

METHODEN – VERFAHREN – ENTWICKLUNGEN

Nachrichten aus dem Statistischen Bundesamt

Ausgabe 2/2003

Das Stichwort

Arbeitsmarkt in Deutschland - Eine telefonische Stichprobenerhebung zur
Arbeitsmarktsituation in Deutschland 3

Methoden der Bundesstatistik – Weiterentwicklung

Die Erweiterung der abgekürzten Sterbetafel in den Altersstufen von 90 bis 100 Jahren..... 6

Entwicklung eines leistungsfähigen Record Linkage Verfahrens 8

Entwurf eines Gesetzes zur Neuordnung der Finanz- und Personalstatistiken..... 9

Standardisierung von Erhebungsunterlagen der amtlichen Statistik 12

BV4.1 - eine neue benutzerfreundliche Software zur Komponentenerlegung
ökonomischer Zeitreihen 14

Veranstaltungen

Wissenschaftliche Tagung „Online-Erhebungen“ 16

12. Wissenschaftliches Kolloquium „Ökonomische Leistungsfähigkeit Deutschlands -
Bestandsaufnahme und statistische Messung im internationalen Vergleich“
am 20. und 21. November 2003 17

Veranstaltungskalender

**ZUMA-Workshop: „Der Mikrozensus als Datenquelle für die empirische Sozial- und
Wirtschaftsforschung: Eine Einführung“ am 17. Februar 2004 18**

**European Conference on Quality and Methodology in Official Statistics (Q2004)
vom 24.-26. Mai 2004 in Mainz..... 19**

Herausgeber: Statistisches Bundesamt, Wiesbaden

Fachliche Informationen

zu dieser Veröffentlichung:

Gruppe IB,
Tel.: +49 (0) 611 / 75 20 77
Fax: +49 (0) 611 / 75 39 50
christian.koenig@destatis.de

Allgemeine Informationen

zum Datenangebot:

Informationsservice,
Tel.: +49 (0) 611 / 75 24 05
Fax: +49 (0) 611 / 75 33 30
info@destatis.de
www.destatis.de

**Veröffentlichungskalender
der Pressestelle:**

www.destatis.de/presse/deutsch/cal.htm

Erscheinungsfolge: (in der Regel) halbjährlich

Erschienen im Dezember 2003

© Statistisches Bundesamt, Wiesbaden 2003

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.

Das Stichwort

Arbeitsmarkt in Deutschland - Eine telefonische Stichprobenerhebung zur Arbeitsmarktsituation in Deutschland

Hintergrund

Im Zuge der zunehmenden Europäisierung und Globalisierung der Wirtschaftsbeziehungen stellen sich für die Arbeitsmarktberichterstattung in Deutschland neue Anforderungen. Neben der bisher dominierenden ausschließlich auf den nationalen Kontext ausgerichteten und auf nationalen sozialrechtlichen Regelungen basierenden Berichterstattung tritt mehr und mehr die international vergleichende Sichtweise in den Vordergrund. Dieser Trend zu einem grenzüberschreitenden Blickwinkel wird durch Benchmarking-Prozesse in der Europäischen Union weiter verstärkt. Als jüngstes Beispiel kann hier der Bericht der „Task Force Beschäftigung“ der Europäischen Union unter Vorsitz von Wim Kok vom November 2003 genannt werden (vgl. http://europa.eu.int/comm/employment_social/employment_strategy/pdf/etf_de.pdf).

Aber auch im Hinblick auf die Aktualität der Daten steigen die Anforderungen. Für eine effiziente Arbeitsmarktpolitik sowie kurzfristige Analysen, Vergleiche und sozialpolitische Programmevaluierungen sind Jahresdaten wenig geeignet. Vielmehr sind Monatsdaten zur Arbeitsmarktentwicklung, die zeitnah zur Verfügung stehen, unabdingbar. Das Zeitfenster zwischen Datenerhebung und Datenbereitstellung wird kleiner.

Labour-Force-Konzept

Eine länderübergreifende Vergleichbarkeit von Arbeitsmarktindikatoren wie die Zahl der Erwerbstätigen, Erwerbslosen und Nichterwerbspersonen erfordert ein standardisiertes und international akzeptiertes Messkonzept, da sonst statistische Artefakte durch unterschiedliche Messverfahren nicht ausgeschlossen werden können.

Die Internationale Arbeitsorganisation (International Labour Organization ILO) in Genf hat auf der 13. Internationalen Konferenz der Arbeitsstatistiker (1982) mit dem Labour-Force-Konzept ein solches standardisiertes Instrumentarium beschlossen. Das Labour-Force-Konzept der ILO wird als internationaler Standard für eine Vielzahl von Bevölkerungsumfragen zur Erwerbsbeteiligung herangezogen. Es hat zum Ziel, alle Erwerbstätigengruppen zu erfassen, unabhängig davon, welchen Stellenwert die Erwerbstätigkeit für den Lebensunterhalt der Person einnimmt. Das Messkonzept unterscheidet zwischen Erwerbstätigen, Erwerbslosen und Nichterwerbspersonen. Nach dem Labour-Force-Konzept kann jeder Person ein eindeutiger Erwerbsstatus zugeschrieben werden (vgl. http://www.destatis.de/download/d/dt_erheb/ilo_res82.pdf). Hier unterscheidet sich das Labour-Force-Konzept von dem in der amtliche Statistik oftmals benutzten "Unterhaltskonzept", das über die Quelle des überwiegenden Lebensunterhalts operationalisiert wird.

„Erwerbstätige“ sind nach dem Labour-Force-Konzept alle Personen ab 15 Jahren, die während eines festgelegten Berichtszeitraums, unabhängig vom zeitlichen Umfang, in einem Beschäftigungsverhältnis stehen, selbständig sind oder als mithelfende Familienangehörige arbeiten. Entscheidend ist, dass sie im Berichtszeitraum mindestens eine Stunde eine auf Erwerb ausgerichtete Tätigkeit ausgeübt haben.

„Erwerbslose“ sind Personen ab 15 Jahren, die im Berichtszeitraum ohne ein auf Erwerb gerichtetes Beschäftigungsverhältnis bzw. nicht selbständig waren, gegenwärtig bzw. innerhalb der nächsten zwei Wochen für eine Beschäftigung verfügbar sind und während der vergangenen vier Wochen aktiv auf der Suche nach einer Beschäftigung waren.

Als „Nichterwerbspersonen“ werden diejenigen Personen bezeichnet, die in der Berichtswoche weder erwerbstätig noch erwerbslos waren.

Piloterhebung nach § 7 Abs. 2 Bundesstatistikgesetz

Die vom Statistischen Bundesamt in Zusammenarbeit mit Infratest Sozialforschung, München, durchgeführte Erhebung „Arbeitsmarkt in Deutschland“ hat das Ziel, die aktuelle Situation und die Entwicklung des deutschen Arbeitsmarktes auf der Basis des Labour-Force-Konzepts der ILO abzubilden und zukünftig eine monatliche international vergleichbare Arbeitsmarktberichterstattung im Kontext der amtlichen Statistik bereitzustellen. Seit April 2003 läuft diese Piloterhebung nach § 7 Abs. 2 Bundesstatistikgesetz (BStatG) mit einem Stichprobenumfang von 10 000 Befragten pro Monat.

Die Erhebung ist als rotierendes Panel konzipiert. Jede Person soll insgesamt sechsmal befragt werden. Dies bedeutet, dass Monat für Monat ein Sechstel der Befragten aus der Erhebung ausscheidet. Um Veränderungsmessungen zwischen den Erhebungsmonaten mit einer für die amtliche Statistik üblichen Qualität durchführen zu können, wird eine monatliche Nettoüberlappung von 70 Prozent angestrebt.

Computer Assisted Telephone Interview (CATI)

Für eine hochaktuelle monatliche Datenerhebung sowie eine zeitnahe Bereitstellung der Ergebnisse sind computerunterstützte telefonische Interviews (Computer Assisted Telephone Interview, CATI-Methode) prädestiniert. In der deutschen amtlichen Statistik gibt es bisher keine systematisch durchgeführten telefonischen Befragungen. Die Piloterhebung leistet damit auch einen bedeutenden Beitrag zur Methodenentwicklung in der amtlichen Statistik.

Vor allem in der angewandten empirischen Sozialforschung hat die CATI-Methode in den vergangenen Jahren einen einzigartigen Bedeutungszuwachs erfahren. Wurde dieses Verfahren bis in die 1980er Jahre hinein in der sozialwissenschaftlichen Fachliteratur noch als qualitativ minderwertige „Quick-and-Dirty-Methode“ bezeichnet, ist es heute zu einem sozialwissenschaftlichen Standardverfahren geworden (vgl. Gabler, Siefried und Sabine Häder (Hrsg.): Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland. Münster 2002). Der starke Bedeutungsgewinn des telefonischen Interviews für die empirische Sozialforschung ist auf eine Reihe von spezifischen Vorteilen dieser Datenerhebungsmethode zurückzuführen. So sind bei verhältnismäßig geringen Kosten viele Kontaktversuche möglich, wodurch der Anteil der nicht erreichten Personen vermindert werden kann. Durch die computerunterstützte Interviewführung ist es möglich, vergleichsweise komplizierte Filterführungen zu durchlaufen. Die Fehlerhäufigkeit durch falsche Filterführungen kann merklich reduziert werden. Da die Antworten der Befragungspersonen während des Interviews direkt in einen Computer eingegeben werden, können zum einen schon während des Interviews Plausibilitätstest durchgeführt werden. Außerdem wird der Rohdatensatz sofort erstellt, was zu einer deutlichen Verringerung der Bearbeitungszeiten führt.

Gabler-Häder-Verfahren

Für die Ziehung von Telefonstichproben ist es erforderlich, dass die Population, aus der die Stichprobe gezogen werden soll, möglichst weitgehend über einen Festnetzanschluss verfügt. Dieses Kriterium kann in Deutschland bei einem Ausstattungsgrad der privaten Haushalte mit Festnetztelefonen von 95 Prozent als erfüllt betrachtet werden (vgl. Statistisches Bundesamt: Ausstattung und Wohnsituation privater Haushalte. Einkommens- und Verbrauchsstichprobe 2003. Wiesbaden 2003, S. 14).

Allerdings hat im Gefolge der Aufhebung der Eintragungspflicht in amtliche Telefonbücher im Jahr 1992 die Zahl der nicht eingetragenen Telefonnummern deutlich zugenommen. Mit dem von Siegfried Gabler und Sabine Häder am Zentrum für Umfragen, Methoden und Analysen (ZUMA) in Mannheim entwickelten „Gabler-Häder-Verfahren“ ist es aber möglich, dieses Problem zu lösen.

Dieses Stichprobenverfahren wird deshalb auch bei der weitaus überwiegenden Zahl von Telefonbefragungen mit Erfolg eingesetzt. Die Grundidee des Gabler-Häder-Verfahrens besteht darin, innerhalb eines Ortsnetzes Intervalle theoretisch möglicher Telefonnummern zu bilden und diese anschließend mit Hilfe eines Telefonnummernverzeichnisses daraufhin zu überprüfen, ob sie mindestens eine gültige Telefonnummer enthalten. Die Intervalle, für die diese Prüfung positiv ausfällt, werden als Stichprobenraum verwendet, aus dem schließlich die Zufallsstichprobe gezogen wird (vgl. Häder, Sabine: Auswahlverfahren bei Telefonumfragen, ZUMA-Arbeitsbericht 1994/03, Mannheim 1994; Häder, Sabine: Telefonstichproben; ZUMA-How-to-Reihe 6, Mannheim 2000).

Stichprobenausschöpfung, Ausfallprozesse und erste Befunde

Die Stichprobenausschöpfung bei den Erstbefragten lag für die Monate April bis November 2003 bei durchschnittlich 44,6 Prozent und liegt damit im Rahmen dessen, was bei sozialwissenschaftlichen Befragungen ohne Auskunftspflicht in Deutschland üblich ist (vgl. Schnell, Rainer: Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen 1999). Bei den Wiederholungsbefragungen konnte im selben Zeitraum eine durchschnittliche Ausschöpfung von 86,7 Prozent erzielt werden. Angesichts der kurzen Feldzeit - die Monatsdaten liegen jeweils spätestens zum siebten Kalendertag des Folgemonats vor - ist dieses Ergebnis positiv zu werten. Die angestrebte Nettoüberlappung von 70 Prozent wurde bislang mit durchschnittlich 76,9 Prozent eingehalten.

Da es sich bei der telefonischen Befragung um eine freiwillige Erhebung handelt, sind Ausfälle (Nonresponse) und damit verbundene Stichprobenverzerrungen unvermeidlich. Vergleiche mit externen Referenzstatistiken belegen jedoch, dass die Verzerrungen dem entsprechen, was aus allgemeinen Bevölkerungsumfragen bekannt ist. Besonders Männer, jüngere und ältere Menschen, Ausländer und Personen mit niedrigem Bildungsniveau sind davon besonders betroffen. Die Analyse von Ausfallprozessen im Längsschnitt zeigen, dass die Stichprobe über diese allgemeinen demografischen Merkmale hinaus eine leichte Verzerrung hinsichtlich des Erwerbsstatus aufweist. So sind Personen, deren Erwerbsstatus sich im Laufe der Befragungen nicht ändert, mit einem etwas höheren Ausfallrisiko behaftet als Personen, die einen Statuswechsel durchlaufen. Die benannten Verzerrungen können jedoch durch das Hochrechnungs- und Gewichtungsverfahren zum Teil behoben werden.

Die bislang vorliegenden Ergebnisse für die Monate April bis November 2003 zeigen, dass die Telefonerhebung den „Rand des Arbeitsmarktes“ stärker abbilden kann, als dies durch bestehende amtliche Statistiken in Deutschland bisher der Fall war. So lassen sich messbedingte Verbesserungen feststellen, die dazu führen, dass geringfügige Beschäftigungen, Erwerbstätigkeiten von Schülern und Studenten und andere Erwerbstätigkeiten am Rand des Arbeitsmarktes sowie die Suche nach diesen Tätigkeiten gut abgebildet werden können.

Weitere Informationen unter: <http://www.destatis.de/arbeitsmarkt>

Thomas Riede, Tel.: +49 (0) 611 / 75 24 33, E-Mail: thomas.riede@destatis.de

Matthias Sacher, Tel.: +49 (0) 611 / 75 25 34, E-Mail: matthias.sacher@destatis.de

Methoden der Bundesstatistik – Weiterentwicklung

Die Erweiterung der abgekürzten Sterbetafel in den Altersstufen von 90 bis 100 Jahren

Die abgekürzte Sterbetafel des Statistischen Bundesamtes

Das Statistische Bundesamt berechnet zur Abbildung der Sterblichkeitsverhältnisse der Bevölkerung und der durchschnittlichen Lebenserwartung in regelmäßigen Abständen Periodensterbetafeln. Im Anschluss an Volkszählungen erstellt das Statistische Bundesamt *allgemeine Sterbetafeln*, die bis zu der Altersstufe von 100 Jahren veröffentlicht und über mathematisch-statistische Verfahren von Zufallsschwankungen und Kohorteneffekten bereinigt werden. In der Zeit zwischen den Volkszählungen berechnet das Statistische Bundesamt *abgekürzte Sterbetafeln*, die mit der Altersstufe von 90 Jahren geschlossen und nicht bereinigt werden. Ausgangspunkt der Berechnung der Sterbetafeln sind die altersspezifischen Sterbewahrscheinlichkeiten, die sich aus den Sterbefällen sowie der durchschnittlichen Bevölkerung in den einzelnen Altersjahren zusammensetzen (s. Sterbeziffernmethode nach Farr). Die geringen Fallzahlen in den hohen Altersstufen und die mit zeitlicher Entfernung zu der letzten Volkszählung zunehmenden Ungenauigkeiten in der Bevölkerungsfortschreibung führen allerdings dazu, dass die Sterbewahrscheinlichkeiten in den *abgekürzten Sterbetafeln* nur bis zu dem Alter von 89 Jahren veröffentlicht werden können. Oberhalb der Altersstufe von 90 Jahren beginnen die berechneten Sterbewahrscheinlichkeiten der Männer unplausibel zu sinken, bei den Frauen aufgrund der höheren Lebenserwartung etwas später. Mit dem Anstieg der Lebenserwartung ergibt sich jedoch aus wissenschaftlicher, wirtschaftlicher und politischer Sicht vermehrt ein Informationsbedarf zur Sterblichkeit und Lebenserwartung der Bevölkerung jenseits des Alters von 90 Jahren. Eine Volkszählung oder ein gleichwertiges Verfahren und damit die Möglichkeit, eine neue *allgemeine Sterbetafel* zu berechnen, ist bislang nicht absehbar. Deshalb werden im Folgenden zwei Verfahren beschrieben, die zusammen eine Ergänzung der Sterbewahrscheinlichkeiten und damit eine Erweiterung der *abgekürzten Sterbetafel* bis zu der Altersstufe von 100 Jahren erlauben.

Die Methode der Extinct Generations

Unter der Voraussetzung einer zuverlässigen Sterbefallregistrierung, die für alle Altersstufen vollständig und nach Alter und Geburtsjahr aufgegliedert ist, können die Lebenden eines Geburtsjahrgangs in den einzelnen Altersstufen über die Anzahl der Gestorbenen zurückgerechnet werden. Die Methode der *Extinct Generations* ist dabei auf die höheren Lebensalter beschränkt, da vorausgesetzt werden muss, dass keine Wanderungen mehr stattfinden. Die ursprünglich von Vincent entwickelte Methode der *Extinct Generations* kann nur bei bereits vollständig ausgestorbenen Geburtsjahrgängen angewendet werden, jedoch lässt sich diese Restriktion umgehen, indem ein als exakt angenommener Bevölkerungsstand einbezogen wird, wie beispielsweise aus der Volkszählung von 1987 (s. Dinkel 1997, S. 496 ff.). Die Sterbewahrscheinlichkeiten der verschiedenen Geburtsjahrgänge werden dann mit der Geburtsjahrmethode ermittelt und durch Umschichtung wieder von der Kohorten- in die Periodenreihung zurückgeführt. Die Methode der *Extinct Generations* kann allerdings nicht vollständig bis zu der Altersstufe von 100 Jahren plausible Sterbewahrscheinlichkeiten liefern, da insbesondere bei den Männern bedingt durch die geringen Fallzahlen vorzeitig Schwankungen auftreten. Eine Möglichkeit, diese Sterbewahrscheinlichkeiten zu ergänzen, ist deshalb die Extrapolation über ein Regressionsmodell.

Das nichtlineare Regressionsmodell

Die Voraussetzung für die Auswahl eines geeigneten Regressionsmodells ist die Kenntnis über den Verlauf der Sterblichkeit in den hohen Altersjahren. Der britische Aktuar Benjamin Gompertz entdeckte bereits 1825, dass nach dem sogenannten „Unfallhügel“ in der Jugend das Risiko zu

sterben eine Exponentialfunktion des Alters ist. Diese Beobachtung, die später auch als „*Law of Mortality*“ bezeichnet wurde, lässt sich allerdings nicht mehr auf die hohen Lebensalter anwenden. Es wurde festgestellt, dass das Risiko zu sterben nach dem Alter von ungefähr 80 Jahren nicht mehr exponentiell ansteigt, sondern sich mit der Abnahme der Zuwachsrates des Sterberisikos verlangsamt (s. Anderson 1999, S. 2 u.a.). Ein geeignetes Regressionsmodell zur Extrapolation der Sterbewahrscheinlichkeiten muss somit die Verlangsamung des Anstiegs des Sterberisikos in den hohen Lebensaltern berücksichtigen. Vor diesem Hintergrund wurden die Modelle von *Kannisto, Heligman & Pollard, Coale & Kisker*, der *modifizierte Gompertz-Ansatz* sowie weitere logistische Funktionen und Polynome höheren Grades (s. Thatcher u.a. 1998, Doray 2002, Statistisches Bundesamt 1991) mit der letzten *allgemeinen Sterbetafel* von 1986/88 und der *abgekürzten Sterbetafel* von 1998/00 getestet. Als zusätzliche Kriterien zur Beurteilung der Eignung der Modelle wurden eine gute Anpassung im Stützbereich (Bestimmtheitsmaß R-Quadrat größer/gleich 0,99), Stabilität bei Variation des Stützbereiches, einfache Handhabung (nicht mehr als drei Regressionskoeffizienten) und Plausibilität zu der letzten *allgemeinen Sterbetafel* 1986/88 festgelegt. Der Stützbereich für die Regression wurde bei der *abgekürzten Sterbetafel* im Altersbereich von 70 bis 89 Jahren variiert, während bei der *allgemeinen Sterbetafel* der Stützbereich von 76 bis 94 Jahren übernommen wurde. Problematisch sind hierbei die sich annähernden Sterbewahrscheinlichkeiten von Männern und Frauen, die in vielen Fällen zu einem sogenannten *Mortality-Crossover* führen, d.h. zu einer Überschneidung der Sterbewahrscheinlichkeiten.

Unter Berücksichtigung der beschriebenen Kriterien hat sich das Regressionsmodell (1) für die Extrapolation der Sterbewahrscheinlichkeiten als geeignet herausgestellt. Für die Auswahl waren insbesondere die Plausibilität zu der allgemeinen Sterbetafel 1986/88 und die Stabilität bei Variation des Stützbereiches ausschlaggebend. In den allgemeinen Sterbetafeln 1960/62, 1970/72 und 1986/88 wurde bisher der *modifizierte Gompertz-Ansatz* (2) verwendet.

$$(1) \quad q_x = 1 - e^{-\left(\sum_{i=0}^2 \beta_i x^i\right)} \quad (2) \quad \ln\left(\ln\left(\frac{1}{1 - q_x}\right)\right) = \sum_{i=0}^2 \beta_i x^i$$

q_x Sterbewahrscheinlichkeit im Alter x

β_i Regressionskoeffizient

Bei dem Regressionsmodell (1) handelt es sich um ein deskriptives Modell, das nur für die Extrapolation des Verlaufs der Sterblichkeit im Intervall von 90 bis 100 Jahren bestimmt ist. Im Gegensatz zu den kausalen Modellen, die auf einer Theorie der Sterblichkeit basieren, versucht das Regressionsmodell (1) nicht den gesamten Verlauf bzw. den Verlauf der Sterblichkeit nach der Jugendphase abzubilden. Das Regressionsmodell (1) wurde auch mit den *abgekürzten Sterbetafeln* für Deutschland seit 1991/93 sowie mit der *allgemeinen Sterbetafel* 1970/72 getestet und konnte auch hier plausible Schätzwerte liefern.

Fazit

Mit der Methode der *Extinct Generations* können die Sterbewahrscheinlichkeiten im Altersbereich von 90 bis 100 Jahren nahezu vollständig empirisch ermittelt werden. Das Regressionsmodell (1) ermöglicht die Schätzung der noch benötigten Sterbewahrscheinlichkeiten, so dass die abgekürzte Sterbetafel bis zu dem Alter von 100 Jahren veröffentlicht werden kann.

Literatur

Anderson, R. N. (1999): Method for Constructing Complete Annual U.S. Life Tables, National Vital Statistics Report Vol.50, Number 6 - United States Life Tables

Dinkel, H. R. (1997): Sterblichkeitsmessung in den obersten Altersstufen: Die Ergänzung abgekürzter bundesdeutscher Sterbetafeln durch die Methode der Extinct Generations, Zeitschrift für Bevölkerungswissenschaft, Jg. 22, Heft 4/1997: S. 493 - 509

Doray, L. G. (2002): Living to Age 100 in Canada in 2000, Society of Actuaries (SOA): Living to 100 and Beyond - Survival at Advanced Ages Symposium 17/01-18/01/02
http://www.soa.org/research/Doray_Final.pdf

Statistisches Bundesamt (1991): Fachserie 1, Bevölkerung und Erwerbstätigkeit, Allgemeine Sterbetafel für die Bundesrepublik Deutschland 1986/88, Wiesbaden

Thatcher, A. R., Kannisto, V., Vaupel, J. W. (1998): The Force of Mortality at Ages 80 to 120, Odense University Press, Denmark
<http://www.demogr.mpg.de/Papers/Books/Monograph5/ForMort.htm>

Matthias Eisenmenger, Tel.: +49 (0) 611 /75 32 75, E-Mail: matthias.eisenmenger@destatis.de

Entwicklung eines leistungsfähigen Record Linkage Verfahrens

Der Wunsch der Wissenschaft, die Einzeldaten der amtlichen Statistik in der empirischen Wirtschafts- und Sozialforschung zu nutzen, wird zunehmend bedeutender. Der vom Gesetzgeber vorgezeichnete Weg ist, der Wissenschaft Einzeldaten in faktisch anonymisierter Form zu erschließen. Daten gelten hiernach als ausreichend anonymisiert, wenn sie nur mit einem „unverhältnismäßig hohen Aufwand“ zugeordnet werden können.

Das Statistische Bundesamt ist an einem internationalen, von der EU-Kommission geförderten Methodenprojekt zur statistischen Geheimhaltung (Computational Aspects of Statistical Confidentiality – CASC) beteiligt, über welches in MVE bereits in den Ausgaben 1/2001 und 2/2002 berichtet wurde. Im CASC-Projekt werden verschiedene Verfahren zur statistischen Geheimhaltung weiterentwickelt und in das Programm-Paket ARGUS eingebunden. ARGUS enthält sowohl ein Programm zur Tabellengeheimhaltung (τ -ARGUS) als auch ein Programm zur Mikrodatengeheimhaltung (μ -ARGUS). Im Statistischen Bundesamt wurde im Rahmen dieser Forschungen bereits das von Sullivan/Fuller eingeführte Verfahren der Anonymisierung via Überlagerung mit Zufallsfehlern für μ -ARGUS weiterentwickelt.

Ein weiterer Beitrag des Statistischen Bundesamtes bestand in der Entwicklung eines Verfahrens zur Simulation des Massenfischzuges, welches ebenfalls in μ -ARGUS implementiert wird. Im Szenario des Massenfischzuges versucht ein Datenangreifer, möglichst viele Merkmalsträger einer externen Datenbank (Zusatzwissen) den Merkmalsträgern einer anonymisierten

Zieldatenbank eindeutig zuzuordnen, um die externe Datenbank um Informationen aus der Zieldatenbank zu erweitern.

Als theoretische Grundlage für das sogenannte Record Linkage Verfahren dienen Algorithmen der Graphentheorie und der mathematischen Optimierung, welche im April 2003 auf einem Workshop zur statistischen Geheimhaltung in Luxemburg durch das Statistische Bundesamt vorgestellt wurden. Das Verfahren wird im Folgenden skizziert.

In der ersten Phase des Record Linkage Verfahrens wird ein Distanzmaß über alle gemeinsamen Merkmale (Schlüsselvariablen) von Merkmalsträgern beider Datenbanken gebildet. In einem realen Angriffsszenario ziehen Datenangreifer einige ausgewählte Merkmale, in welchen sie geringere Abweichungen von den Originaldaten erwarten, anderen weniger zuverlässigen Merkmalen vor. Aus diesem Grunde wird dem Nutzer eine individuelle Gewichtung w_i einzelner Merkmale i – der Einfachheit halber werden die Gewichte auf das Intervall $[0,1]$ normiert – offengelassen. Einer Maximalgewichtung des Merkmals i , d.h. der Wahl des Gewichtes $w_i = 1$, kommt in diesem Kontext die sogenannte „Blockung von Variablen“ gleich. Es ist notwendig für die Zuordnung zweier Merkmalsträger, dass sie in ihren maximal gewichteten Merkmalen (Blockvariablen) exakt übereinstimmen. Bei einer geschickten Wahl von Blockvariablen wird der Rechenaufwand für die Gesamtzuordnung auf der einen und der Anteil der Falschzuordnungen auf der anderen Seite erheblich reduziert. Bei der Minimalgewichtung $w_i = 0$ wird das Merkmal i während des Zuordnungsverfahrens nicht berücksichtigt.

Ziel der zweiten Phase ist es, auf der Basis der ermittelten Distanzen Zuordnungen vorzunehmen. Hierzu wird die Summe der Distanzen aller zu bildenden Zuordnungen (Gesamtabweichung) minimiert. Zum Vergleich wurden neben klassischen Methoden aus der Graphen- und Netzwerktheorie auch in der Literatur bis dato vorgeschlagene, rechentechnisch weniger aufwändige Greedy-Algorithmen in Betracht gezogen, welche einfacher zu implementieren sind und weniger Rechenzeit erfordern, allerdings auf Kosten einer höheren Gesamtabweichung.

Die Ergebnisse der CASC-Forschungen fließen im Statistischen Bundesamt unmittelbar in das Projekt „Faktische Anonymisierung wirtschaftsstatistischer Einzeldaten“ ein, das die statistischen Ämter gemeinsam mit dem Institut für Angewandte Wirtschaftsforschung (IAW), Tübingen, durchführen. In diesem Projekt werden anhand verschiedener Erhebungen Verfahren der Anonymisierung zum einen auf den Erhalt des Analysepotenzials und zum anderen auf das Risiko der späteren Zuordnung (Reidentifikation) von Einzeldaten getestet.

Zur Untersuchung der Schutzwirkung wurde das oben beschriebene Record Linkage Verfahren bereits auf verschiedene Probeanonymisierungen der Kostenstrukturerhebung im Verarbeitenden Gewerbe angewendet.

Dr. Rainer Lenz, Tel.: +49 (0) 611 / 75 26 36, E-Mail: rainer.lenz@destatis.de

Roland Sturm, Tel.: +49 (0) 611 / 75 25 80, E-Mail: roland.sturm@destatis.de

Entwurf eines Gesetzes zur Neuordnung der Finanz- und Personalstatistiken

Wichtige neue statistische Anforderungen der Europäischen Union (EU) erfordern eine Anpassung des Gesetzes über die Statistiken der öffentlichen Finanzen und des Personals im öffentlichen Dienst (Finanz- und Personalstatistikgesetz – FPStatG vom 8. März 2000). Aufgrund der zunehmenden Ausgliederungen von Einrichtungen aus den öffentlichen Haushalten müssen darüber hinaus im FPStatG Vorkehrungen getroffen werden, die Finanzstatistiken als aussagefähige Basisstatistik für den detaillierten Datenbedarf der Volkswirtschaftlichen Gesamtrechnungen des Bundes und der Länder (Staatskonto, Verteilungsrechnung) zu sichern.

Die Finanzstatistiken gewinnen ihre Daten unmittelbar aus den Haushalts-, Kassen- und Rechnungsunterlagen der öffentlichen Verwaltungen, belasten diese damit geringst möglich und sichern zugleich eine hohe Qualität. Im Rahmen der Novellierung sind daher auch die Reformbestrebungen der öffentlichen Haushalte zu einem doppelten Rechnungswesen – insbesondere auf der kommunalen Ebene – zu berücksichtigen. Straffungen des Erhebungsprogramms sollen zeitgleich umgesetzt werden, um die Belastung aus den o.g. neuen Anforderungen wenigstens teilweise abzufangen.

1. Neue finanzstatistische Anforderungen der Europäischen Union

Der von der EU geforderten Berichterstattung über die finanziellen Transaktionen und Bestände (Finanzaktiva und Finanzpassiva) des Staates wird auf der Passivseite bereits bisher durch die vierteljährliche Erhebung der Schulden der öffentlichen Haushalte entsprochen. Gegenüber dem derzeit geltenden Stand muss die Novelle jedoch eine Erweiterung des Berichtskreises dieser Statistik um die Erhebungseinheiten vorsehen, die nach den Regeln der Verordnung (EG) Nr. 2223/96 des Rates vom 25. Juni 1996 zum Europäischen System Volkswirtschaftlicher Gesamtrechnungen auf nationaler und regionaler Ebene in der Europäischen Gemeinschaft (ESVG 95) dem Sektor Staat zuzurechnen sind (ausgegliederte öffentliche Einrichtungen, s.u.2.).

Ab dem Jahr 2005 neu statistisch zu erfassen sind jährlich die Aktivpositionen des öffentlichen Finanzvermögens. Die Neuregelung sieht bei Bund, Ländern, Gemeinden, Gemeindeverbänden und Zweckverbänden, sowie der Sozialversicherung und den staatlichen und kommunalen Fonds, Einrichtungen, Betrieben und Unternehmen die Erhebung des Standes der Finanzaktiva nach Arten zum Jahresende sowie der Veränderungen und Berichtigungen des Bestandes im Laufe des Jahres vor. Auf eine vierteljährliche Erhebung dieser Daten hat die EU in Deutschland verzichtet. Die benötigten Vierteljahresdaten werden auf der Basis der jährlichen Ergebnisse über das öffentliche Finanzvermögen geschätzt.

2. Aktuelle finanzstatistische Erfassung ausgegliederter öffentlicher Einrichtungen

Den starken Ausgliederungstendenzen von öffentlichen Einrichtungen aus den Kernhaushalten von Bund, Ländern und Gemeinden, die nach den Regeln der Verordnung (EG) Nr. 2223/96 des Rates vom 25. Juni 1996 zum Europäischen System Volkswirtschaftlicher Gesamtrechnungen auf nationaler und regionaler Ebene in der Europäischen Gemeinschaft (ESVG 95) dem Sektor Staat zuzurechnen sind, und den dadurch entstandenen Berichtslücken über die öffentlichen Finanzen, wird künftig durch die vierteljährliche Erfassung der Erträge und Aufwendungen der ausgegliederten Einheiten des Staatssektors Rechnung getragen. Dadurch wird die Grundlage für die Integration von Kernhaushalten und ausgegliederten Einheiten geschaffen und das notwendige statistische Gesamtbild über die öffentlichen Finanzen gesichert.

Auch für die Statistikanforderungen der EU über die vierteljährlichen Konten des Staates (Einnahmen und Ausgaben), die von den Volkswirtschaftlichen Gesamtrechnungen erstellt werden, ist diese Änderung sehr wichtig. Die vierteljährliche Kassenstatistik wird durch die Anpassung künftig den Anforderungen der Volkswirtschaftlichen Gesamtrechnungen – insbesondere zur Ermittlung des EU - Stabilitätskriteriums Staatsdefizit - in terminlicher, inhaltlicher und qualitativer Hinsicht voll gerecht. Zudem werden erstmals zeitnah Informationen über die Abgrenzung des Berichtskreises der ausgegliederten Einheiten, die zum Staatssektor gehören, zur Verfügung stehen.

3. Berücksichtigung der Reform des öffentlichen Rechnungswesens

Mit der Novellierung des Finanz- und Personalstatistikgesetzes wird ferner neuen Bestrebungen zur Anpassung des öffentlichen Rechnungswesens an gewandelte Voraussetzungen Rechnung getragen. Bisher erfolgt die finanzstatistische Erfassung der Einnahmen und Ausgaben der öffentlichen Haushalte ausschließlich in einer Gliederung, die dem gemeinsamen haushaltssystematischen Rahmen von Bund und Ländern einerseits sowie den Gemeinden und Gemeindeverbänden andererseits entspricht. Diese Bindung stellt sicher, dass die

finanzstatistischen Daten ohne Zusatzaufwand unmittelbar aus den öffentlichen Haushalts-, Kassen- und Rechnungsunterlagen entnommen werden können. Mit den jüngsten Bestrebungen zur Flexibilisierung des öffentlichen Haushaltswesens, insbesondere aber durch die Modellversuche zu einem neuen doppischen Rechnungswesen, wird die Bindungswirkung des bisherigen haushaltssystematischen Rahmens für die Gliederung der Einnahmen und Ausgaben zunehmend in Frage gestellt.

Für die Kommunen haben sich die Innenminister der Länder bereits auf ein neues einheitliches doppisches Rechnungssystem verständigt, das die Belange der Finanzstatistik und der Volkswirtschaftlichen Gesamtrechnungen berücksichtigt. Die vorteilhafte Bindung der finanzstatistischen Gliederung an den haushaltssystematischen Rahmen bleibt dadurch gesichert. Die neue Regelung im FPStatG lässt künftig auch die Erhebung der kommunalen Finanzen in der Gliederung nach dem Kontenrahmen und Produktrahmen des neuen doppischen Rechnungswesens zu. Auch viele Einrichtungen für Wissenschaft und Forschung haben ein kaufmännisches Rechnungswesen eingeführt, so dass der bisherige haushaltssystematische Bezugsrahmen erweitert werden muss, um künftig die Aufwendungen, Erträge und Investitionsausgaben in adäquater Form statistisch erheben zu können.

Für Zwecke der Kosten und Leistungsrechnung des Bundes und der Länder werden z. Zt. in den Fachgremien ebenfalls einheitliche Konten- und Produktrahmen beraten, die ggf. auch Grundlage für ein doppisches Haushaltswesen einzelner Länder sein können. Nach geltender Regelung des Haushaltsgrundsätzegesetzes (HGrG) sind Bund und Länder aber noch an den bisherigen haushaltssystematischen Rahmen (Gruppierungs- und Funktionenplan) gebunden. Bei Anwendung neuer doppischer Rechnungssysteme ist eine Umsetzung auf die Einnahme- und Ausgabekategorien des Gruppierungs- und Funktionenplan zu gewährleisten. Ob und wann sich ein einheitliches neues doppisches Rechnungswesen für die Haushalte der Länder und auch des Bundes herausbildet, ist z. zt. noch nicht absehbar.

4. Straffungen des bisherigen Erhebungsprogramms

Die Straffung von Teilen des Erhebungsprogramms soll zu Einsparungen und Entlastungen bei den Berichtspflichtigen führen und zumindest teilweise auch die Belastung der statistischen Ämter durch den o.g. Mehrbedarf kompensieren. Die Novelle sieht den Verzicht auf die Erhebungen zur Finanzplanungsstatistik, die Einführung einer flexiblen Abschneidegrenze bei der Gemeindegröße im Rahmen der kommunalen Haushaltsansatzstatistik sowie ein eingeschränktes Erhebungsprogramm insbesondere bei den rechtlich selbstständigen Organisationen für Wissenschaft, Forschung und Entwicklung im Rahmen der Personalstand- und Versorgungsempfängerstatistik vor. Darüber hinaus wird mit der Regelung zur elektronischen Datenübermittlung eine weitere Voraussetzung zur Entlastung der Auskunft gebenden Stellen geschaffen. Sie ermöglicht die Umsetzung der im Rahmen der E – Government - Initiative der statistischen Ämter angestrebten medienbruchfreien Integration der Produktionsprozesse von der Datenerhebung bis zur Informationsverbreitung.

5. Zeitplan der Novellierung

Die Novelle des Finanz- und Personalstatistikgesetzes soll zu Beginn des Jahres 2004 in die parlamentarische Beratung gehen und möglichst zum 30. Juni 2004 verabschiedet werden. Die Inkraftsetzung ist zum 1. Januar 2005 vorgesehen.

*Renate Schulze-Steikow, Tel.: +49 (0) 611 /75 41 66,
E-Mail: renaete.schulze-steikow@destatis.de*

Standardisierung von Erhebungsunterlagen der amtlichen Statistik

Rahmenbedingungen

Die steigende Nachfrage nach statistischen Ergebnissen stellt hohe Anforderungen an die Qualität der zugrunde liegenden Erhebungsunterlagen. Statistische Informationen werden seit jeher von den Nutzern nur akzeptiert, wenn sie Vertrauen in ihre Validität haben. Die Qualität statistischer Informationen hängt wiederum in erheblichem Umfang von der Qualität der eingesetzten Erhebungsunterlagen ab, denn nur wer „richtig“ fragt, kann auch aussagefähige Ergebnisse produzieren. Gut gestaltete Erhebungsunterlagen reduzieren die Belastung der Auskunftgebenden und wirken sich positiv auf deren Auskunftsbereitschaft aus. Dieser Effekt wird durch die Erhöhung des Wiedererkennungswertes von Erhebungsunterlagen der amtlichen Statistik verstärkt. Ein Weg in diese Richtung ist das einheitliche Erscheinungsbild der amtlichen Statistik nach außen.

Bund-Länder-Arbeitsgruppe „Standardisierung von Erhebungsunterlagen“

Die Erarbeitung von Vorschlägen zur Lösung der vorgenannten Anforderungen war das Ziel der Arbeitsgruppe „Standardisierung von Erhebungsunterlagen“, die sich aus Vertretern der Statistischen Ämter des Bundes und der Länder zusammensetzt. Die Arbeitsgruppe hat in den zwei Jahren das Handbuch für die Fragebogenentwicklung überarbeitet und ein System aus Fragebogenmodulen entwickelt. Damit wurden Lösungen erzielt, die eine ...

- einheitliche Gestaltung (Grundlayout) von Erhebungsunterlagen für Bundesstatistiken
 - einheitliche Darstellung der Fragebogeninhalte durch die Vereinbarung bindender Frageformulierungen und –reihenfolgen
 - einfache und schnelle Erstellung von Fragebogenentwürfen
 - bessere Qualität der Fragebogenentwürfe
 - Beleglesefähigkeit aller Fragebogen
 - Erhöhung der Validität (Verringerung von aufwändigen nachträglichen Korrekturen und kostenintensiven Rückfragen bei den Auskunftgebenden)
 - schnellere Bereitstellung der Ergebnisse
- sichern.

Handbuch zur Erstellung von Erhebungsunterlagen in der amtlichen Statistik

Das Handbuch wurde von der Bund-Länder-Arbeitsgruppe „Standardisierung von Erhebungsunterlagen“ überarbeitet und enthält Methoden und Standards für die Gestaltung von Erhebungsunterlagen durch die Statistischen Ämter des Bundes und der Länder. Es wurde im März 2003 von der Amtsleiterkonferenz in Kraft gesetzt. Neben Abschnitten mit grundlegenden Ausführungen zur Erstellung von Erhebungsunterlagen und deren Optimierung erhält das Handbuch auch Detailinformationen, z.B. Checklisten oder Anforderungen der Beleglesung an Erhebungspapiere. Einige Abschnitte des Handbuchs beschreiben bindende Vorgaben bei der Gestaltung von Erhebungsunterlagen, andere Abschnitte, haben den Charakter von Empfehlungen. Diese Aufteilung eröffnet genügend Gestaltungsspielräume, um der Verschiedenartigkeit amtlicher Erhebungen Rechnung zu tragen und landesspezifische Besonderheiten zu berücksichtigen. Zukünftig wird das Handbuch laufend an aktuelle Entwicklungen angepasst.

Module zur Erstellung von Fragebogenentwürfen

Für die Erstellung von standardisierten Fragebogen wurde ein System aus Fragebogenmodulen entwickelt. Die Nutzbarkeit des neuen Systems ist in allen Statistikbereichen möglich und unterstützt notwendige Optimierungen der Arbeitsvorgänge im Rahmen der Statistikprozesse, wie z.B. den verstärkten Ausbau der Belegung und den Faxversand von ausgefüllten Papierfragebogen. Immer knapper werdenden Ressourcen machen diese Maßnahmen notwendig, darüber hinaus helfen die Lösungen, Einsparungen durch Effizienzsteigerungen zu erwirtschaften.

Die Fragebogenmodule wurden mit Hilfe der Textverarbeitungssoftware Microsoft-Word entwickelt. Microsoft-Word ist ein Standardwerkzeug im statistischen Verbund, so dass notwendige länderspezifische Ergänzungen in Fragebogen für Bundesstatistiken, von jedem Statistischen Amt vorgenommen werden können. Weiterhin werden die Statistischen Landesämter in die Lage versetzt, die Wordmodule auch zur Gestaltung von Fragebogen für Länderstatistiken zu nutzen. Mit der Version Word XP werden weitere Funktionalitäten nutzbar sein, welche die zeitgemäße Erstellung von Fragebogen weiter verbessern.

Dem Anwender stehen zur Zeit ca. 20 Dateien mit Vorlagen (Wordmodule) zur Verfügung, mit denen er in der Lage ist, qualitativ hochwertige Fragebogen, einfach und schnell zu entwerfen. Die Wordmodule stellen Vorlagen für immer wiederkehrende Bestandteile von Fragebogen dar und lassen sich beliebig kombinieren und vervielfältigen. Durch die Nutzung dieser Wordmodule hält der Anwender automatisch alle vorgegebenen Standards für die Fragebogenerstellung im statistischen Verbund ein. So können mit der aktuellen Version der Fragebogenmodule belegesefähige Fragebogen effizient erstellt werden. Die Wordmodule sind bereits so ausgereift, dass sie auch für die Erstellung von Fragebogen genutzt werden können, die nicht für die Belegung vorgesehen sind.

Jeder mit diesem modularen System erstellte Fragebogen ist mit anderen in diesem System erstellten Fragebogen kombinier- und/oder erweiterbar. Das modulare System wird laufend durch eine Servicestelle im Statistischen Bundesamt ergänzt, oder es werden fertige Fragebogen als Muster dem System hinzugefügt. Weiterhin stehen Beispiele für jedes Modul zur Verfügung, aus denen der Nutzer ebenfalls wie in einem Baukasten, adäquate Teile in seinen Fragebogen kopieren kann. Um das System so nutzerfreundlich wie möglich zu machen, wurde ein Intranetangebot entwickelt, das allen Statistischen Ämtern ab April 2004 zur Verfügung stehen wird.

Die bisher gemachten Erfahrungen bei mehr als 40 Erhebungen zeigen, dass eine Standardisierung von Fragebogen basierend auf den entwickelten Wordmodulen möglich ist. Wenn auch einige Anwender zum Teil Schwierigkeiten haben, erforderliche Kompromisse für das Fragebogenlayout aufgrund der einzuhaltenden Standards zu akzeptieren, so ist jedoch die Resonanz auf die Nutzung des Werkzeugs durchweg positiv. Die vorliegenden Erfahrungen bestätigen den eingeschlagenen Weg der Standardisierung über wiederverwendbare und kombinierbare Basismodule.

Ausblick

Das Handbuch zur Erstellung von Erhebungsunterlagen in der amtlichen Statistik befasst sich zur Zeit vorrangig mit Vorgaben für Papierfragebogen. Im nächsten Schritt sollen die bestehenden Aspekte zur Gestaltung von elektronischen Fragebogen, z.B. um eigene Erfahrungen aus der Evaluierung von Internetfragebogen und durch Erkenntnisse aus dem Bereich der Online-Forschung erweitert werden.

Der vorhandene Bestand an Wordmodulen genügt, um 70% der Fragebogen nach den standardisierten Layoutvorgaben zu erstellen. Durch die laufende Ergänzung der Wordmodule für fachspezifische Anforderungen und den Ausbau des modularen Systems um Beispielfragebogen, sollen bestehende Lücken weitestgehend geschlossen werden.

Nach der Standardisierung des Fragebogenlayouts richtet sich der Blick stärker als bisher auf die Formulierung von Fragen und Antworten, um die Belastung der Auskunftgebenden weiter zu reduzieren und die Validität der gewonnenen Angaben zu verbessern. Dabei ist in Zukunft auch an eine intensivere Einbeziehung der Auskunftgebenden in die Verbesserung von Fragebogenentwürfen gedacht.

*Hans Joachim Schwamb, Tel.: +49 (0) 611 / 75 24 97,
E-Mail: hans-joachim.schwamb@destatis.de*

BV4.1 - eine neue benutzerfreundliche Software zur Komponentenzerlegung ökonomischer Zeitreihen

Die Komponentenzerlegung und Saisonbereinigung ökonomischer Monats- und Quartalsreihen mit dem Berliner Verfahren (BV) hat in Deutschland eine lange Tradition. Die mathematischen Grundlagen wurden Ende der sechziger Jahre des letzten Jahrhunderts von der Technischen Universität Berlin und dem Deutschen Institut für Wirtschaftsforschung (DIW) entwickelt. Kurze Zeit später (1972) wurde bereits die erste praxistaugliche Version des Verfahrens beim Statistischen Bundesamt eingeführt, um der Öffentlichkeit Informationen zu Entwicklungstendenzen und saisonbereinigten Werten der wichtigsten konjunkturrelevanten Zeitreihen zur Verfügung zu stellen. Seit 1983 wird BV4 eingesetzt, eine methodische Weiterentwicklung des Verfahrens durch das Statistische Bundesamt (s. Nourney, M.: "Umstellung der Zeitreihenanalyse", *Wirtschaft und Statistik* 11/1983, und Nourney, M.: "Seasonal adjustment by frequency determined filter procedures", *Statistical Journal of the United Nations* ECE 2/1984). Sie erwies sich für den Einsatz in der deutschen amtlichen Statistik als sehr geeignet.

Auf dem 54. Weltkongress des Internationalen Statistischen Institutes (ISI) im August dieses Jahres in Berlin hat das Statistische Bundesamt die neue Verfahrensversion BV4.1 vorgestellt. Sie wird im Statistischen Bundesamt ab 2004 eingesetzt werden. Die Verfahrensversion BV4.1 unterscheidet sich von ihrer Vorgängerin durch methodische Verbesserungen bezüglich der Schätzung von ausreißer- und kalenderbedingten Effekten auf die Zeitreihenwerte. Außerdem kann der Anwender selbst Einflussgrößen festlegen, die bei der Analyse berücksichtigt werden sollen (z.B. Sprungstellen, Witterungsgrößen).

Zur Version BV4.1 legt das Statistische Bundesamt nun erstmals auch eine benutzerfreundliche PC-Software für das Berliner Verfahren vor. Damit soll die Verbreitung des Verfahrens - auch international - gefördert werden. Der Release-Kandidat (RC 1) der Software kann vom Statistischen Bundesamt kostenfrei bezogen werden (bv4.1@destatis.de).

Mathematische Grundzüge des Verfahrens BV4.1

Der erste Teil des Verfahrens BV4.1 behandelt die Identifikation von (potentiellen) Ausreißern. Er basiert auf der Annahme, dass die zu analysierende Zeitreihe innerhalb hinreichend kurzer, über den gesamten Analysebereich gleitender Zeitintervalle einer bestimmten festen Länge - der sogenannten Stützbereiche - jeweils die Realisierung eines stationären Normalprozesses ist. Auf diese Weise werden für die den jeweiligen Stützbereichen unmittelbar benachbarten Beobachtungszeitpunkte (Ausreißeridentifizierung vorwärts- und rückwärtsgerichtet) bedingte Verteilungen für die Zeitreihenwerte hergeleitet. Beträgt die Differenz zwischen einem Beobachtungswert und einem derart ermittelten (bedingten) Erwartungswert mehr als ein vom Anwender festgelegtes Vielfaches von dessen (bedingter) Standardabweichung, wird der Beobachtungswert als Ausreißer angesehen.

Sofern erforderlich, erfolgt im nächsten Schritt die integrierte Schätzung von Ausreißern, Kalendereinflüssen und gegebenenfalls Einflüssen von reihenspezifischen, vom Anwender vorzugebenden Variablen. Ausgehend von der Formulierung des erweiterten klassischen

additiven Modells der Komponentenerlegung für eine ökonomische Zeitreihe O (d.h. $O = T + S + KA + A + EX + R$ mit T = Trend-Konjunktur-Komponente, S = Saisonkomponente, KA = Kalenderkomponente, A = an-wenderbestimmte Komponente, EX = Ausreißerkomponente, und R = Restkomponente) als lineares Regressionsmodell

$$O = \sum_{i=1}^h \mu_i T_i + \sum_{i=1}^k v_i S_i + \sum_{i=1}^l \alpha_i KA_i + \sum_{i=1}^m \beta_i A_i + \sum_{i=1}^n \gamma_i EX_i + \varepsilon$$

(T_i = Trend-Konjunktur-Regressoren, S_i = Saisonregressoren, KA_i = Kalenderregressoren, A_i = reihen-spezifische anwenderbestimmte Regressoren, EX_i = reihenspezifische Ausreißer-Dummy-Re-gressoren für die im ersten Schritt identifizierten Ausreißer, ε = Fehlerterm) wird dazu zunächst dessen Transformation durch die bereits bei BV4 bewährte lineare Filterprozedur F zur Trend-Konjunktur- und Saisonbereinigung durchgeführt. Man erhält:

$$F(O) = F\left(\sum_{i=1}^h \mu_i T_i\right) + F\left(\sum_{i=1}^k v_i S_i\right) + \sum_{i=1}^l \alpha_i F(KA_i) + \sum_{i=1}^m \beta_i F(A_i) + \sum_{i=1}^n \gamma_i F(EX_i) + \varepsilon^*,$$

wobei $\varepsilon^* = F(\varepsilon)$ den neuen Fehlerterm bezeichnet. Da unterstellt werden kann, dass gilt

$$F\left(\sum_{i=1}^h \mu_i T_i\right) \approx 0 \text{ und } F\left(\sum_{i=1}^k v_i S_i\right) \approx 0,$$

erfolgt die Schätzung der Parameter α_i , β_i , γ_i und damit der Komponenten KA , A und EX im Rahmen des Modells

$$F(O) = \sum_{i=1}^l \alpha_i F(KA_i) + \sum_{i=1}^m \beta_i F(A_i) + \sum_{i=1}^n \gamma_i F(EX_i) + \varepsilon^{**},$$

und zwar mittels der Methode der kleinsten Quadrate.

Im dritten Teil des Verfahrens erfolgt schließlich die Schätzung der Trend-Konjunktur- und der Saisonkomponente der zu analysierenden Zeitreihe, und zwar anhand der um die geschätzten Ausreißer, Kalendereinflüsse und Einflüsse der anwenderbestimmten Komponente bereinigten Reihe. Dies geschieht wiederum mittels der entsprechenden (linearen) Filter von BV4.

Standardmäßig ist bei BV4.1-Analysen die sogenannte wochentägliche Kalenderbereinigung vorgesehen. Dabei werden acht Kalenderregressoren verwendet, und zwar die Abweichungen der Anzahlen der einzelnen Wochentage Montag bis Samstag, auf die keine Feiertage fallen, die Abweichung der Zahl der Sonntage und die Abweichung der Anzahl der Feiertage in der Beobachtungsperiode von den entsprechenden durchschnittlichen Anzahlen in den der jeweiligen Beobachtungsperiode zugeordneten gleichnamigen Perioden.

Besondere Eigenschaften des Verfahrens

Das Verfahren BV4.1 unterscheidet sich von anderen verbreiteten Zeitreihenanalyseverfahren durch folgende charakteristischen Eigenschaften:

Auch bei sich rasch verändernden Saisoneinflüssen ist durch die Verwendung lokaler Regressionsmodelle eine effiziente Saisonbereinigung sichergestellt.

Die Verläufe der Trend-Konjunktur-Komponenten sind ökonomisch plausibel.

Das Kosten/Nutzen-Verhältnis ist niedrig, denn zur Durchführung qualitativ hochwertiger Analysen benötigt der Anwender weder eine spezielle Schulung oder sogar Expertenwissen noch längere Erfahrungen im Umgang mit dem Verfahren.

Die Analyseergebnisse sind grundsätzlich anwenderunabhängig, da keine Details des Analyseverfahrens festgelegt werden müssen.

Aufgrund der Verwendung linearer Regressionsmodelle gibt es grundsätzlich keine Abweichungen zwischen direkt und indirekt ermittelten Analyseergebnissen von Aggregat-reihen.

Leistungsmerkmale der Software BV4.1

Die Software BV4.1 besitzt folgende wesentlichen Leistungsmerkmale:

- Systemanforderungen: Windows-PC (ab Windows NT 4.0/Windows 98).
- Unterstützung der Ein- und Ausgabeformate EXCEL, ACCESS, SQL-Server and CSV.
- Anwenderfreundliche grafische Benutzerschnittstelle.
- Analysen von Monats- und Quartalsreihen.
- Möglichkeit der Massenproduktion von Zeitreihenanalysen.
- Möglichkeit der grafischen Darstellung von Analyseergebnissen.
- Möglichkeit der Durchführung sukzessiver Analysen, das heißt von Analysen, bei denen die Analysezeiträume sukzessive um jeweils eine Beobachtungsperiode verlängert werden. Diese Option ist nützlich, wenn Revisionen von Analyseergebnissen untersucht werden sollen, die auf das Verfahren BV4.1 zurückzuführen sind.

Hans-Theo Speth, Tel.: +49 (0) 611 / 75 22 26, E-Mail: hans-theo.speth@destatis.de

Veranstaltungen

Wissenschaftliche Tagung „Online-Erhebungen“

Das Statistische Bundesamt veranstaltet seit 1995 gemeinsam mit dem Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V. (ADM) und der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e. V. (ASI) im zweijährigen Turnus gemeinsame wissenschaftliche Tagungen. Thema der fünften gemeinsamen Tagung, die am 26. und 27. Juni 2003 im Statistischen Bundesamt in Wiesbaden durchgeführt wurde, waren „Online-Erhebungen“. Der im Oktober 2003 erschienene Tagungsband dokumentiert die gesamte Veranstaltung.

Im Anschluss an die Einführung in das Thema durch den Moderator (Dr. Manfred Ehling, Statistisches Bundesamt) bilden drei Überblicksreferate einen ersten Themen-Block, die über den aktuellen Stand der Online-Erhebungen in der akademischen Sozialforschung (Dr. Ulf-Dietrich Reips, Universität Zürich), in der kommerziellen Marktforschung (Hartmut Scheffler, TNS EMNID, Bielefeld) und in der amtlichen Statistik (Dr. Thomas Pricking, Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen, Düsseldorf) informieren.

Im Anschluss an die Überblicksreferate erläutert Erich Wiegand vom Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V. (ADM) die Qualitätsstandards und Standesregeln, die für webbasierte Datenerhebungen gelten. Dr. Wolfgang Bandilla vom Zentrum für Umfragen, Methoden und Analysen (ZUMA) in Mannheim stellt die Internetgemeinde als Grundgesamtheit vor. Im direkten thematischen Zusammenhang dazu steht das Referat über die Ziehung von

Stichproben bei Online-Erhebungen von Christiane Heckel von der BIK Aschpurwis + Behrens GmbH in Hamburg.

Frank E. Zander und Thomas Lanninger von der Ipsos Access Panels GmbH in Hamburg informieren über die Möglichkeiten, Access-Panels als Grundlage von Online-Erhebungen zu nutzen. Dr. Michael Bosnjak von der Universität Mannheim illustriert anhand zahlreicher Beispiele die methodischen Möglichkeiten, aktuellen Themen und Erweiterungen bei Online-Fragebogenuntersuchungen. Die Möglichkeiten von Multimedia-Anwendungen bei Online-Erhebungen stellt Dr. Raimund Wildner von der GfK AG in Nürnberg vor. Zum Abschluss informiert der Beitrag von Dr. Frank Knapp von der Psyma Online Research GmbH in Rückersdorf über Online-Erhebungen jenseits der Befragung.

Der Tagungsband ist als Band 7 in der Reihe „Sozialwissenschaftliche Tagungsberichte“ des Informationszentrums Sozialwissenschaften erschienen. Er ist zum Preis von 12 Euro zuzüglich Versandkosten zu beziehen beim Informationszentrum Sozialwissenschaften, Lennestr. 30, 53113 Bonn; Bestellungen können am einfachsten per Internet über den Bestellservice des Informationszentrum Sozialwissenschaften unter „<http://www.gesis.org/Bestellen/IZ/index.htm?order/forschungsberichte.htm>“ aufgegeben werden.

12. Wissenschaftliches Kolloquium - „Ökonomische Leistungsfähigkeit Deutschlands - Bestandsaufnahme und statistische Messung im internationalen Vergleich“ am 20. und 21. November 2003

Das Statistische Bundesamt veranstaltet in Zusammenarbeit mit der Deutschen Statistischen Gesellschaft - Ausschuss Methodik Statistischer Erhebungen - seit 1992 jährlich ein wissenschaftliches Kolloquium, das sich als ein Forum für den Dialog zwischen amtlicher Statistik und ihren wichtigsten Nutzergruppen aus Wissenschaft, Wirtschaft, Politik, Verwaltung und Verbänden etabliert hat.

Nachdem im vergangenen Jahr mit dem Thema „Sozialer Wandel – Daten, Analysen, Gesamtrechnungen“ eine im Schnittpunkt zwischen Sozialwissenschaft und Wirtschaftswissenschaft stehenden Thematik gewählt worden war, rückte das diesjährige Kolloquium „Ökonomische Leistungsfähigkeit Deutschlands - Bestandsaufnahme und statistische Messung im internationalen Vergleich“ rein ökonomische Fragestellungen in den Vordergrund der Betrachtung.

Das Kolloquium wurde von Herrn Prof. Dr. Ullrich Heilemann moderiert, der auch in die Thematik einführte. In dem anschließenden Überblicksreferat analysierte Dr. Hans J. Barth von der Prognos AG in Basel, wo die Deutsche Wirtschaft im internationalen Vergleich nach verschiedensten Kriterien einzuordnen ist. Unter der Überschrift „Wachstum: Wunsch oder Wirklichkeit“ ging Prof. Dr. Michael Hüther von der DekaBank in Frankfurt am Main speziell auf methodische Aspekte und wirtschaftspolitische Folgerungen ein.

Dr. Michael Grömling vom Institut der deutschen Wirtschaft in Köln beantwortete Fragen nach der Messbarkeit von Globalisierungstendenzen der deutschen Wirtschaft. Prof. Dr. Christoph Spengel von der Justus-Liebig-Universität Gießen betrachtete Unternehmensgewinne und -steuerbelastung im internationalen Vergleich und analysierte, inwieweit sie Indikatoren der Leistungsfähigkeit darstellen.

Den feierlichen Abschluss des ersten Tages bildete die Verleihung des Gerhard-Fürst-Preises, mit dem das Statistische Bundesamt jährlich herausragende Arbeiten mit Bezug zur amtlichen Statistik auszeichnet. Mit jeweils mit einem Gerhard-Fürst-Förderpreis für wissenschaftliche Nachwuchskräfte wurden in diesem Jahr die beiden Dissertationen von Dr. Michael Gold

(Universität Lüneburg) „Mikroökonomische Analyse der Arbeitsnachfrage - Eine Untersuchung von Beschäftigungsdynamik und Überstundennachfrage unter dem Einfluss von Anpassungskosten mit deutschen Betriebsdatensätzen“ und von Dr. Michael Niese (Technische Universität Bergakademie Freiberg) „Ursachen von Betriebsschließungen und ihre Konsequenzen für die Marktdynamik – Eine mikroökonomische Analyse von Probezeiten und Todesschatten im Verarbeitenden Gewerbe“ ausgezeichnet. Ebenfalls mit einem Gerhard-Fürst-Förderpreis wurde die Diplomarbeit "Kapitalkosten und Kapitalstrukturen ausgewählter deutscher Unternehmen - eine empirische Untersuchung" von Simon Krotter (Universität Regensburg) prämiert.

Am zweiten Tag des wissenschaftlichen Kolloquiums betrachtete Birgit Mackenthun von der Technische Universität Berlin das Gesundheitswesen als Wachstumsfaktor. Leistungsindikatoren im Bildungsbereich und dabei selbstverständlich auch die PISA-Studie analysierte PD Dr. Stefan C. Wolter von der Schweizerische Koordinationsstelle für Bildungsforschung in Aarau. Die monatliche Erhebung zum ILO-Erwerbsstatus als Instrument der Arbeitsmarktberichterstattung in Deutschland stellte Thomas Riede vom Statistischen Bundesamt dar. Achim Kemmerling von der Freie Universität Berlin betrachtete abschließend, wie sich Brutto- und Nettosozialleistungsquote für die Messung sozialstaatlicher Leistungen eignen.

Christian König; Tel.: +49-(0)611 / 75 20 77, E-Mail: christian.koenig@destatis.de

Veranstaltungskalender

ZUMA-Workshop: „Der Mikrozensus als Datenquelle für die empirische Sozial- und Wirtschaftsforschung: Eine Einführung“ am 17. Februar 2004

Neben den gängigen großen sozialwissenschaftlichen Erhebungen wie ALLBUS, SOEP und Wohlfahrtssurvey steht der Wissenschaft seit einiger Zeit mit dem Mikrozensus auch die größte laufende amtliche Haushaltsstichprobe zur Verfügung. Der Mikrozensus ist in Form von 70%-Substichproben für die Forschung zugänglich. 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 unter der Leitung von Helga Christians informiert über Nutzungsbedingungen, Erhebungskonzept und Frageprogramm des Mikrozensus und gibt einen Überblick zum Analysepotenzial. Darüber hinaus wird in das umfangreiche Serviceangebot zum Mikrozensus eingeführt.

Im Zentrum steht die Vermittlung von anwendungsorientiertem Basiswissen sowohl für interessierte Wissenschaftler und Studenten, die bisher noch nicht mit Mikrozensusdaten gearbeitet haben, als auch für Forscher mit ersten Erfahrungen. Der Workshop wird in enger Kooperation mit dem Statistischen Bundesamt (Gruppe IX B – Mikrozensus, Arbeitskräftestichprobe, Haushalt und Familie) durchgeführt.

Anmeldeschluss ist der 15.01.2004, der Teilnahmebeitrag beträgt € 60, für Studenten: € 40. Die Teilnehmerzahl ist auf 20 Personen begrenzt.

Interessenten werden gebeten, sich beim ZUMA-Tagungssekretariat (workshop@zuma-mannheim.de, Tel.: 0621-1246-221) anzumelden.

European Conference on Quality and Methodology in Official Statistics (Q2004) vom 24.-26. Mai 2004 in Mainz

Vom 24. bis 26. Mai 2004 veranstaltet das Statistische Bundesamt im Kurfürstlichen Schloss zu Mainz die European Conference on Quality and Methodology in Official Statistics (Q2004). Die mit Unterstützung von Eurostat durchgeführte Veranstaltung bildet den Auftakt zu einer Reihe zweijährlicher Europäischer Konferenzen zu Fragen der Qualität und der Methoden der amtlichen Statistik. Ziel ist es, ein regelmäßiges Forum für methodische Fragen und neue Entwicklungen der amtlichen Statistik Europas zu schaffen. Die Veranstaltung baut auf den positiven Erfahrungen der International Conference on Quality in Official Statistics (Q2001) auf, die 2001 in Stockholm stattgefunden hat.

Zu Q2004 werden etwa 400 Teilnehmerinnen und Teilnehmer vorwiegend aus Europa und Nordamerika erwartet. Die Themen der Veranstaltung umfassen alle qualitätsbezogenen Fragen der amtlichen Statistik. Abgedeckt werden neben den Prozessen der Erstellung amtlicher Statistiken auch Fragen der Messung der Qualität von Statistik sowie des Qualitätsmanagements. Am ersten Tag der Veranstaltung stehen zudem drei eintägige parallele Seminare zu den Themen „Survey Quality“, „Quality Management in Statistical Institutes“ und „Variance Estimation in Complex Surveys“ auf dem Programm. Als Kursleiter konnten international bekannte Experten gewonnen werden. Konferenzsprache ist Englisch.

Inhaltlich wird die Veranstaltung von einem Programmkomitee vorbereitet, das sich aus Vertretern von insgesamt sieben nationalen statistischen Ämtern, Eurostat und zwei Vertretern der Wissenschaft zusammensetzt. Der Vorsitz liegt gemeinsam bei Eurostat und dem Statistischen Bundesamt. Im Rahmen der Konferenz werden zudem die Ergebnisse des europäischen Projektes DACSEIS (Data Quality in Complex Surveys within the New European Information Society) vorgestellt. Ein Call for Papers fand eine außerordentlich positive Resonanz. Derzeit stellt das Programmkomitee das Konferenzprogramm zusammen, das Ende Januar 2004 im Internet veröffentlicht wird.

Neben näheren Informationen zur Konferenz und zum Programm ist auch das Anmeldeformular im Internet unter <http://q2004.destatis.de> verfügbar. Für weitere Informationen wenden Sie sich bitte an die Vorsitzenden des Programmkomitees, Werner Grünwald (Tel. +352-4301-33280; E-Mail: werner.grunewald@cec.eu.int) und Thomas Körner (Tel. +49-(0)611 / 75-4413; E-Mail: thomas.koerner@destatis.de).