DESTATIS
wissen.nutzen.

METHODS – APPROACHES – DEVELOPMENTS
Information of the German Federal Statistical Office

Edition 2/2009

The catchword

Methods of federal statistics – Further development

Publications

Events

The catchword

Results of the microcensus interviewer survey

In complex surveys such as the microcensus, a number of factors may influence the response behaviour of participants and thereby bias the survey's results. The field work of interviewers constitutes an integral part of the survey. Its quality is influenced both by the organisation of the survey at the statistical offices of the Länder, for instance by the recruitment and coaching of interviewers, the conducting of training measures, technical organisation etc., and by the work of the interviewers themselves.

This is why the interviewers of 13 statistical offices of the Länder were asked to fill in a standardised questionnaire about their experience with microcensus field work. The interviewer survey, conducted between December 2007 and January 2008, was a joint project of the Federation and the Länder. The high response rate of more than 80% shows that interviewers are willing to comment on their work and on the quality and organisation of the microcensus.

This article presents the results of the interviewer survey and reports on actions to further improve the quality of the microcensus, some of which have been implemented already.

The work of a microcensus interviewer

Generally, the statistical offices of the Länder have personal data of their interviewers. However, an overview of the overall structure of microcensus interviewers in Germany did not exist so far. The data of the interviewer survey give a detailed picture of the composition of the interviewer staff throughout Germany.

A total of 1,740 interviewers[1] filled out the survey. On average, they have done microcensus field work for twelve years in the old Länder and about nine years in the new Länder and Berlin, gaining substantial experience and practice in interviewing respondents.

The composition of the interviewer staff differs between the western and the eastern part of Germany with respect to several characteristics. While the great majority of interviewers in the old Länder (72%) is in employment and conducts interviews as a part time job, this holds for less than half of the interviewers in the new Länder (42%). Retired persons are employed rather frequently as interviewers in the new Länder. They represent about one third of the interviewer staff in the new Länder and Berlin, while the corresponding group in the former territory of the Federal Republic accounts for only 13% of interviewers. There are also more unemployed persons among the interviewers in the new Länder (9%) than in the old Länder (1%).

The average workload of interviewers in Germany is 21 households per month and largely depends on the way in which the statistical offices of the Länder handle this matter; depending on the organisational framework and the number of interviewers employed, the work packages of interviewers may vary considerably.

Interviewers' job satisfaction is high. Depending on the Land, between 85% and 100% are satisfied or very satisfied with their work as interviewers in general. The cooperation with the statistical offices of the Länder is also regarded as positive by nearly all respondents (99%).

The high job satisfaction is also reflected by interviewers' assessment of the remuneration received for their work. A vast majority (former territory of the Federal Republic: 77%, new Länder and Berlin: 71%) considers the compensation as adequate or quite adequate.

Assessment of microcensus field work

---

[1] January 2008, excl. data for Saarland.

Interviewers see some potential for improvement of field work, above all in the sphere of methodology. Both a simpler wording of questions and a shortening of the questionnaire may facilitate the work of interviewers in the opinion of 35% and 23% of respondents. 18% would be happy with a simplified handling of the laptop application and another 14% stated that more information about changes in the list of questions would facilitate their work. As many as 39% of respondents did not see any need for improvement at all (former territory of the Federal Republic: 35%, new Länder and Berlin: 50%).

In general, interviewers have only few problems with the technical aspects of conducting microcensus interviews. The share of those reporting difficulties with entering the data on the laptop, with filtering and with plausibility checks is very small. The technical cooperation with the statistical offices of the Länder also works very well; most interviewers did not have problems worth mentioning with the transfer of data or with the downloading of work packages and updates. In contrast, more than half of the respondents often or sometimes had problems with the coding of answers which is necessary, for instance, when occupations and branches of economic activity have to be indicated.

As a rule, interviews shall be conducted personally with the support of a laptop (CAPI)[1]. Only in few cases do interviewers resort to the paper questionnaire (8%) and/or conduct telephone interviews (7%).

Handling the list of questions

All in all, the comprehensibility of the questions was regarded as positive: 79% of interviewers thought that the questions could easily be understood. Accordingly, interviewers rarely reported that respondents had refused to answer questions because of complicated wording: this had not happened at all to 57% of the interviewers, and only seldom to 33% of them.

The fact that many questions appear to be similar and that the differences between questions do not become clear was regarded as more of a problem than their wording; this refers for example to the questions on employment. 57% of the interviewers often or sometimes experienced situations where respondents had the impression that redundant questions were asked as the answer had already been given before. Therefore, in practice not all of the questions are asked in all interviews. 26% of the interviewers reported that they do not always ask all the questions when they think they already know the answer from the course of the interview.

Many questions of the laptop program have not been optimised for personal interviews and are based on the paper survey. For that reason, most of the interviewers often or sometimes rephrase questions so that respondents can understand them more easily. A total of 88% of interviewers said they had altered the wording of the questions to facilitate understanding, 83% of them had changed the wording of the question only slightly and named the essential terms, 65% had given their own examples and 5% had rephrased the questions entirely.

Surprisingly, the experience gathered with handling the list of questions did not play any part in this. Experienced microcensus interviewers had rephrased questions as seldom or as often as interviewers who had been active in interviewing for a few years only.

Interviewers feel that the households surveyed do not have problems with the majority of the questions. They sometimes received negative or critical remarks when asking about employment status. More often there is criticism in connection with questions about secondary employment and marginal work. Three fourths of the interviewers had already heard negative or critical remarks in that block of questions. Questions relating to income and subsistence are considered most problematic. Almost all interviewers (99%) had been faced with displeasure when asking these questions.

---

[1] Computer Assisted Personal Interview.

Only rarely do participants cut the interview off once it has started. 72% of the interviewers had never seen an interview being broken off by the participant. Of the 28% of interviewers who had conducted an interview which could not be finished, 71% reported that it had been terminated while questions concerning subsistence and income were being asked.

Sabine Köhne-Finster: tel:+49(0)228 99 643 87 11, e-mail: sabine.koehne-finster@destatis.de

Gesine Güllner, e-mail: guellner@uni-bonn.de

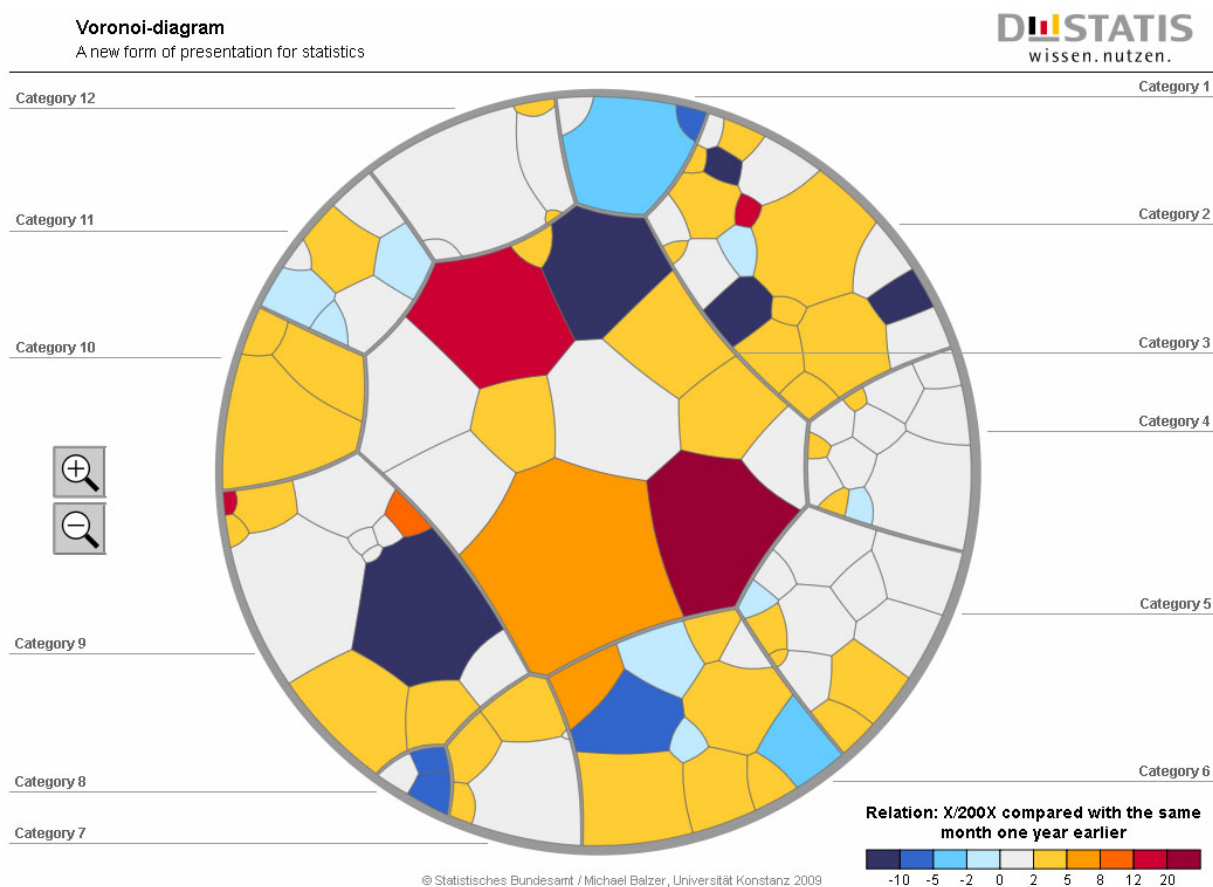# Methods of federal statistics – Further development

Price kaleidoscope: a new visualisation tool for price statistics

Technical background

The basket of the consumer price index consists of about 700 goods. For each type of goods such as books, cinema tickets or petrol, the price development is calculated. Then the overall consumer price index is a weighted mean of the price development recorded for all the roughly 700 types of goods.

For outsiders it is sometimes difficult to comprehend that price developments are aggregated so much as to be expressed in a single figure, the consumer price index. The correlation between the basket of goods and the price development was made perfectly clear in a Sunday edition and on the homepage of the New York Times in May 2008. With the help of a Voronoi diagram, as it is called, the weights and the price developments of the product groups in the basket were visualised with a kind of "kaleidoscope" in an aggregated form. The amounts of the weights in the goods basket were represented by the size of organically shaped segments within a circle. The price changes recorded for the individual groups of goods were reflected by colour codes. The online version provided for an interactive change between the static presentation on a coarser and a more refined level of display.

That form of presentation was adapted for the price development in Germany. Products are shown according to the Classification of Individual Consumption by Purpose (COICOP) commonly used in European price statistics. Both a print and an online version are planned. In the online version, 107 groups of goods will be presented, with the option to zoom down to 667 types of goods.

Voronoi-diagram
A new form of presentation for statistics

Relation: X/200X compared with the same month one year earlier

© Statistisches Bundesamt / Michael Balzer, Universität Konstanz 2009

On the one hand, the Voronoi diagram provides a quick overview of how significant the groups of goods are and of their price developments. On the other, those interested may gain very detailed insights into the price development in certain areas. The aim is an intuitive grasp of the complex index construction in price statistics. So the Voronoi diagram may help to reduce mistrust due to a lack of understanding.

It also makes it easier to grasp the structure of the COICOP classification. It is a special feature of that breakdown that there is no aggregate "services" position, the individual services are allocated to their purpose instead. For example, the costs of the service "laying and fixing of floor coverings" is part of the purchase prices of carpets and other floor coverings. In the online version of the Voronoi diagram, alternative aggregates can be displayed by clicking on individual buttons. In addition to "services" these are aggregates such as "energy" or "consumer durables".

Technical implementation

The form of representation known as Voronoi diagram is based on a method developed by the mathematicians Thiessen and Voronoi for the decomposition of geometric spaces into regions determined by a specified number of points in space (called centres). Each point of a region is nearer to its centre than to all other centres. Depending on the locations of the centres, segments of different sizes emerge. For decomposition with pre-determined area sizes, in contrast, the locations of the centres and thus the locations of the regions have to be calculated. A technique which may be used to do so was published, among others, in "Michael Balzer, Oliver Deussen: 'Voronoi Treemaps', IEEE Symposium on Information Visualization, 2005". Therefore, the Voronoi diagram was implemented at the Federal Statistical Office in co-operation with Michael Balzer of Konstanz University. Mr. Balzer has developed a tool to calculate the locations and forms of the segments within a given polygon. In the case of price statistics, the form started with is a circle. The size of the segments is proportional to the weighting of the types or groups of goods in the basket. Several levels can be represented hierarchically and thus be nested.

The result of that area calculation is processed further and published in an interactive SVG file. SVG (Scalable Vector Graphics) is a standard used to describe two-dimensional vector graphics. SVG files may be scaled without quality losses and, using script languages, they may be extended by functionalities enabling user interaction.

In order to show not only the weighting of the groups of goods by means of the size but also make the corresponding price change visually comprehensible, the latter is coded by different colours. Intensive or dark colours signify stronger price changes, with shades of red representing price increases and shades of blue standing for price decreases. Areas without or with only minor price changes are shown in neutral colours. The user is given detailed information on the name of the group of goods, its weight(ing) and price change when he or she crosses an area with the mouse. The required detailed information is then provided by a tool tip. When one zooms into the diagram, part of that information is already displayed for selected groups of goods and with the highest level of detail.

Tanja Raschke, tel: +49(0)611 / 75 27 93, e-mail: tanja.raschke@destatis.de

Nadya Konrad, tel: +49(0)611 / 75 47 48, e-mail: nadya.konrad@destatis.de

Manuel Wirsing, tel.: +49(0)611 / 75 94 42, e-mail: manuel.wirsing@destatis.de

The new IT system of consumer price statistics

With the regular revision of the consumer price index for Germany at the beginning of 2008, a new IT system for consumer price statistics was introduced, too. Major characteristics of that new system when compared to the previous solutions are centralised production and data storage of the decentralised consumer price statistics, complete documentation of the metadata and methodical improvements in index calculation.

Changing over to centralised production and data storage leads first of all to better efficiency in the process of statistics production. External data transfers, which involve the risk of interface-related losses, are largely avoided and the time and costs required for changing and maintaining the all-German standard steering files are reduced. Especially in German consumer price statistics, which is characterised by a complex interaction of decentralised and centralised price collection, those measures can markedly improve data quality. Also, centralised production harmonises the calculation methods and eliminates previous degrees of freedom regarding the input of specific control variables which are important for calculation.

What should particularly be mentioned in the context of the complete documentation of metadata of consumer price statistics is the variables for a detailed description of the individual survey items as well as the information on the various reporting units. This ensures that product switches are easier to identify and can be included with the correct price into the index calculation. In the medium term, detailed product descriptions will allow performing regional price comparisons.

When we examine the system of index calculation in German consumer price statistics, the steps required to get from the collected raw data to the consumer price index can be subdivided into three levels.
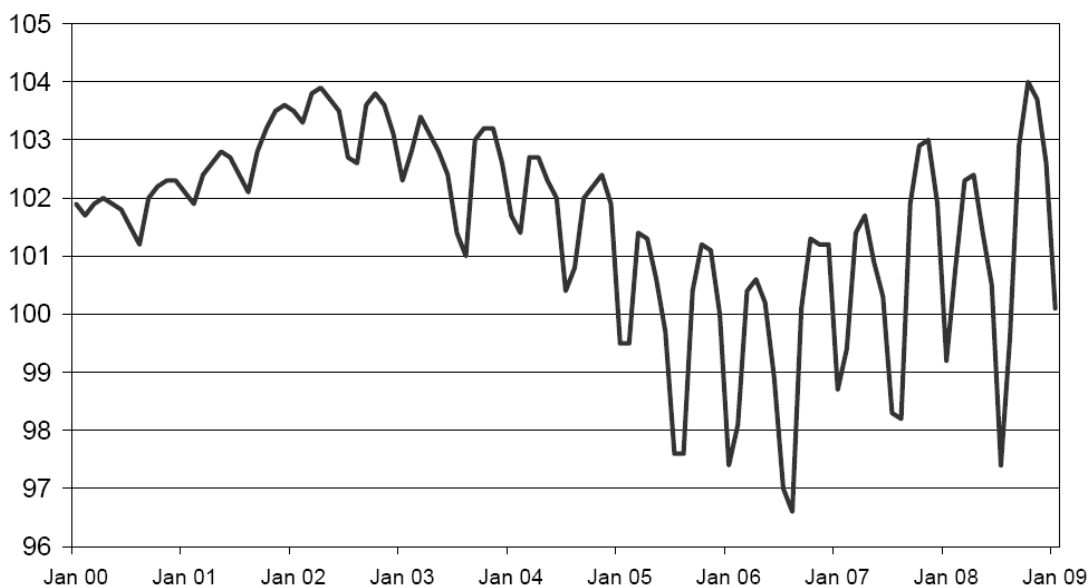
The first level of index calculation refers to the processing of the collected raw data on a product to ensure the comparability over time. With the huge number of monthly price data recorded, quantitative and qualitative changes or temporary non-response occur all the time. Price processing is controlled by coding through the price collectors, which has to be confirmed by staff at the statistical offices of the Länder. In the new IT system, a distinction is made between coding of reporting units, of products and of prices. A chosen combination of such codings will define the calculation routine used to process the collected raw data. Contrary to the previous approach of creating comparability between the currently collected price and the price of the previous period in case of quantitative or qualitative changes by calculating a fictitious comparable price of the previous month, the new IT system always converts the price collected in the current period into a so-called processed price, which refers to the quantity and quality of the sample defined in the base period and finally is included in the calculation of the consumer price index. Consequently, the implicit weight of the original target sample is maintained.

In case of temporary non-response for individual products or entire reporting units, the new IT system updates the processed prices of the previous period for the relevant items by using the price trend of the relevant product group in the current period. By changing over to centralised production and data storage, it is possible in the calculation of the updating factors required here to include the prices of the items collected centrally (e.g. prices of the mail order business).

A specific problem occurs in the reconversion of special offer prices after non-response or product switches. Especially in the clothing sector, such constellations are frequently observed (changes in range of products after seasonal sale). As in those cases the distinctive variables of the product no longer offered and those of the replacing product often are not comparable with each other, the price series must be chain-linked. Chain-linking normal prices with special offer prices is problematic because the temporary character of a special offer is not taken into account and, instead, is shown as a permanent price reduction. In the past, this involved a higher risk of the index trend being in part biased downwards. Every month and for every product, the new IT system stores not only the processed price but also a so-called normal price; the latter is

obtained by eliminating the existing special offer structures. This allows automating the reconversion of special offer prices into a normal price level even in the case of non-response or product switches. Examining the index trend for "clothing and footwear" before and after the introduction of the new IT system[1] illustrates the influence of the new approach. Due to the above problem involved in chain-linking after product switches, the index of the sector examined showed a clearly negative trend up to, and into 2006. Thanks to the methodical modifications it has been possible to stop that trend and to eliminate the biases of the index for clothing and footwear:

clothing and footwear (2005 = 100)

At the second level of index calculation, the prices processed at the first level are taken to calculate elementary indices. In the new system, the delimitation of the elementary indices by COICOP 10-digit items[2] and Land (federal state) has been extended to include the dimension of the outlet type. A total of eight outlet types are distinguished (department store; self-service department store; supermarket; discount supermarket/specialised self-service store; specialist shop; other retail; services/rent; mail order business/online business). At the same time, the calculation of municipality average prices is abandoned. This allows better implicit weighting of the various products within the limits of the elementary indices.

At the third level of index calculation, the elementary indices are aggregated to obtain the systematic and non-systematic aggregates and the consumer price index for Germany. For the aggregation of the elementary indices, explicit weighting of the various outlet types has been introduced at the level of Länder. The weights per product group are based on the relevant importance for consumption in the relevant Land.

By abandoning municipality average prices and simultaneously introducing explicit weighting of outlet types it is now possible to base the number of price observations within the limits of elementary indices on the volatility of the price trend. If the price fluctuation within the limits of

---

[1] The new IT system was introduced at the beginning of 2008. Within the scope of the regular revision of consumer price statistics performed at that time and putting it on base 2005 = 100, the consumer price index has been calculated according to the new system as from 2005. However, due to insufficient information contained in the old data, the new method of calculation has fully been applied only from January 2007.

[2] COICOP = Classification of Individual Consumption by Purpose. A COICOP 10-digit item refers to the lowest level of collection in consumer price statistics, that is an individual product.

an elementary index is large, the number of prices to be collected tends to be larger than for small fluctuations.

Florian Burg, tel: +49(0)611 / 75 41 32, e-mail: florian.burg@destatis.de

Daniel Seeger, tel: +49(0)611 / 75 42 31, e-mail: daniel.seeger@destatis.de

Morbidity in out-patient medical care in Germany – a pilot project

Background and objectives of the pilot project

In Germany, reliable population-related data on morbidity in medical care of the German health system are available primarily in in-patient diagnoses and incapacity for work diagnoses. Contrary to that, morbidity in out-patient medical care has been a less developed field of analysis; up to now there has been no routine reporting on out-patient diagnoses and services. However, co-operation between medical associations, statutory health insurance institutions, regional associations of statutory health insurance physicians and research institutes shows a markedly growing interest in processing and using data on out-patient care in Germany. Within the scope of developing a Community strategy of health policy, European health statistics focuses on opportunities and limits of national morbidity statistics in order to make a contribution to describing the health status of the European population. In this context, the Federal Statistical Office (Destatis) participates in the pilot project on "morbidity statistics" commissioned by the Statistical Office of the European Communities (Eurostat). The pilot project focuses on identifying and evaluating potential sources of diagnosis-specific morbidity data regarding out-patient care in Germany. This is done against the background of a rather long-term and steady supply of reliable data. Based on the data sources considered as adequate, the frequencies of selected diseases (European Shortlist, version: March 2007) in out-patient medical care is estimated in relation to the population. The objective of the entire research process is to check and enhance the methodology defined at the European level as a guideline concept[1] by the Morbidity Statistics Development Group (MSDG).

Data sources

The main data stock of the pilot project is formed by data on the invoicing of services provided by statutory health insurance physicians. By evaluating such data it is possible to show specific morbidity indicators of the diagnoses contained in the European Shortlist – although only for statutory health insurance and accompanied by a detailed discussion of data quality. Potential additional sources are documentation systems of research networks, disease-related registers and the health surveys of the Robert Koch Institute. However, it is hardly possible to obtain from such sources diagnosis-specific information that is concretely associated with the out-patient medical care sector and allows performing population-related analyses. Co-operation with various private health insurance companies provides the option of covering the privately insured patients with regard to out-patient medical care.

Morbidity and diagnosis data of statutory health insurance physicians

The patient-related diagnosis data on the invoicing of medical care services provided by statutory health insurance physicians in Germany are used in the pilot project to show the health or illness status as documented in out-patient care. More than 85% of the physicians working on an out-patient basis in Germany are actors in medical care of statutory health insurance physicians and bill their services with the relevant regional association of statutory health insurance physicians

---

[1] Eurostat/ MSDG (2007): Principles and guidelines for diagnosis-specific morbidity statistics. 23 April 2007. Version 1.1.

(RASHIP). Such RASHIP invoicing data refer to all patients in statutory health insurance and thus cover some 90% of the population.

Until 2007, invoicing data referring to groups of physicians were analysed by the institution keeping the data, which is the Federal Association of Statutory Health Insurance Physicians (NASHIP) or the Central Research Institute of Ambulatory Health Care in Germany (CI).[1] Processing the patient-related diagnosis data for the whole of Germany is something new both for those institutions and for the data landscape in Germany. This is why the CI performed test studies as part of the Eurostat project on out-patient morbidity in Germany in order to prepare the analysis of the all-German data and to optimise the data quality.

As the data are related to patients, it is possible to show individual treatment careers across groups of physicians. Duplication or multiple counts of diagnoses caused by multiple patient-physician contacts – which is typical of case-related invoicing data – can thus be avoided, while multiple counts of natural persons cannot: Patients changing from one health insurance institution to another within the system of statutory health insurance will get a new insured person code. As a consequence, a new patient is created in the RASHIP data set.

For the presentation of (invoiced) morbidity in out-patient medical care based on the anonymised process data, two central indicators are used in the pilot project:

- The prevalence describes the number of cases of illness in a population either at a given point in time (point prevalence) or in relation to a defined period (period prevalence).

- The incidence describes the number of new cases of illness in a population.

In order to obtain valid estimates of prevalence and incidence on the basis of out-patient invoicing data, it is necessary to distinguish between chronic and acute illness, so that overestimation and underestimation of morbidity can be avoided. The working definition of the CI used in the pilot project, which is the so-called M2Q criterion (occurrence of the same diagnosis per patient in at least two quarters of the year), identifies chronic diseases only on the basis of the frequency of the patient-related diagnosis documentation in the reference year. In addition, the analysis includes the year preceding the reference year as a control year, so that calculating especially incidences of diagnoses in the reference year is more valid.

Restrictions

The invoicing data are meant to document medical treatment occasions relevant for invoicing.[2] So their primary purpose is not to describe morbidity in relation to the population. Consequently, discussing the information value of indicators for morbidity estimation with regard to a medical care sector is crucial. For example, there may be considerable time intervals between the beginning of a patient's illness and the diagnosis made by a physician treating the patient on an out-patient basis. However, the actual date of a patient falling ill is not a compulsory element of the invoicing documents. Only when including the relevant patient record and/or the patient's data as a person insured might it be possible to approximate the data analyses to the actual date when the patient fell ill. Evaluating those data stocks cannot be implemented at present.

Possible solutions

In this context, the diagnosis quality is one of the crucial points of consistent morbidity statistics. The RASHIPs apply internal verification procedures to the out-patient diagnosis data with regard to subject-matter and calculatory correctness, to plausibility and to cost effectiveness. In

[1] cf. Zentralinstitut für die kassenärztliche Versorgung der BRD: Morbiditätsanalyse. Datenbasis: Dateninhalte.

http://www.zi-berlin.de/morbilitaetsanalyse/dateninhalt.php, 09.05.2008

[2] cf. Kerek-Bodden, H./ Heuer, J./ Brenner, G./ Koch, H. / Lang, A. (2005), p.40; see also: Art. 295 Social Code V, para. 3: main diagnosis = treatment occasion on treatment day.

addition, the CI intends to improve the quality assurance methods. Drug prescriptions, other (disease-specific) data sources such as registers and health surveys as well as the comparison of patient records of the software used in practices of physicians are core elements of validation methods which in current scientific studies are compared with samples of invoicing data. Various research projects carried out in recent years show population-related analyses of the epidemiology of specific clinical pictures. In the pilot project, studies focusing on out-patient medical care (or sections such as family doctors) are listed in a description matrix and are allocated to the diagnosis groups of the European Shortlist. Disease-specific morbidity estimates obtained from those projects might be relevant for an analysis by groups of physicians of the all-German invoicing data for a comparison of various prevalence values in out-patient medical care of the German health system. However, such comparisons cannot replace a comprehensive validation of the invoicing diagnosis data of medical care by statutory health insurance physicians.

In January 2009 another reform of the German health system entered into force. Now both the remuneration of statutory health insurance physicians in Germany and the payments made between the statutory health insurance institutions under the risk structure compensation scheme are based on the morbidity structure of the persons insured under statutory health insurance schemes. So it is expected that there will be a further growing interest in the processing and validation of diagnosis data and in the introduction of diagnosis coding guidelines for out-patient medical care.

Susann Behrendt, tel: +49(0)228 / 99 643 8176, e-mail: susann.behrendt@destatis.de

Using regular expressions in the business register, taking as an example automatic coding of legal forms

An enterprise's legal form is an important evaluation variable of the statistical business register. But the legal form is important not only for evaluations but also for the identification of enterprises where data from various administrative sources are processed in the statistical business register.

According to the German Commercial Code, registered business persons, partnerships, corporations and co-operatives are legally required to include the legal form in their company name. Using string search patterns (regular expressions), it is possible to search for such legal form designations in the company names and thus to improve the coding of the legal form.

Automated legal form coding through regular expressions is used in the business register to check and correct the stock of legal form codes and for new entries from administrative sources. In addition to using regular expressions for legal form coding, that tool can also be applied in any type of search for string patterns and text editing.

Legal forms of enterprises in the statistical business register

The legal form of an enterprise is an important evaluation variable for the statistical business register – in the following referred to as business register. Also, the legal form is an important additional identification variable of enterprises. As there is no standard identifier available for combining the different administrative data which are used to maintain the business register (administrative sources), the information from the various administrative sources is combined in the business register, among other things, through the company name. Consequently, the correct legal form is crucial for identification in case of similar company names.

The main administrative source for the legal form of enterprises in the business register currently is the VAT records. A new source for the legal form in the business register has been opened up with the electronic trade register. In many cases, however, the legal form of an enterprise has to be searched and revised in the business register. A possible solution here is using the company name in the business register as another source for coding the legal form.

Using regular expressions for legal form coding

Regular expressions describe in a specific way text patterns used to check and change files that contain strings[1]. Using such string search patterns, it is possible to search for legal form designations in the company name (in the following referred to as name string). The legal form designations found can then be used to code the legal form in the business register. This makes the company name another systematic source for legal form coding in the business register. Regular expressions allow especially the qualification of legal forms which must be contained in the name, such as for registered business persons, corporations, partnerships and co-operatives[2]. Automated legal form coding by means of string search patterns also allows searching for legal forms (for example, "non-profit limited liability company") which are not contained in such detail in the administrative sources.

For the definition of the search patterns, it is first of all necessary to define which legal forms are to be searched for and in what variants the legal forms occur in the company names. In that way, different search patterns are defined for the individual legal forms contained in the business register.

Quality of legal form coding through regular expressions

From the viewpoint of those developing regular expressions, the correctness and efficiency of regular expressions are most important. Correctness and efficiency mean that the string search pattern will quickly and precisely find the right hit and will not have any wrong hits[3]. A regular expression must not react sensitively to variations in the search pattern such as upper and lower case letters, blanks or different connectors. Another aspect of correctness is the extent to which the hits lead to coding a legal form. In the best case, every string search pattern will lead to a clear legal form.

The time aspect is relevant especially for program optimisation. What is more interesting in subject-matter terms is the fact that, by using the regular expressions, the program correctly identifies the legal forms according to the specified rules. Due to the legal requirement, company names referring to business persons, partnerships, corporations and co-operatives should contain legal form designations. Thus it is possible to judge the regular expressions on the basis of the completeness with which they identify legal form designations. Despite the legal obligation to include a legal form designation in the company name, it is not possible in all cases, even for partnerships and corporations, to find the legal form through regular expressions. This may be due to incomplete search patterns. Another source of error is the incompleteness of the administrative data used, i. e. the company name in the administrative file does not include the legal form designation. As the business register is updated by administrative files, the quality of the administrative data used is reflected in the content of the business register.

Therefore, an important aspect for users of regular expressions is the reliability with which the legal form is identified in the company name, i. e. the fact of whether the regular expressions recognise the legal form designation in the name string in cases where the name string contains a legal form. In order to be able to give information on the reliability, a sample of the legal form

---

[1] Friedl, Jeffry, "Reguläre Ausdrücke" , O'Reilly, 3rd edition, 2008 and also Stubblebine, Tony, "Reguläre Ausdrücke – kurz und gut", O'Reilly Taschenbibliothek, 2nd edition, 2008.

[2] Pursuant to secton 19 of the German Commercial Code in the adjusted version published in the Federal Law Gazette, part III, no. 4100-1, last amended by section 17 of the act of 21 December 2007 (Federal Law Gazette I, p. 3089).

[3] Friedl, Jeffry, "Reguläre Ausdrücke" , O'Reilly, 3rd edition, 2008, p. 227.

codings of the program were checked for incorrectly[1] coded units. It turned out that the program recognised the correct legal form in 97.5% of the cases where the legal form was contained in the name string. Where the name string did not contain the legal form, this was also very reliably identified by the program (99.8% of the cases).

Patrizia Mödinger, tel: +49(0)611 / 75 46 07, e-mail: patrizia.moedinger@destatis.de

National implementation of a revised Classification of Economic Activities (WZ 2008)

1. General notes on the introduction of the new Classification of Economic Activities WZ 2008

Since January 2009, the national Product Classification for Production Statistics, edition 2009 (GP 2009) has replaced the former edition 2002 (GP 2002). At the same time, the Classification of Economic Activities, edition 2008 (WZ 2008) is gradually replacing the edition of 2003 (WZ 2003). The WZ 2008 is based on the European Statistical Classification of Economic Activities NACE Rev. 2, which has been in force since January 2008 and has replaced the Statistical Classification of Economic Activities NACE Rev. 1.1 .

Classifications of economic activities must be adjusted from time to time, so that the statistics produced on their basis can adequately reflect reality that has changed in economic and technical terms. So, such adjustments also mean that statistical units may now be classified under other economic sectors. This generally involves breaks in statistical time series, which make it more difficult to analyse statistical results. Changing over to the new classification thus required careful preparation and an approach co-ordinated between the statistics concerned.

This article briefly outlines the activities required for implementing the revised classification, regarding the changeover of the business register for statistical purposes (URS) as well as double processing and recalculation of statistics.

2. Changeover of the business register

The business register for statistical purposes (URS) represents the population of all economically active units and allocates them to an economic activity. It provides the relevant surveys with a frame for sampling and it is also a separate evaluation tool. So, comprehensive adjustments had to be performed also in the URS in the context of changing over to the new classification. According to EU specifications, the information on economic activity in the URS had to be recoded in terms of time so that generally the groups of units reporting to structural statistics can be shown according to WZ 2008 for reporting periods from 1 January 2008 and those reporting to short-term statistics from 1 January 2009.

Consequently, all units contained in the URS were given a WZ-2008 code in early 2008 in addition to the existing WZ-2003 code. Generally, various cases had to be distinguished depending on the existing WZ-2003 code: Within the conversion key, there were WZ-2003 headings that could exactly be allocated to a WZ-2008 heading (1:1 allocations). For them, automated conversion could be applied. The same applies to headings for which several codes of the WZ 2003 were allocated to one new heading of the WZ 2008 (n:1 allocations). However, within the conversion key there were also many conversions for which the activity code could not

---

[1] Errors of first kind (wrong – positive) and errors of second kind (wrong – negative).

clearly be allocated to one new WZ-2008 heading (1:n allocations), so that a different approach had to be applied here.
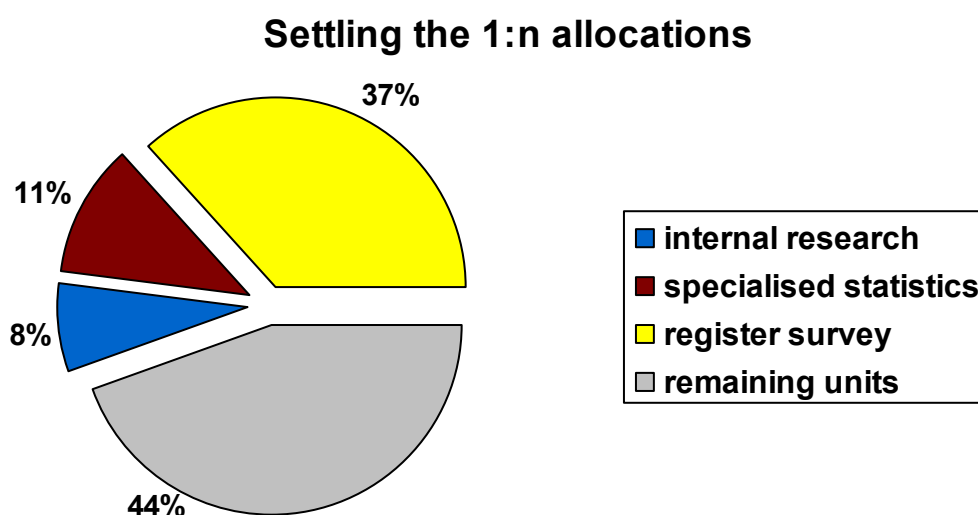
The URS in Germany currently contains some 3.5 million economically active enterprises. For about 1.9 million enterprises, no clear conversion was possible. What was done first of all in those cases was purely automated allocation to WZ-2008 headings. Such allocation according to the main activity (main-activity conversion key) was developed by a specialist committee consisting of experts of the statistical offices of the Federation and the Länder. For each of the relevant WZ-2003 codes, a WZ-2008 heading was chosen to which most units of the WZ-2003 code were expected to be allocated. So, this step allowed correct automated allocation of a large number of units. After that provisional automated conversion, those allocations by main activity were to be settled as part of the continuous URS maintenance. For reasons of capacity and workload, however, it was not possible to directly settle all those cases. For the first partial follow-up, three options were available:

a: Check of the WZ-2008 allocation for units reporting to continuous statistical surveys by experts responsible for the production of the relevant statistics,

b: internal research (e. g. internet, trade register) by register staff, and

c: a register survey using questionnaires and online methods.

The units included in the register survey were selected according to their importance for the relevant economic activities. The enterprises questioned were those for which 80% of the turnover and/or of the employees subject to social insurance contributions referred to their Land and to their WZ heading and which could be allocated to two different WZ-2008 two-digit headings according to the classification conversion key.

With those methods, a qualified WZ-2008 code was allocated to about one million units. The other approximately 900,000 units were smaller units which were not specifically examined for cost-benefit reasons. The main-activity conversion key ensures a minimum level of quality here.

The WZ-2003 code is maintained unchanged in the URS as a statistical variable – though for a specific reference day –, so that the units can be represented both by the old and the new classification. The chart below shows the shares in settling the 1:n allocations.

Settling the 1:n allocations

37%

11%

8%

44%

■ internal research
■ specialised statistics
□ register survey
□ remaining units

Once the settlement was finished, the results obtained were integrated into a revised main-activity conversion key, which is applied to the units not contacted. This will further improve the quality of the URS and of the WZ 2008 maintained in it.

Further information on the conversion processes is available at:

http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Navigation/Klassifikationen/UmstellungWZ2008,templateId=renderPrint.psml__nnn=true

3. Double processing and back calculation of statistics

In order to make things easier for users in the context of the conversion of the classification of economic activities, double processing and back calculation of statistics and time series must be performed. Double processing is the processing of individual statistical data both by WZ 2003 and WZ 2008, whereas back calculation refers to the recoding of statistical time series, that is of aggregates. Generally, at the European level, double processing by the old and the new classification is required for all structural statistics for reference year 2008.

Back calculation is relevant especially for short-term statistics. Results of short-term statistics will be shown by WZ 2008 for the first time for reference year 2009. To ensure parallel procedures all over the EU, a Commission Regulation on back calculation (EC No 472/2008) specifies the requirements to be fulfilled by time series calculated backwards.

According to those specifications, 2005 is the first base year in which the new classification of economic activities is to be applied to short-term statistics. The periods of back calculation differ considerably between the various statistics. For some statistics, back calculation is performed only for a few years, while for the indices in mining and manufacturing, new orders, turnover and production, back calculation is done back to 1991. Performing back calculations also required adjusting the weighting systems of the composite indices.

National accounts will be the last area to convert its results as part of the revision planned for 2011.

Converting classifications is a highly time-consuming and work-intensive process which requires a high degree of co-ordination. The entire process from the first discussions to the concrete implementation of the new classification WZ 2008 took several years.

Ines Rafael, tel: +49(0)611 / 75 22 84, e-mail: ines.rafael@destatis.de

Dr. Bernd Waldmüller, tel: +49(0)611 / 75 35 97, e-mail: bernd.waldmueller@destatis.de

StatVG – a new processing procedure for the production surveys in manufacturing

1. Background

The monthly production survey (MP) and the quarterly production survey (VP) are part of a system of statistics produced in manufacturing as well as in mining and quarrying[1]. The decentralised statistics of that system comprise not only the production surveys but also the monthly report on local units (MBB), the annual report on local units (JBB), the annual report on multi-unit enterprises (JBMU) and the annual investment survey (IV); surveys conducted centrally by the Federal Statistical Office are the annual cost structure survey (KSE) and the four-yearly survey of materials and commodities received (MWE).

The production surveys are conducted in producing local units of enterprises of the manufacturing sector with generally 20 or more active persons and in producing local units of the

---

[1] For simplification, only the term of manufacturing will be used in this article.

manufacturing sector of enterprises of other economic branches. Local units reporting to the MP are those reporting also to the MBB. That are local units with 50 or more active persons in September of the previous year. The production surveys cover on a monthly basis the production for sale by quantity and value and, for specific products, also the quantity of the production for further processing in a breakdown by the Product Classification for Production Statistics, edition 2009 (GP 2009). For local units with fewer than 50 active persons (status: September of the previous year), the same variables are covered quarterly. The groups of respondents for the monthly and quarterly survey are defined at the beginning of any new reference year on the basis of data from the MBB, the JBB and the JBMU as well as from the business register.

2. Previous processing procedure

Until the end of reference year 2008, what was available for handling the MP and VP was the mainframe procedures MPROD and VPROD, which covered data capture, data editing and tabulation. The two production surveys were processed on the basis of ADABAS databases in connection with Natural dialogue programs.

The methods used to capture the paper questionnaires were conventional manual data capture, optical document reading, and capture through the input masks of the dialogue application provided by the MPROD and VPROD procedures. Data that were not directly entered through the dialogue application could be scanned through a batch process which, after scanning, included data editing and compiling the resulting data in a print list. Online reports were delivered to the relevant dialogue procedure as flat data records through the KonVertCenter. Incorrect and implausible data in the reports were usually corrected in a dialogue either during capture through the dialogue application or after scanning the data.

The procedures allowed handling the data of a reference period. Therefore, the procedures included specific routines (batch processes) for starting the handling of data of a reference period and for finishing it, including the automated estimation of lacking data; those routines performed changes in the databases and produced processing logs or lists of entries performed.

After data editing, the plausible microdata were further processed to prepare data delivery to the Federal Statistical Office and the production of tables. For those work steps, too, batch processes were available, which could be started through the dialogue application. The tables were created by means of various Assembler programs. The work steps for primary and secondary confidentiality, which had to be performed before creating the publication tables for the VP, were included in the VPROD procedure within the tabulation program runs.

3. Developing a new procedure – StatVG

The reasons for the changeover and for programming a new processing procedure were the following:

- The tabulation programs and parts of the processing programs of the old procedure had been programmed in Assembler. According to a decision adopted by the information technology working group (AKIT) in June 2003, those procedures had to be switched to modern technology.

- The old procedures did not meet the basic requirement for centralised production and data storage (ZPD) because they were not multitenant.

- For a long time already, experts had demanded extensions to better support the work processes in processing the production surveys. Those demands referred to functions regarding search, processing process control and extended granting of rights as part of user administration.

- The Product Classification for Production Statistics underlying the production surveys was changed to the edition 2009 as from reference month January 2009.

Due to the large number and complexity of the changes required in the old procedures and the coming changeover to the new version of the underlying product classification and taking account of cost-benefit aspects, an entirely new program has been created which is based on the currently preferred hardware and software standards in the co-ordinated system of the Federation and the Länder. *Information und Technik Nordrhein-Westfalen (IT.NRW)* has developed an integrated procedure (both for MP and VP) as a client server solution to capture and process the statistical data. StatSpez[1]-supported tabulation and production of totals records have been added.

That new client-server application will also be used as a framework application, into which the other decentralised statistics in manufacturing mentioned above will be integrated in the future (MBB, JBB, JBMU, IV). Such an integrated procedure for all decentralised statistics of manufacturing is useful not only for a standard handling of the application but also because of the strong interlinkage of the statistics.

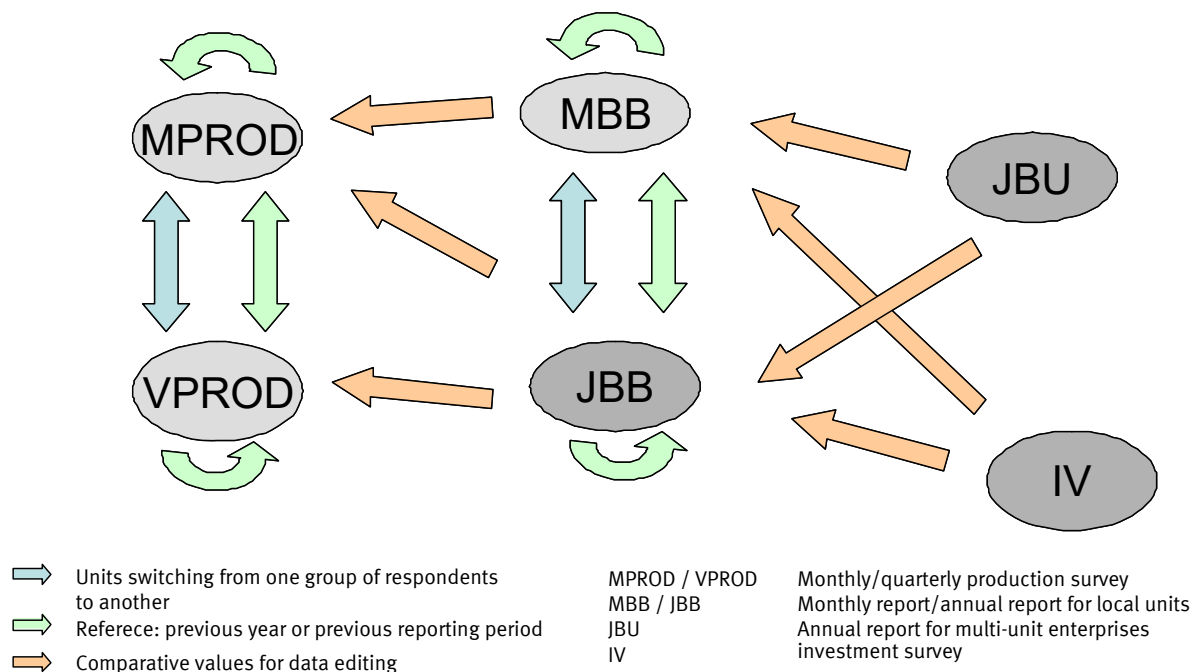|	
Units switching from one group of respondents to another	MPROD / VPROD — Monthly/quarterly production survey
Referece: previous year or previous reporting period	MBB / JBB — Monthly report/annual report for local units
Comparative values for data editing	JBU — Annual report for multi-unit enterprises
	IV — investment survey

Chart 1: Interdependencies between the statistics in manufacturing (extract)

The production surveys and the statistics intended for future integration into the procedure are part of the decentralised statistics in manufacturing which are entirely processed at the statistical offices of the Länder by currently some 150 staff members. As regards the operation of the new application, the changeover from decentralised solutions to centralised IT production and data storage created new organisational and technical framework conditions which had to be taken into account by IT.NRW during development.

Creating a new program had several goals, which are explained below:

1. Technical goals

---

[1] StatSpez is a software for the **Stat**istical **Spec**ification of Tables developed in official statistics and applied in the co-ordinated system of the statistical offices of the Federation and the Länder.

According to a decision taken in the co-ordinated system, the Assembler programs contained in the old procedures of the production surveys had to be replaced by modern software technologies (Java-based client-server procedure), which in the long term can be maintained by the staff in the statistical offices. Program maintenance is facilitated especially by the use of standard development tools (OVIS framework of the co-ordinated statistical system). With regard to the changes concerning the operation of co-ordinated procedures, it was demanded to have a multitenant application that can be used within the scope of ZPD.

Another goal was that program development should take account of standard software for individual work steps (SteP tools (SteP = standardisation of the production process)) that has been developed or co-ordinated in the co-ordinated system and harmonises both the work processes and the IT system architectures of the applications – at least to the extent allowed by the development status of the various tools. Therefore, the possible uses of the following tools were examined as part of the requirement analysis: survey database, input database, classification server, PL Editor, register systems and StatSpez as a tool for tabulation. The verification code for data editing that is created in the specialised unit by means of the PL Editor has been integrated into the new application. Also, the integration of the classification server on the basis of the system used in 2007 has been prepared. By the end of the project, access to the survey database as a central place for data editing verification codes will be implemented.

2. Organisational goals

Altogether, the new procedure is expected to allow better support of the work processes in processing the statistical data received. Therefore, functions of the old procedure were adjusted when programming the new procedure (export of the provisional data) and new functions were added (table of key results), the accessibility of information was improved by integrating new navigation tools and additional comparative data were included for data editing.

3. Ergonomic goals

The new application has a graphic user surface which is tailored to user needs and, consequently, is easy to use. Changing over from mainframe applications to the client-server solution has been accomplished by users without major readjustment problems.

4. Economic goals

Most of the goals described under the above items include economic aspects. The main purpose of more effective IT support for the activities in the statistical processing process is timely, high-quality and efficient statistics production. As better maintenance of the programs used is now possible, the programming efforts required for program maintenance during utilisation of the application can be reduced. Also, taking account of the specific requirements for ZPD meets the prerequisite for cross-agency operation of the application. It is not possible yet to assess the extent to which those goals have been achieved. However, centralised operation has proved very useful especially in the initial phase because, when problems occurred, it allowed much faster reaction than had been possible before.

The software package developed by IT.NRW mainly consists of two parts: The client-server application for data processing and data editing, based on JAVA and the OVIS framework, and a number of StatSpez programs for further processing and tabulation of the edited data.

1. Processing application

The range of functions of the processing application is based on the functions available in the old application for capture and editing of the statistical data received. In addition, the following extensions and adjustments have been implemented:

- Extension of the control functions for the processing process,

- automated conversion of the data reported by units switching from one group of respondents to another (not data loss for such units),

- change in user guidance, so that the reporting unit rather than the work step to be performed is the starting point when handling the data received,

- concept of rights in user administration adjusted to the concept of the existing MBB,

- change in the functions regarding the estimation of lacking data, so that provisional estimations, which are not stored in the database, and final estimations can be performed,

- provision of tables with key results for assessing data quality,

- separate administration of pre-filled contents for paper and IDEV questionnaires,

- integration of a maintenance function for the steering file regarding the product classification.

When programming the dialogue application, care was also taken to develop largely general solutions for program functions which were expected to be needed in the same form also for processing the other decentralised statistics in manufacturing. That approach makes it easier to integrate other statistics into the StatVG application.

2. Programs for further processing and tabulation

The StatSpez programs for further processing of edited data mainly cover the range of functions available in the old procedures for the production of working tables and publication tables, for primary confidentiality and for the provision of various data. Using production specifications and production orders, the required work steps are automated as much as possible in order to support the activities in the specialised units.

4. Changing over to the new procedure

The first version of the application was put into operation at the beginning of 2009. It is provided centrally by IT.NRW and is operated under ZPD conditions. All statistical offices of the Länder have changed over from the old procedures to the new application. As is probably the case for any new procedure, problems occurred in the initial phase which, however, were quickly solved, so that the production index for January 2009 was published in early March as scheduled.

Dr. Michael Forster, tel: +49(0)211 / 9449 5518, e-mail: michael.forster@it.nrw.de

Doreen Haase, tel: +49(0)/211 / 9449 5177, e-mail: doreen.haase@it.nrw.de

Ingo Wagner, tel: +49(0)611 / 75 47 46, e-mail: ingo.wagner@destatis.de

List of Municipalities Information System (GV-ISys):
New regional code and web page

A major goal of official statistics is that all data collected and processed are available and supplied in a timely manner in the form of tables, databases and master files for further internal processing and for utilisation by external users such as ministries, political parties as well as economic, scientific and opinion research institutions. To achieve that goal, and taking account of the rapid development in automated data processing and in the programming languages used, it was just a matter of time when the space-related regional classification of Germany, which is a cross-section task of official statistics, would change over from centralised to decentralised maintenance.

By putting into operation the newly developed List of Municipalities Information System (*Gemeindeverzeichnis-Informationssystem – GV-Isys*) in the statistical offices of the Federation and the Länder, the project was implemented in mid-2007 and it was finished for the time being when a web page was set up in early 2009.

With the introduction of the new database system, which is available centrally at the Federal Statistical Office and is designed for decentralised access, the statistical offices of the Länder for the first time are able to directly access the regional classification of their administrative levels with all regional units and the relevant variables. The statistical offices of the Länder are now able to enter, on their own authority, their territorial changes adopted and to be implemented.

For the first time in 2009, the twelve-digit regional code is used in GV-ISys as the main identifier of municipalities. Its code digits 1 – 5 and 10 – 12 can be used to derive the eight-digit Official Municipality Code used in many statistics. In code digits 6 – 9 the regional code contains the allocation of a municipality to an association of municipalities. This harmonises the coding of the level of associations of municipalities, which in the past differed between Länder, with a view to comparable values of regional units in member states; in the context of an increasing number of enquiries by national and international institutions – especially by the Statistical Office of the European Communities (Eurostat) – such harmonisation has become necessary.

What has been achieved as a result is the complete representation of the level of associations of municipalities, consisting of municipalities belonging or not belonging to an association and of areas not attached to a municipality, because the municipalities in those Länder which so far have not had the level of associations of municipalities are treated as separate associations with one member municipality. The regional units of the level of associations of municipalities are distinguished by the sixth digit of the code (the "t" characteristic), with "0" indicating a municipality not belonging to an association of municipalities, "5" a municipality belonging to an association and "9" an area not attached to a municipality. The 9-digit code of associations of municipalities and the 12-digit regional code are structured as follows:

Association of municipalities

Type*)	L	L	R	K	K	t	V	V	V
1)	X	X	X	X	X	0	0	0	0
2)	X	X	X	X	X	0	0	1	0
3)	X	X	X	X	X	5	7	0	1
4)	X	X	X	X	X	9	4	4	4

Municipality

L	L	R	K	K	t	V	V	V	G	G	G
X	X	X	X	X	0	0	0	0	0	0	0
X	X	X	X	X	0	0	1	0	0	1	0
X	X	X	X	X	5	7	0	1	1	2	3
X	X	X	X	X	9	4	4	4	4	4	4

*) 1) Town not attached to an administrative district, 2) Municipality not belonging to an association of municipalities, 3) Municipality belonging to an association of municipalities, 4) Area not attached to a municipality.

LL=*Land*, R=*Regierungsbezirk* (administrative region), KK=*Kreis* (administrative district), t=characteristic, VVV=association of municipalities, GGG=municipality

In rows 1) and 2) the table shows that for a town not attached to an administrative district an association code of "0000" is now indicated. For a municipality not belonging to an association the three digits of the municipality code are repeated in code digits 7 to 9. The purpose is to illustrate now the all-encompassing representation of the association level where a municipality not belonging to an association is treated as a separate association of municipalities. Other, e.g. Land-specific, distinctions between associations of municipalities and their member municipalities can be made using "text characteristics".

With the new all-German standard coding of the level of associations of municipalities, it is possible to compare administrative territorial classifications of similar sizes. Such classifications can further be distinguished by the non-administrative breakdowns available in GV-ISys such as breakdowns in terms of settlement structure, which may result in areas with comparable degrees of urbanisation. For example, a direct comparison at municipality level between the Land of Hessen (with 426 politically independent municipalities and some 6 million inhabitants) and the Land of Rheinland-Pfalz (with 2,306 municipalities and some 4 million inhabitants) does not seem useful. If, however, we examine the 163 associations in Rheinland-Pfalz, with an average 25,000 inhabitants, and the data of Hessian municipalities (an average of 14,000 inhabitants), we will get results that are better comparable.

Another goal in the context of the new List of Municipalities Information System is to provide direct online access to selected public contents of the GV-ISys, so that better up-to-dateness and flexibility is ensured for access in terms of the most recent territorial status. Implementing that project had to be postponed for technical reasons. However, an intermediate solution for data provision has been found that is acceptable and feasible both for internal and external data users: Based on its long-term experience with the data requests from the statistical departments and from external consumers, the specialised unit responsible for the List of Municipalities at the Federal Statistical Office has compiled a comprehensive range of data for online access; the relevant products are available for free download. That range comprises the search for selected municipalities, complete regional breakdowns with all politically independent municipalities and their superior administrative levels such as association of municipalities, administrative district, administrative region and Land, and a multitude of various standard tables covering a wide range of topics such as breakdowns in terms of settlement structure and regional development, allocation to constituencies in Bundestag and Landtag elections, travel areas, etc. Of course, we continue pursuing the original goal, which is direct online access to the GV-ISys data in order to further enhance the up-to-dateness of the data provided.

Another important tool for monitoring regional trends is the territorial changes which now are performed monthly in a decentralised way by the statistical offices of the Länder (see above). The territorial changes – consisting of partial exclusions/partial inclusions, complete inclusions into existing or new municipalities, code and name modifications – are documented within GV-ISys and made available in the form of tables. Hence those tables contain important information on municipalities to be included, to be created, to be dissolved or on municipalities excluding parts of themselves, on changes in associations of municipalities and their member municipalities as well as on changes in the name of regional units. Territorial changes are always provided for the ending quarter.

Harry Krajzar, tel: +49(0)611/75 28 75, e-mail: harry.krajzar@destatis.de

The German export performance and how it is determined

Components of the export performance indicator

Between 2000 and 2008, German exports were up by two thirds. On average, this was an annual increase of nearly 7%. However, these figures do not provide any information on the development of the market share of German goods in the volume of imports of the partner countries. This development is measured by an indicator which is referred to as export performance.

The indicator used by the Federal Statistical Office measures the German market share as the share of German exports to the partner countries in the total of imports of these partner countries, each in the reference period t compared to base year 0.

German export performance: $$ExPerf_t^{DE} = \frac{Ex_t^{DE} \Big/ \sum_i Im_t^i}{Ex_0^{DE} \Big/ \sum_i Im_0^i}$$

where: $ExPerf_t^{DE}$ = German export performance in period t

$Ex_{t,0}^{DE}$ = German exports to the partner countries in periods t, 0

$Im_{t,0}^i$ = Total imports of the partner countries in periods t, 0

The measure of the market share used by the OECD is the ratio between exports to the partner countries and the export market of the country whose performance is to be determined. The export market is defined as the total of weighted imports of the partner countries. To determine, for instance, the German export performance, current imports are weighted with the German market share in the base year, i.e. the share of German exports to each partner country in the base year in the total imports of each of these countries.

$$ExM_t^{DE} = \sum_i \frac{Ex_{i0}^{DE}}{Im_0^i} Im_t^i$$

where: ExM_t^{DE} = German export market in period t

Ex_{i0}^{DE} = German exports to partner country i in period 0

$Im_{0,t}^i$ = Total imports of country i in period 0 or t

OECD export performance: $$ExPerf_t^{DE} = \frac{Ex_t^{DE}}{ExM_t^{DE}}$$

For reasons of simplification, not all of the nearly 200 partner countries are covered. The analysis rather includes only the relations with the major partner countries.

Development of the export performance

The interest focuses on analysing changes in the export performance. In the context of the German index system, the change in total German exports is put into relation to the change in the total imports of the countries which, from a German point of view, are the most important

partners. As regards the OECD system, the relation is established to the development of the German export market.

Development of the export performance based on the German definition:

$$\Delta ExPerf_t^{DE} = \frac{Ex_t^{DE}\Big/Ex_{t-1}^{DE}}{\sum_i \mathrm{Im}_t^i \Big/ \sum_i \mathrm{Im}_{t-1}^i} - 1$$

Development of the export performance based on the OECD definition:

$$\Delta ExPerf_t^{DE} = \frac{Ex_t^{DE}\Big/Ex_{t-1}^{DE}}{ExM_t^{DE}\Big/ExM_{t-1}^{DE}} - 1$$

In accordance with the OECD definition, this ratio can also be expressed as the relation between the size of the export market, its development and the development of total exports. To this end, the size of the export market[1] is multiplied by the difference of export growth and export market growth.

$$\Delta ExPerf_t^{DE} = \frac{ExM_{t-1}^{DE}}{ExM_t^{DE}}\left(\frac{Ex_t^{DE}}{Ex_{t-1}^{DE}} - \frac{ExM_t^{DE}}{ExM_{t-1}^{DE}}\right)$$

Development of the German export performance in relation to individual partner countries

To measure the change in the export performance with regard to individual partner countries, the change in German exports to a given country is put into relation to the change in the total imports of that partner country. The two definitions of export performance coincide in this respect:

$$\Delta ExPerf_t^{DE,FR} = \frac{Ex_t^{FR}\,\mathrm{Im}_{t-1}^{FR}}{Ex_{t-1}^{FR}\,\mathrm{Im}_t^{FR}} - 1$$

where $\Delta ExPerf_t^{DE,FR}$: Change in the German export performance in relation to France between period t-1 and t

$Ex_{t,t-1}^{FR}$: German exports to France in periods t, t-1

$\mathrm{Im}_{t,t-1}^{FR}$: French imports in periods t, t-1

Assessing the export performance

In line with the German definition, the development of the German export performance is considered to be good if the development of total German exports is more favourable than the development of the total imports of the partner countries. In other words, the German share in the import markets of the partner countries increases in period t compared to the previous period t-1.

---

[1] Measured as the export market size of the previous year t-1 relative to the reference year t.

Based on the OECD definition, the German export performance shows a positive trend if the development of total German exports exceeds the growth of the German export market. The index trend will be even more positive if, on the one hand, the increase in exports is higher than the export market growth and, on the other, the export market decreases compared to the previous year.

Practical aspects of the export performance

The indicator does not exclusively reflect German competitiveness; it can rather be affected by external factors, too. Declining trends could also be due to increasing crude oil prices as the price elasticity of the crude oil demand is rather small. Rising crude oil prices raise the value share of this raw material in the total imports of the partner countries at the expense of other import goods. As Germany is not a major crude oil exporter, a shift of the above type could deteriorate the German export performance.

In addition to the export data of the country concerned, the total import figures of the partner countries are needed for the actual index computation. For reasons of better comparability and consistency of the partner country data, it is recommended to obtain these data from one and the same international source because the data supplies of the countries concerned can be expected to be based on a uniform methodology. Also, the data are denominated in the same currency, and comparable official exchange rates are used for their conversion. In this way, the problematic impact of the assumption made for indicator use can be lessened to a certain extent. The assumption is that the total of imports of a given country corresponds to the total of exports of the partner countries to that country.[1] This is an abstracted approach to the problem of mirror exercise differences (discrepancies), different evaluations of imports and exports (CIF/FOB problem) and exchange rate effects.

Silke Gehle-Dechant, tel.: +49(0)611/75 34 92, e-mail: silke.gehle@destatis.de

Foreign trade by enterprise characteristics

The trade relations between countries are an important issue in the discussion about globalisation. As a matter of fact, however, the extent of these external economic ties varies between the individual branches of the economy. There are goods which are typically imported (e.g. raw materials) and other goods which are typically exported by a country (e.g. machinery).

Obtaining branch-specific information presupposes a link between foreign trade statistics and business data. After almost ten years of carrying out preparatory work in the European Community and conducting project studies and voluntary analyses, which the Federal Statistical Office has been involved in right from the beginning, such linkage has now been established by both the new EU regulation regarding the compilation of Community Statistics on external trade with non-member countries and the amended Intrastat regulation, which require all Member States to provide an annual sectoral analysis of foreign trade results.[2] The goal of establishing the above link is to supplement the conventional commodity-related variables of foreign trade statistics by business-related variables such as economic activity, number of persons employed and turnover. As a result, information can for instance be obtained on the motor vehicle

---

[1] Durand, Martine; Simon, Jacques; Webb, Colin: OECD's Indicators of International Trade and Competitiveness; OECD Economics Department, Working Papers, No. 120, Paris 1992, p. 32

[2] See Article 6, 2. of Regulation (EC) No 471/2009 of the European Parliament and of the Council of 6 May 2009 on Community statistics relating to external trade with non-member countries and repealing Council Regulation (EC) No 1172/95 and also Article 12, 4. of Regulation (EC) No 222/2009 of the European Parliament and of the Council of 11 March 2009 amending Regulation (EC) No 638/2004 on Community statistics relating to the trading of goods between Member States.

industry's share in total German exports, the number of persons employed in enterprises in that industry or the proportion of movements of goods undertaken by the ten biggest enterprises.

As early as in 2002, the Federal Statistical Office conducted a feasibility study on linking foreign trade and business data in the framework of a Eurostat project for the reference year 2000. The study provided a basis for the Standardisation Exercise. In the context of this Exercise, a group of Member States has, since 2005, provided Eurostat with annual standardised data evaluations regarding the business structure in the foreign trade area. The list of variables has been extended gradually over time. Right from the beginning, the Federal Statistical Office has co-operated in the Eurostat Trade Register – Globalisation working group. In 2008, the latter was renamed Linking Trade and Business Statistics working group and the range of its members enlarged considerably.[1] The data supplies to Eurostat now prescribed by the relevant regulations will be based on the above preliminary work.

So far, German foreign trade statistics have provided sectoral data exclusively for the trade with other EU Member States. The reason is that there has been no common identifier with the business register which, however, would have been required in relation to extra-Community trade statistics. The different data situation in intra-Community and extra-Community trade statistics is due to the fact that German foreign trade statistics are based on two data sources. These are, on the one hand, the intra-Community trade system for the trade with other Member States of the European Union (EU) and, on the other, the extra-Community trade system regarding the trade with third, i.e. non-member countries. To compile intra-Community trade statistics, data are collected directly from businesses whose imports and exports each exceed a certain value-related threshold in the reference year[2]. The proportion of trade exempted from declaration is included as an estimate in the national foreign trade result. As regards the extra-Community trade system, the declarations submitted for customs clearance purposes are still evaluated as a basis for the relevant statistics.

Since both the monthly data of intra-Community trade statistics and the business register include turnover tax (VAT) codes as a common element, a link can be established with the detailed business register data and structural information be derived on the businesses engaged in foreign trade operations. Based on the new regulation, which requires the customs authorities to provide the national statistical agencies with adequate identification codes for the economic operators, Germany will be in a position to supply sectoral data on the foreign trade with non-member countries (extra-Community trade), too.[3] Due to the strong EU internal market orientation, however, the relevant results so far obtained exclusively for the intra-Community trade in goods have covered not less than about three fifths of the total of German imports and exports.[4]

The data for reference year 2006, which were transmitted to Eurostat in the context of the Standardisation Exercise 2008 in due time in February 2009, for the first time included information on the number of enterprises and foreign trade turnover by number of persons employed for businesses whose volume of trade was below the declaration threshold (EUR 300,000 for each direction of trade in reference year 2006). Based on the information stored in the intra-Community trade register on all firms with foreign trade activities, the data on businesses whose trade volume is below the declaration threshold can be included in the analysis, too.

Sabine Allafi, tel.: +49(0)611/75 29 93, e-mail: sabine.allafi@destatis.de

---

[1] In addition to Destatis, the authorities responsible for the compilation of foreign trade statistics of the Member States Denmark, Italy, Lithuania, Latvia, the Netherlands, Portugal, Romania, Finland, Slovenia and Sweden currently attend the annual meetings in Luxembourg.

[2] Since 1 January 2009, this threshold has been EUR 400,000.

[3] See footnote 1.

[4] See Allafi, Sabine: Gesamtentwicklung des deutschen Außenhandels 2007 weiterhin positiv, Wirtschaft und Statistik 5/2008, pp. 409 ff.

New EU legal acts for foreign trade statistics

The Intrastat basic regulation (Regulation (EC) No 222/2009), which covers the foreign trade with other EU Member States, took effect retroactively as from 1 January 2009, while the Extrastat basic regulation regarding the trade with non-member countries outside the European Union shall be put into force from 1 January 2010. These regulations amend or replace the previous Regulations (EC) No 638/2004 (Intrastat) and (EEC) No 1172/1995 (Extrastat).

The goal of revising the above two regulations was to make their wording clearer, simpler and easier to understand. The regulations lay down that, based on a linkage with the business register, foreign trade statistics shall provide annual information on the distribution of foreign trade by economic sectors and enterprise size classes without placing an additional burden on respondents. In addition, the Extrastat data available shall be used to compile biennial statistics on the international importance of the euro compared to other international trading currencies.

Another goal is to improve data quality. In particular, surveys conducted among businesses for the purposes of intra-Community trade statistics are characterised by continuous non-response. Data on the trade in goods have therefore to be estimated and added for businesses which fail to submit their declaration or do not keep the relevant deadlines or which are exempted from the obligation to submit a declaration. The measures taken in this respect include the preparation of criteria for measuring quality, the compilation of an annual quality report by the national statistical offices and authorising Eurostat to assess the quality of the data received.

Extrastat

New customs declaration procedures with potentially significant effects on statistics have required an amendment of the Extrastat regulation. The new declaration procedures (regarding Single Authorisation, Centralised Clearance or Self-Assessment) have led to a separation of the movement of goods and the declaration thereof as the customs declaration can now be submitted in another Member State than the state into which a commodity is actually imported or from which it is exported. Another option is that businesses only have to store the relevant data in their accounting systems (Self-Assessment). Extra-Community trade statistics is a secondary statistics which is based on the data provided by the customs administration (data collection office) in the form of what could be called a copy of the original declaration. Hence it faces the problem that the data required for the compilation of national foreign trade statistics may either be available in another Member State which is not concerned by the given physical movement of goods or be not available at all if a declaration is not requested for customs purposes.

The problem in this context is the gap between the European orientation of the customs administration and the national orientation of the Member States' foreign trade statistics as part of the national balance-of-payments statistics and national accounts. And another general problem of statistics becomes apparent in this context, too. While, on the one hand, the data quality requirements have increased, on the other, the response burden is to be reduced by cutting bureaucracy. Due to simplification measures in the customs area, part of the administrative source used for the purposes of foreign trade statistics has been lost. For this reason, the relevant trends and effects on statistics should be closely monitored in this field. Political players should be made aware of potential effects at an early stage and attempts be made to counter losses of quality which might be the result of such trends.

Hence the new version of the statistical basic regulation must make sure that the data of a customs declaration which are relevant for statistical purposes will be channelled into the EU Member State where they are required for the compilation of national foreign trade statistics. As regards the Self-Assessment procedure, the regulation lays down that the businesses themselves shall provide the data needed.

Since the Rotterdam effect, as it is called, causes a misallocation of statistical data which might be faced in a similar way in relation to the new customs declaration procedures, too, the new basic regulation also provides for a reduction of this effect. The problem arises, for instance, in

cases where goods are subject to customs clearance procedures in big sea ports and thus are statistically recorded in the Member State of the port although they are only moved through that country (transit operation), while the actual Member State of destination or of export is another country. This trend of inflating the trade of the Member States at the EU external borders shall be countered.

Intrastat

In accordance with the attempt to avoid an unnecessary burden on businesses caused by statistical obligations, the goal of the amended version of the Intrastat basic regulation is to further reduce the burden on respondents in the short run. At the same time, however, noticeable losses in the quality of the statistical data shall be avoided.

Based on an exemption threshold of EUR 400,000, nearly 90% of all businesses engaged in intra-Community trade have been exempted from the obligation to provide information in Germany. The number of businesses that are required to give information amounts to about 65,000. Their effort in relation to the monthly declaration of their intra-Community trade operations in a detailed commodity-related and regional breakdown accounts for more than 40% of the overall response burden caused by economic statistics as a whole.

To reduce the burden on businesses, the minimum coverage rate for arrivals of goods will now be 95% instead of 97%. As far as dispatches of goods are concerned, however, the coverage threshold will continue to be 97%. The difference in the required minimum coverage rates regarding the two directions of trade is intended. There are plans to introduce other medium-term to long-term simplifications (such as the single flow system) whose implementation should not be hindered by wrong decisions. The single flow system would only record dispatches of goods, while information on arrivals would be derived from the dispatch data of the relevant partner countries. If, however, the information on arrivals were produced from the dispatch data of the partner countries, the minimum coverage rate of the latter would have to be high so that the "receiving" countries with low declaration thresholds would not suffer from considerable information losses. For this reason, minimum coverage will remain at 97% regarding dispatches of goods.

In Germany, the room created by the reduction of minimum coverage can be used for raising the declaration threshold after an amendment of the national Foreign trade statistics implementing regulation (AHStatDV).

In the future, the Commission will be in a position to flexibly "adapt the minimum Intrastat coverage to technical and economic developments" while, in doing so, it will have to take into account the given quality indicators and standards. This will require an amendment to the relevant EU implementing regulation. In the whole process, however, the Commission will be subject to what is called a regulatory procedure with scrutiny. This means that scrutiny will be exercised by the bodies involved in the preparation of legislation (European Council and European Parliament) before such a measure can be taken.

Silke Gehle-Dechant, Tel.: +49(0)611/75 34 92, E-Mail: silke.gehle@destatis.de

Publications

Data Report 2008: the newly published social report for Germany

The Data Report is a co-production of the Federal Statistical Office (Destatis), the Social Science Research Center Berlin (WZB) and the German Social Science Infrastructure Services Association (GESIS). With its combination of data from official statistics and from social research, it has become a standard work of German social reporting by now. The Data Report offers a wealth of information on the social and economic situation in Germany and in Europe and illustrates people's subjective perception and assessment of their quality of life.

The 2008 edition provides important background information concerning the reforms currently under discussion: How has the labour market developed over the past years? How do people assess their state of health and how much do they spend on health? How much expenditure is spent on the social security systems and who receives benefits from these systems?

In the new Data Report, the results of the Federal Statistical Office and the social research institutes have been restructured and rearranged to achieve a closer thematic relationship. A colour guide system helps to distinguish the contributions by the Federal Statistical Office and by the social research institutes.

The Data Report is available for free download from the Federal Statistical Office's website (http://www.destatis.de). Against payment of a nominal fee of EUR 4 plus forwarding costs, it can be ordered in book form from the Federal Agency for Civic Education (www.bpb.de) and the SFG Servicecenter Fachverlage (destatis@s-f-g.com).

Ilka Willand, tel: +49(0)611/75 45 45, e-mail: ilka.willand@destatis.de

New European handbook on the application of quality adjustment methods in European price statistics

In the last few years, a key activity of the consumer price statistics unit of the Federal Statistical Office was to manage the joint European CENEX HICP Quality Adjustment project promoted by Eurostat. The result of the project is a handbook of nearly 300 pages to be published soon under the title "Handbook on the application of quality adjustment methods in the harmonised index of consumer prices" in the publication series *Statistik und Wissenschaft* (Statistics and Science) of the Federal Statistical Office. The handbook is designed for price statisticians in the European Statistical System. Its goal is to provide them with practical assistance in introducing and developing quality adjustment methods and, at the same time, to contribute to a harmonised application of quality adjustment methods in Europe. Another important function of the handbook is to implement the new Commission Regulation (EC) No 1334/2007 with respect to the application of quality adjustment methods.[1]

Subject background

The Harmonised Index of Consumer Prices (HICP) is compiled in a decentralised manner by the national statistical institutes in Europe. Its compilation is governed by a number of regulations of the European Council and the European Commission. One of these regulations stipulates that the

---

[1] Commission Regulation (EC) No 1334/2007 includes precise definitions of the terms "consumption segment", "quality change" and "quality adjustment". In addition, it lays down the principles of using consumption segments and handling quality adjustment and it describes the minimum standards for replacements and quality adjustment.

HICP shall be a *Laspeyres-type index*[1]. Basically, this means that the indicated price trend shall refer to the products of a basket which is typical for the consumption expenditure of economic operators in a past base period.

However, the application of the above principle poses a problem if the products offered change considerably over the index period, for instance, due to technical innovation. In that case, the products originally selected might either no longer reflect the buying behaviour of consumers in a representative manner or they might cease to be available in shops. To apply the Laspeyres principle in a dynamic world of consumption, so-called *consumption segments* are therefore defined. A consumption segment includes a range of products which have common properties and serve a similar consumption purpose. The consumption segment offers a certain freedom of choice regarding the selection of products to be used for price measurement and permits price statisticians to respond to changes in the products offered.

The supply prices selected for price observation are to represent the consumption segment to the best possible extent regarding product type, outlet type and region. This requirement does not only refer to the first selection of product-offers, it rather extends over the whole index period. Due to the dynamic development of the markets, the product models which lose their representativity must thus be replaced by new models in due time to ensure, at any time, the representativity of the products selected.

However, a comparison of the supply prices of two successive months can only provide informative results if the quality of the products is maintained at a constant level. If the quality of the goods changes, so-called *quality adjustment* measures must be undertaken. Frequently, such adjustment uses as a basis the consumers' preparedness to pay for certain quality properties. How much would consumers for instance be ready to pay for a prolonged (rechargeable) battery life of notebooks or for larger TV screen sizes? The relevant expenditure is considered as the 'monetary value of quality changes' in the price comparison. The focus of the new handbook is the approach to be adopted in calculating quality-adjusted consumer price indices. It is illustrated using product examples from various areas such as computers, washing machines, television sets, used and new cars, books and software.

New form of co-operation

The handbook on the application of quality adjustment methods in European price statistics is the result of a pilot project based on a new form of co-operation within the European Statistical System, which was developed some years ago. The project team included price statisticians from Germany, Belgium, Ireland, the Netherlands, Austria, Portugal and Sweden. The project team was assisted and guided by a steering committee whose members comprised representatives from the statistical offices of Finland, Great Britain and Italy and from the European Central Bank and Eurostat.

The basic idea of this form of co-operation is to involve statistical experts from various countries in the development of certain methods and, subsequently, to provide all statistical institutes in the European Union with the results they have achieved. The idea of this type of co-operation was raised as early as in the Palermo Action Plan of the European Statistical System in 2002. On its basis, a Task force led by the Federal Statistical Office was established, which suggested the so-called Centres and Networks of Excellence (CENEX) to function as co-operation networks. The Task force proposed price statistics as the subject of a pilot project which was implemented under the guidance of the Federal Statistical Office in 2006.

The principles of co-operation have been further developed. CENEX was renamed ESSnet (European Statistical System networks) and, as a model, integrated into the Community Statistical Programme 2008-2012, which specifies important milestones for the medium-term and long-term development of European official statistics. Recently, co-operation networks like

---

[1] See Article 9 of Council Regulation (EC) No 2494/95 concerning harmonized indices of consumer prices.

the ESSnet were also included in the revised Regulation of the European Parliament and the Council on European Statistics and have thus gained even more importance in the context of the European Statistical System. The above regulation lays down, in a detailed manner, the framework for the development, production and publication of European statistics.

Stefan Linz, tel.: +49(0)611/75 26 59, e-mail: stefan.linz@destatis.de

Events

8th scientific conference: "Non-reactive survey methods"

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 years' intervals. This year's eighth joint conference was held in Wiesbaden on 25 and 26 June 2009 and was devoted to the subject of "Non-reactive survey methods". The conference was moderated by Prof. Dr. Frank Faulbaum, Chairman of the Working Party of Social Science Institutes (ASI).

After Roderich Egeler, President of the Federal Statistical Office, had welcomed the conference participants, Professor Faulbaum provided an introduction to the conference theme in his overview paper and described the historical development of non-reactive methods.

Then, Hartmut Scheffler, Chairman of the Association of German Market and Social Research Institutes (ADM), discussed the possibilities and limits of data collection by other means than polls.

Franz-Josef Kilzer of TNS Infratest presented a report on behaviour studies at the point of sale.

The possibilities of integrating biometric data were illustrated by Prof. Dr. Rainer Schnell of Duisburg-Essen University.

Following these contributions, Dr. Holger Heidrich-Riske of the Federal Statistical Office in Wiesbaden explained the possibilities of using geographical information systems for official statistics.

The first conference day was concluded by Dr. Andreas Czaplicki of the Institute for Market Research in Leipzig who talked about the use of GPS data, by means of which the routes taken by visitors at Leipzig Zoo were analysed.

At the beginning of the second day, Dr. Tanja Hackenbruch of GfK Telecontrol AG in Bern presented a paper describing the technique of collecting data with the help of Mediawatch.

Prof. Dr. Martin Welker of Macromedia University for Media and Communication (MHMK) in Munich provided information on the uses of log file analyses.

Afterwards, Andrea Maldonado of the Federal Statistical Office in Wiesbaden gave an account of how the register of addresses and buildings was set up for the census 2011.

At the close of the conference, PD Dr. Ulf-Dietrich Reips of Zurich University presented a paper entitled "Brave new world of research", which examined future trends of non-reactive survey methods.

The conference volume containing all the contributions will be published as part of the GESIS series of conference reports. It will presumably become available in the autumn of 2009. The volumes for the three previous joint scientific conferences on "online surveys", "data fusion and data integration" and "quality management and quality assurance" were also published in this series. For information about these conference volumes please visit:

http://www.gesis.org/en/product-catalogue/?vgesissecen=-2,-9_-8&order=_title.en

Christian König, tel.: +49(0)611/75 20 77, e-mail: christian.koenig@destatis.de