

METHODS – APPROACHES – DEVELOPMENTS

Information of the German Federal Statistical Office

Edition 1/2006

The catchword

Experimental computation of an index of perceived inflation (IWI)..... 3

Methods of federal statistics - Further development

Anonymity of survey units and statistical confidentiality in digital maps of official statistical data 5

Flash estimation of the quarterly gross domestic product..... 7

New statistics of public finance in Germany - increased relevance and timeliness 9

Using administrative data for surveys of the cattle population 12

Publications

Volume on the Scientific Conference „Data fusion and data integration“ 15

Events

Fourth user conference on "Research with the microcensus: Analyses of the social structure and labour market" 16

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 information

on the data supply:

Information service,
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

Periodicity: (generally) twice a year

Published in 2006

© Statistisches Bundesamt, Wiesbaden 2006

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

Experimental computation of an index of perceived inflation (*Index der wahrgenommenen Inflation - IWI*)

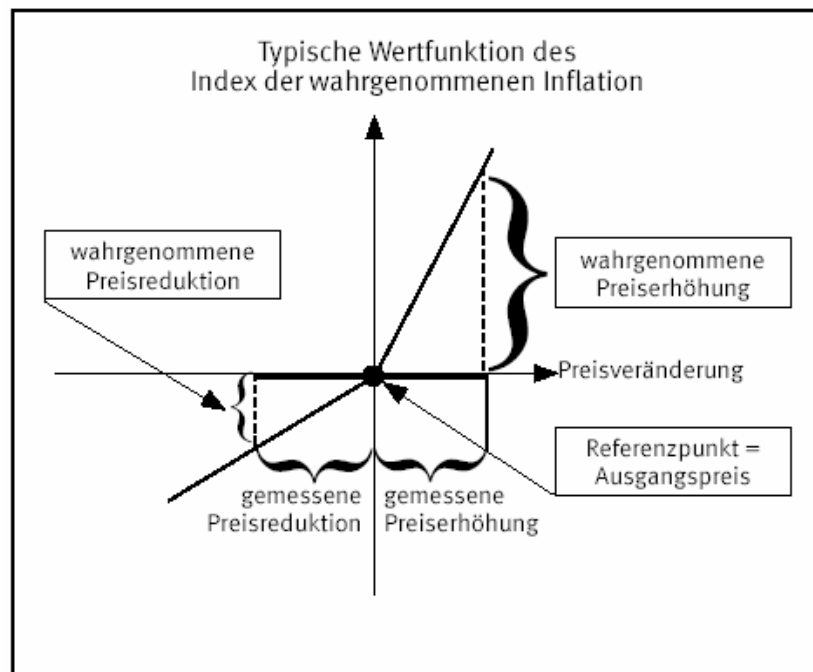
After the introduction of euro banknotes and coins in January 2002, official statistics faced the phenomenon that the change of consumer prices as recorded by statistics differed considerably from the inflation perceived by the population. While the measured rate of inflation showed a very moderate trend, the general public was convinced that the euro had led to massive price rises. Such divergence between “perceived” and measured inflation was not only observed in Germany, it emerged in all European Union countries in which the changeover to euro banknotes and coins had been performed. Public indignation about the high perceived inflation was directly followed by indignation about the statements of official statistics contradicting that perception.

In a joint project of the Federal Statistical Office and Prof. Dr. Hans Wolfgang Brachinger of Université de Fribourg (Switzerland), an attempt was made to explain that divergence between perceived and measured inflation through a scientific model and to illustrate it by means of the official figures used to calculate the German consumer price index.

The “theory of inflation perception” of Prof. Brachinger comprises three major elements:

1. Based on the Prospect Theory of Kahnemann and Tversky¹, it is assumed that the prices of goods with which the consumer is confronted are coded as profits and losses in relation to goods-specific reference prices and are perceived asymmetrically. In that context, price rises are rated higher than corresponding price reductions. Such different rating is reflected by the so-called “loss aversion parameter” which in any case is larger than 1. In the model calculations, loss aversion parameters of 1.5, 2 and 2.5 were assumed.

Chart 1 (chart 2 from WiSta 9/2005, p. 1004)



¹ Kahnemann, D./Tversky, A.: “A Prospect Theory: An Analysis of Decision under Risk” in *Econometrica* 47 (1979), No. 2, pp. 263 et seqq.; Tversky, A./Kahnemann, D.: “Loss Aversion in Riskless Choice: A Reference-Dependent Model” in *The Quarterly Journal of Economics* 106 (1991), No. 4, pp. 1039 et seqq.

2. It is assumed that a major factor influencing the perception of inflation is how often the consumer is confronted with a price change. In consumer price statistics, relative price changes are aggregated to a rate of price change on the basis of the relevant expenditure shares², whereas the weights used by the IWI are frequencies of purchase. The most important groups of goods for the consumer price index (CPI) in Germany are net rents exclusive of heating expenses, petrol, travels, and car purchases, whereas for the IWI it is the daily newspaper, cigarettes, draught beer and rolls.
3. The third major element is the definition of a reference period for the price. The usual presentation of inflation rates in official statistics refers to a price that had to be paid exactly a year earlier. The reference period for the price³ is irrelevant for the CPI result because it is just a basis for adjustment. For the IWI, the price reference period is very important because it is used to rate prices as profits or losses and to attach a loss aversion parameter to them. In the model calculations, various alternatives of that “anchor” were implemented, that is the average price of the last two years, the average price of the last year before the introduction of euro banknotes and coins, and an average price of all years from the last DM year to the year preceding the IWI calculation.

The results of the calculations are largely in line with expectations. The IWI generally is above the CPI results, and the more clearly so the larger the chosen loss aversion parameter and the longer the average time lag between the reference period and the reporting period. There are, however, some things to be noted:

1. In 1999 and 2000, the IWI is below the CPI results although the price rises were rated with a loss aversion parameter of higher than 1. At that time, there was extreme price competition among discounters, the main reason being probably that a foreign industry giant was trying to break into the German market. Such competition has a particular impact on prices of cheap and frequently bought goods.
2. In 2001 and 2002, the IWI is in part considerably above the CPI results. The peak is in January 2002, that is exactly the month when the euro notes and coins were introduced. Although the two curves start to diverge already one year before the introduction of euro notes and coins, the trend of the curve can be used to explain why the euro is perceived as having an upward effect on the price trend.
3. From the middle of 2002, the IWI and CPI curves have been reconverging.⁴

Conclusion: Calculating an IWI is a suitable tool to explain in scientific terms the divergence between “perceived” and measured inflation as a result of the introduction of the euro notes and coins. A causal connection between the introduction of euro notes and coins and particular price rises cannot be observed with that approach.⁵

² Expenditure of households on a specific group of goods, measured by total consumption expenditure of all households in the economic territory.

³ The computation of the German consumer price index is currently based on price base 2000.

⁴ Of course, the IWI curves on average are steeper than the CPI curve, and the more so the higher the loss aversion parameter is chosen.

⁵ There were too many special factors that influenced the price trend around the time of the introduction of euro notes and coins. Factors to be mentioned here are the end of the price war among discounters, the tobacco tax rise in January 2002, and the beginning winter in Southern Europe which led to a scarcity in supply of fresh fruit and vegetables.

For more detailed information on the project please refer to two papers published in *Wirtschaft und Statistik*, no. 9 /2005. In his invited paper “The euro pushing up inflation? Perceived inflation in Germany”, Prof. Dr. Hans Wolfgang Brachinger introduces his theory of inflation perception and presents major results. In their paper “Measuring perceived inflation in Germany: Determination of purchase frequencies by the Federal Statistical Office”, Dr. Sabine Bechtold, Günther Elbel and Heinz-Peter Hannappel explain how the frequency weights were calculated for IWI computation.

Günther Elbel, tel: +49-(0)611 /75 26 27, e-mail: guenther.elbel@destatis.de

Methods of federal statistics – Further development

Anonymity of survey units and statistical confidentiality in digital maps of official statistical data

Introduction

At present, the statistical offices of the Federation and the Länder publish most of their results of surveys and calculations in the form of tables. Maps as an alternative basis for the presentation and analysis of official statistical data are not used often for purposes of publication. Their use is usually limited to selected information of regional statistics from the Land level down to municipality level (e. g. www.destatis.de/onlineatlas/ and www.statistik.baden-wuerttemberg.de/intermaktiv/).

Maps are better suited than tables to illustrate statistical items with spatial distribution patterns. In a digital form, maps are also suitable for spatial evaluations and the production of new statistical information through linking various map strata from different work areas (e. g. from statistics and from land surveying) in so-called geo-information systems (GIS). However, the quality and information value of such analyses strongly depends on the degree of detail of the space-related statistical data.

As mentioned above, the spatial resolution of currently published maps of official statistics is rather small, at best going down to municipality level. Especially for analyses, it would however be desirable to have a degree of detail going down below that administrative level. However, such maps cannot be created at present because data of official statistics in Germany cannot be allocated to smaller spatial units. For that purpose, information would have to be available on where exactly a survey unit is located (e. g. company building or residential building). If that were available, the statistical information could exactly be allocated in spatial terms.

Due to the legal provisions applicable, it is generally not allowed for official statistics to permanently store, and use for publications, the information on the location of a survey unit (the address). However, over the last few years, there has been a sharp rise in the general demand for data that can be presented and analysed in spatial terms. This also refers to official statistical data which generally have a clear relation to space regarding their survey basis. Official statistics therefore endeavours to arrive at a situation in the medium term where it would be allowed at least to permanently store – in especially protected computer systems – addresses of buildings with statistical data on the companies or persons residing there. For the publication of statistical data or results of spatial analyses in maps, however, it will be necessary to apply methods ensuring the anonymity of survey units and meeting the statistical confidentiality requirements.

Anonymity of survey units in digital maps

Addresses of survey units – and their values – can be shown as dots in a digital map (chart 1a). However, that kind of presentation is not suitable for reasons of statistical confidentiality, further processing in the GIS and possibilities of cartographic design. It is better to combine the dot-shaped information to form grid cells (chart 1b) and to present only the aggregated values of the survey units in those cells (chart 1c). Consequently, individual survey units with their exact values can no longer be located.

For anonymisation, it would of course also be possible for the values of survey units shown as dots (cf. chart 1a) to be aggregated in municipality areas (chart 1d). In that case, however, the possible degree of detail and the density of information that could be presented would be much lower than in a grid map (chart 1c), and there would be no additional benefit compared with the current form of presentation of tabular aggregated statistical data in administrative units. Also, processing digital map information on the basis of irregular administrative units is highly complex in terms of methodology. Grid maps, however, are easy to compile and to analyse.

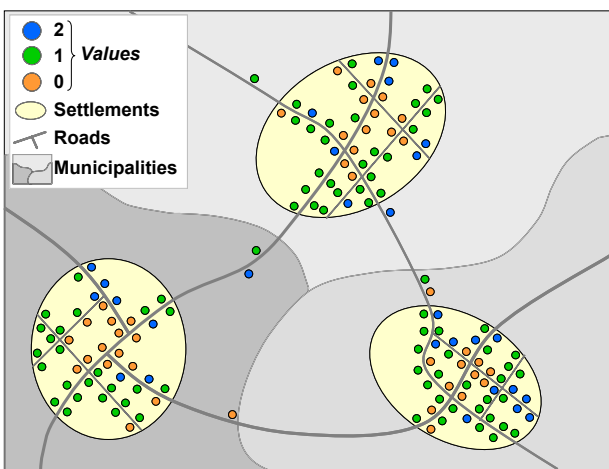


Chart 1a: Values of survey units in a fictitious landscape section.

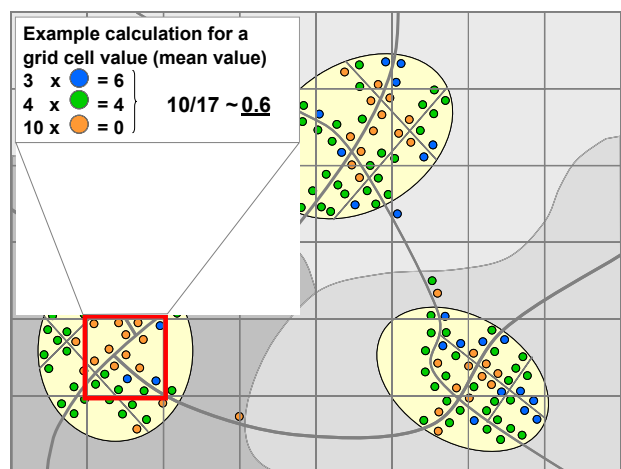


Chart 1b: Aggregation of values of survey units in grid cells.



Chart 1c: Rounded mean values of values in grid cells.

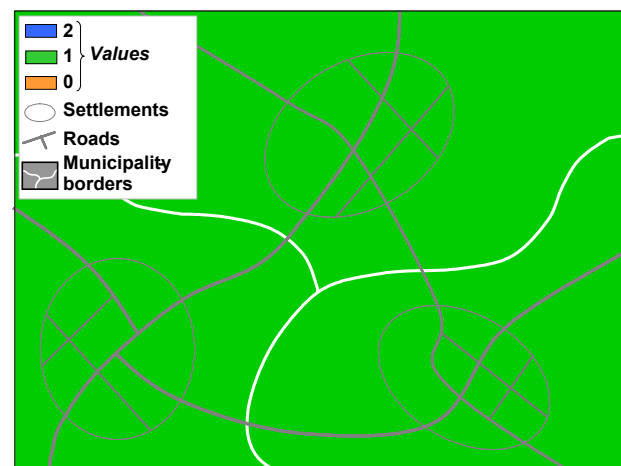


Chart 1d: Rounded mean values of values in municipalities

Statistical confidentiality in grid maps

In official statistics, a distinction is made between primary and secondary confidentiality in tabular publications. In primary confidentiality, it is checked whether the number of survey units whose values contribute to the table value is below a specific threshold or whether individual values of survey units stand out. If one of the two conditions applies, the relevant table value is not indicated (suppressed). In secondary confidentiality, it is checked whether suppressed table values can possibly be calculated or estimated by means of row, column or table totals. If this is possible, further values of the table must be suppressed. As another measure, table data can be presented in less detail by classifying the data.

Grid maps can be considered as tables with two spatial breakdowns, that is one in a north-south direction and another in a west-east direction. Therefore, the above-described methods of ensuring statistical confidentiality in tables generally can be applied to this type of maps. As an alternative to the methods developed for tables, it might be possible to apply methods which so far have not become established in official statistics because most of the statistical data are published in the form of tables. This includes, for example, filter methods which are used in digital processing of satellite and aerial photographs and methods of geostatistics which are suitable for spatial analysis of digital maps.

To ensure the statistical confidentiality in grid maps while adhering to the rules and procedures applying to tables, the Federal Statistical Office has already performed methodical studies. The application of filter methods and geostatistical methods for confidentiality and anonymisation purposes is still in an experimental phase.

Dr. Martin Szibalski, tel: +49-(0)611 /75 36 76, e-mail: martin.szibalski@destatis.de

Flash estimation of the quarterly gross domestic product

Background

In view of the advancing European integration and an increasingly harmonised economic, finance, and monetary policy in the euro area, there is a growing demand for informative statistics at the European level. Improved reliability and timeliness are considered as major quality criteria in further developing the European Statistical System (ESS). The demand on the part of political decision-makers in the European Union (EU) for earlier availability of quarterly data on the gross domestic product (GDP) and its major aggregates has grown also from the aspect of comparability with data of non-European economies.

Currently, Germany releases quarterly GDP data (including changes from the preceding quarter and the relevant quarter of the previous year) 45 days after the end of the reference quarter, thus making an important contribution to estimating the GDP growth for the euro area and the total EU. More detailed breakdowns of the GDP by economic sectors and for the use aggregates can currently be provided by Germany about 55 days after the end of the quarter.

Goal and progress of the study

Against that background, the Federal Statistical Office started in 2002 to perform studies – with the support of Eurostat – on a more rapid estimation of the quarterly GDP. The goal of the study is to find out whether more up-to-date gross domestic product data can be estimated as early as about 30 days after the end of the quarter with sufficient reliability and accuracy. The results can also provide a basis for supplying more detailed GDP breakdowns by major aggregates as early as 45 days after the end of the quarter.

At the end of September 2003, a first project report was available, followed by the second in February 2005. The studies continue, and it is expected that in early 2006 a third report will present the new findings on flash estimation, especially in the context of the large-scale revision of national accounts in 2005.

The activities performed in the context of the feasibility study on GDP flash estimation can roughly be subdivided into four work packages:

- examining the estimation methods of other national statistical institutes,
- developing a suitable forecasting model (“three-pillar forecast”),
- testing the developed forecasting model under “real” conditions,
- estimating the efforts required and the accuracy of the model when applied in permanent operation.

The activities centre on developing the forecasting model and on checking it during practical tests. The purpose of such practical tests is to judge the statistical quality, the procedure and the organisation of the estimations. What is utilised for that model is the experience acquired in examining the estimation methods of other national statistical institutes. Estimating the efforts required is necessary to determine the capacities needed for a possible permanent operation. The accuracy of the model provides information needed to decide about future permanent operation.

The three-pillar model for short-term estimation of the quarterly GDP

The GDP flash estimation method that was developed during the study and has now been elaborated is based on three pillars. These constitutive elements are

- pillar 1: econometric forecast,
- pillar 2: expert forecast,
- pillar 3: co-ordinated forecast.

The econometric forecast provides purely quantitative estimates of the price-adjusted (real) original values of the aggregates of the production and uses sides of the GDP on the basis of data available 27 days after the end of the quarter. Adding up the estimates of the production-side aggregates provides a production-side estimate of the price-adjusted GDP. The same applies to the components of the uses side.

In methodical terms, the forecasting methods used are ARIMA models. Depending on the data situation, those methods known from time series analyses are used to forecast monthly or quarterly indicators which are then used to update the trend of an aggregate or to forecast the aggregate itself. The selection of ARIMA models and their specification are based on ex-post forecasts for the last twelve quarters or 36 months.

The expert forecast consists of the estimations performed by the specialised units responsible for the production side and the uses side. What is allowed here is both quantitative estimations, i.e. calculations, and qualitative information, i.e. subjective assessments on the real quarterly trend (compared with the same quarter a year earlier).

The co-ordinated forecast is based on estimates of the econometric forecast and the expert forecast. In a multi-stage procedure, the differing estimates of the aggregates, components and the GDP are taken to produce a co-ordinated forecast of the real year-on-year trend.

Overall assessment and outlook

Altogether, the developed method of the three-pillar model for a GDP flash estimate to be available already about 30 days after the end of the reference quarter can be considered as feasible from the aspects of both subject-matter and organisation.

It has not been possible yet to take a decision on the permanent application of the method, including continuous publication of such an early estimate of the quarterly GDP in Germany. A possible start date for permanent use depends mainly on the more difficult data quality assessment after the large-scale national accounts revision was performed in Germany in April 2005. So far, the methodical changes implemented as part of the revision have not fully been taken account of in the estimation method. Solutions for the changeover of computations to the previous year's price basis (chain indices) and for the changed recording of financial intermediation services indirectly measured (FISIM) are currently developed and tested on the basis of the revised traditional quarterly accounts. Whether and when the estimation method developed in the project will be applied in permanent operation will depend on the results of those studies and on the availability of capacities needed.

For further information on the GDP flash estimation please refer to the relevant paper in the periodical *Wirtschaft und Statistik* (Hartmann, N./Schmidt, J./Oltmanns, E.: *Schnellschätzungen für das vierteljährliche Bruttoinlandsprodukt. Ergebnisse einer Machbarkeitsstudie*, in: *WiSta* 7/2005, pp. 682 et seqq., to be obtained from the Statistics Shop of the [Federal Statistical Office](#)) and the latest project report (*Statistisches Bundesamt: Machbarkeitsstudie zu Schnellschätzungen für das vierteljährliche Bruttoinlandsprodukt*, Wiesbaden: 2005 [copy]). The latter may be obtained upon request from the author.

Dr. Erich Oltmanns, tel: +49-(0)611 / 75 20 87, e-mail: erich.oltmanns@destatis.de

New statistics of public finance in Germany – increased relevance and timeliness

The former scope of finance statistics

In Germany, the statistics of public finance (finance statistics) have the important function of producing an overall picture of public finance by using the data of the largely independent budgets of the Federation, Länder, municipalities and social insurance institutions. The purpose of this overall picture is to serve as a basis for central decisions in the spheres of financial, economic and monetary policies and in specialised areas of politics. The finance data are directly incorporated into the general government account in national accounts and thus help to ensure the international comparability of public finance within the European Economic and Monetary Union.

The computer-based direct collection of finance data from public budgetary, cash and accounting files provides a differentiated insight into public finance. The very detailed classification of public revenue and expenditure by economic and financial types in the quarterly cash statistics permits a timely assessment of the structure of public finance and its overall economic effects. The statistics reveal in particular whether the general government's revenue was not sufficient to cover expenditure in a specific period and to what extent saved funds (reserves) or even external funds had to be used (financial deficit). The additional functional breakdown of all revenue and expenditure recorded in the annual statistics of accounts shows the functions for which general government funds were used and the domains in which general government generated its own revenue. Data on the debts of public budgets are of utmost importance for assessing public finance, especially within the framework of the EU Stability Pact. Statistical information on the

amount and composition of debts (in terms of types), on maturity periods of loans and on debt maturities reveals the constraints imposed on future budgets by debt servicing and the ensuing loss of room for financial manoeuvre with respect to future functions.

Overall, the data of finance statistics have so far ensured a rather good analysis of public finance in the year under review and of certain debt-related effects in subsequent years.

Fundamental changes in public budget management

At the present time, fundamental changes are under way in public budget management in Germany due to shrinking financial resources, but also because of scientific impulses and incentives from abroad. In 2003, the Interior Ministers of the Länder agreed that the entire budget management in the municipalities and associations of municipalities should gradually switch over to modern double-entry accounting. Initial approaches towards double-entry accounting have also been made in the Länder budgets. To an increasing extent, specific facilities and functions are separated out from the traditional budgets of the Federation, Länder and municipalities and are operated as independent entities with a commercial accounting system, although they still form part of the general government sector as defined by the European System of Accounts. The amendments to the Law on Finance and Personnel Statistics of June 2005 make allowance for this development and will improve the statistical reporting system on public finance in Germany:

- New system of double-entry accounting to provide the basis for finance statistics

Since 2005, the municipalities in North-Rhine Westphalia have been authorised to apply the new system of double-entry accounting, and from 2006 the municipalities in Lower Saxony, Hesse and Saxony-Anhalt will follow suit. Other Länder are set to authorise their municipalities in 2007 to switch over to the new accounting system. During a transition period, the municipalities are free to decide when to switch over to double-entry accounting. Due to the extensive and costly preparations (training of staff, procurement of adequate software, first valuation of assets), the transition will only be possible on a step-by-step basis. At first, municipalities with the traditional cash-based cameralistic accounting of revenue and expenditure will form the majority.

Nevertheless, the necessary adjustments have been made to the statistics of municipal public finance to enable the municipalities from 2006 onwards to directly report the requested finance data according to the new classification for the framework of accounts and products. During the initial phase, with many municipalities still using the traditional accounting system, only specific accounting components of the information reported by municipalities with double-entry accounting can be used for finance statistics (see figure). The financial account contains all inpayments and outpayments and thus provides an interface for the new framework of accounts. This interface permits to compare the revenue and expenditure data with those of municipalities maintaining a cash-based cameralistic accounting system and to combine all the finance data to provide an overall picture of municipal public finance in a specific Land or in Germany as a whole. In consultation with the Interior Ministries of the Länder, the framework of accounts has been designed in such a way as to enable the municipalities to supply the information required for debt statistics and for the new statistics of public financial assets directly from their double-entry accounting files (see figure 1; the blue-coloured area indicates the section of public finance that has hitherto been presented in statistics). When double-entry accounting has been introduced in all municipalities and associations of municipalities at the end of the transition period, the accounting system will also make it possible – with no extra effort or expense:

- to directly extract data on municipal tangible assets in a detailed structural breakdown for statistical purposes,
- to determine other liabilities apart from debts and to complement the statistics on the liabilities side of the municipal balance sheet, and

- to cover and present municipal finance data according to the new consumption-of-resources concept.

By establishing three components in the new municipal accounting system, i.e. the asset and liability account, income account and financial account, the Interior Ministries of the Länder created a useful symbiosis between cash-based cameralistic and double-entry accounting. For this reason, the coverage and presentation of revenue and expenditure at the time of their entering or leaving cash holdings should be retained so as to benefit from this symbiosis. On the one hand, it permits to produce an integrated statistical picture of municipal assets and the resources received and used. On the other, it allows to present the generation of and changes in liquidity according to causes.

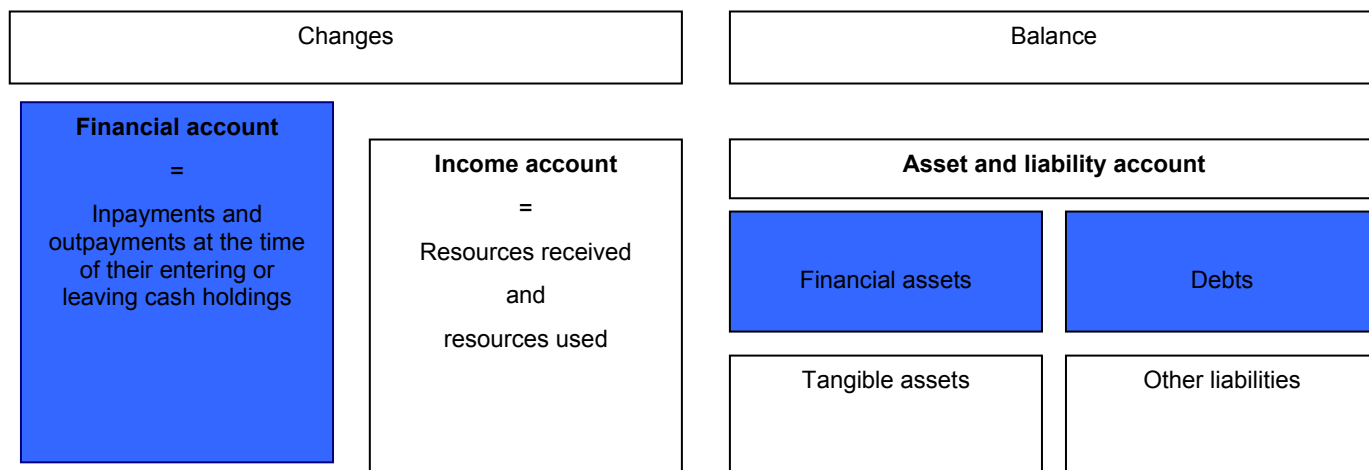


Figure 1: Statistical presentation of public finance

At present, consultations are also held among the Federal and Länder Ministries of Finance to set up a consistent framework of accounts and a consistent product plan. However, the process of reform is progressing quite slowly. Unfortunately, the common draft for a framework of accounts does not comprise an integrated liquidity account (financial account). Hence, only an ex-post analysis will permit to determine the liquidity status by way of the cash flow derived from the income account.

- Up-to-date reporting on the financial management of separated-out public facilities

Based on federal legislation (Hospital Financing Law) and legal initiatives in specific Länder (waste-water disposal in Rhineland-Palatinate), public facilities have been separated out from the core budgets of the Länder and municipalities since the mid-seventies to be operated under their own commercial accounting system. The increasing pressure on public budgets, combined with general efforts towards independent and economically efficient action, has enhanced the process of separating facilities out from a wide variety of public sector domains. Nevertheless, this trend shows a very different intensity in the individual Länder. The annual balance sheet statistics on public funds, institutions and businesses with an accounting system of their own have been extended to cover all public domains. As a result, the financial data of the separated-out facilities can be reintegrated into government finance statistics, thereby ensuring a high degree of comparability between the public functions. The majority of the facilities separated out from the public budgets are not part of the general government sector according to the definition of the European System of Accounts (ESA 95) (see figure 2). Consequently, the current statistical picture of finance in the core section of general government has remained unaffected as a whole. However, the most recent wave of separating facilities out from the Länder budgets has concerned agencies (institutions of higher education, road construction authorities, statistical offices) that are part of the general government sector as defined by ESA 95, and therefore required swift action by public finance statisticians. Amendments to the Law on Finance and

Personnel Statistics in 2005 created the basis for collecting finance data (revenue, expenditure, debts) on separated-out facilities of the general government sector also with a quarterly periodicity so that the statistics will continue to provide an up-to-date overall picture of government finance.

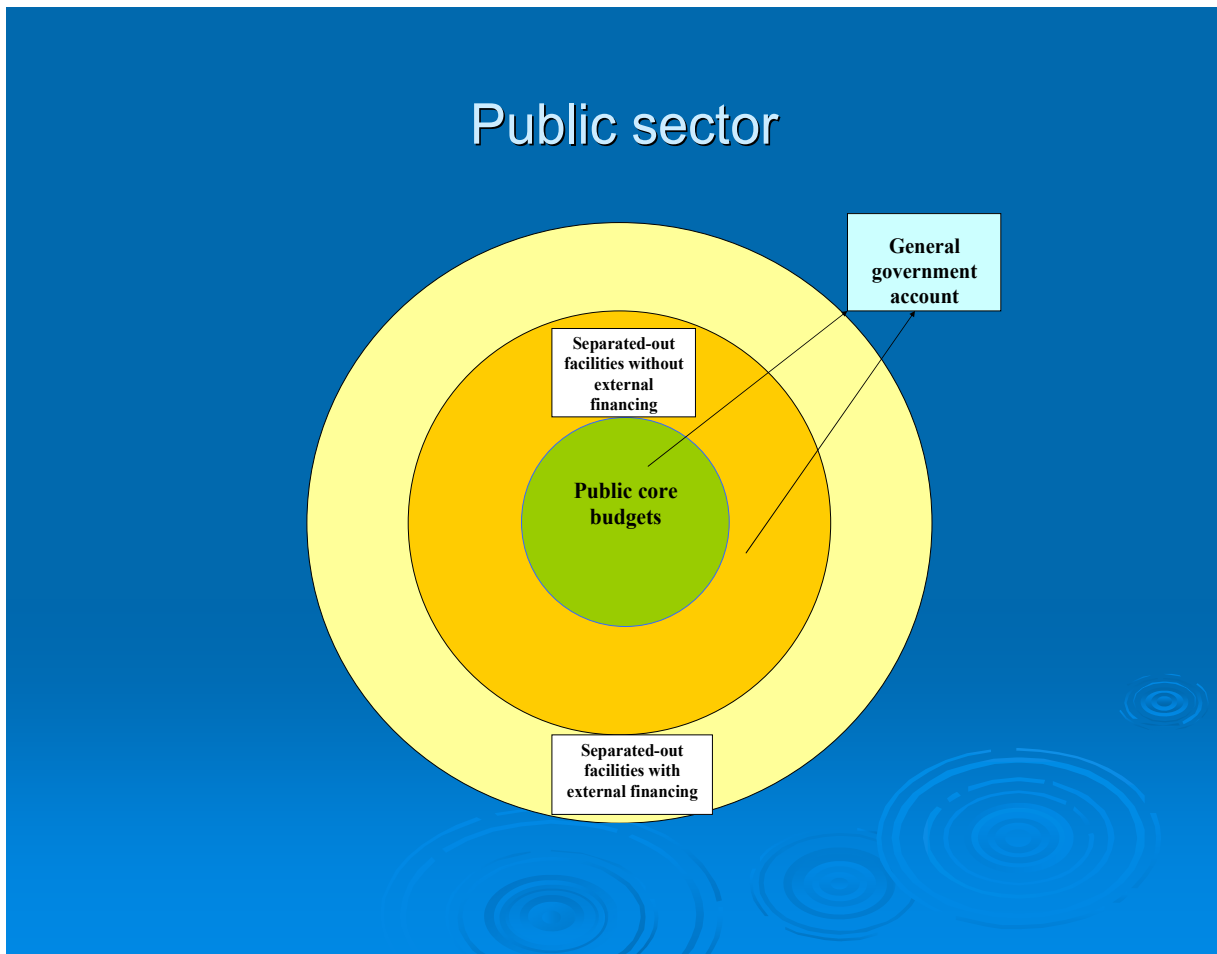


Figure 2: Representation of the public sector in finance statistics

On the whole, the amendments to the Law on Finance and Personnel Statistics have led to significant improvements in the statistical reporting on public finance. The picture of public finance has been complemented by the important component of public assets, and information on the financial management of the public facilities that have been separated out from the federal, Länder and municipal budgets will be available on a timely basis. In mid-2006, a detailed description of the reporting system for finance statistics will be published in the journal "Wirtschaft und Statistik" (Economy and Statistics).

Hans Rehm, tel: (+49-611) 75-4174, e-mail: hans.rehm@destatis.de

Using administrative data for surveys of the cattle population

The Federal Statistical Office and the statistical offices of the Länder are expected to use their resources for producing statistics efficiently and to impose as little burden as possible on the respondents. One means of saving resources and reducing the burden on respondents is to use administrative data. The Identification and Information System for Animals [(*Herkunftssicherungs- und Informationssystem für Tiere* (HIT))] is one of the sources for such administrative data. In the medium term, it may be possible to substitute these data for the questions covering variables on cattle in the survey of livestock populations. This contribution describes the methodology to be applied in implementing the transition. The planned methods have been developed through a number of methodological studies conducted at the Federal Statistical Office.

The bases of the survey of livestock populations

The survey of livestock populations records the number of animals and several other variables for various kinds of animals for production as of a specific reference day. As regards cattle, the survey covers the variables 'age, by age class' and 'intended use'. The statistics of livestock populations are based on a decentralised survey which is conducted by the statistical offices of the Länder with compulsory response for farmers. However, not all livestock farmers are obliged to provide information. Compulsory response only applies to agricultural holdings with a utilised agricultural area (UUA) of at least 2 ha on the reference day, a livestock of at least 8 heads of cattle (or 8 pigs, 20 sheep or 200 heads of poultry), 30 acres of specialist crops, or 3 acres of crops under glass. The reference days are 3 May and 3 November each year.

Description of the cattle database

The Identification and Information System for Animals stores a number of variables on cattle from birth via movement to death in a central database. The database is operated by the relevant regional agencies of the Länder. All the data are held centrally by the Bavarian State Ministry for Agriculture and Forestry in Munich.

Every keeper of cattle in Germany is obliged to report changes in his stock of cattle within 7 days. These reports may be submitted by post, telephone or internet. In addition to the reason for reporting (birth, death or movement of an animal), the database records the date of the event, the reporting date, the reporting channel and the identification number of the relevant holding. Hence, later corrections are possible and fully traceable, and the reports take effect retroactively from the date of the event.

The structure of the HIT database is linked to individual animals. All the data on each individual animal, which can be clearly identified through its ear tag number, are stored in the database. Consequently, each animal can be attributed at any time to an agricultural holding, which, in turn, is clearly defined by its identification number. The holding identification number in the HIT database contains a regional code which permits to identify all keepers of cattle in a Land.

Methodology

As described above, all keepers of cattle in Germany are obliged to register their stock of cattle in the HIT database. In addition to the respondents for the survey of livestock populations, this applies to small agricultural holdings which do not exceed the thresholds outlined above and to keepers of cattle who do not operate an agricultural holding (e.g. cattle traders, slaughterhouses). This means that the number of cattle keepers registered in the HIT database is larger than the number recorded by official statistics, which ensures a complete representation of the cattle population in Germany.

Nevertheless, the database does not contain all the variables required for the survey of livestock populations. What is lacking in particular is information on the intended use of female bovine animals aged over 1 year, which is needed for purposes of statistical processing.

As part of the required transition work, it transpired soon that it would not be possible to include data on the use of each animal in the HIT database. The ensuing reporting requirements would dramatically increase the burden on the respondents and therefore are not acceptable.

To reduce the costs and efforts involved in data maintenance, an indirect approach has been suggested under which the intended use is not indicated for each individual animal. This approach makes use of the variable 'breed', which is the only variable in the database that can be linked to the intended use. Based on the characteristics of the different breeds of cattle, the following groups of cattle can be distinguished: dairy cattle (animals producing a lot of milk), beef cattle (animals that gain a lot of weight in relation to the amount of feed offered but produce less milk) and dual-use breeds (animals producing a lot of milk and gaining quite a lot of weight). The studies conducted have shown however that the intended use cannot be adequately ascertained through the variable 'breed' alone, because dual-use breeds play an important role in some regions of Germany where they are or can be used for milk and beef production. Hence, an extended approach is necessary to determine the animals' intended use.

The extended approach starts from the assumption that agricultural holdings show an increasing tendency towards specialisation, i.e. agricultural holdings practise only a few or just one of the wide variety of types of farming. This holds true also for cattle farming which can be categorised into the following types: dairy farming, cattle rearing and cattle fattening. Besides these specialist types of farming, there are agricultural holdings that practise any conceivable combination of the various types of farming. When considering specialist agricultural holdings which practise just one type of farming (these account for about 80% of all cattle-keeping holdings), the intended use of cows (female animals after first calving) can be derived directly. As regards holdings with several types of farming, the intended use can be determined through a matrix combining the type of farming with the kind of breed.

The holding's type of farming is not indicated in the HIT database, either. However, under Article 24b of the Ordinance on Livestock Movement all agricultural holdings must notify the relevant authority if they keep cattle and immediately report any changes. This means that the required variable is available in administrative files and in principle can be used.

Female bovine animals are categorised into cows, heifers (female animals aged 2 years and over before first calving) and female animals aged 1 but less than 2 years. The HIT database does not reveal whether a female animal has had her first calf until a birth registration for a calf is submitted with the mother's ear tag number. The obligation to register, however, only applies to calves that do not die or are not culled before their 7th day of life. The number of such calves is quite considerable so that, due to the inherent systematic error, the recorded number of cows is too small and that of heifers too large. The time of first calving usually is between the 24th and 36th month of age. Since the number of older heifers is irrelevant in terms of quantity, heifers aged over 36 months can be counted as cows. This method is a reliable compensation for the missing number of births.

For females that have not had a calf and are aged 1 year or over, the intended use depends on what further use is made of them. A rather constant proportion of such animals is intended for slaughter, while the remaining animals are intended for breeding or other use. Due to the rather constant relation between animals for slaughter and animals for production, the relevant shares of each group can be estimated using a coefficient. This coefficient can be defined as the proportion of animals slaughtered during the last reference period in the total number of the relevant stock of cattle. If this coefficient is multiplied by the number of relevant animals on the reference day, the number of animals intended for slaughter can be calculated.

The HIT database does contain information on slaughtering. Consequently, the above coefficients can be directly computed from the HIT database by determining the number of slaughtered animals among the total cattle population on a given date (e.g. 1 year prior to the survey date).

Findings

Using the methodology described above, several comparative studies have been conducted on the basis of data from livestock surveys (May 2001, May 2003, Nov. 2004) and the corresponding HIT data. The studies were aimed at reproducing the results of the surveys of

livestock populations by means of the HIT data. The degree of compatibility between the livestock surveys and the HIT database that has been revealed by the studies is considered to be appropriate.

The number of cattle, in a breakdown by age and sex, can be directly derived from the HIT database and largely corresponds to the results of the livestock surveys for the respective units. As regards the intended use of cows, there is a strong compatibility with the results of the livestock surveys.

To collect data on the use of cattle, the livestock survey covers the intended use of the animals. However, the prognosis made at the time of the survey is rather uncertain because the intended use can change after the relevant information has been entered in the questionnaire.

Consequently, the comparison of results on these animals shows higher differences because these uses are determined by means of retrospective coefficients.

Outlook

The prerequisite for applying the methodology outlined above is that the type of farming is recorded in the HIT database as an auxiliary variable. Without this variable, it is not possible to obtain the information on the use of female animals that is not included in the HIT database. As soon as the required legislation has been adopted and the variable 'type of farming' has been integrated in the database with an adequate degree of quality, the survey of livestock populations (cattle population) is to be replaced by using the HIT database.

For further information on this project, please refer to the journal "Economy and Statistics" (*Wirtschaft und Statistik*), no. 8/2004, p. 845.

<http://www.ec.destatis.de/csp/shop/sfg/bpm.html.cms.cBroker.cls?cmspath=struktur,vollanzeige.csp&ID=1014757>

Dr. Matthias Walther, tel: (+49-611) 75-8654, e-mail: matthias.walther@destatis.de

Publications

Volume on the Scientific Conference "Data fusion and data integration"

Since 1995, the Federal Statistical Office has organised joint scientific conferences in cooperation 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-year intervals. The subject of the sixth joint scientific conference, which was held on 30 June and 1 July 2005, was "Data fusion and data integration". The conference was moderated by Hartmut Scheffler, Chairman of the Association of German Market and Social Research Institutes (ADM).

After Johann Hahlen, President of the Federal Statistical Office, had welcomed the conference participants, Mr. Scheffler provided an introduction to the conference theme in his opening paper.

The first thematic block of the conference consisted of two contributions on the techniques of data fusion and data integration. PD Dr. Susanne Rässler from the Institute for Employment Research (IAB) in Nuremberg provided an overview of the techniques and application areas of data integration and data fusion. The contribution by Michael Wiedenbeck from the Centre for

Survey Research and Methodology (ZUMA) in Mannheim focussed in particular on the techniques of data fusion.

The second thematic block of the conference was devoted to data fusion in research on media use. Uwe Czaia from CZAIA Marktforschung GmbH in Bremen described the application of data fusion in media analysis. Haluk Akinci, Dr. Jörg Hagenah and Prof. Dr. Heiner Meulemann from the Teaching and Research Centre for Media Science (MLFZ) of the Economics and Social Science Department of Cologne University then presented their contribution on "Data synopsis and data fusion for media use in Germany since 1972".

Using results of the census test, Hans-Gerd Siedt from the Federal Statistical Office in Wiesbaden illustrated the effects of duplicates on the quality of the population registers. The first day of the conference was concluded with the paper on "Data matching: integration of opinion survey results and business data" by Dr. Stefan Tuschl from TNS Infratest GmbH in Munich.

The agenda for the second day started with questions concerning the integration of micro-geographical and regional information.

Dr. Raimund Wildner from GfK AG in Nuremberg described the integration of opinion survey data and micro-geographical information.

PD Dr. Jürgen H.P. Hoffmeyer-Zlotnik from the Centre for Survey Research and Methodology (ZUMA) in Mannheim presented a contribution on the substitution of opinion survey data by regional information.

The thematic block devoted to applications of data fusion and data integration ended with the presentation by Prof. Dr. Jürgen Krause from the Social Science Information Centre (IZ) in Bonn on "The integration of social science textual information and data from various sources by users".

Concluding the scientific conference, Erich Wiegand from the Association of German Market and Social Research Institutes (ADM) addressed questions concerning data protection in cases of data fusion and data integration and explained the relevant rules of conduct.

The conference volume on "Datenfusion und Datenintegration" was edited by Christian König, Matthias Stahl and Erich Wiegand and has been published (only in German) as Volume 10 of the series "Tagungsberichte" of the Informationszentrum Sozialwissenschaften (Social Science Information Centre) (ISBN: 3-8206-0148-1). It can be ordered at a price of EUR 10.- plus forwarding charges from the Informationszentrum Sozialwissenschaften, Lennestr. 30, 53113 Bonn, Germany. 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/forschungsuuebersichten.htm>.

Christian König, tel: (+49-611) 75-2077, e-mail: christian.koenig@destatis.de

Events

Fourth user conference on "Research with the microcensus: Analyses of the social structure and labour market"

On 12 and 13 October 2005, the fourth user conference on "Research with the microcensus: Analyses of the social structure and labour market" was convened in Mannheim. This conference, which is organised regularly by the Centre for Survey Research and Methodology (ZUMA) in cooperation with the Federal Statistical Office, is aimed at researchers whose present or future

work involves the de facto anonymised data (Scientific Use Files) of the microcensus. The user conference has by now become a forum not only for the presentation and discussion of research results and for an exchange of experience among researchers, but also for an intensive exchange of views with the statistical offices as data producers. Twenty contributions covering a broad range of content-related and methodological research activities were presented at the conference to an audience of more than 60 participants.

As regards *content-related* issues, the focus of the 2005 microcensus user conference was on educational and gender-specific aspects of the labour market and on the potential and possible limitations of the microcensus as a data source for analysing specific social situations. It was a pleasant development that, based on the increasing availability of microcensus data for various years, nearly all contributions discussed cross-cutting aspects in a comparison over time. Many contributions also addressed specific methodological issues and problems arising from the work with microcensus data.

The *methodological* focus of the conference was on microcensus panel data, which will be available as a Scientific Use File in the near future. The team members of the joint methodological project on "Processing and provision of the microcensus as a panel sample" (Federal Statistical Office, Land Office for Data Processing and Statistics of North Rhine-Westphalia, Free University of Berlin, ZUMA) provided the research community for the first time with a comprehensive account of their research results.

The first content-related session was devoted to the subject "Labour market and gender". *Esther Geisler* and *Michaela Kreyenfeld* (Max Planck Institute for Demographic Research) described determinants for the employment of mothers in eastern and western Germany. The contribution by *Franziska Schreyer* (Institute for Employment Research of the Federal Employment Agency) focused on the labour market for women with university degrees in non-gender typical subjects (e.g. engineering, physics and computer science). On the basis of a three-country comparison (Great Britain, France, Germany), *Patrick Puhani* (Darmstadt Technical University) examined the impact of changing the field of study on gender-related pay gaps in the period from 1970 to 2000.

During the second block of sessions, the activities of the joint methodological project on establishing a Scientific Use File of the microcensus panel were presented. This panel file represents a major extension of the uses of microcensus data. Covering the reference dates of the period from 1996 to 1999, it will presumably be made available to the research community in 2006. *Holger Breiholz* (Federal Statistical Office) described the panel features of the microcensus in his introductory contribution. On the basis of content-related analyses, several subsequent papers discussed the advantages offered by a microcensus panel file. *Robert Herter-Eschweiler* (Federal Statistical Office) dealt with types of employment and familial living arrangements, while *Bernhard Schimpl-Neimanns* (ZUMA) examined educational careers from the start of the last three years of grammar school to Abitur (general university entrance qualification). *Sabine Böttcher* (Halle-Wittenberg University) presented her study on mobility processes at the end of working life. All the above contributions also discussed the problems involved with a set of panel data. This referred, first of all, to systematic nonresponse due to the lack of geographical mobility data in the microcensus, which is designed as an area sample. The ensuing effects on transition analyses of family and employment issues were examined by *Sylvia Zühlke* and *Michael Konold* (Land Office for Data Processing and Statistics, North Rhine-Westphalia), while the contribution by *Edin Basic* (Free University of Berlin) focused on the analysis of incomplete contingency tables. *Sandra Rohloff* (Federal Statistical Office) and *Ivo Marek* (Free University of Berlin) addressed the crucial issue of the construction and the properties of expansion factors for the microcensus panel. *Ulrich Rendtel* (Free University of Berlin) summed up the project results achieved so far by posing the question "Does the microcensus provide an adequate data basis for longitudinal analyses?" and reached the conclusion that the microcensus panel will be more than just an independent and valuable data set for research. The study of methodological

aspects, such as the non-coverage of geographically mobile households or persons, also supplies useful information on the quality of the data and thus considerably increases the value of the "regular" cross-sectional microcensus data.

The third thematic block centred on the potential and limitations of the microcensus as a data source for the analysis of specific social situations. Once again it became evident that the big size of the sample offers considerable advantages where the analysis of specific small populations is concerned. *Herbert S. Buscher* (Institute for Economic Research, Halle) presented a contribution on "Precarious income situations in eastern and western Germany". *Jürgen Schiener* (Mainz University) described a project using the microcensus for the first time as a "Data source for social reporting on disabled persons". The questions included in the 1999 microcensus on long-term nursing care were examined by *Uta Ziegler* (Rostock University) in her paper on "Household forms among the elderly population in Germany with special emphasis on persons requiring long-term care". The contribution by *Doris Bardehle* (Land Institute for Public Health Care) and *Paul Berke* (Land Office for Data Processing and Statistics, North Rhine-Westphalia) centred on the usability of the microcensus for health monitoring.

The conference closed with analyses focusing on "Labour market and education". *Rainer Frietsch* (Fraunhofer Institute, Karlsruhe) presented a paper on the development of private returns on education between 1970 and 2004. The contribution by *Christiane Mück* (Oldenburg University) and *Karin Mühlenbein* (Hamburg University) on "Crowding out and innovation" highlighted the reaction of the German labour market to the university expansion in the 1990s. *Daniel Lois* (Technical University of Chemnitz) studied the measurement of continuing vocational training for older persons employed, while *Helmut Rudolph* and *Kerstin Blos* (Institute for Employment Research) discussed different concepts for surveying unemployment (ILO unemployment rate versus registered unemployment as recorded by the Federal Employment Agency). Finally, *Uwe Fachinger* (Centre for Social Politics, Bremen) and *Dieter Bögenhold* (Free University of Bolzano) presented their research results on the "New economy in the services society".

It is not possible to discuss the multitude of findings here in detail. Interested readers are requested to refer to the presentations and papers published on the internet. Very lively discussions were held over two days, for instance on measuring income in the microcensus and the use of these income data in analyses, on discontinuities in series of microcensus trend data due to changes in variables (e.g. changes in questions), on the quality of the data and on other subjects. These discussions certainly enhanced the knowledge of all participants about microcensus data. However, another contributory factor was that researchers from a variety of disciplines – demography, labour market research, economics, statistics and social research – came together at the conference. Ultimately, this is a direct consequence of the broad design of the microcensus survey programme, but also an effect of the increasing use of the Scientific Use Files of the microcensus in empirical research.

The contributions presented at the fourth user conference and at earlier conferences can be found on the internet at <http://www.gesis.org/Dauerbeobachtung/GML/Service/Veranstaltungen/index.htm>.

Holger Breiholz, tel: (+49-611) 75-8711, e-mail: holger.breiholz@destatis.de