

# METHODS – APPROACHES – DEVELOPMENTS

## Information of the German Federal Statistical Office

**Edition 1/2004**

### **The Catchword**

Labour Market in Germany - a telephone sample survey on the labour market situation in Germany..... 3

### **Methods of federal statistics – Further development**

Method for complementing the abridged life table in the range of ages from 90 to 100 ..... 5

Elaborating a powerful record linkage technique ..... 8

Bill to amend the Law on Public Finance and Service Personnel Statistics..... 9

Standardising survey documents in official statistics ..... 11

BV4.1 - a new user-friendly software for decomposing economic time series ..... 13

### **Events**

Scientific conference on online surveys..... 15

12th Scientific Colloquium "Germany's Economic Efficiency - Stocktaking and Statistical Measurement in International Comparison" on 20 and 21 November 2003 ..... 16

### **Calendar of events**

European Conference on Quality and Methodology in Official Statistics (Q2004) in Mainz from 24 - 26 May 2004..... 17

**Published by:** Statistisches Bundesamt (Federal Statistical Office), Wiesbaden

**Subject-related information**  
on this publication:

Division IB, Christian König  
tel: (+49-611) 75 20 77  
fax: (+49-611) 75 39 50  
christian.koenig@destatis.de

**General informationen**  
on the data supply:

Information sservice,  
tel: (+49-611) 75 24 05  
fax: (+49-611) 75 33 30  
info@destatis.de  
www.destatis.de

**Release calendar**  
**of the Press Office:**

[www.destatis.de/presse/deutsch/cal.htm](http://www.destatis.de/presse/deutsch/cal.htm)

Periodicity: (generally) twice a year

Published in April 2004

© Statistisches Bundesamt, Wiesbaden 2004

Reproduction and free distribution, also of parts, for non-commercial purposes are permitted provided that the source is mentioned. Any distribution, also of parts, via electronic systems/data media requires prior permission. All other rights reserved.

## The Catchword

### Labour Market in Germany - a telephone sample survey on the labour market situation in Germany

#### Background

Along with the growing Europeanisation and globalisation of economic relations, labour market reporting in Germany has to meet new requirements. So far, reporting based only on the national context and on national provisions in terms of social law has been predominant, whereas now aspects of international comparability are becoming ever more important. That trend towards considering international aspects is further enhanced by benchmarking processes within the European Union. The most recent example here is the report of the European Union Employment Task Force chaired by Wim Kok of November 2003 (cf. [http://europa.eu.int/comm/employment\\_social/employment\\_strategy/pdf/etf\\_en.pdf](http://europa.eu.int/comm/employment_social/employment_strategy/pdf/etf_en.pdf)).

However, requirements are becoming more demanding also in terms of data timeliness. For efficient labour market policies and short-term analyses, comparisons and socio-political programme evaluations, annual data are less suitable. What is indispensable here is monthly data on the labour market development that are available in a timely manner. The time window between data collection and data supply is becoming smaller.

#### Labour force concept

Cross-country comparability of labour market indicators such as the number of persons in employment, unemployed, and non-active persons requires a standardised and internationally accepted concept of measurement; without such a concept, it cannot be ruled out that statistical artefacts are created by different measuring methods.

At the 13<sup>th</sup> International Conference of Labour Statisticians (1982), the International Labour Organization (ILO) in Geneva adopted such a standardised set of tools, that is the labour force concept. As an international standard, the ILO labour force concept is applied to a multitude of population surveys on labour force participation. Its goal is to cover all groups of persons in employment, irrespective of the importance of such employment for the subsistence of the person concerned. The concept of measurement distinguishes between employed, unemployed, and persons not economically active. According to the labour force concept, a clear employment status can be allocated to any person (cf. [http://www.destatis.de/download/d/dt\\_erheb/ilo\\_res82.pdf](http://www.destatis.de/download/d/dt_erheb/ilo_res82.pdf)). Here, the labour force concept differs from the subsistence concept that is frequently used in official statistics and is operationalised through the main source of subsistence.

The "employed", according to the labour force concept, are all persons from the age of 15 who, during a defined reference period and irrespective of the time volume, are in paid employment, are self-employed or unpaid family workers. The decisive criterion here is that they performed some gainful activity for at least one hour within the reference period.

The "unemployed" are persons from the age of 15 who, in the reference period, did not have any gainful employment, were not self-employed, are available for employment now or within the next two weeks, and have been actively searching for employment for the last four weeks.

The "population not economically active" refers to persons who, in the reference week were neither employed nor unemployed.

## **Pilot survey according to Article 7, paragraph 2 of the Federal Statistics Law**

The goal of the survey "Labour Market in Germany" conducted by the Federal Statistical Office in co-operation with *Infratest Sozialforschung*, Munich is to represent the current situation and development on the German labour market on the basis of the ILO labour force concept and to offer in the future monthly and internationally comparable labour market reports in the context of official statistics. Since April 2003, the pilot survey according to Article 7, paragraph 2 of the Federal Statistics Law has been conducted with a sample size of 10,000 respondents per month.

The survey is designed as a rotating panel. Every person is to be interviewed six times altogether. This means that every month a sixth of the respondents will leave the survey. In order to be able to measure changes between the survey months, with a quality customary in official statistics, the aim is to achieve a monthly net overlap of 70 percent.

## **Computer Assisted Telephone Interview (CATI)**

For highly up-to-date monthly data collection and early provision of results, computer assisted telephone interviews (CATI method) are the obvious choice. So far, telephone interviews have not systematically been conducted in German official statistics. The pilot survey thus also contributes considerably to further methodological development in official statistics.

Especially in applied empirical social research, the CATI method has considerably gained in importance over the last few years. While until the 1980s, the social sciences literature referred to it as a low-quality quick-and-dirty method, it has now become a standard method in social sciences (cf. Gabler, Siegfried and Sabine Häder (ed.): *Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland*. Münster 2002). The fact that telephone interviews have rapidly grown in importance in empirical social research is due to a number of specific benefits provided by that data collection method. It is relatively low-cost and allows a large number of attempted contacts, thus reducing the share of persons not contacted. Computer assisted interviewing allows to follow rather complex filter sequences. The frequency of errors caused by faulty filter guidance can considerably be reduced. As the responses given by the persons interviewed are entered direct into a computer during the interview, plausibility checks can be performed even during the interview. Also, the raw data set is created immediately, which clearly reduces processing times.

## **Gabler-Häder method**

For drawing telephone samples, it is necessary that a maximum share of the population from which the sample is to be drawn can be reached via the fixed-line telephone network. With 95% of households in Germany being equipped with fixed-line telephones, that criterion can be considered to be fulfilled (cf. *Statistisches Bundesamt: Ausstattung und Wohnsituation privater Haushalte. Einkommens- und Verbrauchsstichprobe 2003*. Wiesbaden 2003, p. 14).

However, as a consequence of the obligation to be registered in official telephone directories having been abolished in 1992, the number of telephone numbers not registered has considerably increased. However, the "Gabler-Häder method" developed by Siegfried Gabler and Sabine Häder at the *Zentrum für Umfragen, Methoden und Analysen (ZUMA)* in Mannheim offers a solution to the problem. Therefore, that sampling method is successfully applied to the large majority of telephone interviews. The basic idea underlying the Gabler-Häder method is to form, within a local telephone area, intervals of theoretically possible telephone numbers and to check with the help of a telephone directory whether they include at least one valid telephone number. The intervals for which the check is positive are used as a sample space from which the random sample is then drawn (cf. Häder, Sabine: *Auswahlverfahren bei Telefonumfragen, ZUMA-Arbeitsbericht 1994/03*, Mannheim 1994; Häder, Sabine: *Telefonstichproben*; ZUMA How-to series 6, Mannheim 2000).

## Sample coverage, non-response processes and first findings

Sample coverage among those interviewed for the first time was an average 44.6% for the months of April to November 2003, which is within the range of what is common for voluntary socio-scientific surveys in Germany (cf. Schnell, Rainer: *Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen*. Opladen 1999). For repeated interviews over the same period, an average coverage of 86.7% was achieved. Considering the short field time - the monthly data are available not later than on the seventh calendar day of the subsequent month - that achievement is quite positive. With an average 76.9%, the envisaged net overlap of 70% has so far been maintained.

As the telephone interview is a voluntary survey, non-response and consequent sample biases are inevitable. Comparisons with external reference statistics show, however, that the biases correspond to what is known from general population surveys. This refers in particular to men, younger and older people, foreigners and persons with a low educational level. Longitudinal analyses of nonresponse processes show that, beyond such general demographic characteristics, the sample involves a slight bias regarding the employment status. Persons whose employment status remains unchanged during the interviews show a slightly higher nonresponse risk than persons undergoing status change. However, such biases can in part be eliminated through the expansion and weighting process.

The results so far available for the months of April to November 2003 show that the "margin of the labour market" can be better represented by the telephone survey than by the official statistics existing in Germany. Improvements are observed that are due to different measurement and that allow good representation of marginal part-time work, employment of pupils and students and other employment at the margin of the labour market as well as the search for such activities.

*Further information is available at: <http://www.destatis.de/arbeitsmarkt>*

*Thomas Riede, tel: (+49-611) 75 24 33, e-mail: [thomas.riede@destatis.de](mailto:thomas.riede@destatis.de)*

*Matthias Sacher, tel: (+49-611/ 75 25 34, e-mail: [matthias.sacher@destatis.de](mailto:matthias.sacher@destatis.de)*

## Methods of federal statistics – Further development

### Method for complementing the abridged life table in the range of ages from 90 to 100

#### The abridged life table of the Federal Statistical Office

The Federal Statistical Office, at regular intervals, constructs period life tables to provide information on the mortality pattern and average life expectancy of the population. In the wake of population censuses, the Federal Statistical Office draws up *complete life tables*, which are published for the individual ages up to 100 and on which mathematical-statistical methods are used to adjust them for random variations and cohort effects. In the interim between population censuses, the Federal Statistical Office compiles *abridged life tables* which stop at age 90 and remain unadjusted. The construction of life tables is based on the age-specific probabilities of dying, which are calculated by taking into account the number of deaths and the average number of population at individual ages (cf. *Farr's* death rate method). The small numbers at high ages and the growing inaccuracies in current population statistics as the time interval to the last population census increases are the reasons why, in the *abridged life tables*, the probabilities of dying can only be published up to age 89. Above age 90, the probabilities of dying calculated for males start showing an implausible decline which, due to women's higher life expectancy,

appears somewhat later for females. From scientific, economic and political points of view, however, the general rise in life expectancy creates an increased demand for information on the mortality pattern and life expectancy of the population aged above 90 years. Since there is no immediate prospect of another population census or an equivalent method of collecting data, it is not possible to construct a new *complete life table*. Therefore, the following paragraphs describe two procedures which together make it possible to complement the probabilities of dying and to expand the *abridged life table* up to age 100.

### **The method of extinct generations**

With a reliable system of death registration that completely covers all years of age and provides data in a breakdown by age and year of birth, the number of survivors to a birth cohort can be calculated backwards for the individual ages via the number of deaths. The method of *extinct generations* is, however, confined to higher years of age because migration must be ruled out. The method of *extinct generations* as originally proposed by *Vincent* can only be applied to extinct birth cohorts. Yet, this restriction can be circumvented by including a number of population which is assumed to be exact, such as the number recorded by the last population census of 1987 (cf. Dinkel 1997, pp. 496 f.). The probabilities of dying for the individual birth cohorts are then determined by applying the year-of-birth method developed by *Becker-Zeuner* and are afterwards rearranged from a cohort series into a period series. But the method of *extinct generations* cannot provide plausible probabilities of dying for all ages up to 100 because in particular the probabilities for males are subject to premature variations due to the small numbers. One possibility of complementing these probabilities of dying is to extrapolate them by using a regression model.

### **The non-linear regression model**

The precondition for choosing an adequate regression model is to know the development of mortality at high ages. The British actuary *Benjamin Gompertz* discovered as early as in 1825 that, following an increase due to accidents in youth, the risk of dying becomes an exponential function of age. This observation, which was later called the "*Law of Mortality*", can no longer be applied to high years of age. It has been found that the risk of dying after the age of about 80 tends to increase slower as its growth rate decelerates (cf. Anderson 1999, p. 2 et al.). Hence, it is important that an adequate regression model for extrapolating the probabilities of dying should make allowance for the decelerating growth rate of the risk of dying at high ages. Against that background, the models of *Kannisto*, *Heligman & Pollard*, *Coale & Kisker*, the *modified Gompertz approach*, other logistic functions and high-order polynomials (cf. Thatcher et al. 1998, Doray 2002, Statistisches Bundesamt 1991) were tested on the last *complete life table* of 1986/88 and on the *abridged life table* of 1998/00. The following additional criteria for assessing the models' adequacy were laid down: goodness of fit of the input data (coefficient of determination  $r$  larger than or equal to 0.99), stability against variation of the input data, easy handling (not more than three regression coefficients) and plausibility in relation to the last *complete life table* of 1986/88. In the *abridged life table*, the input data for the regression were varied within the range of ages from 70 to 89, while the input data for ages 76 to 94 were taken on in the *complete life table*. Problems arose in this context because of the approximating probabilities of dying for males and females, which in many cases resulted in a *mortality crossover*.

Taking into account the above-described criteria, regression model (1) proved to be adequate for extrapolating the probabilities. In particular, its plausibility in relation to the *complete life table* of 1986/88 and its stability against variation of the input data were the decisive reasons for choosing this model. In the *complete life tables* of 1960/62, 1970/72 and 1986/88 the *modified Gompertz approach* (2) had been used.

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

$q_x$  Probability of dying at age x

$\beta_i$  Regression coefficient

The regression model (1) is a descriptive model designed only for extrapolating the course of mortality in the range of ages from 90 to 100. Contrary to causal models, which are based on a theory of mortality, the regression model (1) does not attempt to depict the entire development of mortality or the course of mortality after youth. The regression model (1) was also tested on the *abridged life tables* for Germany starting from 1991/93 and on the *complete life table* of 1970/72 and provided plausible estimates for these life tables as well.

### Conclusion

The method of *extinct generations* makes it possible to empirically determine the probabilities of dying for nearly the full range of ages from 90 to 100. The regression model (1) permits to estimate the additionally required probabilities so that the *abridged life table* can be published up to age 100.

### Bibliography

*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, Jahrgang 22, Heft 4/1997: pp. 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](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: (+49-611) 75-3275, e-mail: matthias.eisenmenger@destatis.de*

## Elaborating a powerful record linkage technique

The Federal Statistical Office is participating in an international methodological project on statistical confidentiality (Computational Aspects of Statistical Confidentiality – CASC), which is supported by the EU Commission, a project on which MAD already reported in issues 1/2001 and 2/2002. The CASC project is destined to further develop various statistical confidentiality procedures and to integrate them into the ARGUS program package. ARGUS contains both a program on tabular data confidentiality ( $\tau$ -ARGUS) and a program on microdata confidentiality ( $\mu$ -ARGUS).

A contribution of the Federal Statistical Office consisted in the development of a technique for simulating database cross matches, which is equally implemented in  $\mu$ -ARGUS. In a database cross match scenario, an attacking data intruder tries to assign as many external database records as possible (extra knowledge) unambiguously to the anonymised target database records in order to enlarge the external database by information mined from the target database.

The so called record linkage technique is theoretically based on algorithms of the graph theory and mathematical optimisation, which the Federal Statistical Office introduced to a workshop on statistical confidentiality in Luxembourg in April 2003. The technique is outlined as follows.

In a first phase, the record linkage technique generates a distance measure covering all common (key) variables of the records in the two databases. As in a real attack scenario data intruders tend to prefer a few selected variables, supposed to include less deviations from the original data, to other, less reliable variables, we decided to leave it to the user to assign concrete weights  $w_i$  to variables  $i$ , although, for the sake of simplicity, standardised weight intervals of [0.1] were laid down. In this context, giving a variable  $i$  the maximum weight, i.e. choosing a weight of  $w_i = 1$ , is equal to the so called "blocking of variables". For two records to be assigned clearly to each other, it is necessary that they coincide exactly in their maximum-weighted variables (block variables) as, naturally, categorical variables are rather used as block variables. With a skilful choice of block variables, it is possible to considerably reduce the general computing expenditure involved in assignment operations and to substantially lower the proportion of wrongly assigned variables as well. A variable  $i$ , to which the minimum weight  $w_i = 0$  has been assigned, is left out of consideration during the assignment procedure.

The objective of the second phase is to make assignments of records on the basis of identified distances. For that purpose, we minimise the sum of distances for all assignments to be made (total deviation). For the purpose of comparison, we have, in addition to the classical methods of the graph and network theory, also considered Greedy algorithms to the extent that they have been suggested in literature so far, as they are easier to implement and require less computing time, although, on the other hand, they have the disadvantage of causing total deviation to rise.

The Federal Statistical Office incorporates the results of CASC research activities directly into a project on "De facto Anonymisation of Business Microdata", jointly undertaken by the statistical offices and the Institute for Applied Economic Research (IAW), Tübingen. Based on data from various surveys, this project is supposed to test anonymisation techniques both in terms of the analytical potential maintained and a possible risk for the microdata to be reassigned (reidentified) at a later time.

To examine the protective effect of anonymisation, the record linkage technique described above was applied to various data - anonymised for test purposes - of the cost-structure survey in manufacturing.

*Dr. Rainer Lenz, tel: (+49-611) 75 26 36, e-mail: rainer.lenz@destatis.de*

*Roland Sturm, tel: (+49-611) 75 25 80, e-mail: roland.sturm@destatis.de*



## **Bill to amend the Law on Public Finance and Service Personnel Statistics**

It has become necessary to adapt the Law on Public Finance and Service Personnel Statistics (Public Finance and Service Personnel Statistics Law - FPStatG - of 8 March 2000) to new and important statistical requirements of the European Union (EU). In addition, with public budgets outsourcing an increasing number of institutions, the bill (FPStatG) will have to make sure that statistics on public finance will provide a full and reliable database, which will meet the detailed statistical requirements for national accounts (general government account, distribution of income account) of the Federation and the Länder. Statistics on public finance are based on data directly taken from budgetary documents, cash reports and accounting records of public administrations, a procedure, which reduces the statistical burden on these administrations to a possible minimum and guarantees a high quality standard at the same time. Thus, amending the law, it will also be necessary to take into account the reformatory endeavours of public budgets on their way towards double-entry accounting, in particular, on the municipal level. At the same time, provision has been made for a face-lifting of the survey programme, which will offset at least part of the burden caused by the new requirements mentioned above.

### **1. New requirements of the European Union on public finance statistics**

On the liabilities side, the quarterly survey of the debts of public budgets already provides the data the EU is requesting on financial transactions and stocks (financial assets and liabilities) in the government sector. However, unlike the provisions which are presently in force, the amended provisions will have to make sure that units that in accordance with Council Regulation (EC) No. 2223/96 of 25 June 1996 on the European system of national and regional accounts in the Community (ESA 95) need to be assigned to the government sector (outsourced public institutions, please, refer to item 2.) are added to the group of reporting units for this survey.

In 2005, a new statistical data collection is to be introduced, which will record the individual items of public financial assets. The new stipulations make provision for a survey to collect end-of-year data on financial assets by types as well as flow data on changes and adjustments in assets made in the course of the year from the Federation (central government), the Länder, the municipalities, both local authority and special-purpose associations as well as from the social insurance system and from governmental and municipal funds, institutions, establishments and enterprises. The quarterly data required by the EU will be estimated on the basis of the annual results on public financial assets.

### **2. Topical financial statistics on outsourced public institutions**

Presently there is a strong tendency to outsource public institutions from the core budgets of the Federation, the Länder and the municipalities, which according to Council Regulation (EC) No. 2223/96 of 25 July 1996 on the European system of national and regional accounts in the Community (ESA 95) need to be assigned to the government sector. This trend is causing gaps in the data reported on public finance. In future, these gaps will be closed by a quarterly survey of the revenue and expenses of units outsourced from the government sector. That will provide a basis for integrating core public budgets and budgets of outsourced units, making sure that a complete statistical picture of public finance is obtained, as required.

This change is equally important to meet the statistical requirement of the EU regarding quarterly statistics on government sector accounts (revenue and expenditure) compiled in the context of national accounts. In future, the adapted quarterly survey for the collection of cash statistics will fully meet the requirements of national accounts in terms of timeliness, subject and quality, in particular, regarding the data needed to determine the EU stability criterion - government budget deficit. In addition, this is the first time that early information will be available on how the group of outsourced units belonging to the government sector is actually delimited.

### **3. Making allowance for the reform of public accountancy**

The amending of the Public Finance and Service Personnel Statistics Law also makes allowance for the endeavours to adapt public accountancy to new conditions. So far, public finance statistics on revenue and expenditure of public budgets have been obtained exclusively in a breakdown which fully corresponds to the common budget classification framework used by the Federation and the Länder, on the one hand, and by the municipalities and local authority associations, on the other hand. This linkage makes sure that statistics on public finance can directly be extracted from public budgetary documents, cash reports and accounting records without any extra effort. Recently, the effort to make public budget accountancy more flexible and, in particular, the pilot trials undertaken to test a new double-entry accounting system adversely affected the classification framework for the breakdown of revenue and expenditure of the public budget, which was binding in the past.

The Ministers of the Interior of the Länder have already adopted a new standard double-entry accounting system for municipalities, which meets the requirements of public finance statistics and national accounts. The new system maintains the useful practice of linking the breakdown of public finance statistics to the public budget classification framework. But, in addition, the Public Finance and Service Personnel Statistics Law (FPStatG) will include a new provision, which will make allowance for the municipalities' public finance data to be also collected in compliance with the classification of accounts and products used by the new double-entry accounting system. As many science and research institutions have introduced commercial accounting systems as well, there is an obvious necessity of extending the presently used public budget classification framework in order to maintain the capability of collecting adequate statistics on expenses, revenues and investment outlays in the years to come.

At present, expert groups concerned with the cost and results accounting system used by the Federation and the Länder are also discussing the subject of a standard framework of accounts and products classifications, which, possibly, could serve as well as a basis for a double-entry accounting system to be used by Länder. However, the Federation and the Länder are still obliged to abide by the present public budget classification framework (classification by object and by function), as stipulated by the Law on the Basic Rules Governing Budget Legislation (HGrG). When they apply the rules of a new double-entry accounting system, they have to make sure that the collected data is converted into the categories of revenue and expenditure as listed in the classification by object and by function. At present it is not foreseeable yet whether and when a new standard double-entry accounting system will be available for the budgeting purposes of the Länder and the Federation.

### **4. Face-lifting of the former data collection programme**

A face-lifting effort regarding parts of the data collection programme is to bring down the cost on respondents and to reduce their statistical burden and also to compensate at least some of the extra burden incurred by statistical offices due to the additional data requirements described above. Thus, the amendment provides for action to be taken as follows: to discontinue collection of budgetary planning statistics, to establish a flexible and size-related cut-off limit for municipalities liable to collect data for municipal budget planning statistics, and to envisage a reduced data collection programme, in particular, for legally independent organisations of science, research and development which are liable to supply data for personnel and pension recipients statistics. The planned use of electronic data transfer is another factor which will reduce the response burden. It will make it possible for the statistical offices to put through, in the context of the e-government initiative, a philosophy aimed at integrating the production process from data collection to data dissemination in a way that is free of disruptions in media usage.

## 5. The time-table for the amendment

Parliament is supposed to discuss the amendments to the Law on Public Finance and Service Personnel Statistics in early 2004 and to adopt them by 30 June 2004, if possible. The bill is to be enacted as of 1 January 2005.

*Renate Schulze-Steikow, tel: (+49-611) 75 41 66, e-mail: reate.schulze-steikow@destatis.de*

## Standardising survey documents in official statistics

### Situation

The increasing demand for statistical results calls for a high quality of the survey documents used to obtain the data. Statistical information has always been accepted by the users only if they trust its validity. In turn, the quality of statistical information depends largely on the quality of the survey documents used, as only the “right” questions can provide informative results. Well-designed survey documents reduce the burden on the respondents and have a positive impact on their readiness to provide information. This effect can still be intensified by an increased recognition value of the survey documents used in official statistics. One possibility to achieve this goal is the uniform corporate design of official statistics as seen from outside.

### Joint Federation and Länder Working Group on the “Standardisation of Survey Documents“

It was the aim of the Joint Working Group on the “Standardisation of Survey Documents”, which is composed of representatives of the statistical offices of the Federation and the Länder, to provide suggestions as to how the above requirements could be met. In the course of two years, the Working Group revised the Survey Document Preparation Manual and developed a system of questionnaire modules. Thus, solutions could be found which ensure

- a uniform design (basic layout) of survey documents used for federal statistics
- a uniform presentation of the questionnaire contents by agreeing on a binding wording of questions and the order in which they are listed
- an easy and rapid preparation of draft questionnaires
- an improved quality of draft questionnaires
- the readability of all questionnaires by optical document reader
- an increase in validity (reducing time-consuming subsequent corrections and costly queries at the respondents)
- a faster provision of results.

### Official Statistics Survey Document Preparation Manual

The Manual was revised by the Joint Federation and Länder Working Group on the “Standardisation of Survey Documents“. It contains methods and standards for designing survey documents to be used by the statistical offices of the Federation and the Länder. In March 2003, it was put into force by the Conference of the Heads of the Statistical Offices. Apart from sections providing basic explanations on how to prepare survey documents and to optimise them, the Manual contains detailed information such as check lists or requirements of optical document reading the survey documents have to meet. Some sections of the Manual describe binding rules which have to be observed in designing survey documents, other sections provide mere recommendations. This structure provides sufficient flexibility so that the diversity of official surveys can be taken into account as well as specific features of individual Länder. In future, the Manual will be adjusted to current developments on a continuous basis.

## Modules for preparing draft questionnaires

A system of questionnaire modules has been developed for the preparation of standardised questionnaires. The new system can be used in all areas of statistics. It supports necessary optimisations of operations forming part of statistical processes, e.g. the more frequent use of optical character readers and the sending of completed paper questionnaires by fax. Shrinking resources call for these measures, and the solutions also help to reduce costs through increased efficiency.

The questionnaire modules were developed using the word-processing software Microsoft Word. Microsoft Word is a standard tool at the statistical offices so any Land statistical office can make additions to federal statistics questionnaires which may be necessary on account of specific Land features. Moreover, the statistical offices of the Länder can use the Word modules also for designing questionnaires employed to collect Länder statistics. The Word XP Version will provide additional functions further improving the state-of-the-art preparation of questionnaires.

Presently, the user can draw upon some 20 files containing templates (word modules). On this basis, he can easily and rapidly design questionnaires of a high quality. The Word modules are templates of frequently used parts of questionnaires, they may be combined and copied as required. By using these Word modules, the user automatically adheres to all standards established for the preparation of questionnaires within the joint statistical system of the Federation and the Länder. For example, questionnaires to be read by optical character readers can be produced efficiently with the present version of the questionnaire modules. The Word modules are already sophisticated enough to be used also to prepare questionnaires not intended for optical character reading.

Each questionnaire produced on the basis of this modular system can be combined and/or supplemented with other questionnaires prepared within the system. The modular system is constantly being supplemented by a service unit at the Federal Statistical Office, or questionnaires ready for use are added to the system as samples. Furthermore, examples of each module are available adequate parts of which the user can also copy into his questionnaire like from a construction kit. In order to make the system as easy to use as possible, an Intranet page was designed all statistical offices will be able to draw upon as of April 2004.

The experience gathered in the course of more than 40 surveys shows that it is possible to standardise questionnaires on the basis of the Word modules developed. Although some users find it difficult to accept compromises regarding the questionnaire layout which are necessary due to the standards to be maintained, once they use the tool their responses are all positive. The experience gathered so far confirms that the right path has been chosen, namely standardisation via basic modules that can be combined with each other and used again.

## Outlook

At present, the Official Statistics Survey Document Preparation Manual predominantly deals with specifications for paper questionnaires. In a next step, the existing aspects of electronic questionnaire design shall be supplemented, for example, by experience we gathered in evaluating Internet questionnaires and by findings from the field of online research.

The existing stock of Word modules is sufficient to prepare 70% of the questionnaires according to the standardised layout specifications. To fill existing gaps as far as this is possible, the Word modules are continuously being supplemented to meet the specific requirements of specialised statistics, and examples of questionnaires are incorporated into the modular system.

After standardising the questionnaire layout, more attention is paid to the wording of questions and answers in order to reduce the burden on the respondents still further and to improve the validity of the data collected. It is intended to have respondents participate more in the improvement of draft questionnaires in future.

*Hans Joachim Schwamb, tel: (+49-611) 75 24 97, e-mail: [hans-joachim.schwamb@destatis.de](mailto:hans-joachim.schwamb@destatis.de)*

## BV4.1 – a new user-friendly software for decomposing economic time series

In Germany, the decomposition and the seasonal adjustment of monthly and quarterly economic time series by means of the Berlin procedure (BV) have a long tradition. The mathematical bases were developed in last century's end sixties at Berlin Technical University and the German Institute for Economic Research (DIW). Shortly afterwards (in 1972), the first feasible version of the procedure was already introduced at the Federal Statistical Office in order to provide the general public with information on trends and seasonally adjusted values of major time series relevant to the economic situation. Since 1983, the BV4 version of the procedure has been used, being a further development of the procedure in terms of methodology by the Federal Statistical Office (s. Nourney, M.: "Umstellung der Zeitreihenanalyse", in: *Wirtschaft und Statistik* 11/1983, and Nourney, M.: "Seasonal adjustment by frequency determined filter procedures", in: *Statistical Journal of the United Nations ECE* 2/1984). That version proved to be quite suitable for use in German official statistics.

At the 54th World Session of the International Statistical Institute (ISI) in Berlin in August 2003, the Federal Statistical Office introduced the new BV4.1 version of the procedure. It will be used at the Federal Statistical Office from 2004. As compared with its predecessor, the BV4.1 version has been improved methodologically regarding the estimation of outlier and calendar effects. Furthermore, the user himself can specify explanatory variables which are to be considered in the analysis (e.g. level shifts, weather-related variables).

For the first time, the Federal Statistical Office also provides an easy-to-use PC software for the Berlin procedure, i.e. its BV4.1 version. This is to promote the dissemination of the procedure - also at international level. The release candidate (RC 1) of the software can be obtained free of charge from the Federal Statistical Office (bv4.1@destatis.de).

### Mathematical bases of the BV4.1 procedure

Part one of the BV4.1 procedure deals with the identification of (potential) outliers. It is based initially on the assumption that within sufficiently short moving time intervals of fixed length - the so-called basic spans - the time series is the realisation of a normal stationary process. That way the conditional distributions of the observations (data) of the time series directly left and/or right outside the particular basic spans (backward and forward identification of outliers) are determined. If the difference between an observation and a thus determined (conditional) expected value exceeds a certain user-defined multiple of its (conditional) standard deviation, then the observation is regarded as an outlier.

If required, the next step is an integrated estimation of outliers, calendar effects and the effects of series-specific user-defined variables. Based on the general additive model for time series decomposition (i.e.  $O = T + S + KA + A + EX + R$  where  $O$  = time series,  $T$  = trend-cycle component,  $S$  = seasonal component,  $KA$  = calendar component,  $A$  = user-defined component,  $EX$  = outlier component, and  $R$  = residual component), this is done using the following linear regression model

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

( $T_i$  = trend-cycle regressors,  $S_i$  = seasonal regressors,  $KA_i$  = calendar regressors,  $A_i$  = series-specific user-defined regressors,  $EX_i$  = series-specific outlier dummy regressors according to the outliers identified in the first part of the procedure,  $\varepsilon$  = error term), but transformed by the linear filter procedure  $F$  of BV4 for trend-cycle and seasonal adjustment. The result is:

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

with  $\varepsilon^* = F(\varepsilon)$  indicating the new error term. As it can be assumed that the following holds true

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

parameters  $\alpha_i$ ,  $\beta_i$ ,  $\gamma_i$  and thus the components  $KA$ ,  $A$  and  $EX$  are estimated by the model

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

using the ordinary least squares criterion.

In part three of the procedure, the trend-cycle component and the seasonal component of the time series to be analysed are estimated, using the time series adjusted for outliers, calendar effects and effects of the user-defined variables. This is done with the respective (linear) BV4 filters.

Generally, the BV4.1 modelling approach for the calendar component is based on eight calendar regressors, namely the series of the differences between the monthly or quarterly numbers of the individual weekdays Monday through Saturday which are not public holidays, the number of Sundays, the number of public holidays and the corresponding average numbers regarding all periods of the same name.

### **Specific features of the procedure**

The BV4.1 procedure is distinguished from other commonly used decomposition methods by the following characteristic features:

Even in the case of rapidly changing seasonal influences, efficient seasonal adjustment is ensured through the use of local regression models.

Trend cycles are depicted plausibly from an economic point of view.

The cost-benefit ratio is low as the user does not need special training or even expert knowledge nor long-term experience with the procedure to make high-quality analyses.

On principle, the analysis results do not depend on the respective user since there is no need to determine any details of the procedure.

As linear regression models are used, on principle there are no differences between indirect and direct analysis results of aggregate series.

### **Features of the BV4.1 software**

The essential features of the BV4.1 software are:

- System requirements: Windows PC (from Windows NT 4.0/Windows 98).
- Support of the file formats EXCEL, ACCESS, SQL-Server and CSV.
- User-friendly graphical user interface.
- Analyses of monthly and quarterly series.
- Possibility of mass production of time series analyses.
- Various possibilities of graphic evaluations of analysis results.

- Possibility to execute so-called successive analyses, i.e. analyses where the analysis spans are extended gradually by one additional period. This option is useful for examining revisions of analysis results originating from the BV4.1 procedure itself.

*Hans-Theo Speth, tel: (+49-611) 75 22 26, e-mail: hans-theo.speth@destatis.de*

## Events

### Scientific Conference on Online Surveys

Since 1995, the Federal Statistical Office has organised joint scientific conferences in co-operation with the Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM - Association of German Market and Social Research Institutes) and the Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI - Working Party of Social Science Institutes) at two-years' intervals. The subject of the fifth joint conference held at the Federal Statistical Office in Wiesbaden on 26 and 27 June 2003 were online surveys. The conference volume published in October 2003 documents the entire event.

Following an introduction into the subject by the presenter (Dr. Manfred Ehling, Federal Statistical Office), three papers provide an overview of a first set of subjects, informing about the state of the art in online surveys in academic social research (Dr. Ulf-Dietrich Reips, University of Zürich), in commercial market research (Hartmut Scheffler, TNS EMNID, Bielefeld) and in official statistics (Dr. Thomas Pricking, Land Office for Data Processing and Statistics of North Rhine-Westphalia, Düsseldorf).

Subsequent to these papers, Erich Wiegand of the Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (ADM) explains the quality standards and professional ethics of web-based data surveys. Dr. Wolfgang Bandilla of the Zentrum für Umfragen, Methoden und Analysen (ZUMA - Centre for Survey Research and Methodology) in Mannheim presents the internet community as a population in statistical terms. Directly related is the paper on drawing samples for online surveys of Christiane Heckel of BIK Aschpurwis + Behrens GmbH in Hamburg.

Frank E. Zander and Thomas Lanninger of Ipsos Access Panels GmbH in Hamburg provide information on the use of access panels as a basis of online surveys. Giving numerous examples, Dr. Michael Bosnjak of Mannheim University illustrates methodological possibilities, topical issues and expansions with regard to online surveys employing questionnaires. The use of multimedia applications in online surveys is described by Dr. Raimund Wildner of GfK AG in Nuremberg. Finally, the contribution of Dr. Frank Knapp of Psyma Online Research GmbH in Rückersdorf informs about online surveys beyond enquiries.

The conference volume has been published as volume 7 of the series "Sozialwissenschaftliche Tagungsberichte" of the Informationszentrum Sozialwissenschaften (Social Science Information Centre). It can be ordered at € 12 plus forwarding charges from the Informationszentrum Sozialwissenschaften, Lennestr. 30, 53113 Bonn; orders may most conveniently be placed via the Internet ordering service of the Informationszentrum Sozialwissenschaften at "<http://www.gesis.org/Bestellen/IZ/index.htm?order/forschungsberichte.htm>".

## **12th Scientific Colloquium “Germany’s Economic Efficiency – Stocktaking and Statistical Measurement in International Comparison“ on 20 and 21 November 2003**

In co-operation with the Ausschuss Methodik Statistischer Erhebungen (Committee on the Methodology of Statistical Surveys) of the Deutsche Statistische Gesellschaft (German Statistical Society), the Federal Statistical Office has organised annual scientific colloquiums since 1992, which have established themselves as a platform for the dialog between the bodies of official statistics and major groups of users from science, business, politics, administration and associations.

After a subject at the intersection of social sciences and economics had been chosen for 2002, namely “Societal Change – Data, Analyses, Accounts”, the 2003 Colloquium on “Germany’s Economic Efficiency – Stocktaking and Statistical Measurement in International Comparison“ centred on issues of a purely economic nature.

The Colloquium was chaired by Prof. Dr. Ullrich Heilemann, who also provided an introduction into the topic. Then Dr. Hans J. Barth of Prognos AG in Basel analysed where the German economy has its place in international comparison on the basis of rather different criteria. In his contribution entitled “Growth – wishful thinking or reality”, Prof. Dr. Michael Hüther of DekaBank in Frankfurt/Main commented specifically on methodological aspects and conclusions for economic policy.

Dr. Michael Grömling of the Institut der deutschen Wirtschaft Köln answered questions relating to the measurability of globalisation trends in the German economy. Prof. Dr. Christoph Spengel of the Justus-Liebig-Universität in Gießen considered the profits and tax burden of enterprises in international comparison and analysed to what extent they may serve as efficiency indicators.

The first day ended with a ceremony celebrating the presentation of the Gerhard Fürst Award, which is granted each year by the Federal Statistical Office for outstanding work on issues related with official statistics. The Gerhard Fürst Promotion Prize for junior academic staff was awarded this year to Dr. Michael Gold (University of Lüneburg) for his dissertation entitled “Mikroökonomische Analyse der Arbeitsnachfrage - Eine Untersuchung von Beschäftigungsdynamik und Überstundennachfrage unter dem Einfluss von Anpassungskosten mit deutschen Betriebsdatensätzen“ and to Dr. Michael Niese (Freiberg University of Mining and Technology) for his dissertation on “Ursachen von Betriebsschließungen und ihre Konsequenzen für die Marktdynamik – Eine mikroökonomische Analyse von Probezeiten und Todesschatten im Verarbeitenden Gewerbe“. The Gerhard Fürst Promotion Prize was granted also to Simon Krotter (University of Regensburg) for his diploma dissertation on “Kapitalkosten und Kapitalstrukturen ausgewählter deutscher Unternehmen - eine empirische Untersuchung“.

On the second day of the Scientific Colloquium, Birgit Mackenthun of Berlin Technical University dealt with the health system as a growth factor. Performance indicators in the educational system and naturally also the PISA study were analysed by PD Dr. Stefan C. Wolter of the Schweizerische Koordinationsstelle für Bildungsforschung (Swiss Coordination Centre for Research in Education) located in Aarau. The monthly survey of the ILO activity status as an instrument of labour market monitoring in Germany was presented by Thomas Riede of the Federal Statistical Office. Finally, Achim Kemmerling of Berlin Freie Universität discussed in how far gross and net social expenditure ratios are suitable for measuring welfare state payments.

*Christian König, tel: (+49-611) 75 20 77, e-mail: christian.koenig@destatis.de*



## Calendar of events

### **European Conference on Quality and Methodology in Official Statistics (Q2004) in Mainz from 24 - 26 May 2004**

From 24 to 26 May 2004, the Federal Statistical Office will hold the European Conference on Quality and Methodology in Official Statistics (Q2004) in the *Kurfürstliches Schloss* in Mainz. The event supported by Eurostat will be the first in a series of biennial European conferences on issues of quality and methodology of official statistics. The goal is to create a regular forum for methodological issues and new developments in European official statistics. The event builds on the positive experience of the International Conference on Quality in Official Statistics (Q2001) held in Stockholm in 2001.

About 400 participants, mainly from Europe and North America, are expected for Q2004. The topics of the Conference will embrace all quality-related aspects of official statistics. Topics will encompass not only the processes of official statistics production but also issues of measuring statistics quality and quality management. The first day of the conference will also include three parallel one-day seminars on the subjects „Survey Quality“, „Quality Management in Statistical Institutes“, and „Variance Estimation in Complex Surveys“. The seminars will be conducted by internationally renowned experts. The conference language is English.

The content of the Conference will be prepared by a programme committee consisting of experts from a total of seven National Statistical Institutes and Eurostat as well as two university experts. The event will be chaired jointly by Eurostat and the Federal Statistical Office. What will also be presented at the conference is the results of the European project DACSEIS (Data Quality in Complex Surveys within the New European Information Society). A call for papers met with an extraordinarily positive response. The programme committee is currently compiling the conference programme, which will be published in the Internet by end of January 2004.

Further information on the Conference and its programme as well as the registration form are available on the Internet at <http://q2004.destatis.de>. For any further information on the conference please contact the programme committee chairmen, Werner Grünewald (tel. +3 52 / 43 01 / 33 28 0; e-mail: [werner.gruenewald@cec.eu.int](mailto:werner.gruenewald@cec.eu.int)), or Thomas Körner (tel. +49 (0) 611 / 75 44 13; e-mail: [thomas.koerner@destatis.de](mailto:thomas.koerner@destatis.de)).