert werden, da eine größere Handelsspanne hinzukommt.

Bei Anwendung der Approximation muss das Ziel daher sein, die großen europäischen Großhändler einzubeziehen. Für eigene Erhebungen mit Auskunftspflicht kommen aber nur die in Deutschland angesiedelten kleineren Großhändler in Frage. Um Daten der europäischen Großhändler zu verwenden, musste daher auf Marktforschungsdaten zurückgegriffen werden. Die Marktforschungsunternehmen erhalten die Daten direkt aus den Warenwirtschaftssystemen der befragten Unternehmen und bezahlen hierfür. Damit haben sie in diesem Fall bessere Möglichkeiten der Datenbeschaffung als die amtliche Statistik. Die Marktforschungsdaten werden durch eigene Erhebungen im Bereich der Direktvermarktung ergänzt.

Die Produktauswahl hat sich letztlich danach gerichtet, welche Daten auf der erzeugernahen Großhandelsstufe von den Marktforschungsinstituten zu bekommen waren. Dabei handelt es sich um Preise und Produktmerkmale für die folgenden neun Produkte:

Übersicht 2: Produkte für die hedonische Preismessung bei EDV-Investmentsgütern

(1) Desktop-Computer	(6) Multifunktionsdrucker
(2) Notebooks	(7) Prozessoren
(3) Server	(8) Speicherchips
(4) Laserdrucker	(9) Festplatten
(5) Tintenstrahldrucker	

## 2.4 Preisrepräsentanten für Gütergruppen

Die Güter der Erzeuger-, Einfuhr-, Ausfuhr- und Großhandelsverkaufspreisstatistik werden nach dem Güterverzeichnis für Produktionsstatistiken, Ausgabe 2002 (GP 2002), klassifiziert. Der EDV-Bereich umfasst im Güterverzeichnis im Wesentlichen die Abteilungen (Zweisteller) 30 „Büromaschinen, Datenverarbeitungsgeräte und -einrichtungen“ sowie 32 „Nachrichtentechnik, Rundfunk- und Fernsehgeräte sowie elektronische Bauelemente“. Bei den Güterarten des Güterverzeichnisses im EDV-Bereich handelt es sich um „Sammelposten“, die viele unterschiedliche Produkte umfassen können. Bei den Datenverarbeitungsgeräten gibt es zum Beispiel die GP-Nummer 3002 12 000 mit folgender Bezeichnung:

„Tragbare digitale automatische Datenverarbeitungsmaschinen mit einem Gewicht von 10 kg oder weniger, die mindestens eine Zentraleinheit, eine Tastatur und einen Bildschirm enthalten.“

Dazu gehören einerseits die bekannten Notebooks, andererseits gibt es im Erzeugerbereich noch eine Reihe anderer Produkte, auf die diese Beschreibung ebenfalls zutrifft (z. B. Handheld-Computer oder spezielle Messgeräte). Die Produkte aus Übersicht 2 dienen deshalb als Repräsentanten für die verschiedenen Güterarten des Warenkorbes.

Die bisherigen Preismeldungen werden dort, wo Repräsentanten verfügbar sind, vollständig durch diese ersetzt. Durch die Verwendung der Repräsentanten ergibt sich, dass in der Breite weniger Gütervarianten, dafür aber in der Tiefe mehr Produkte in die Statistiken eingehen. Zum Bei-

spiel werden bei der Güterart „Desktop-Computer“ momentan nur neun Produkte, darunter neben Desktops auch sehr spezielle Produkte, wie zum Beispiel Patientenmonitore, erfasst. Nach der Umstellung auf die hedonische Preismessung würden für diese Gütergruppe nur noch reine Desktop-Computer herangezogen, hiervon allerdings etwa 130 unterschiedliche Produkte.

## 2.5 Warenkorbgewichte der betroffenen Statistiken

EDV-Investitionsgüter spielen insbesondere für die Einfuhr eine bedeutende Rolle. Im Warenkorb des deutschen Einfuhrpreisindex nehmen die EDV-Investitionsgüter ein Gewicht von insgesamt 8,8% ein. Aber auch für die Ausfuhr und den Großhandel bilden diese Produkte mit jeweils etwa 5% des Warenkorbes einen wichtigen Güterbereich. Der Erzeugerpreisindex umfasst insgesamt ein sehr breites Spektrum verschiedener Güterarten, sodass die EDV-Investitionsgüter hier nur ein Gewicht von 1,3% ausmachen. Die Güterarten der Erzeuger-, Einfuhr-, Ausfuhr- und Großhandelsverkaufspreisstatistik, die nun durch hedonische Indizes repräsentiert werden, sind in Tabelle 1 mit ihren Anteilen am jeweiligen Warenkorb aufgeführt.

## 3 Hedonische Methode

### 3.1 Imputationsmethode

Zur Berechnung der hedonischen Preisindizes werden die monatlichen Daten zu Preisen und Qualitätsmerkmalen der

Tabelle 1: Hedonische Preisindizes und deren Gewichte im Bereich der EDV-Investitionsgüter  
Warenkorbbanteile in Prozent

Nr. der Klassifikation GP <sup>1)</sup> bzw. WZ <sup>2)</sup>	Name des hedonisch berechneten Preisindex	Erzeugerpreise gewerblicher Produkte	Einfuhrpreise	Ausfuhrpreise	Großhandelsverkaufspreise
30 02 12	Tragbare digitale automatische Datenverarbeitungsmaschinen .....	0,305	0,522	0,354	
30 02 13	Stationäre digitale automatische Datenverarbeitungsmaschinen .....	0,048	0,181	0,044	
30 02 14	Digitale Datenverarbeitungsmaschinen in Form von Systemen .....	0,463	0,286	0,093	
30 02 15	Digitale Datenverarbeitungsmaschinen in Form von Systemen, andere .....		0,572	0,454	
30 02 16	Ein- oder Ausgabeeinheiten, auch wenn sie in einem gemeinsamen Gehäuse Speichereinheiten umfassen .....	0,033	1,206	0,471	
30 02 17	Speichereinheiten .....		1,024	0,426	
30 02 18	Andere Geräte für die Datenverarbeitung		0,271	0,090	
30 02 19	Teile und Zubehör für automatische Datenverarbeitungsmaschinen und ihre Einheiten, a.n.g. ....	0,234	2,114	1,146	
32 10 6	Elektronische integrierte Schaltungen und zusammengesetzte elektronische Mikroschaltungen (Mikrobausteine) ..	0,223	2,653	1,964	
WZ 51.64	Großhandel mit Datenverarbeitungsgeräten, peripheren Einheiten und Software .....				4,948
	EDV-Investitionsgüter insgesamt ...	1,306	8,829	5,042	4,948

1) Güterverzeichnis für Produktionsstatistiken (GP 2002). – 2) Klassifikation der Wirtschaftszweige, Ausgabe 1993.

verschiedenen Produkte jeweils in drei Unterstichproben A, B und C eingeteilt:

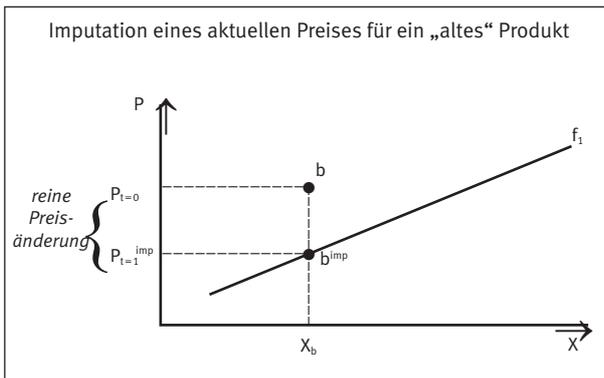
Übersicht 3: Teilstichproben

Teilstichprobe	Basismonat	...	Aktueller Monat
A: unveränderte Produkte	Produkt vorhanden	...	Produkt vorhanden
B: alte Produkte	Produkt vorhanden	...	-/-
C: neue Produkte	-/-	...	Neues Produkt

Die „unveränderten“ Produkte waren sowohl im Basismonat ( $t=0$ ) als auch im aktuellen Monat ( $t=1$ ) in gleicher Form am Markt erhältlich. Für diese Produkte ist keine Qualitätsbereinigung notwendig, denn die reine Preisentwicklung kann direkt beobachtet werden. Die „alten“ Produkte gehörten im Basismonat zur Stichprobe, sind jedoch in der aktuellen Periode nicht mehr beobachtbar. Für diese Produkte wird in der aktuellen Periode ein Preis imputiert, der unter den gegenwärtigen Marktbedingungen für ein Produkt mit der alten Qualität durchschnittlich zu zahlen wäre. Die „neuen“ Produkte waren im Basismonat noch nicht am Markt erhältlich, wurden aber in der aktuellen Periode neu beobachtet. Für diese neuen Produkte wird ein Basispreis imputiert, der zu den Marktverhältnissen der Basisperiode für das Produkt mit der aktuellen Qualität hätte gezahlt werden müssen. Die gesamte Indexreihe für die betreffende Produktart wird dann als Mittelwert der Preisänderungen bei den unveränderten, alten und neuen Produkten berechnet.

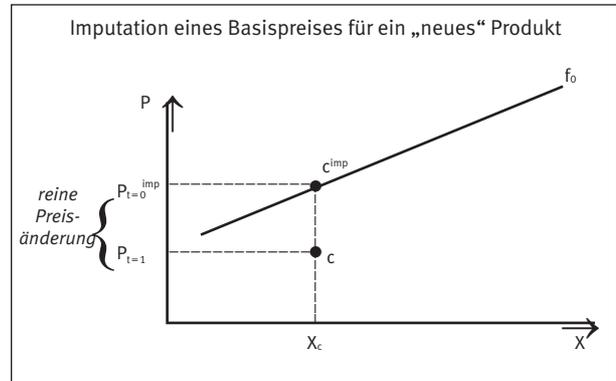
Bei der Imputation der aktuellen Preise für alte Produkte und der Basispreise von neuen Produkten wird die Regressionsanalyse angewendet. Im Schaubild 2 ist das Prinzip an einem Beispiel für ein „altes“ Produkt der Teilstichprobe B dargestellt. Die Regressionsfunktion  $f_1$  wurde ausschließlich mit Daten der aktuellen Periode berechnet. Sie gibt für alternative Qualitätsniveaus  $X$  den Preis an, den die Abnehmer in der aktuellen Periode für die jeweilige Qualität im Durchschnitt zu zahlen bereit sind. Das Produkt  $b$  ist in der Basisperiode in die Stichprobe gelangt, war jedoch in der aktuellen Periode nicht mehr am Markt erhältlich. Für die aktuelle Periode wird daher ein Produktpreis geschätzt und imputiert. Die Differenz zwischen dem imputierten aktuellen Preis  $P_{t=1}^{imp}$  und dem tatsächlichen Basispreis  $P_{t=0}$  für das Produkt  $b$  entspricht der reinen Preisänderung für dieses Beispielprodukt zwischen der aktuellen und der Basisperiode.

Schaubild 2



Ebenso wird verfahren, um den Basispreis für ein „neues“ Produkt zu schätzen, wobei hier die Regressionsfunktion ausschließlich mit den Daten der Basisperiode berechnet wird.

Schaubild 3

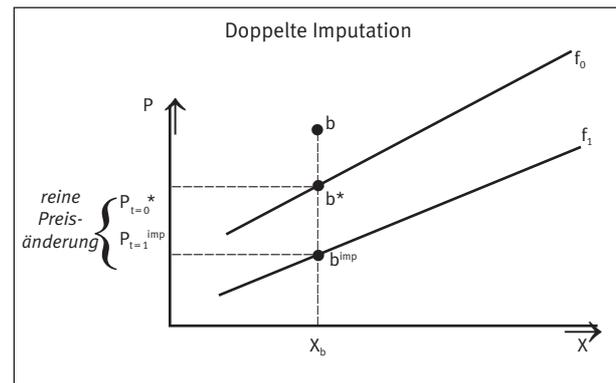


### 3.2 Doppelte Imputation

In den Schaubildern 2 und 3 wurde zur Erläuterung gezeigt, dass der imputierte Preis der einen Periode jeweils mit dem tatsächlich beobachteten Preis der anderen Periode verglichen werden kann. Für die praktische Anwendung ist es jedoch vorteilhaft, den imputierten Preis mit einem ebenfalls geschätzten Preis zu vergleichen.

Bei dieser Vorgehensweise wird auch für die Periode, in der ein Preis vorhanden ist, eine Regression berechnet und ein Preis geschätzt. Dieser ersetzt den direkt beobachteten Preis, was den Vorteil hat, dass Ausreißer ausgeglichen werden. Die Vorgehensweise ist im Schaubild 4 für den Fall der Imputation eines aktuellen Preises dargestellt.

Schaubild 4



### 3.3 Indexberechnung

Der Index für eine Produktart wird als geometrischer Mittelwert der Preisreihen aus den Teilstichproben A, B und C berechnet. Die Indexberechnung ist in der Tabelle 2 schematisch dargestellt.

Die beiden Produkte der Teilstichprobe A sind in beiden Perioden ohne Qualitätsänderungen am Markt erhältlich, sodass die Preisentwicklung in dieser Teilstichprobe direkt gemessen werden kann. Die Produkte der Teilstichprobe B

Tabelle 2: Indexberechnung

Produkt	Teilstichprobe	Preise		Einzelpreisentwicklung
		Basisperiode	aktuelle Periode	
1	A	$P_{t=0}$	$P_{t=1}$	$\hat{P}_1 = P_{t=1} / P_{t=0}$
2		$P_{t=0}$	$P_{t=1}$	$\hat{P}_2 = P_{t=1} / P_{t=0}$
3	B	$P_{t=0}^*$	$P_{t=1}^{imp}$	$\hat{P}_3 = P_{t=1}^{imp} / P_{t=0}^*$
4		$P_{t=0}^*$	$P_{t=1}^{imp}$	$\hat{P}_4 = P_{t=1}^{imp} / P_{t=0}^*$
5	C	$P_{t=0}^{imp}$	$P_{t=1}^*$	$\hat{P}_5 = P_{t=1}^* / P_{t=0}^{imp}$
6		$P_{t=0}^{imp}$	$P_{t=1}^*$	$\hat{P}_6 = P_{t=1}^* / P_{t=0}^{imp}$

sind in der aktuellen Periode nicht mehr verfügbar, sodass ein aktueller Preis imputiert wird. Die Basispreise werden ebenfalls durch Schätzwerte ersetzt, um Verzerrungen durch Ausreißerwerte zu vermeiden. Für die Teilstichprobe C werden die fehlenden Basispreise imputiert und die aktuellen Preise durch Schätzwerte ersetzt.

Mit Hilfe der Imputation können für die Teilstichproben B und C konsistente Preisreihen erzeugt werden, die sich auf eine gleich bleibende Produktqualität beziehen. Der gesuchte Preisindex wird dann als Mittelwert aller Einzelpreisentwicklungen der Teilstichproben A, B und C berechnet:

$$(1) \text{ Preisindex für eine Güterart} = \prod_{i=1}^n \hat{P}_i^{g_i}$$

mit

$g_i$  = Gewichtungsfaktor des Produktes  $i$  ( $\sum_i g_i = 1$ )

$\hat{P}_i$  = Preisentwicklung gegenüber der Basisperiode bei Produkt  $i$

Die einzelnen Preisreihen werden mit der Umsatzbedeutung gewichtet, welche die Produkthersteller haben. Ist ein Hersteller mit mehreren Produkten in der Stichprobe vertreten, so wird sein Umsatzanteil gleichmäßig auf die Produkte verteilt. Besser geeignet wäre eine genaue Gewichtung auch der verschiedenen Produkte eines Herstellers nach deren Marktanteilen; diese Information ist jedoch nicht verfügbar.

### 3.4 Regressionsanalyse

Zur Berechnung der Regressionsgleichung wird die doppelt logarithmische Funktionsform verwendet. Diese Funktionsform ist im Vergleich zur linearen oder semilogarithmischen Form geeigneter, da unter anderem nicht-lineare Zusammenhänge und Zusammenhänge zwischen den einzelnen Merkmalen besser abgebildet werden können. Für Desktop-Computer wird beispielsweise folgende Regressionsfunktion verwendet:

$$(2) \ln(P) = \beta_0 + \beta_1 \cdot \ln(\text{Leistung}) + \beta_2 \cdot D_1\text{RAM} + \dots + \beta_4 \cdot D\text{ Graphik} + \beta_5 \cdot D_1\text{Marke} + \dots + \beta_9 \cdot D_1\text{Prozessor} + \dots + \beta_{11} \cdot D\text{ Software} + \varepsilon$$

mit

Leistung = Leistung des Prozessors in MHz

$D_i\text{RAM}$  = Dummyvariablen für verschiedene Arbeitsspeichergößen ( $i = 1, 2$ )

$D\text{ Graphik}$  = Dummyvariable für Graphikkarte

$D_j\text{Marke}$  = Dummyvariablen für verschiedene Herstellermarken ( $j = 1, \dots, 4$ )

$D_k\text{Prozessor}$  = Dummyvariablen für verschiedene Prozessortypen ( $k = 1, 2$ )

$D\text{ Software}$  = Dummyvariable für Betriebssystemtyp

$\varepsilon$  = Zufallsvariable

Der Preis eines Desktop-Computers wird mit Hilfe der Merkmale Prozessorleistung und -typ, Größe des Arbeitsspeichers, Graphikkarte, Marke und Betriebssystem erklärt. Dabei geht die Größe des Arbeitsspeichers nicht als metrische Variable ein, sondern als Dummyvariable, da es für die Arbeitsspeichergröße nur wenige Ausprägungen gibt. Die Größe der Festplatte wurde nicht einbezogen, da sie sehr stark mit der Größe des Arbeitsspeichers korreliert ist und der Festplattenumfang zudem für die Kaufentscheidung zusehends an Bedeutung verliert. Über die Dummyvariable  $D\text{ Graphik}$  wird abgebildet, ob der PC eine Graphikkarte besitzt oder ob die Graphik über einen Chip auf dem Motherboard gesteuert wird.

Die Auswahl der Merkmale wird monatlich angepasst. In der Tabelle 3 sind beispielhaft die Ergebnisse einer Regressionsanalyse für Desktop-Computer im Februar 2004 angegeben. Das Referenzmodell, bei dem alle Dummyvariablen den Wert Null annehmen, besitzt einen Intel Pentium 4 Prozessor, einen Arbeitsspeicher mit einer Größe von 512 MB, stammt von der Marke N. N. und verfügt über Windows 2000 als Betriebssystem. Der Koeffizient für die Leistung lässt sich im doppelt logarithmischen Modell als Elastizität interpretieren: Weist der Computer eine 1% höhere Prozessor-

Tabelle 3: Regressionsergebnisse Februar 2004 für Desktop-Computer im Einfuhrpreisindex<sup>1)</sup>

Variable	Parameterschätzer	Standardfehler	t-Wert	Varianzinflationsfaktor
Absolutglied .....	-5,667	2,82	-2,01	0
$\ln(\text{Leistung})$ .....	1,567	0,35	4,42	1,87
$D_1\text{RAM (128 MB)}$ ...	-0,512	0,16	-3,14	1,88
$D_2\text{RAM (256 MB)}$ ...	-0,089	0,05	-1,65	2,13
$D\text{ Graphik}$ .....	-0,183	0,07	-2,75	2,18
$D_1\text{Marke}$ .....	-0,196	0,05	-3,68	1,46
$D_2\text{Marke}$ .....	0,162	0,05	3,27	1,55
$D_3\text{Marke}$ .....	0,215	0,07	3,06	1,71
$D_4\text{Marke}$ .....	0,296	0,11	2,74	1,61
$D_1\text{Prozessor}$ .....	-0,271	0,06	-4,74	1,47
$D_2\text{Prozessor}$ .....	-0,153	0,09	-1,79	1,69
$D\text{ Software}$ .....	0,110	0,05	2,18	1,67

1) Bestimmtheitsmaß  $R^2 = 81\%$ .

geschwindigkeit auf, so zahlen die Käufer in der Regel einen etwa 1,6% höheren Kaufpreis.

Der Erklärungsgehalt dieses Modells für Desktop-Computer ist mit einem Bestimmtheitsmaß von 81% sehr hoch. Für die anderen Produkte zeigen sich ähnlich hohe Werte (Notebooks 70%, Server 88%, Laserdrucker 93%, Tintenstrahldrucker 65%, Multifunktionsdrucker 73%, RAM 80%, Festplatten 95%, Prozessoren 86%). Der Varianzinflationsfaktor als Maß für Multikollinearität nimmt hier wie in allen anderen Regressionsmodellen meist Werte unter bzw. nur knapp über zwei an.

#### 4 Indexergebnisse

In den ersten Spalten der Tabelle 4 ist für den Bereich der EDV-Investitionsgüter, für den Preisindizes in Zukunft hedonisch berechnet werden, die Preisänderung im April 2004 gegenüber Januar 2004 angegeben. Die Indizes mit hedonischer Qualitätsbereinigung zeigen durchweg stärkere Preisrückgänge als Matched-Model-Preisindizes an, da auch indirekte Preiseffekte gemessen werden können, die auf Qualitätsverbesserungen bei in etwa konstanten Preisen beruhen.

Die Implementierung der hedonischen Qualitätsbereinigung hat deutliche Auswirkungen insbesondere auf den Gesamtindex der Einfuhrpreisstatistik (rechte Seite der Tabelle 4).

Tabelle 4: Hedonische Indizes und Matched-Model-Indizes für EDV-Güter sowie Gesamtindizes  
Preisänderung April 2004 gegenüber Januar 2004 in %

Statistik	EDV-Investitionsgüter (GP 3002 und 32106)		Gesamtindex	
	Matched-Model	hedonisch	Matched-Model	hedonisch
Einfuhrpreisindex ...	+0,5	-5,5	+1,8	+1,2
Ausfuhrpreisindex ...	-1,5	-4,7	+0,9	+0,7
Erzeugerpreisindex ..	-3,7	-5,4	+1,1	+1,0

Der gesamte Ausfuhrpreisindex und der gesamte Erzeugerpreisindex zeigen für den betrachteten Zeitraum bei hedonischer Berechnung der Teilindizes für EDV-Investitionsgüter eine nur etwas schwächere Preissteigerung an als beim bisherigen Verfahren, was auch auf das geringere Gewicht der EDV-Investitionsgüter bei der Ausfuhr und im Erzeugerbereich zurückzuführen ist.

Beim Großhandelsverkaufspreisindex, für den die Ergebnisse ausschließlich nach Wirtschaftszweigen und nicht nach Güterarten ausgewiesen werden, liegt der hedonisch berechnete Index im gleichen Zeitraum um 0,7 Prozentpunkte niedriger als nach bisheriger Methodik.

Hochgerechnet auf ein Jahr ergeben die hedonisch berechneten Preisentwicklungen bei den EDV-Investitionsgütern Preisrückgänge von 20% und mehr pro Jahr. Solche Preisrückgänge lassen sich bei EDV-Produkten üblicherweise beobachten, wenn Qualitätsbereinigungsverfahren angewendet werden, da auf diese Weise der technische Fortschritt besser abgebildet wird. So sind die Preise für Perso-

nalcomputer im deutschen Verbraucherpreisindex in den Jahren 2002 und 2003 jährlich im Durchschnitt um 26% gefallen. Der englische harmonisierte Verbraucherpreisindex für Datenverarbeitungsgeräte weist im selben Zeitraum eine Jahresrate von durchschnittlich etwa - 23% auf. Der amerikanische Erzeugerpreisindex für EDV-Güter ist in den Jahren 2002 und 2003 im Durchschnitt jährlich um 22% gefallen.

#### 5 Traditionelle Verfahren der Qualitätsbereinigung

Hedonische Indizes unterscheiden sich von den Matched-Model-Indizes dadurch, dass bei der hedonischen Preisermessung eine Qualitätsbereinigung durchgeführt wird. Neben der hedonischen Qualitätsbereinigung gibt es aber auch traditionelle Qualitätsbereinigungsverfahren, die unter Umständen zu sehr ähnlichen Ergebnissen führen können. Bei früheren Analysen im Bereich des Verbraucherpreisindex zeigten sich für die Produkte „Computer“ und „Neuwagen“ nur geringe Unterschiede im Ergebnis zwischen der hedonischen Berechnung und der dort zuvor bereits angewandten traditionellen Qualitätsbereinigung. Die Einführung hedonischer Methoden hatte daher im Verbraucherpreisindex deutlich schwächere Auswirkungen auf die gemessene Preisentwicklung als jetzt im Erzeuger-, Einfuhr-, Ausfuhr- und Großhandelsbereich, wo bisher ein Matched-Model-Index ohne systematische Qualitätsbereinigung berechnet wurde.

#### 6 Hedonische Preise und Messung des Wirtschaftswachstums

Die Qualitätsbereinigung der Preisstatistik spielt auch eine wichtige Rolle bei der Messung des realen Bruttoinlandsprodukts. Im Rahmen der Deflationierung wird das nominale Bruttoinlandsprodukt mit Hilfe von Preisindizes in reale Größen umgerechnet. Die Umrechnung erfolgt im Prinzip dadurch, dass die Nominalwerte durch die entsprechenden Preisindizes dividiert werden. Je niedriger die Preisentwicklung bei einer Produktkategorie ausfällt, desto höher sind die gemessenen realen Umsatzzuwächse dieser Branche. Sinkende oder schwächer steigende Preise aufgrund eines verbesserten Qualitätsbereinigungsverfahrens der Preisstatistik finden daher auch ihren Niederschlag in der Messung des realen Wirtschaftswachstums.

In den Vereinigten Staaten werden hedonische Methoden schon seit längerer Zeit zur Qualitätsbereinigung in der Preisstatistik eingesetzt. Der geringere Anstieg der Preisindizes im Vergleich zum Matched-Model-Index führte dort rechnerisch zu einer höheren Zuwachsrates des realen Wirtschaftswachstums. Aufgrund der unterschiedlichen Wirtschaftsstrukturen in den Vereinigten Staaten und in Deutschland sind diese Effekte bei der Einführung hedonischer Methoden in Deutschland jedoch nicht in ähnlichem Umfang zu erwarten. Insbesondere werden in Deutschland EDV-Investitionsgüter in deutlich größerem Umfang eingeführt als in den Vereinigten Staaten, sodass bezüglich der

Wirkung auf das Bruttoinlandsprodukt der positive Effekt auf die „realen“ Investitionen zu einem Großteil von einem entgegengesetzten Effekt auf die preisbereinigten Importe kompensiert wird. Allerdings dürfte es in Deutschland zu einer Verlagerung zwischen den Verwendungsaggregaten des Bruttoinlandsprodukts kommen, das heißt mehr Investitionsdynamik und geringere Wachstumswirkung des Außenbeitrags. [u](#)

Dipl.-Kauffrau Gudrun Eckert

## Preise im Mai 2004

Im Mai 2004 zeigte sich auf den unterschiedlichen Wirtschaftsstufen ein uneinheitliches Bild: Der Index der industriellen Erzeugerpreise erhöhte sich um 1,6% (April 2004: +0,9%), der Index der Großhandelsverkaufspreise sogar um 3,6% (April 2004: +2,4%). Der Index der Einzelhandelspreise stieg um 0,7% (April 2004: +0,4%), der Verbraucherpreisindex lag um 2,0% (April 2004: +1,6%) über dem Niveau des Vorjahres.

Im Vormonatsvergleich zogen sowohl die industriellen Erzeugerpreise (April 2004: +0,4%) als auch die Großhandelsverkaufspreise (April 2004: +0,4) um 0,5% an. Der Index der Einzelhandelspreise lag im Mai 2004 ebenso um 0,2% über dem Niveau des Vormonats (April 2004: +0,0%) wie der Verbraucherpreisindex (April 2004: +0,3%).

Die steigenden Erzeuger- sowie Verbraucherpreise sind zum Teil auf höhere Energiepreise zurückzuführen. Die Erzeugerpreise für Energie haben sich gegenüber Mai 2003 um durchschnittlich 3,9% erhöht. Einen besonders starken Preisschub gab es bei den Mineralölerzeugnissen (+12,9% gegenüber Mai 2003). Darin spiegeln sich die stark gestiegenen Weltmarktpreise für Erdöl wider.

Auch die Verbraucherpreise für Heizöl und Kraftstoffe sind überdurchschnittlich stark gestiegen. Sie lagen im Mai 2004 um 12,7% über dem Niveau des Vorjahres. Ohne Heizöl und Kraftstoffe hat sich der Verbraucherpreisindex im Vergleich zum Vorjahr nur um 1,5% erhöht.

Der Index der Erzeugerpreise gewerblicher Produkte<sup>1)</sup> lag im Mai 2004 um 1,6% höher als im Mai 2003. Im März hatte die Jahresveränderungsrate noch bei +0,3%, im April bei +0,9% gelegen. Im Vergleich zum Vormonat ist der Index im Mai 2004 um 0,5% gestiegen.

Bei den stark gestiegenen Preisen für Mineralölerzeugnisse gab es im Einzelnen folgende Preisveränderungen im Jahresvergleich: Kraftstoffe +11,6% (darunter Benzin: +11,7%, Diesel: +11,6%), leichtes Heizöl +23,0%, schweres Heizöl +12,9%, Flüssiggas +21,5%.

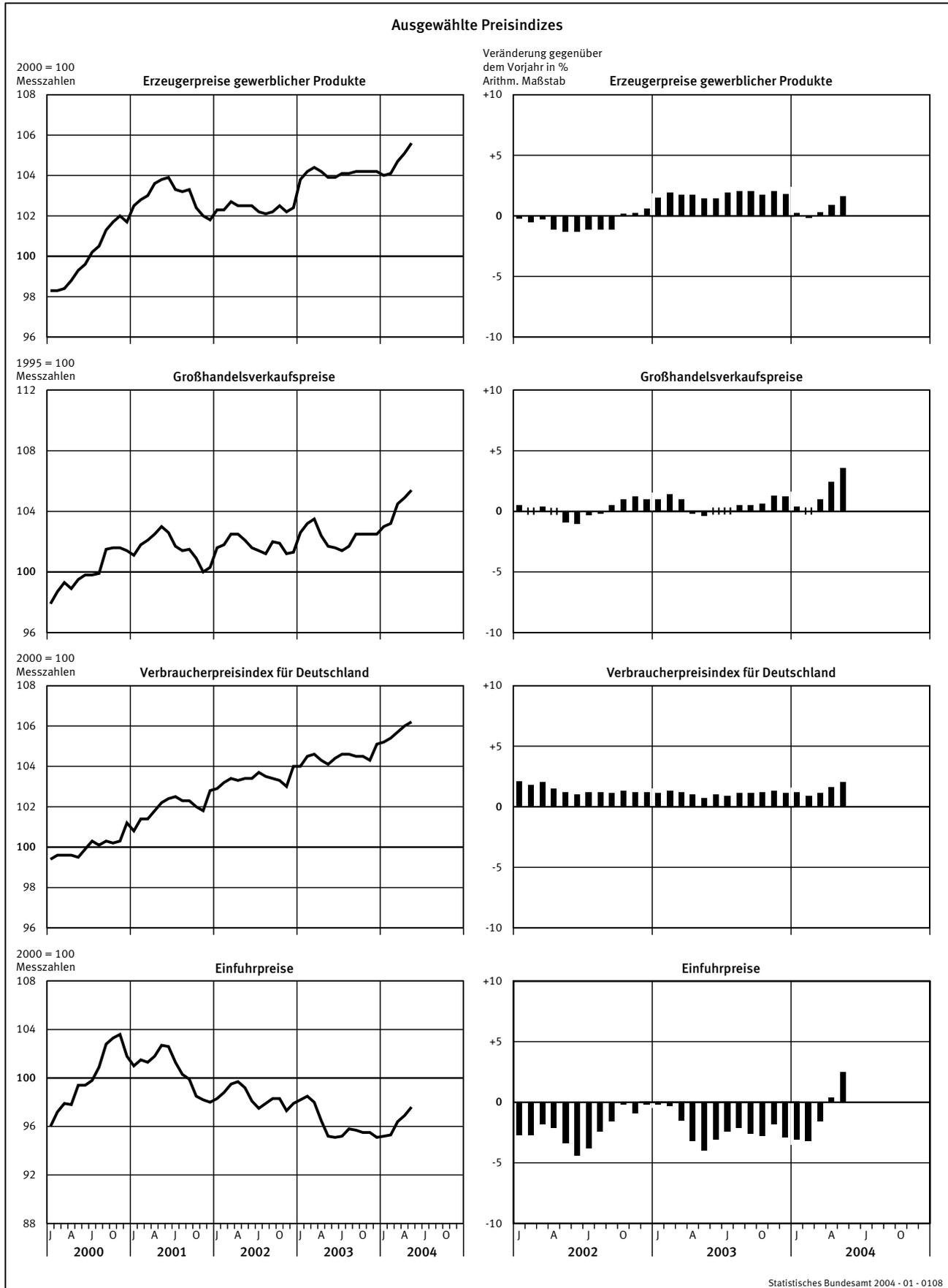
Ohne Energie hat der Erzeugerpreisindex um 1,0% über dem Vorjahresniveau gelegen. Vorleistungsgüter verteuerten sich binnen Jahresfrist um durchschnittlich 2,0%, Investitionsgüter um 0,2% und Konsumgüter um 0,9%.

	Veränderungen Mai 2004 gegenüber	
	April 2004	Mai 2003
	%	
Index der Erzeugerpreise gewerblicher Produkte		
insgesamt .....	+ 0,5	+ 1,6
ohne Mineralölerzeugnisse .....	+ 0,3	+ 1,1
Mineralölerzeugnisse .....	+ 4,9	+ 12,9
Index der Großhandelsverkaufspreise .....	+ 0,5	+ 3,6
Index der Einzelhandelspreise .....	+ 0,2	+ 0,7
Verbraucherpreisindex		
insgesamt .....	+ 0,2	+ 2,0
ohne Heizöl und Kraftstoffe .....	-	+ 1,5
Heizöl und Kraftstoffe .....	+ 5,4	+ 12,7
ohne Saisonwaren .....	+ 0,2	+ 2,0
Saisonwaren .....	+ 2,5	+ 2,9

1) Im Berichtsmonat Mai 2004 wurden die Preisindizes für EDV-Investitionsgüter für den Erzeugerpreisindex sowie für den Index der Großhandelsverkaufspreise erstmals mit Hilfe der so genannten hedonischen Methode berechnet (siehe dazu den Beitrag von Linz, S./Behrmann, T./Becker, U.: „Hedonische Preismessung bei EDV-Investitionsgütern“ in dieser Ausgabe auf S. 682 ff.

Schaubild 1

Ausgewählte Preisindizes



Statistisches Bundesamt 2004 - 01 - 0108

Ausgewählte Preisindizes

Jahr Monat	Erzeuger- preise gewerblicher Produkte <sup>1)</sup>	Großhandels- verkaufs- preise <sup>1)</sup>	Einzel- handels- preise <sup>2)</sup>	Verbraucher- preis- index
2000 = 100				
1998 D .....	98,0	96,1	99,7	98,0
1999 D .....	97,0	95,3	99,9	98,6
2000 D .....	100,0	100,0	100,0	100,0
2001 D .....	103,0	101,6	101,1	102,0
2002 D .....	102,4	101,8	101,8	103,4
2003 April ..	104,2	102,4	102,1	104,3
Mai ...	103,9	101,7	102,0	104,1
Juni ...	103,9	101,6	102,1	104,4
Juli ...	104,1	101,4	101,9	104,6
Aug. ...	104,1	101,7	101,7	104,6
Sept. ...	104,2	102,5	102,0	104,5
Okt. ...	104,2	102,5	102,1	104,5
Nov. ...	104,2	102,5	102,1	104,3
Dez. ...	104,2	102,5	102,1	105,1
2004 Jan. ...	104,0	103,0	102,2	105,2
Febr. ...	104,1	103,2	102,1	105,4
März ..	104,7	104,5	102,5	105,7
April ...	105,1	104,9	102,5	106,0
Mai ....	105,6	105,4	102,7	106,2
Veränderungen gegenüber dem jeweiligen Vormonat in %				
2003 April ...	- 0,2	- 1,1	- 0,2	- 0,3
Mai ...	- 0,3	- 0,7	- 0,1	- 0,2
Juni ...	-	- 0,1	+ 0,1	+ 0,3
Juli ...	+ 0,2	- 0,2	- 0,2	+ 0,2
Aug. ...	-	+ 0,3	- 0,2	-
Sept. ...	+ 0,1	+ 0,8	+ 0,3	- 0,1
Okt. ...	-	-	+ 0,1	-
Nov. ...	-	-	-	- 0,2
Dez. ...	-	-	-	+ 0,8
2004 Jan. ...	- 0,2	+ 0,5	+ 0,1	+ 0,1
Febr. ...	+ 0,1	+ 0,2	- 0,1	+ 0,2
März ..	+ 0,6	+ 1,3	+ 0,4	+ 0,3
April ...	+ 0,4	+ 0,4	-	+ 0,3
Mai ....	+ 0,5	+ 0,5	+ 0,2	+ 0,2
Veränderungen gegenüber dem entsprechenden Vorjahreszeitraum in %				
1998 D .....	- 0,4	- 2,0	+ 0,4	+ 0,9
1999 D .....	- 1,0	- 0,8	+ 0,2	+ 0,6
2000 D .....	+ 3,1	+ 4,9	+ 0,1	+ 1,4
2001 D .....	+ 3,0	+ 1,6	+ 1,1	+ 2,0
2002 D .....	- 0,6	+ 0,2	+ 0,7	+ 1,4
2003 April ..	+ 1,7	- 0,1	- 0,2	+ 1,0
Mai ...	+ 1,4	- 0,4	- 0,1	+ 0,7
Juni ...	+ 1,4	-	+ 0,3	+ 1,0
Juli ...	+ 1,9	-	+ 0,3	+ 0,9
Aug. ...	+ 2,0	+ 0,5	+ 0,4	+ 1,1
Sept. ...	+ 2,0	+ 0,5	+ 0,5	+ 1,1
Okt. ...	+ 1,7	+ 0,6	+ 0,7	+ 1,2
Nov. ...	+ 2,0	+ 1,3	+ 0,8	+ 1,3
Dez. ...	+ 1,8	+ 1,2	+ 0,7	+ 1,1
2004 Jan. ...	+ 0,2	+ 0,4	+ 0,3	+ 1,2
Febr. ...	- 0,1	-	- 0,1	+ 0,9
März ..	+ 0,3	+ 1,0	+ 0,2	+ 1,1
April ...	+ 0,9	+ 2,4	+ 0,4	+ 1,6
Mai ....	+ 1,6	+ 3,6	+ 0,7	+ 2,0

1) Ohne Umsatzsteuer. – 2) Einschl. Umsatzsteuer; einschl. Kraftfahrzeughandel und Tankstellen.

Die Stahlpreise befinden sich weiterhin auf dem höchsten Niveau seit 1989. Hauptursache für die drastischen Preissteigerungen ist ganz wesentlich die erhöhte weltweite Nachfrage nach Stahl. Die Preise für in Deutschland produzierten und abgesetzten Walzstahl stiegen im Mai 2004 gegenüber dem Vormonat nur noch um durchschnittlich 1,4% (April 2004 gegenüber März 2004: + 8,2%), lagen aber um 17,6% über dem Niveau vom Mai 2003. Einzelne Stahl-

sorten verteuerten sich binnen Jahresfrist noch deutlich stärker: Betonstahl um 78,7%, Walzdraht um 46,5% und schwere Profile aus unlegiertem Stahl um 21,3%. Bei den in Deutschland produzierten Rohstoffen für die Metallherzeugung sind die Preisanstiege für Steinkohle und Steinkohlenbriketts (+ 28,6% gegenüber Mai 2003) und für metallische Sekundärrohstoffe (+ 20,0%) hervorzuheben.

Weitere überdurchschnittliche Preiserhöhungen im Jahresvergleich gab es im Mai 2004 bei folgenden Gütern: Drahtwaren (+ 42,9%), Kupfer und Kupferhalbzeug (+ 29,5%), Blei, Zink und Zinn und Halbzeug daraus (+ 18,7%), Futtermitteln für Nutztiere (+ 16,1%), Tabakerzeugnissen (+ 12,3%), Rohren aus Eisen und Stahl (+ 10,8%), Polyvinylchlorid (+ 10,5%), pflanzlichen und tierischen Ölen und Fetten (+ 8,2%) sowie elektrischem Strom (+ 5,5%).

Billiger als vor Jahresfrist waren im Mai 2004 unter anderem: Datenverarbeitungsgeräte und -einrichtungen (- 8,5%), anorganische Grundstoffe und Chemikalien (- 7,5%), nachrichtentechnische Geräte und Einrichtungen (- 7,2%), Klebstoffe und Gelatine (- 6,8%), Fischerzeugnisse (- 6,5%) sowie synthetischer Kautschuk (- 5,1%).

Der *Index der Großhandelsverkaufspreise* lag im Mai 2004 um 3,6% über dem Vorjahresstand. Dies war die höchste Veränderung gegenüber einem Vorjahresmonat seit April 2001 (+ 3,6%). Im März 2004 und im April 2004 hatten die Jahresveränderungsraten + 1,0% bzw. + 2,4% betragen. Gegenüber April 2004 erhöhte sich der Großhandelspreisindex um 0,5%.

Besonders stark stiegen im Mai 2004 unter anderem die Preise des Großhandels mit Erzen, Eisen, Stahl, Nicht-Eisen-Metallen und Halbzeug (+ 21,5% zum Vorjahr bzw. + 5,1% zum Vormonat), mit Tabakwaren (+ 13,7% zum Vorjahr bzw. + 4,4% zum Vormonat) sowie mit festen Brennstoffen und Mineralölherzeugnissen (+ 10,4% zum Vorjahr bzw. + 5,2% zum Vormonat). Auch bei den Preisen des Großhandels für Getreide, Saaten und Futtermittel war ein starker Anstieg gegenüber Mai 2003 (+ 19,4%) zu beobachten. Dagegen waren im Vorjahresvergleich Preisrückgänge beim Großhandel mit pharmazeutischen Erzeugnissen und medizinischen Hilfsmitteln (- 6,4%) sowie mit Büromaschinen (- 6,3%) zu verzeichnen.

	Veränderungen Mai 2004 gegenüber	
	April 2004	Mai 2003
	%	
Großhandel mit		
Landwirtschaftlichen Grundstoffen und lebenden Tieren .....	- 3,4	+ 10,6
Nahrungsmitteln, Getränken und Tabakwaren .....	- 0,1	+ 2,3
Gebrauchs- und Verbrauchsgütern .....	- 0,4	- 1,7
Rohstoffen, Halbwaren, Altmaterial und Reststoffen .....	+ 2,6	+ 9,9
Maschinen, Ausrüstungen und Zubehör .....	- 0,4	- 2,4
Sonstiger Großhandel .....	-	+ 3,1

Für die gewerblichen Erzeugerpreise sowie für die Großhandelsverkaufspreise ergaben sich in Deutschland folgende

größere Veränderungen gegenüber dem Vormonat bzw. dem entsprechenden Vorjahresmonat:

	Veränderungen Mai 2004 gegenüber	
	April 2004	Mai 2003
	%	
<b>Erzeugerpreise gewerblicher Produkte</b>		
Betonstahl .....	+ 0,6	+ 78,7
Walzdraht .....	+ 4,3	+ 46,5
Drahtwaren .....	+ 7,7	+ 42,9
Kupfer und Kupferhalbzeug .....	- 4,0	+ 29,5
Steinkohle und Steinkohlenbriketts .....	+ 5,8	+ 28,6
Leichtes Heizöl .....	+ 5,9	+ 23,0
Flüssiggas .....	+ 0,9	+ 21,5
Schwere Profile aus unlegiertem Stahl .....	+ 3,6	+ 21,3
Metallische Sekundärrohstoffe .....	- 3,6	+ 20,0
Blei, Zink und Zinn und Halbzeug daraus .....	+ 0,9	+ 18,7
Walzstahl .....	+ 1,4	+ 17,6
Futtermittel für Nutztiere .....	- 0,2	+ 16,1
Schweres Heizöl .....	+ 7,8	+ 12,9
Tabakerzeugnisse .....	+ 0,1	+ 12,3
Motorenbenzin .....	+ 5,3	+ 11,7
Dieselmotorenkraftstoff .....	+ 4,7	+ 11,6
Rohre aus Eisen und Stahl .....	+ 2,6	+ 10,8
Polyvinylchlorid .....	+ 1,3	+ 10,5
Pflanzliche und tierische Öle und Fette .....	- 0,8	+ 8,2
Elektrischer Strom .....	+ 1,9	+ 5,5
Synthetischer Kautschuk .....	+ 1,6	- 5,1
Fischerzeugnisse .....	- 0,4	- 6,5
Klebstoffe und Gelatine .....	- 0,3	- 6,8
<b>Nachrichtentechnische Geräte und Einrichtungen .....</b>		
Einrichtungen .....	- 0,1	- 7,2
<b>Anorganische Grundstoffe und Chemikalien ..</b>		
Anorganische Grundstoffe und Chemikalien ..	+ 0,9	- 7,5
<b>Datenverarbeitungsgeräte und -einrichtungen</b>		
Datenverarbeitungsgeräte und -einrichtungen	- 0,4	- 8,5
<b>Großhandelsverkaufspreise</b>		
<b>Großhandel mit:</b>		
Erzen, Eisen, Stahl, NE-Metallen und Halbzeug	+ 5,1	+ 21,5
Getreide, Saaten und Futtermitteln .....	- 3,1	+ 19,4
Tabakwaren .....	+ 4,4	+ 13,7
Festen Brennstoffen und Mineralölerzeugnissen	+ 5,2	+ 10,4
Pharmazeutischen Erzeugnissen und medizinischen Hilfsmitteln .....	- 1,8	- 6,4
Büromaschinen .....	- 1,3	- 6,3

Der *Index der Einzelhandelspreise* stieg im Mai 2004 im Vergleich zum Vorjahreszeitraum um 0,7%. Im März 2004 hatte die Jahresteuersatzrate + 0,2%, im April 2004 + 0,4% betragen. Im Vergleich zum Vormonat erhöhten sich die Einzelhandelspreise im April 2004 um 0,2%.

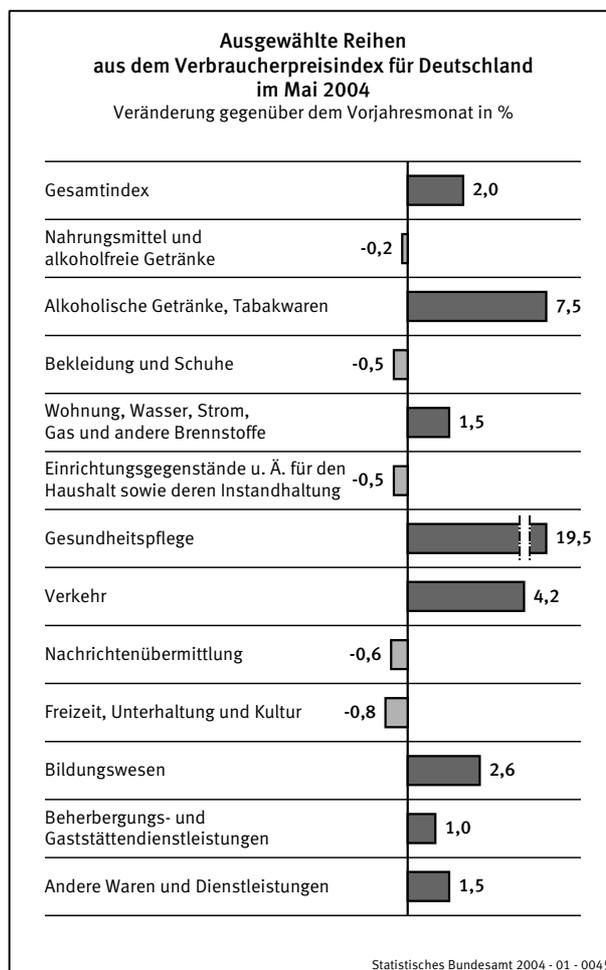
Der *Verbraucherpreisindex für Deutschland* ist im Mai 2004 gegenüber Mai 2003 um 2,0% gestiegen. Im März bzw. April 2004 hatte die Jahresveränderungsrate bei + 1,1% bzw. + 1,6% gelegen. Gegenüber April 2004 sind die Verbraucherpreise um 0,2% gestiegen.

Die deutlich höhere Jahresteuersatzrate im Mai ist im Wesentlichen auf die Preisentwicklung für Mineralölerzeugnisse zurückzuführen. Im Mai 2004 lagen die Preise für leichtes Heizöl um 19,2% und für Kraftstoffe um 11,6% (darunter Normalbenzin: + 11,7%; Superbenzin: + 11,4%; Diesel: + 11,8%) höher als im Jahr zuvor. Im Vergleich zum Vormonat kostete leichtes Heizöl 6,9% mehr und die Kraftstoffpreise stiegen um 5,2% an.

Wie in den Vormonaten wirkten sich auch im Mai 2004 die Maßnahmen der Gesundheitsreform und die Erhöhung der

Tabaksteuer im Verbraucherpreisindex aus. Im Vergleich zum Vorjahr stiegen die Preise für Gesundheitspflege um 19,5% (zum Vormonat: + 0,1%). Die Preise für Tabakwaren legten binnen Jahresfrist um 12,2% zu, blieben gegenüber dem Vormonat aber stabil.

Schaubild 2



Ohne die Sondereinflüsse bei den Preisen für Mineralölprodukte, Tabakwaren und Produkte der Gesundheitspflege erhöhten sich die Verbraucherpreise nur moderat um 0,5% (siehe dazu den Themenkasten in diesem Aufsatz).

Günstig haben sich aus Sicht der Verbraucher im Vergleich zum Vorjahr die Preise für Nahrungsmittel entwickelt (- 0,2%). Billiger wurden insbesondere Obst (- 1,4%) und Gemüse (- 0,5%); darunter Lauch: - 25,7%; Kopfsalat: - 18,8%). Deutlich billiger als im Vorjahr ist weiterhin Bohnenkaffee (- 6,0%). Im Vergleich zum April 2004 stiegen die Preise für Nahrungsmittel dagegen wieder leicht an (+ 0,3%), insbesondere Obst (+ 2,1%; darunter Weintrauben: + 8,9%) und Gemüse (+ 0,8%) wurden teurer.

Gemildert wurde der Indexanstieg durch die Preisentwicklung für Pauschalreisen. Diese verbilligten sich binnen Jahresfrist um 1,1% und gegenüber dem Vormonat um 1,4%.

Weiterhin deutlich geringere Preise im Jahresvergleich wiesen Informationsverarbeitungsgeräte (- 14,3%), Telefon-

**Allgemeine Preisentwicklung nach wie vor stabil**

Die Inflationsrate ist seit Jahresbeginn 2004 aufgrund der Wirkung von administrativen Maßnahmen sowie durch die Ölpreisentwicklung relativ stark angestiegen. Rechnet man jedoch diese Einflüsse aus dem Verbraucherpreisindex heraus, so verbleibt im Mai 2004 eine Preisänderungsrate von nur 0,5 % gegenüber dem Vorjahr. Über die letzten drei Monate hinweg war diese verbleibende Inflationsrate stabil.

In der Tabelle ist jeweils der Inflationsbeitrag der Indexpositionen Tabakwaren, Gesundheitspflege und Heizöle/Kraftstoffe dargestellt. Angegeben ist jeweils der Beitrag dieser Positionen zur Veränderungsrate des Verbraucherindex gegenüber dem Vorjahr in Prozentpunkten.

**Inflationsbeitrag verschiedener Indexpositionen 2004**  
Anteil an der Vorjahresrate des Verbraucherpreisindex in Prozentpunkten, Verbraucherpreisänderung in Prozent

Inflationsbeitrag in Prozentpunkten	Januar	Februar	März	April	Mai
Tabakwaren .....	0,0	0,0	0,3	0,3	0,3
Gesundheitspflege .....	0,6	0,6	0,6	0,7	0,7
Heizöle, Kraftstoffe .....	-0,2	-0,4	-0,3	0,1	0,5
Verbleibende Inflation .....	0,8	0,7	0,5	0,5	0,5
Verbraucherpreisänderung insgesamt in Prozent .....	1,2	0,9	1,1	1,6	2,0

Die Umsetzung der ersten Stufe der Gesundheitsreform für gesetzlich Krankenversicherte führt seit Beginn des Jahres zu einer um 0,6 Prozentpunkte höheren Preisänderung gegenüber dem Vorjahr. Preistreibend waren hier Maßnahmen wie die Praxisgebühr für Arztbesuche, Wegfall der Zuzahlungen für Brillengläser, Änderungen der Zuzahlungen für Arzneimittel und Krankenhausaufenthalte usw. Ein weiterer Schritt der Gesundheitsreform zum 1. April 2004 – die Regelung der Kostenübernahme für nicht verschreibungspflichtige Medikamente – wirkte sich ebenfalls leicht preis-

treibend aus, sodass der Inflationsbeitrag der Indexpositionen zur Gesundheitspflege auf 0,7 Prozentpunkte gestiegen ist.

Seit der Tabaksteuererhöhung im März 2004 kommt eine Preiserhöhung der Tabakwaren hinzu, die einen Inflationsanteil von 0,3 Prozentpunkten umfasst. Die Tabaksteuererhöhung selbst machte rein rechnerisch 0,2 Prozentpunkte aus. Die verbleibenden 0,1 Prozentpunkte des Inflationsbeitrags der Tabakwaren gehen auf zeitgleiche Preiserhöhungen der Tabakbranche zurück.

Zusammen mit den Maßnahmen im Bereich der Gesundheitspflege ist demnach die Inflationsrate im Mai 2004 etwa zur Hälfte durch administrative Neuregelungen verursacht. Diese Preiserhöhungen wirken sich jeweils ein Jahr lang auf die Inflationsrate aus: Ab dem Zeitpunkt der Maßnahme verschiebt sich das Preisniveau nach oben und die Preise werden dann in den folgenden 12 Monaten mit dem niedrigeren Preisniveau des Vorjahresmonats verglichen, welches vor der Maßnahme gegeben war.

Ein weiterer Grund für das Anziehen der Jahresteuersrate ist seit April 2004 die Entwicklung des Ölpreises. Anfang letzten Jahres war kurz vor Kriegsbeginn im Irak ein sehr hoher Ölpreis gegeben, der im März 2003 dann plötzlich abfiel. Die dann wieder folgenden Ölpreiserhöhungen konnten in Deutschland noch lange durch die Euro-Aufwertung abgefedert werden, sodass sich Anfang dieses Jahres noch immer rückläufige Preise für Heizöl und Kraftstoffe gegenüber dem Höchststand im Februar 2003 zeigten. Durch die konjunkturelle Belebung der Weltwirtschaft und die politischen Unsicherheiten im Nahen Osten hat sich das Bild jetzt jedoch wieder umgekehrt. Seit April 2004 übertrifft das Ausmaß der Ölpreissteigerungen den preisdämpfenden Wechselkurseffekt, sodass jetzt eine preistreibende Wirkung auf den Gesamtindex gegeben ist. Im Mai 2004 war ein Viertel der gesamten Inflationsrate auf den Inflationsbeitrag der Indexpositionen Heizöl und Kraftstoffe zurückzuführen.

und Telefaxgeräte (- 11,6%) sowie Foto- und Filmausrüstungen (- 7,3%) auf.

**Verbraucherpreisindex für Deutschland**  
auf Basis 2000 = 100

	Veränderungen Mai 2004 gegenüber	
	April 2004	Mai 2003
	%	
Gesamtindex .....	+ 0,2	+ 2,0
Nahrungsmittel und alkoholfreie Getränke ....	+ 0,2	- 0,2
Alkoholische Getränke, Tabakwaren .....	+ 0,1	+ 7,5
Bekleidung und Schuhe .....	- 0,4	- 0,5
Wohnung, Wasser, Strom, Gas usw. ....	+ 0,2	+ 1,5
Einrichtungsgegenstände, Apparate, Geräte und Ausrüstungen für den Haushalt u. Ä. ....	- 0,1	- 0,5
Gesundheitspflege .....	+ 0,1	+ 19,5
Verkehr .....	+ 1,2	+ 4,2
Nachrichtenübermittlung .....	- 0,3	- 0,6
Freizeit, Unterhaltung und Kultur .....	- 0,3	- 0,8
Bildungswesen .....	-	+ 2,6
Beherbergungs- und Gaststättendienstleistungen .....	+ 0,3	+ 1,0
Andere Waren und Dienstleistungen .....	+ 0,1	+ 1,5

Der für europäische Zwecke berechnete harmonisierte Verbraucherpreisindex für Deutschland hat sich im Mai 2004 gegenüber Mai 2003 um 2,1% erhöht. Im März hatte die Jahresveränderungsrate bei 1,1%, im April bei 1,7% gelegen. Im Vergleich zum Vormonat stieg der Index um 0,2%. [u](#)

# ÜBERSICHT

## über die im laufenden Jahr erschienenen Textbeiträge

	Heft	Seite
<b>Allgemeines, Methoden, Klassifikationen</b>		
Revision von Wirtschaftsklassifikationen bis 2007 – ein Zwischenbericht .....	4	381
Twinning-Projekt mit dem polnischen Statistikamt: ein Beispiel für bilaterale Verwaltungspartnerschaft .	2	123
Internationale Harmonisierung von Statistiken – Grundlagen und Beispiele aus dem Bereich der Haushaltsstatistiken .....	1	32
Selbstbewertungen im Qualitätsmanagement .....	2	131
Studie über die „Bedeutung der Belastung der Wirtschaft durch amtliche Statistiken“ läuft an .....	1	43
Maße für die faktische Anonymität von Mikrodaten .....	6	621
Geheimhaltungsmethoden auf dem Prüfstand – eine Analyse anhand der Umsatzsteuerstatistik .....	6	639
Schätzungen von europäischen Struktur- und Beschäftigungsindikatoren .....	4	404
Erwerbstätigkeit in Deutschland .....	2	139
Arbeitsmarkt in Deutschland – erster Baustein der neuen ILO-Statistik .....	2	148
Arbeitsmarkt in Deutschland – Analysen zu Ausfallprozessen in der Piloterhebung zum ILO-Erwerbsstatus .....	6	649
Produktbegleitende Dienstleistungen und ihre statistische Erfassung .....	3	269
Novellierung des Verkehrsstatistikgesetzes .....	3	259
Vom Preisindex für die Lebenshaltung zum Verbraucherpreisindex .....	1	11
Index der Erzeugerpreise gewerblicher Produkte (Inlandsabsatz) auf Basis 2000 .....	1	84
Die Indizes der Außenhandelspreise auf Basis 2000 .....	5	511
Zur Anwendung hedonischer Methoden beim Häuserpreisindex .....	5	525
Hedonische Preismessung bei EDV-Investmentsgütern .....	6	682
Sozialrechnungsmatrix für Deutschland .....	6	605
<b>Bevölkerung</b>		
Wanderungen zwischen Deutschland und den mittel- und osteuropäischen EU-Beitrittsländern .....	3	280
Ergebnisse des Mikrozensus 2003 .....	6	663

	Heft	Seite
<b>Wahlen</b>		
Europawahl 2004 .....	4	391
<b>Erwerbstätigkeit</b>		
Schätzungen von europäischen Struktur- und Beschäftigungsindikatoren .....	4	404
Erwerbstätigkeit in Deutschland .....	2	139
Arbeitsmarkt in Deutschland – erster Baustein der neuen ILO-Statistik .....	2	148
Arbeitsmarkt in Deutschland – Analysen zu Ausfallprozessen in der Piloterhebung zum ILO-Erwerbsstatus .....	6	649
Ergebnisse des Mikrozensus 2003 .....	6	663
<b>Unternehmen und Arbeitsstätten</b>		
Konzept und Ergebnisse der §7-Erhebung zum Unternehmensbegriff in der Statistik .....	2	155
Informationstechnologie in Unternehmen .....	5	530
Ergebnisse der Strukturhebung im Dienstleistungsbereich 2001 .....	2	163
Produktbegleitende Dienstleistungen und ihre statistische Erfassung .....	3	269
Insolvenzen in Deutschland 2003 .....	4	410
<b>Land- und Forstwirtschaft, Fischerei</b>		
Landwirtschaft in Deutschland 2003 .....	2	173
Weinbau und Weinstatistiken in Deutschland .....	3	288
<b>Produzierendes Gewerbe</b>		
Erneuerbare Energien 1991 bis 2003 .....	5	538
<b>Binnenhandel, Gastgewerbe, Tourismus</b>		
Online-Erhebungen im Binnenhandel und Gastgewerbe .....	1	53
Entwicklung im Großhandel im Jahr 2003 .....	3	302
Entwicklung im Einzelhandel im Jahr 2003 .....	3	308
Entwicklung im Gastgewerbe im Jahr 2003 .....	3	315
Inlandstourismus 2003: Stabilisierung trotz ungünstiger Rahmenbedingungen .....	4	422
<b>Außenhandel</b>		
Stand der Automatisierung von Intrastat in Europa .....	1	57
Außenhandel mit der Europäischen Union .....	3	322
Außenhandel mit den GUS-Staaten .....	2	184
Außenhandel 2003 nach Ländern .....	4	433
<b>Verkehr</b>		
Novellierung des Verkehrsstatistikgesetzes .....	3	259
Kombinierter Verkehr 2002 .....	1	63
Eisenbahnverkehr 2003 .....	5	543
Unternehmen der Binnenschifffahrt 2002 .....	6	673
Gewerblicher Luftverkehr 2003 .....	4	442
Unfallgeschehen im Straßenverkehr 2003 .....	5	548
<b>Gesundheitswesen</b>		
Gesundheitsausgaben und Gesundheitspersonal 2002 .....	4	449
Behinderung und Gesundheit .....	4	462
<b>Sozialleistungen</b>		
Behinderung und Gesundheit .....	4	462
Ergebnisse der Sozialhilfe- und Asylbewerberleistungsstatistik 2002 .....	2	192
Leistungen der Jugendhilfe für junge Volljährige .....	4	468

	Heft	Seite
<b>Finanzen und Steuern</b>		
Personalausgaben der Gebietskörperschaften .....	3	331
Versorgungsempfänger des öffentlichen Dienstes am 1. Januar 2003 .....	1	74
Kommunal финанzen 2004 .....	5	560
Öffentliche Finanzen 2003 .....	4	477
<b>Wirtschaftsrechnungen</b>		
Ausstattung mit Gebrauchsgütern und Wohnsituation privater Haushalte .....	2	209
<b>Preise</b>		
Vom Preisindex für die Lebenshaltung zum Verbraucherpreisindex .....	1	11
Index der Erzeugerpreise gewerblicher Produkte (Inlandsabsatz) auf Basis 2000 .....	1	84
Die Indizes der Außenhandelspreise auf Basis 2000 .....	5	511
Zur Anwendung hedonischer Methoden beim Häuserpreisindex .....	5	525
Hedonische Preismessung bei EDV-Investitionsgütern .....	6	682
Preisentwicklung im Jahr 2003 .....	1	94
Preise im Januar 2004 .....	2	228
Preise im Februar 2004 .....	3	341
Preise im März 2004 .....	4	483
Preise im April 2004 .....	5	575
Preise im Mai 2004 .....	6	690
<b>Volkswirtschaftliche Gesamtrechnungen</b>		
Sozialrechnungsmatrix für Deutschland .....	6	605
Bruttoinlandsprodukt 2003 .....	1	19
<b>Umwelt</b>		
Wasser- und Abwassersituation in den deutschen Flussgebieten 2001/2002 .....	5	563
<b>Gastbeiträge</b>		
10 Vorschläge zur Weiterentwicklung der statistischen Informationsinfrastruktur aus Sicht der deutschen Industrie .....	1	104
Anpassungskosten: eine Ursache für Überstunden .....	2	234
Wanderungen zwischen Deutschland und den mittel- und osteuropäischen EU-Beitrittsländern .....	3	280
Soziale und kulturelle Grundlagen der Kriminologie – der Beitrag der Kriminalstatistik .....	3	346
Schwache Betriebe – starker Markt .....	4	489
Kapitalkosten und Kapitalstrukturen ausgewählter deutscher Unternehmen – eine empirische Untersuchung .....	5	581
Arbeitsmarkt in Deutschland – Analysen zu Ausfallprozessen in der Piloterhebung zum ILO-Erwerbsstatus .....	6	649



Neuerscheinungen<sup>1)</sup> vom 23. Mai 2004 bis 26. Juni 2004

<b>● Zusammenfassende Veröffentlichungen</b>		EUR [D]	Fachserie 16: Löhne und Gehälter	EUR [D]
Wirtschaft und Statistik, Mai 2004	13,75		Reihe 2.1 Arbeiterverdienste im Produzierenden Gewerbe, Januar 2004	14,50
Ausgewählte Zahlen für die Bauwirtschaft, Februar 2004	18,70		Reihe 2.2 Angestelltenverdienste im Produzierenden Gewerbe; Handel; Kredit- und Versicherungsgewerbe, Januar 2004	17,-
In the Spotlight: Energy and Environment	10,-			
Sozio-ökonomisches Berichtssystem für eine nachhaltige Gesellschaft			<b>Fachserie 17: Preise</b>	
Band 4: Sozio-ökonomische Input-Output-Rechnung 1998	18,80		Reihe 1 Preisindizes für die Land- und Forstwirtschaft, November 2003	4,47
Band 5: Analyse von Lebenszyklen – Ergebnisse des 4. und 5. Weimarer Kolloquiums	14,80		Reihe 1 Dezember 2003	4,47
Statistik und Wissenschaft, Band 1: MIKAS – Mikroanalysen und amtliche Statistik	24,80		Reihe 2 Preise und Preisindizes für gewerbliche Produkte (Erzeugerpreise), April 2004	7,45
			Reihe 7 Verbraucherpreisindizes für Deutschland, April 2004 (Eilbericht)	3,-
			Reihe 7 April 2004	10,79
			Reihe 7 Mai 2004 (Eilbericht)	3,-
<b>● Fachserien</b>			<b>Fachserie 18: Volkswirtschaftliche Gesamtrechnungen</b>	
Fachserie 1: Bevölkerung und Erwerbstätigkeit			Reihe 3 Vierteljahresergebnisse der Inlandsproduktberechnung, 1. Vierteljahr 2004	8,80
Heft 2 Wahl der Abgeordneten des Europäischen Parlaments aus der Bundesrepublik Deutschland am 13. Juni 2004: Vorläufige Ergebnisse nach kreisfreien Städten und Landkreisen	11,20			
Fachserie 2: Unternehmen und Arbeitsstätten			<b>● Elektronische Veröffentlichungen</b>	
Reihe 4.1 Insolvenzverfahren, Dezember und Jahr 2003	5,65		Todesursachenstatistik 2002	75,-
Fachserie 4: Produzierendes Gewerbe			Außenhandel nach Waren und Ländern, Februar 2004	25,-
Reihe 4.1.1 Beschäftigung und Umsatz der Betriebe des Verarbeitenden Gewerbes sowie der Gewinnung von Steinen und Erden, Januar 2004	14,-		März 2004	25,-
Reihe 4.3 Kostenstruktur der Unternehmen des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden 2002	18,70		<b>● Bund/Länder-Veröffentlichungen</b>	
Reihe 8.1 Eisen und Stahl (Eisenerzbergbau, eisenschaffende Industrie, Eisen-, Stahl- und Tempergießerei), 4. Vierteljahr 2003	12,49		Kulturfinanzbericht 2003	15,-
Reihe 8.2 Düngemittelversorgung, 1. Vierteljahr 2004	2,67			
Fachserie 7: Außenhandel				
Reihe 1 Zusammenfassende Übersichten für den Außenhandel, 2003 (Vorläufige Ergebnisse)	14,-			
Reihe 1 März 2004	10,20			
Fachserie 14: Finanzen und Steuern				
Reihe 2 Vierteljährliche Kassenergebnisse der öffentlichen Haushalte, 1. bis 4. Vierteljahr 2003	10,79			

**Neueröffnung des Statistik-Shops des Statistischen Bundesamtes**

Ab sofort steht nahezu das gesamte Angebot an Standardveröffentlichungen des Statistischen Bundesamtes im neu organisierten Statistik-Shop online zur Verfügung oder kann online bestellt werden:

[www.destatis.de/shop](http://www.destatis.de/shop)

Gleichzeitig werden alle aktuellen **Fachserien**, deren Neuerscheinungen bislang an dieser Stelle aufgeführt waren, in elektronischer Form als PDF- oder Excel-Dateien zum **kostenfreien** Download im Statistik-Shop bereitgestellt.

**Veröffentlichungskalender für Pressemitteilungen**

Das Statistische Bundesamt gibt die Veröffentlichungstermine wichtiger wirtschaftsstatistischer Pressemitteilungen in einem Jahresveröffentlichungskalender, der wöchentlich präzisiert wird, bekannt.

Der Kalender kann unter der Internetadresse <http://www.destatis.de/presse/deutsch/cal.htm> abgerufen werden.

<sup>1)</sup> Zu beziehen durch den Buchhandel oder über den Vertriebspartner: SFG Servicecenter Fachverlage, Part of the Elsevier Group, Postfach 4343, 72774 Reutlingen, Telefon + 49 (0) 7071/93 53 50, Telefax + 49 (0) 7071/93 53 35, E-Mail: [destatis@s-f-g.com](mailto:destatis@s-f-g.com). Preise verstehen sich ausschließlich Versandkosten.