

# METHODS – APPROACHES – DEVELOPMENTS

## Information of the German Federal Statistical Office

**Edition 1/2010**

### The catchword

Quality aspects in foreign trade statistics at the European level ..... 3

### Methods of federal statistics – Further development

Online questionnaire applied for the first time for the general information of the sample survey of income and expenditure 2008 ..... 5

Developing an algorithm to calculate in air transport statistics the air kilometres travelled in the German air space ..... 10

Study on the use of multiple imputation methods in wholesale trade statistics ..... 16

Implementing the new FATS regulation without causing an additional burden ..... 17

Using hot-deck procedures in foreign trade statistics ..... 20

Revision of the UN methodological manual for foreign trade statistics ..... 21

Simplification of intra-Community trade statistics: current situation and trends ..... 23

Quality assurance in public finance statistics ..... 25

Agreed earnings online: The agreed earnings database is a new basis for analyses and activities in the statistics of agreed earnings ..... 28

### Events

18th Scientific Colloquium: “Information visualisation – Graphical processing and statistical data analysis” ..... 30

Gerhard Fürst Award 2009 ..... 31

**Published by:** Federal Statistical Office, Wiesbaden

**Subject-related information**

On this publication:

Division I B,  
tel: +49 (0) 611 / 75 20 77  
fax: +49 (0) 611 / 75 39 50  
[christian.koenig@destatis.de](mailto:christian.koenig@destatis.de)

**General information**

on the data supply:

Information service,  
tel:+49 (0) 611 / 75 24 05  
fax: +49 (0) 611 / 75 33 30  
[www.destatis.de/kontakt/](http://www.destatis.de/kontakt/)

**Release calendar  
of the Press Office:**

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

Periodicity: (generally) twice a year

Published in June 2010

© Federal Statistical Office, Wiesbaden 2010

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

## The catchword

### Quality aspects in foreign trade statistics at the European level

For some time already, European foreign trade statistics has been dealing with the issue of quality. Since 2001, there has been a Eurostat working group on quality whose task is to achieve sustainable improvement in the quality of foreign trade data in the EU member states and a maximum harmonisation of the methods applied. In 2005 the European Statistics Code of Practice was published, in which all institutions involved in the European Statistical System (ESS) have committed themselves to complying with 15 principles. Subsequently, the quality criteria specified in the Code of Practice have systematically been transferred to, and concretised for foreign trade statistics. Today there is a comprehensive quality management system which not only defines quality criteria and indicators but also monitors compliance with them in separate quality reports and, since recently, has also assessed them in assessment reports. In addition, a detailed quality handbook is being prepared, which will reflect the progress made in all quality-related activities and will provide recommendations and methodical descriptions.

From the very beginning, the goal of quality assurance and improvement has involved the simplification and modernisation of foreign trade statistics. This was mainly aimed at intra-EU trade and, only to a small extent, to extra-EU trade. Based on the general quality standards and recommendations of the ESS and the Code of Practice, foreign trade statistics has set up manifold quality indicators of its own for all main quality goals and has included some of them as binding (minimum) quality standards in the legal regulations. More demanding proposals for improvement are made to the member states as (non-binding) recommendations for the future. In the entire co-ordination process, the general direction is not defined by Eurostat alone but is discussed and agreed on in the working group on quality in which Germany and 13 other member states are represented.

In 2007 the reduction of asymmetries in the mirror trade statistics of the EU partner countries<sup>1)</sup> was included as another goal in addition to general quality improvement. A goal which not only has a direct impact on quality but also is intended to take the ESS a step closer to the ambitious long-term goal of introducing the single-flow method<sup>2)</sup>.

To achieve its quality goals, the working group on quality has introduced some tools which will support its activities of measuring, documenting and assessing quality and the faster reduction of asymmetries in the mirror statistics:

- Quality report

Since 2004 an annual quality report on foreign trade statistics in the EU member states has been compiled. It is modified and extended every year and summarises all major quality indicators regarding relevance, timeliness and punctuality, accuracy, accessibility and clarity, comparability and coherence, while adding explanatory texts. In the 2009 quality report, further issues have been added, which will have to be concretised by relevant indicators. The quality report currently is the most important source of information on the quality standards of foreign trade statistics in the EU and is available free of charge to all users on the Eurostat website.

- Assessment reports

Since 2009 the quality of foreign trade statistics in the EU member states has annually been assessed by Eurostat in specific assessment reports. A distinction is made between individual

---

1) The imports registered in a country as coming from a partner country should mirror the exports registered in the partner country as going to the other country and vice versa. Where this is not the case, the term of asymmetries is used.

2) In the single-flow method, only one direction of trade is covered (exports), while the other direction (imports) is obtained through the export data of the trading partners.

assessment reports for each country and a summary assessment report showing all countries together without mentioning individual member states. In both reports, a detailed pattern is applied to assess and rate the compliance with legal quality requirements and recommendations as well as the methods and procedures applied. Grading systems ranging from 1 to 4 for the legal requirements and from A to C for meeting the quality recommendations are applied. While the summary report should be available to all member states, the dissemination of the individual reports is limited to the countries concerned for the time being. All data contained in the assessment reports refer to the situation and the results of the last but one year, as is the case for the quality report.

- Quality handbook

For 2010 Eurostat is planning the publication of a comprehensive quality handbook. It will be a kind of compendium on the issue of quality containing, among other things, a description of the entire quality framework and all relevant quality documents and indicators. Also, it will document recommendable methods and procedures regarding supplementary estimations, revisions, reduction of asymmetries, confidentiality and checking procedures and reflect the progress made in quality as shown in the quality reports and assessment reports.

- Annual reconciliation exercises aimed at reducing major asymmetries

As commissioned by Eurostat, reconciliation exercises have been performed for the first time this year as a standardised procedure to identify, examine and eliminate the largest asymmetries in intra-EU trade at a detailed commodity level (8-digit commodity codes). The main asymmetries are selected by Eurostat according to specific criteria<sup>1)</sup> and transmitted to the member states. Then the member states have to check, and eliminate as far as possible the discrepancies in co-operation with their partner countries. Germany participated in 25 of the total of 69 reconciliation exercises performed in 2009. Those reconciliation exercises will now be performed annually in order to identify in the long term in particular persistent discrepancies and those that are due to methodology. In addition, for continuous monitoring of the asymmetries, Eurostat regularly provides monitoring reports to all member states in order to inform them about their largest discrepancies in the annual and monthly data.

The future activities of the working group on quality will focus not only on the further development of the various quality tools but also on the recommendation of suitable and harmonised procedures and methods regarding data revision, the treatment of confidential data and supplementary estimations in case of non-response or late data deliveries. The exchange of experience planned for those topics will, in the end, lead to defining recommendable procedures and methods or best practices which will be included in the quality handbook. Describing new quality indicators for some additional issues, which are fixed on the basis of general ESS standards, will be another function of the working group on quality. Those issues are the trade-off between conflicting quality goals (e.g. accuracy versus timeliness or accuracy versus reducing burdens on respondents), taking account of user interests, cost-benefit ratio of the data collection systems, burden on respondents as well as data protection, confidentiality and data security.

All quality aspects described are part of a complex framework that is to guarantee a high quality level in foreign trade statistics within the European Statistical System and will certainly be considered as exemplary for other statistical areas.

Joseph Steinfeld, tel: +49 (0) 611 / 75 84 74, e-mail: [joseph.steinfeld@destatis.de](mailto:joseph.steinfeld@destatis.de)

---

1) Annual discrepancy of more than EUR 600 million, important as a share in total discrepancies, persistent discrepancy, i.e. occurring in a similar size in the previous or subsequent year.

## Methods of federal statistics – Further development

### Online questionnaire applied for the first time for the general information of the sample survey of income and expenditure 2008

Within the scope of the sample survey of income and expenditure (EVS), about 60,000 households in Germany are requested every five years to supply data on their income and expenditure, capital formation, equipment with consumer durables and housing situation. Participation of households in that survey is voluntary. The EVS provides an important data basis for assessing the income and consumption situation of the various groups of the population. The EVS covers all private households at their main place of residence whose monthly household net income is less than EUR 18,000. At the beginning of any EVS survey year, the households willing to participate get the questionnaire on “general information” containing questions on the household composition, the socio-economic variables of the household members, the equipment with consumer durables and the housing situation. In a second document, the households are requested to provide data on financial and non-financial assets. In another survey document (the “household book”), the households record any monthly income, that is separately by income types, and any monthly expenditure. Expenditure is collected according to pre-defined expenditure categories. The household book is kept for three months and the participating households are distributed over the four quarters of the year. In addition, every fifth participating household keeps a “detailed diary” in which it records its expenditure made in a given month on food, beverages and tobacco in a detailed manner by quantity and price. The months of detailed recording, too, are distributed evenly over the year. For the current tenth EVS conducted in 2008, the households had the opportunity for the first time in a voluntary household survey of that size to answer the questions on “general information” in an online questionnaire instead of the paper questionnaire.

Online surveys and online reporting were introduced in official statistics in 2000, starting with parts of business statistics. Within the scope of the feasibility study “Online LWR 2004” performed in 2004, it was examined to what extent online surveys can be applied in a useful way also in official social statistics, where private households are questioned, and in voluntary surveys. The continuous household budget surveys (LWR) are surveys conducted annually on income and expenditure of private households. In terms of survey content, the LWR are similar to the EVS. Altogether, some 8,000 households of all population groups participate in the LWR, with the exception of households of self-employed and farmers and households with a monthly household net income of EUR 18,000 or over. In the feasibility study “Online LWR 2004”, the survey part on “general information” was offered as an online questionnaire in addition to the paper questionnaire. Compared with the household book, the “general information” can be implemented as a user-friendly online questionnaire because all the answers are standardised and that part of the survey does not take much time. The results on response rates and on how the online respondents assessed the online questionnaire as well as the feedback given by the statistical offices on how the survey proceeded led to an altogether very positive assessment of the survey tool “online questionnaire” in that household survey. Analysing the data quality and the socio-economic composition of online respondents provided highly positive results, too, so that the online questionnaire has permanently been applied in the LWR since 2005.

Due to the positive experience acquired in the LWR, the survey part on “general information” was offered as an online questionnaire also in the EVS 2008. Data that could be supplied online by the households to the statistical offices of the Länder, in addition to the “general information”, is their address and declaration of willingness to respond. Data capture, checking and transmission are done by means of the IDEV procedure (internet-based data collection of the statistical offices of the Federation and the Länder). Using the ID “evs2008” and the password “evs2008”, households could log on via the internet to supply their address and the data of the declaration of willingness to respond. The data contained in the declaration are the data required for selecting the household (household size, number of children in the household, social status of the main income earner, household net

income) and the information on whether or not the household is ready to respond online. Those data were checked for completeness and were exported in the process of online responding. Before sampling, when reading the addresses into the EVS data capture program, the addresses were checked for valid data and addresses transmitted several times were filtered out by means of names and postal codes. After sampling, the addresses of all sample households wishing to complete the questionnaire online were read into the IDEV program. This includes households wishing to supply the “general information” online although they had not supplied their address online. The statistical offices of the Länder sent IDs and passwords (generated by the IDEV program) to the online respondents so that they could log on for online responding. Online respondents also had the option to cancel the process of online responding. In that case, the paper questionnaire was sent to them.

Technical requirements permitting households to respond online are a PC with Microsoft Internet Explorer (version 6.0 or more recent) or Mozilla Firefox (from version 1.0) as a browser software. For help texts and other documents in the pdf format, Adobe Acrobat or AdobeReader are required. Using their ID and password, households get direct to the first page of the questionnaire; no intermediate steps such as registration are required. The questionnaire is designed as a coloured index tab structure. It is nearly identical to the paper version of the questionnaire. Respondents can go direct to the individual questions in the online questionnaire, using the computer mouse or the tab key, but they are not automatically led to the next question. The integrated plausibility checks ensure that the data are completely transmitted and that no questions are left out. The household must stay online during the entire response process because otherwise no interactive information boxes can be used. However, the explanatory notes on the questions – which in some cases are quite lengthy – can be displayed as an overall document in a separate index tab. Households may save their data at any time, for example, to interrupt completing the questionnaire. Once entirely completed, the forms are sent by the households to the relevant Land statistical office. Households may save or print a transmission record showing the data. Also, they may send the questionnaire several times to correct or add data. At the statistical offices of the Länder, the online documents are received as xml files and are read into the capture program. For the subsequent steps of data processing it is irrelevant whether the data of a household have been transmitted offline or online. As the online reports are checked only for completeness in the IDEV program, data editing is performed in the capture program, as it is done for the data of offline households. Data received more than once are treated in the capture program, and the first online report that has been read in is used for subsequent handling. Any further reports are shown side by side in one form with all boxes. It is then possible to take over the second report, or just individual boxes of the second report, instead of the first report.

20% of all households which participated in the EVS 2008 made use of the possibility to send the “general information” online.

*Carola Kühnen, tel: +49 (0) 611 / 75 88 49, e-mail: [carola.kuehnen@destatis.de](mailto:carola.kuehnen@destatis.de)*

## Analysing the health questions in the microcensus by social class

### Objective

Differences in health-related findings between social classes are well known and there is much empirical evidence for them, e.g. based on the telephone health survey conducted by Robert-Koch-Institut. Considering this, it appears even more necessary to perform a class-specific analysis of the health questions in the microcensus, too. As a continuous representative survey involving the obligation to provide information, the microcensus has very high response rates, even for the health questions, which are answered on a voluntary basis.

The index used most frequently and taking account of the three dimensions of income, education and occupational status to represent “social gradients” is the social class index. The connection between social and health inequality is treated by various theoretical models and index approaches, cf. e.g. [2, 4, 5]. For a criticism of the class model as a concept for studying health inequality cf. e.g. [6] and for a criticism of the model of index construction cf. recommendation by Dulon et al. [7].

Many models of index construction are based on a method by Scheuch [1]. Essentially, they are based on adding up points awarded for the three dimensions on a multi-stage scale. Recent recommendations of the *AG Epidemiologische Methoden* (working group on epidemiologic methods) are based on the demographic standards of the *Arbeitskreis Deutscher Marktforschungsinstitute* (Association of German Market and Social Research Institutes), the *Arbeitsgemeinschaft Sozialwissenschaftlicher Institute* (Working Party of Social Science Institutes) and the Federal Statistical Office [8]. The analysis of the health questions of the microcensus 2005 as described here was performed on the basis of the score group method of TNS-Infratest, which was also applied in the second national food consumption survey published in 2008 [9], and on the basis of the social class index construction method of the telephone health survey 2003 (GSTel03) of Robert-Koch-Institut [3].

Both social class index construction methods and their methodical implementation using microcensus data are described below.

### Social class index construction in the microcensus on the basis of the TNS-Infratest method

Household net income, level of school education and vocational training of the respondent and occupational status of the household’s main income earner are the items to which points are attributed which are then added up. Then the persons are allocated to one of five social classes, that is lower class (4 - 8 points), lower middle class (9 – 11 points), middle class (12 - 15 points), upper middle class (16 - 19 points) and upper class (20 - 25 points).

Transferring the index dimensions to the microcensus catalogue of questions required several steps, which are briefly explained here for every dimension. Detailed information may be obtained from the authors.

The income is included in the index as household net income. The number and age of household members is irrelevant here. In case of non-response, it was assumed that the household net income corresponds to the average household net income of the old Länder (5 points) or the new Länder (4 points), depending on the place of residence.

For school education and vocational training, far-reaching consistence with microcensus answer categories was achieved. Compared with the dimensions of income and occupational status, the number of non-response cases was smaller for the educational dimension (under 1%). For this reason, and because the household members can differ between classes only in this item, a person is allocated to a social class only if data on the highest school qualification obtained are available.

As was the case in TNS-Infratest, the occupational status examined was that of the main income earner rather than that of the respondent. Where no information was available, 3 points were attributed.

## **Social class index construction in the microcensus on the basis of the telephone health survey 2003**

In the telephone health survey 2003 (GSTel03), too, the social class is covered through a multi-dimensional index, which had been developed for the federal health survey 1998 [3]. The dimensions of education, household net income and occupational status are the same as those of TNS-Infratest, although there are differences in details. According to the total score achieved, the participants are allocated to three (instead of five) social classes, that is lower class (3-8 points), middle class (9-14 points) and upper class (15-21 points).

Where the value is lacking for a specific dimension, the arithmetic mean of the values for the other two variables is used. Where more than one value is lacking, the recommendations on demographic standards are followed, that is the index value is not calculated and the person is excluded from analysis. This leads to sample losses of about 25%.

Another difference compared with the class concept of TNS-Infratest is the fact that the occupational status achieved by any individual is analysed rather than the occupational status of the main income earner as a measure for all household members. So, compared with the Infratest approach, it is the individual development rather than the family environment which is considered as decisive.

### **Comparing the stratification concepts and results**

In both concepts there is a predominance of the lower class(es), which is more distinct in GSTel03 than in TNS-Infratest (chart 1).

When showing the average body mass index (BMI) by classes, a continuous improvement of the situation from the lower class to the upper class is observed in both methods (chart 2). In GSTel03, however, the lower classes have slightly lower BMI values than in TNS-Infratest. This is due to the fact that persons for whom two or three values of the class index variables are lacking are excluded.

As regards the share of smokers, the gradient between classes is less distinct (chart 3). As a consequence, the results of the two methods differ even in terms of trend. While, in TNS-Infratest, an increase in the share of smokers is observed from the lower to the middle class, the opposite is true for GSTel03. So, in some cases, the method of class index construction can exert considerable influence on the results.

### **Sources**

1. Scheuch EK (1970) Sozialprestige und soziale Schichtung. In: Glass DV, König R (ed.) Soziale Schichtung und soziale Mobilität. Kölner Zeitschrift für Soziologie und Sozialpsychologie, special issue 5, Westdeutscher Verlag Opladen, pp. 65-103
2. Dulon M (2002) Ist der Mikrozensus geeignet, soziale Ungleichheit zu messen?, Eine Untersuchung zur Berechnung des Sozialschichtindex auf Grundlage der Mikrozensus-Daten. Universität Bielefeld, Fakultät für Gesundheitswissenschaften
3. Lampert T, Kroll LE (2006) Die Messung des sozioökonomischen Status in sozialepidemiologischen Studien. In: Richter M, Hurrelmann K (ed.) Gesundheitliche Ungleichheit - Theorien, Konzepte und Methoden, VS Verlag für Sozialwissenschaften Wiesbaden, pp. 297-319
4. Hoffmeyer-Zlotnik JHP (1998) Beruf und Stellung im Beruf als Indikatoren für soziale Schichtung. In: Ahrens W, Bellach BM, Jöckel KH (ed.) Messung soziodemographischer Merkmale in der Epidemiologie. RKI-Schriften 1/1998, MMV Medizin Verlag GmbH München, pp. 54-64

5. Grüttner U, Bloomfield K, Kramer S et al. (2006) Die Konstruktion eines empirisch bestimmten Sozialschichtindex mittels optimaler Skalierung am Beispiel von Deutschland. Das Gesundheitswesen vol. 68: pp. 116-122
6. Steinkamp G (1993) Soziale Ungleichheit, Erkrankungsrisiko und Lebenserwartung: Kritik der sozialepidemiologischen Ungleichheitsforschung. Soz-Präventivmed 38: pp. 111-122
7. Dulon M, Bardehle D, Blettner M (2003) Zur Messung der sozialen Ungleichheit im Mikrozensus und Bundes-Gesundheitssurvey, Assessing Social Inequality in Microcensus Data and German National Health Examination Survey. Georg Thieme Verlag, Stuttgart
8. Heller G, Müller U (1998) Die Standarddemographie nach ADM-ASI-Statistisches Bundesamt. In: Ahrens W, Bellach BM, Jöckel KH (ed.) Messung soziodemographischer Merkmale in der Epidemiologie, RKI-Schriften 1/1998, MMV Medizin Verlag GmbH München, pp. 47-53
9. Max Rubner-Institut, Bundesforschungsinstitut für Ernährung und Lebensmittel (2008) Nationale Verzehrsstudie II - Ergänzungsband zum Ergebnisbericht, Teil 1, Ausgewählte Ergebnisse nach Schichtindex. Karlsruhe

Chart 1: Distribution of the microcensus population aged 15 years and over among the social classes according to TNS-Infratest and GSTel03

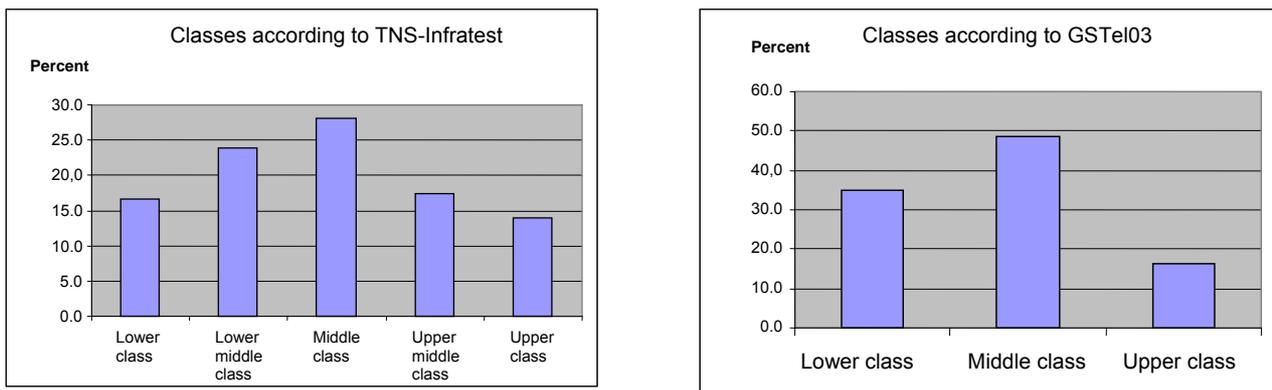


Chart 2: Average body mass index (kg/m<sup>2</sup>) of the microcensus population aged 18 years and over in the social classes according to TNS-Infratest and GSTel03

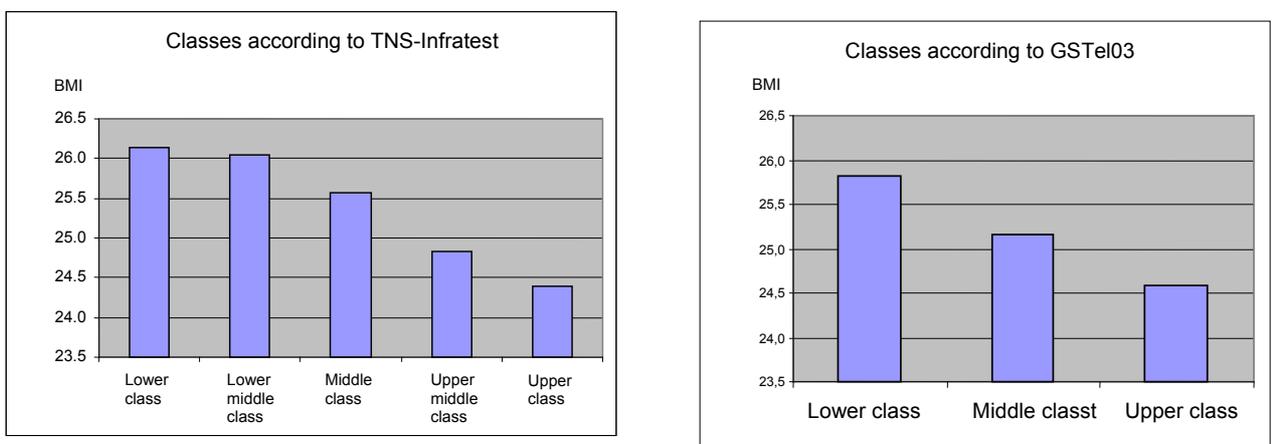
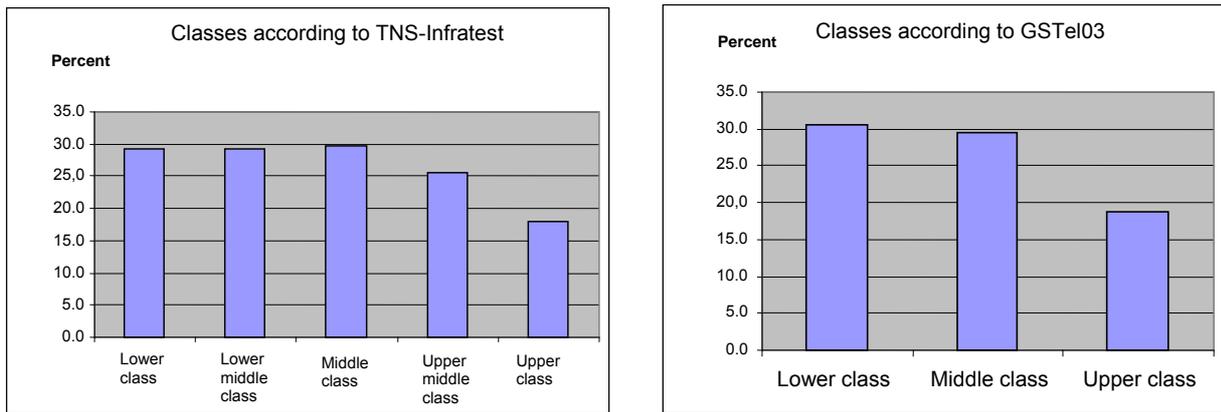


Chart 3: Share of smokers in the microcensus population aged 15 years and over in the social classes according to TNS-Infratest and GSTel03



Hans-Jürgen Heilmann, tel: +49 (0) 228 / 99 643 8145, e-mail: [hans-juergen.heilmann@destatis.de](mailto:hans-juergen.heilmann@destatis.de)

Silke Mardorf, tel: +49 (0) 511 / 168-4 69 66, e-mail: [Silke.Mardorf@hannover-stadt.de](mailto:Silke.Mardorf@hannover-stadt.de)

Winfried Eberhardt, tel: +49 (0) 228 / 99 643 8119, e-mail: [winfried.eberhardt@destatis.de](mailto:winfried.eberhardt@destatis.de)

## Developing an algorithm to calculate in air transport statistics the air kilometres travelled in the German air space

Apart from information on the number of passengers and freight or mail quantities transported, the distance over which persons or goods are transported is an important statistical figure in air transport statistics. Information on distances is needed, among other things, for performance indicators and intermodal comparisons. In this context, the air kilometres travelled in the German air space are relevant. Within the scope of the current reprogramming of air transport statistics, such kilometre data are recalculated.

### Bases

The length of the shortest air route between an airport of departure and an airport of destination is referred to as A kilometres. The part of the route which is over German territory is called B kilometres. While A kilometres can be calculated rather easily from the geographic co-ordinates of the airports of departure and destination, determining B kilometres is more complicated. To do so, the form of the earth is approximated by a sphere and, in addition, it is assumed that the routes of the airplanes run along the shortest line between the airports of departure and destination.

The shortest route between two points on a sphere is on a great circle and is referred to as orthodrome. Great circles divide a sphere into two equal halves, and their central point is identical to that of the sphere. The orthodrome between two points on a sphere is obtained by examining the plane containing the sphere's central point and the two given points. Its intersection with the sphere's surface is a circle located on the sphere's surface and containing the two given points. The shortest route between the two points on that circle is also the orthodrome on the sphere (see chart 1).

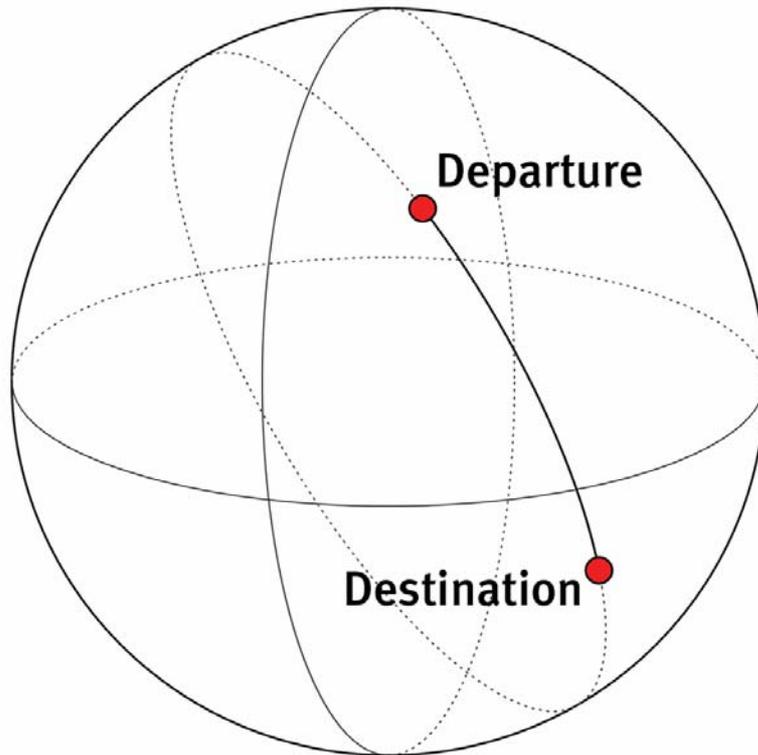


Chart 1: An orthodrome is the shortest route between two points on the surface of a sphere

### Calculating the A kilometres

Under the above conditions, the length of the air route (A kilometres) can be calculated by means of the co-ordinates of the airports of departure and destination.

According to Weisstein<sup>1</sup>, the following holds for the distance  $d$  between two points on the sphere's surface:

$$d = r \cdot \arccos(\cos \delta_A \cdot \cos \delta_B \cdot \cos(\lambda_A - \lambda_B) + \sin \delta_A \cdot \sin \delta_B) \quad (1)$$

$r$ : Radius of the earth; when using the GRS80 ellipsoid of the WGS84 reference system, there holds for the radius of a sphere of the same volume  $r = 6,371,000.785\text{m}$

$\delta_A, \delta_B$ : latitudes of the airport of departure and destination (in radian measure)

$\lambda_A, \lambda_B$ : longitudes of the airport of departure and destination (in radian measure)

### Determining the orthodrome to calculate the B kilometres

The B kilometres refer to the part of an air route that is over German territory. To calculate that distance, the intersection between the air route and the border of the German territory is to be determined. Applying the co-ordinates of the intersection and of the German airport to formula 1 then provides the B kilometres (cf. chart 2).

---

<sup>1</sup> Weisstein, Eric W. "Great Circle." From MathWorld--A Wolfram Web Resource. <http://mathworld.wolfram.com/GreatCircle.html>

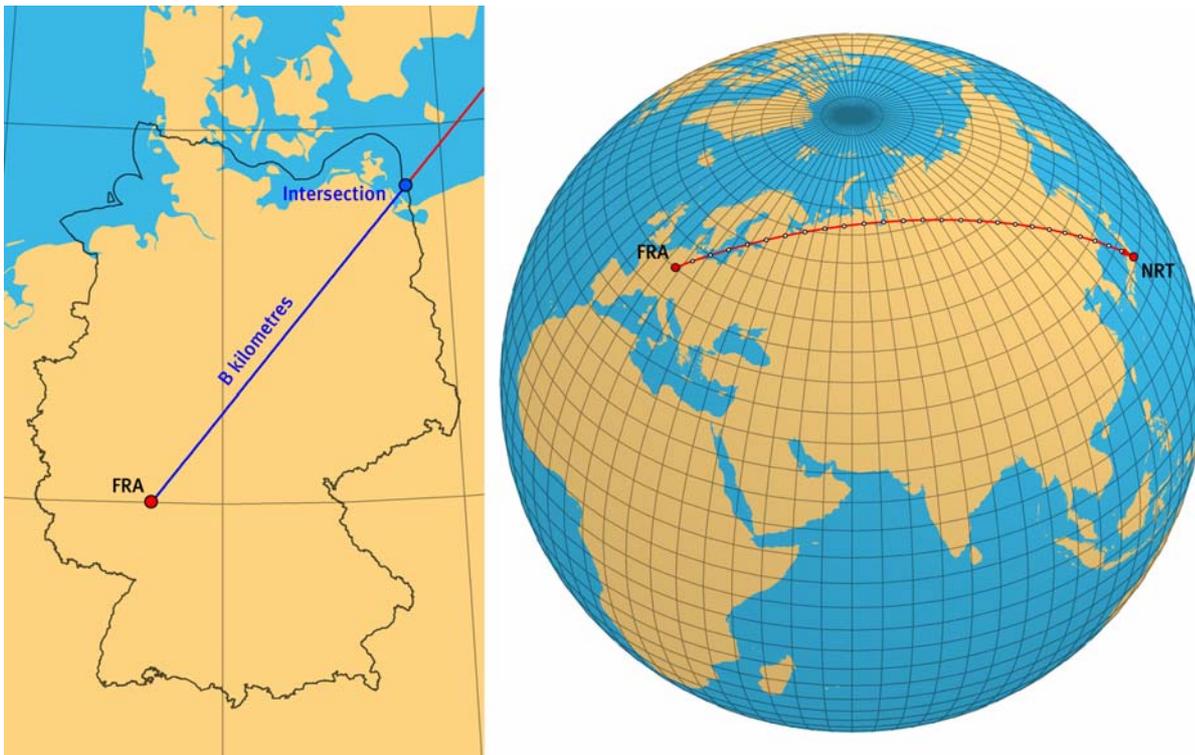


Chart 2: B kilometres can be calculated by means of the co-ordinates of the German airport and the intersection between air route and national border (blue point, left-hand picture). Before doing so, the geometry of the air route must be described through a line based on the co-ordinates of several points (right-hand picture, using the example of the orthodrome between Frankfurt airport, FRA, and Tokyo airport, NRT).

A variety of software tools are available to calculate the intersection<sup>1</sup> between the national border and the air route. Such intersecting of two geometric objects is a standard function of geo-information systems and it is also available as class libraries for various programming languages. This always requires that the geometries of the two objects to be intersected are available in the form of lines defined by co-ordinates. While the borders of the territory of the Federal Republic of Germany are defined by the official geometry of the Federal Agency for Cartography and Geodesy<sup>2</sup>, the geometry of an air route must be calculated for any combination of airports of departure and destination. For that purpose, the co-ordinates of a defined number of points on the great circle between the airports must be determined. By connecting neighbouring points, the geometry of the orthodrome is obtained as a line.

The co-ordinates of the points are calculated in the three-dimensional Cartesian co-ordinate system. So, the geographic co-ordinates of the relevant airports of departure and destination must be transformed into Cartesian co-ordinates. For the conversion of spherical co-ordinates ( $r$ ;  $\theta$ ;  $\varphi$ ) into Cartesian co-ordinates ( $x$ ;  $y$ ;  $z$ ), there holds:

$$x = r \cdot \sin \theta \cdot \cos \varphi \tag{2}$$

$$y = r \cdot \sin \theta \cdot \sin \varphi \tag{3}$$

$$z = r \cdot \cos \theta \tag{4}$$

1 In practice, several intersections may be obtained if the air route intersects the German national border more than once. For the sake of simplicity, this paper refers to one intersection only.

2 The product used is VG250 (administrative borders 1:250.000) of the Federal Agency for Cartography and Geodesy, status 31 December 2008, ellipsoid WGS84, date WGS84. The national territory comprises the administrative border of the Federal Republic of Germany plus the 12-mile sea zone and the German sub-area of the Lake of Constance.

There are:

r: Radius of the earth<sup>1</sup>

$\theta$ : Polar angle, measured from the pole (Z axis in the Cartesian co-ordinate system). For the latitude  $\beta$  there holds:  $\beta = 90^\circ - \theta$  or  $\theta = 90^\circ - \beta$

$\varphi$ : Angle of azimuth: Angle between the positive axis X and  $r_{xy}$ , counted anti-clockwise from 0 to  $2\pi$  ( $0^\circ$  to  $360^\circ$ ).

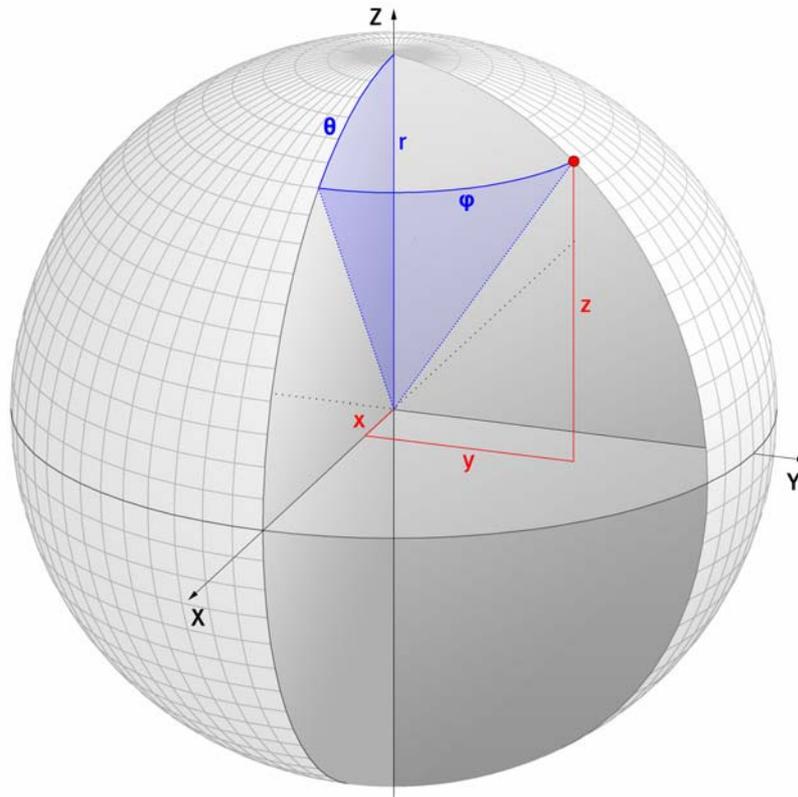


Chart 3: Transformation of the geographic or spherical co-ordinates (blue) into a three-dimensional Cartesian co-ordinate system (red)

Calculating the points is based on the construction of the orthodrome as outlined in the section on “bases” (see chart 4). The distribution of the points on the orthodrome is defined by a subdivision of angle  $\omega$  between the position vectors of the points of departure and destination (A and B). In the case examined, an equidistant distribution is created. For that purpose,  $\omega$  is subdivided into n equal angle increments (see chart 4). The intersections between the arms thus created and the sphere’s surface provide the orthodrome points (white point in chart 4).

First of all, the angles  $\omega$  and  $\alpha$  enclosed by the position vectors  $\vec{a}$  and  $\vec{b}$  as well as by  $\vec{a}$  and the departure-destination vector  $\vec{d}$  (required in formula (8)) are calculated.

<sup>1</sup> As the points’ geographic co-ordinates to be calculated are independent of the size of the sphere, it can be assumed without restrictions in the following that  $r = 1$  .

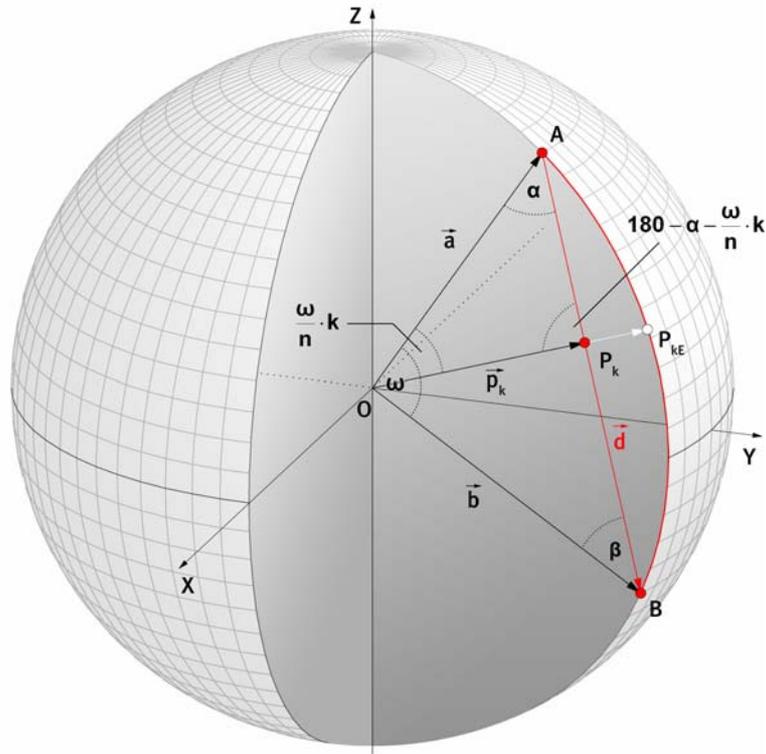


Chart 4: Calculation of evenly distributed points on an orthodrome

As the points A, B and O (origin in the co-ordinate system) form a plane triangle, the angles can be calculated using the formulas of trigonometry in the plane. Therefore, according to the cosine rule for general plane triangles, there holds:

$$\cos \omega = \frac{|\vec{a}|^2 + |\vec{b}|^2 - |\vec{d}|^2}{2 \cdot |\vec{a}| \cdot |\vec{b}|} = \frac{2 - |\vec{d}|^2}{2} = 1 - \frac{|\vec{d}|^2}{2} = 1 - \frac{d_1^2 + d_2^2 + d_3^2}{2} \quad (5)$$

For angle  $\alpha$  there also follows:

$$\cos \alpha = \frac{|\vec{d}|^2 + |\vec{a}|^2 - |\vec{b}|^2}{2 \cdot |\vec{d}| \cdot |\vec{a}|} = \frac{|\vec{d}|^2 + 1 - 1}{2 \cdot |\vec{d}|} = \frac{|\vec{d}|}{2} = \frac{\sqrt{d_1^2 + d_2^2 + d_3^2}}{2} \quad (6)$$

Then  $\omega$  is subdivided into  $k=0, 1, \dots, n$  equal angle increments. The intersection between any angle increment and  $\vec{d}$  provides a point  $P_k$  (see chart 4). By extending its position vector  $\vec{p}_k$  to the length of the spherical radius, the point  $P_{kE}$ , projected to the sphere's surface, is obtained.

First, for the linear equation of vector  $p_k$ , there holds:

$$\vec{p}_k = \vec{a} + \lambda_k \cdot \vec{d} \quad (7)$$

and there holds:  $|\vec{a}| = r = 1, 0 \leq \lambda_k \leq 1$ , especially  $\vec{p}_0 = \vec{a}$  and  $\vec{p}_1 = \vec{b}$ .

As there also follows from the sine rule for plane triangles:

$$\frac{\lambda_k \cdot |\vec{d}|}{|\vec{a}|} = \frac{\lambda_k \cdot |\vec{d}|}{1} = \frac{\sin\left(\frac{\omega}{n} \cdot k\right)}{\sin\left(180 - \alpha - \frac{\omega}{n} \cdot k\right)}, \quad (8)$$

$\lambda_k$  can be calculated and, consequently,  $P_k$  can be obtained.

By normalising vector  $\vec{p}_k$  to its length, point  $P_{kE}$  on the sphere's surface is obtained (see chart 4 and note r = 1 according to footnote 4):

$$\vec{P}_{kE} = \frac{\vec{P}_k}{|\vec{P}_k|}, \quad (9)$$

The three-dimensional Cartesian co-ordinates thus obtained for point  $P_{kE}$  must finally be transformed into spherical or geographic co-ordinates (see chart 3). For the angle of azimuth  $\varphi$  there holds:

$$\varphi_H = \arctan\left(\frac{y}{x}\right) \quad (10a)$$

and it should be noted that:

$$\varphi = \begin{cases} \varphi_H, & \text{for } x > 0 \\ \text{sign}(y) \frac{\pi}{2}, & \text{for } x = 0 \\ \varphi_H + \pi, & \text{for } x < 0 \text{ and } y \geq 0 \\ \varphi_H - \pi, & \text{for } x < 0 \text{ and } y < 0 \end{cases} \quad (10b)$$

The angle of azimuth  $\varphi$  (measured from 0 to  $2\pi$ ) corresponds to the longitude, and for  $\varphi > 180^\circ$  ( $\pi$ ) there holds:

$$\varphi_{\text{longitude}} = \varphi - 360^\circ \quad (11)$$

For the polar angle  $\theta$ , there holds:

$$\theta = \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \quad (12)$$

From that, there follows for latitude  $\beta$  (measured from the equator):

$$\beta = 90^\circ - \theta = 90^\circ - \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \quad (13)$$

The geographic co-ordinates thus obtained for the n points form the line of the orthodrome. That geometry can then be used to calculate the co-ordinates of the intersection with the national border (see chart 2). That intersection and the co-ordinates of the German airport then provide the B kilometres when applying formula (1).

Arndt von Eschwege, tel: +49 (0) 611 / 75 41 69, e-mail: [arndt.eschwege@destatis.de](mailto:arndt.eschwege@destatis.de)

Dr. Holger Heidrich-Riske, tel: +49 (0) 611 / 75 27 46, e-mail: [holger.heidrich-riske@destatis.de](mailto:holger.heidrich-riske@destatis.de)

## Study on the use of multiple imputation methods in wholesale trade statistics

For the purpose of data editing with regard to the annual survey in wholesale trade and commission trade<sup>1</sup>, a fully automated data editing method is currently being examined for feasibility.

### Fully automated data editing

Fully automated data editing comprises error identification and error correction. In error identification, the data fields of an incorrect data record are identified which cause the error in the data record. A lacking value is considered as an error, too. The values of those data fields are deleted during error correction of a fully automated data editing method. Then new values have to be estimated for those data fields. Estimations of lacking values are also referred to as imputations. This means that imputations must be generated for the data fields causing the error. It is currently being examined whether multiple imputation methods are suited as an error correction method.

### Multiple imputation

In multiple imputation, for every lacking value,  $m$  values are generated as estimates for the lacking value. Common values of  $m$  are between 3 and 10. This means that, after multiple imputation,  $m$  completed data sets are available. Those  $m$  completed data sets are then analysed separately, as usual. An example of an analysis is table expansion and the relevant error calculation. This produces  $m$  analysis results, which are combined through simple formulas, e.g. the arithmetic mean, to form a result.

Imputations are estimations. This means especially that imputing values involves an additional uncertainty, i.e. a larger standard deviation, of the estimator in question.

The advantage of multiple imputation over “simple” imputation is that, under certain conditions, multiple imputation produces an unbiased variance estimator. That variance estimator is easily calculated.

An example of a “simple” imputation method, which does not produce an unbiased variance estimator if it is applied  $m$  times, is a simple hot-deck method<sup>2</sup>.

### What is the reason for multiple imputation of wholesale trade statistics?

The data underlying wholesale and commission trade are obtained by means of a sample. So the expanded values of the published tables are estimates of totals of the population of wholesale trade and commission trade. For those estimators, error calculations are published, too.

So the issue to be examined with regard to using multiple imputation methods in wholesale trade statistics is whether the statistical results obtained after multiple imputation are “similar” to those obtained in a survey without errors (complete data set). Statistical results are similar if the relative deviation between the estimator after multiple imputation and the estimator of the data set without error is “small” and if the confidence interval of the estimator after multiple imputation includes the confidence interval of the estimator of the data set without error and is not “much” larger.

### Software

The first software tested to study the use of multiple imputation methods in wholesale trade statistics was the SAS macro IVEware<sup>3</sup>. The method implemented in IVEware is based on Bayes

---

1 The annual survey in wholesale trade and commission trade is part of the annual survey in domestic trade and the hotel and restaurant industry.

2 A method achieving that is called “proper”. See in “Multiple Imputation for Nonresponse in Surveys“ (1987), Donald B. Rubin, Wiley Series in Probability and Mathematical Statistics; in “Analysis of Incomplete Multivariate Data“ (1997), Schafer, J., Chapman & Hall; or in “Missing Data: Our View of the State of the Art”, Schafer, J., Graham, J. (2002), Psychological Methods, Vol. 7, No. 2, pp. 147-177.

3 The software may be downloaded from [www.isr.umich.edu/src/smp/ive](http://www.isr.umich.edu/src/smp/ive) .

regressions performed sequentially<sup>1</sup>. That method of creating the multiple imputations is flexible, i.e. variables with different types of distribution can be imputed and restrictions can be set for the imputations. That imputation method can produce good imputation results if the regression models describe the variables in a sufficiently good way<sup>2</sup>.

## Study and results

The test data used were correct data records that had entirely been observed. Values lacking in those data were generated at random. Then the lacking values were multiply imputed using IVEware. The parameters in IVEware were varied, so that several multiple imputation results were available. This allowed assessing the algorithm's behaviour. Assessing the quality of the results was mainly based on comparing the resulting expansions, the relevant variances and the error calculation with the original values.

After transformation, 19 of the total of 26 variables of wholesale trade statistics show a type of distribution that is supported by IVEware. The imputation results obtained are good for 8 of those 19 variables and satisfactory for 5 variables. The other 6 variables could not successfully be imputed and, therefore, were not used to create the imputations of the other variables.

First results of the study show that, actually, good imputation results can be obtained by means of IVEware. The less positive results for some variables might be improved by choosing a different transformation of variables and by modifying the regression models<sup>3</sup>.

Unfortunately, however, the regression parameters calculated by IVEware during the iterations are not available. That would be desirable because those parameters would allow documenting the algorithm's behaviour through the iterations and checking the quality of the resulting regression models. So it would be possible to assess more accurately the imputation results.

A program generating imputations through sequential Bayes regressions, as is done by IVEware, and providing the calculated parameters, which is not done by IVEware, is the R program "mi"<sup>4</sup>. This is why it is intended to use the R program "mi" for further studies.

*Katrin Schmidt, tel: +49 (0) 611 / 75-23 35, e-mail: [katrin.schmidt@destatis.de](mailto:katrin.schmidt@destatis.de)*

## Implementing the new FATS regulation without causing an additional burden

In July 2007, the Regulation on Community statistics on the structure and activity of foreign affiliates (FATS – Foreign Affiliates Statistics)<sup>5</sup> entered into force. The FATS regulation defines both the statistical units which form the basis for the statistics and the variables (incl. their breakdown and periodicity) to be collected by the EU Member States. In Germany, the Federal Statistical Office is responsible for compiling statistics on foreign-controlled enterprises in the domestic territory (Inward FATS), while the Deutsche Bundesbank (central bank of the Federal Republic of Germany) produces statistics on affiliates abroad (OutwardFATS).

---

<sup>1</sup> See "A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regressions" (2001) by Trivellore E. Raghunathan, James M. Lepkowski, John Van Hoewyk and Peter Solenberger.

<sup>2</sup> See e.g. "Far from Normal – Multiple Imputation of Missing Values in a German Establishment Survey" (2009) by Jörg Drechsler.

<sup>3</sup> For the first studies, the covariables were integrated into the model in a linear way only. For example, no interactions were modeled.

<sup>4</sup> When the studies started, the R program "mi" did not exist yet.

<sup>5</sup> Regulation (EC) No 716/2007 of the European Parliament and of the Council of 20 June 2007 on Community statistics on the structure and activity of foreign affiliates (OJ No L 171, p. 17), last amended by Commission Regulation (EC) No 747/2008 of 30 July 2008 (OJ No L 202, p. 20).

As regards InwardFATS, results are compiled for enterprises of business economy. Data are collected for important variables of structural statistics such as turnover, gross value added at factor cost and number of persons employed in a geographical breakdown by origin of the foreign ultimate controlling Institutional unit (UCI).<sup>1</sup> Regarding Sections C, D and F of NACE Rev. 1, results are additionally produced for the following variables: total intra-mural R & D expenditure and total number of R & D personnel.

The FATS regulation provides for both feasibility studies with a pilot character and regular annual data supplies from reference year 2007. The Federal Statistical Office took part in a feasibility study on the breakdown of InwardFATS enterprises by size classes for reference year 2006. In the context of the feasibility study, valuable methodological insights were gained for future data processing. In addition, first results could be compiled. The Federal Statistical Office will provide the first regular results for reference year 2007 in spring 2010. In methodological terms, however, data processing for the purpose of this supply of regular results will differ from processing in the context of the feasibility study<sup>2</sup>.

The InwardFATS data are processed centrally at the Federal Statistical Office. Although detailed results have to be compiled, an additional burden on enterprises due to the FATS regulation could have been avoided. The reason is that the business register for statistical purposes, a commercial database and results of the annual structural surveys have been used as data sources to produce InwardFATS. The business register provides the basis for covering all economically active enterprises in the sectors of economic activity that are of relevance to InwardFATS. However, it reflects neither affiliations with groups of enterprises nor control relations between enterprises. This type of information cannot be obtained from the administrative sources that are used to input data into the business register<sup>3</sup>. For this reason, the data needed are acquired from a commercial data provider. The information on control relations is maintained in an enterprise group database which is linked to the business register. As is illustrated in Figure 1, InwardFATS now link the information on participation in enterprises that is available in the business register with different individual data that were collected for the purposes of structural business statistics. This approach can be taken because the provisions of the FATS regulation are largely in line with those of the SBS regulation<sup>4</sup>. Thus additional surveys which would place a burden on businesses do not have to be carried out. All in all, data of ten different primary surveys are used which reflect the activities of non-financial business economy in Germany: the annual structural surveys of enterprises. In part, however, these surveys differ considerably as regards their design and ambition. Hence the quality of InwardFATS results varies depending on the sectors of economic activity, size classes and variables.

---

<sup>1</sup> Further variables include the following: number of enterprises, production value, total purchases of goods and services, purchases of goods and services purchased for resale in the same condition as received, personnel costs and gross investment in tangible goods.

<sup>2</sup> See Feuerhake, J./Untz, K.: Inward-FATS – Auslandskontrollierte Unternehmen in Deutschland 2006, WiSta 7/2009, pp. 676ff., for explanatory notes on this methodology.

<sup>3</sup> See Sturm, R./Tümmler, T./Opfermann, R.: Unternehmensverflechtungen im statistischen Unternehmensregister, WiSta 8/2009, pp. 764ff.

<sup>4</sup> Regulation (EC) No 295/2008 of the European Parliament and of the Council of 11 March 2008 concerning structural business statistics (OJ No L 97, p. 13).

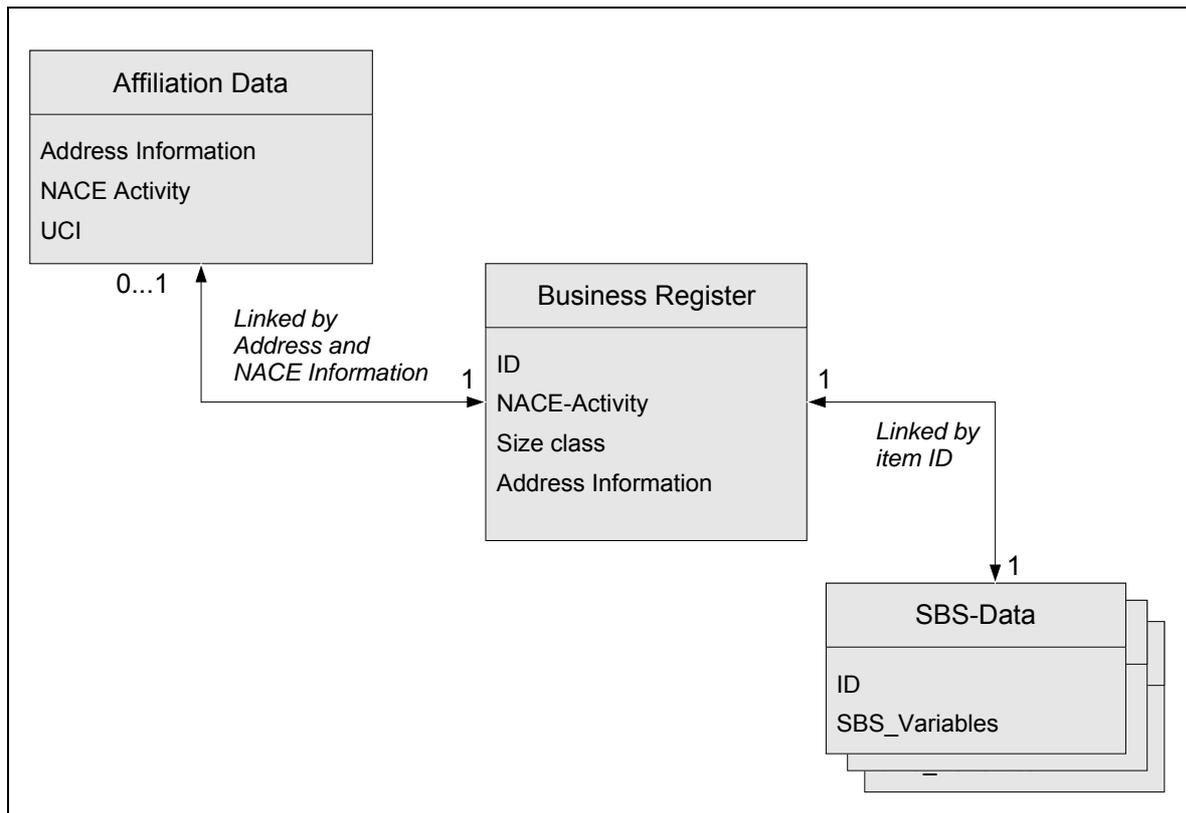


Figure 2: Relations between InwardFATS data sources

For the purposes of the current processing operations, all InwardFATS enterprises that can be identified in the business register and the enterprise group database are included in the population. The information which becomes available for an identified InwardFATS enterprise from primary surveys or administrative sources for the relevant reference year is adequately assigned to it. As regards the remaining units, the data for the relevant variables are estimated based on information on both size class and economic activity. By means of this imputation, the information to be supplied for the variables concerned is generated for all enterprises marked as InwardFATS units. The adjusted methodology has the advantage that all foreign-controlled enterprises, including those with an ‘exotic’ origin of controlling unit, can actually be covered although they are not included in the structural sample surveys. However, this method cannot be used for providing data on enterprises that are controlled from Germany. Information on those enterprises is obtained by taking the difference between the relevant results of the SBS surveys and the InwardFATS results. Actually, the results for NACE J are an exception as they are produced on the basis of another methodology. Currently, the Federal Statistical Office does not have any information of its own on financial services. The main provider of these data is the Deutsche Bundesbank which supplies results for this sector in a tabulated form with an integrated primary confidentiality scheme.

As a matter of fact, the discussion about improvements and adjustments of the methodological approach will continue in the future. And the harmonisation efforts undertaken at the European level will also improve the quality of InwardFATS in the medium term. In line with the FATS regulation, quality reports that are harmonised across Europe are submitted for the production of various results. Another issue of high priority is the present development of a European register of multinational enterprise groups (EuroGroups Register). This platform provides a Europe-wide, co-ordinated and up-to-date data basis which integrates data coming from commercial sources and from the national business registers. In the medium term, this data basis can be used to produce information on enterprise interrelationships. In this context, too, the objective will be to compile results without placing an additional burden on businesses.

More detailed information and publications on foreign-controlled enterprises can be accessed at:

<http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Navigation/Statistiken/UnternehmenGewerbelsolvenzen/Auslandsunternehmen/Auslandsunternehmen.psm1>

Jörg Feuerhake, tel: +49 (0) 611 / 75 41 16, e-mail: [joerg.feuerhake@destatis.de](mailto:joerg.feuerhake@destatis.de)

Kirsten Untz, tel: +49 (0) 611 / 75 25 70, e-mail: [kirsten.untz@destatis.de](mailto:kirsten.untz@destatis.de)

## Using hot-deck procedures in foreign trade statistics

To increase the efficiency of data processing in foreign trade statistics, a selective approach to processing data declarations was introduced in the context of implementing the ASA system (System of automated processing in foreign trade statistics) in early 2009<sup>1</sup>. Based on specific thresholds, the declarations which have a larger impact on the results are identified for each commodity code and, if necessary, have to be corrected manually<sup>2</sup>. As a rule, however, declarations that are of rather minor importance for the overall result are corrected automatically.

As a result of using the new ASA processing system, the requirements of declaration-related automatic error correction have increased. Due to prioritising declarations on the basis of specific thresholds for each commodity code, a relatively high value (compared to the values declared for other commodity codes) may be of rather low importance for the overall result of the commodity code concerned and will consequently be corrected automatically. Besides, the automatic error correction procedure is only run towards the end of data processing for a given month to give the data managers the opportunity to deal with some few data records of lesser importance, too. Hence, there is little time after the end of automatic error correction to discover and correct possibly remaining errors before detailed results will be published. For this reason, the automatic error correction processes of the new ASA system must ensure that both missing or incorrect values will be replaced by adequate ones and the data declared for a given month will indeed be largely plausible after automatic error correction.

Before the introduction of the ASA system, deterministic error correction rules were exclusively used to automatically correct categorical variables. Actually, that correction method can only be applied properly to inaccurate data records in which the value of a given variable can unmistakably be defined by the values of other variables<sup>3</sup>. However, such clear relations between variables are rare in foreign trade statistics. Also, it is often not clear which of the variables has caused the error in an inaccurate combination of variable values. In particular, the values of certain categorical variables such as “country of origin/country of destination” can often not be determined without difficulty by means of other variables. Considering the variability of the values is also difficult. As a matter of fact, defining the rules for correcting errors is a complex task, too. It requires to take into account that a change in the value of a variable may cause new errors due to an improper combination with the value of another variable. If the error correction rules are not consistent, there will be the danger both of a considerably prolonged automatic correction process and of obtaining implausible data.

Since the ASA system was introduced, a nearest neighbour hot deck method has therefore been used in the automatic correction of categorical variables. Deterministic error correction rules are

---

1 See Blang, D.: Neuausrichtung der Aufbereitung der Außenhandelsstatistik, *Wirtschaft und Statistik* 12/2006, pp. 1247 ff.

2 See Helmert, T.: Implementing selective editing and imputation methods in foreign trade statistics, Working Paper 43, UN/ECE Work Session on Statistical Data Editing, Neuchâtel (Schweiz), 2009, (<http://www.unece.org/stats/documents/2009.10.sde.htm>)

3 See Giles, P./Patrick, C.: Imputation Options in a Generalized Edit and Imputation System, *Survey Methodology*, Vol. 12, No. 1, 1986, pp. 49 ff.

only applied if the value an incorrect variable should have is definitely clear and if the relevant correction does not cause new errors. In the context of a hot deck approach, an incorrect data record (recipient record) is corrected by imputing the variable values from a correct data record (donor record). As regards foreign trade statistics, the donor record to be used for a recipient record is identified using the nearest neighbour approach. This means that the data record which is “most similar” to the recipient record will be selected as the donor record. With few exceptions, selecting potential donors is restricted to correct data records for the same commodity code. The following distance function is used to determine the degree of similarity between a recipient record  $X$  and a potential donor record  $Y$ :

$$D_{XY} = D(X, Y) = \sum_{k=1}^r w_k |x_k - y_k|$$

The value of the distance function is the sum of the weighted deviations of the variable values ( $x_k$ ) of the recipient record from the variable values ( $y_k$ ) of the potential donor record. As for categorical variables, the distance value is 0 for matching variable values ( $x_k = y_k$ ), while it is 1 for differing variable values ( $x_k \neq y_k$ ). Regarding numerical variables, the deviation is made comparable by means of the mean values in the relevant commodity code. Weighting parameter  $w_k$  is used to determine the relevance of the variables for the distance function. This also allows to control, to a limited extent, the error correction procedure based on the hot deck approach. In the course of automatic error correction, changes to a valid value, for instance, of the important categorical variable “country of origin/country of destination” should be avoided in foreign trade statistics. For this reason, weighting parameter  $w_k$  has a relatively high value for this variable so that a potential donor record whose variable value coincides with that of the recipient record often has the lowest distance value, too, and thus will be identified as the donor record.

The new design of automatic error correction ensures that very few data records will remain implausible after that correction and the quality of the imputed values of categorical variables will be convincing. However, applying the nearest neighbour hot deck procedure has also increased the importance of the other processing steps. In particular, error detection and manual data processing have to be of high quality so that a sufficient number of good quality donor records will be in place as a precondition for proper error correction based on the hot deck procedure. Currently, the effects are being studied which automatic error correction based on the nearest neighbour hot deck procedure has on the results obtained. A goal of this analysis is to further enhance the above method.

Thomas Helmert, tel.: +49 (0) 611 / 75 33 49, e-mail: [thomas.helmert@destatis.de](mailto:thomas.helmert@destatis.de)

## Revision of the UN methodological manual for foreign trade statistics

Revising the UN recommendations for foreign trade statistics (International Merchandise Trade Statistics, Concepts and Definitions) became necessary as the latest 1998 version was no longer in line with the new requirements caused by globalisation effects and changed user needs. The revision of the recommendations has been based on a process of worldwide consultation with the national institutions responsible for compiling foreign trade statistics. In addition to the Concepts and Definitions, the Compiler Manual is to be revised, too.

Based on a proposal of the UN Statistical Division (UNSD), the UN Statistical Commission commissioned the above revision in February 2008. A draft version is to be submitted to the UN Statistical Commission by 2010.

### Approach/procedure

Assisted by an Expert Group composed of members from developed and developing countries and also international organisations, UNSD first prepared a provisional draft of the revised Concepts and Definitions. After an opening meeting at the UN, the Expert Group, of which the Federal Statistical Office has been a member, organised various virtual meetings to continue its work online. Subject-related work was carried out in two stages. First the central issues of foreign trade statistics were discussed, then recommendations were worked out in relation to all individual topics. The Expert Group prepared a provisional draft regarding both components of work. A questionnaire was used to discuss the content of the provisional draft in the context of several rounds of worldwide consultation. In this way, each country had the opportunity to agree to or reject the proposals or to submit suggestions for modification. UNSD co-ordinated, considered and expressed in a text form that would be acceptable to a majority the proposals made by the members of the Expert Group and, subsequently, by the statistical agencies of the countries not represented in that group. In October 2009, UNSD presented the results of the second worldwide round of consultation. All proposals were adopted overwhelmingly. However, the recommendations regarding the partner countries for which data are to be collected were only passed by a three-quarter majority vote<sup>1</sup>. As regards the next steps, the draft version will be submitted to the Statistical Commission and the Concepts and Definitions will be explained in more detail by revising the manual.

### **Need for adjustment**

A revision of the Concepts and Definitions became necessary in view of the changes observed in foreign trade, such as an increasing globalisation of the production and distribution processes, expansion of intra-firm trade, and transactions with bundled goods and services components.

Besides, the methodological framework and legal environment for collecting foreign trade data have changed in the process of harmonising and simplifying customs procedures and adjusting them to increased security requirements. Also, countries which are members of a customs union have increasingly used data sources other than customs declarations. And user requirements have changed, too. The System of National Accounts and the Balance of Payment Manual 6 have undergone changes in the same way as other user needs. The latter include a rising demand for more detailed and promptly available data. Furthermore, it became necessary to add some clarification to existing concepts and to improve the readability of the Concepts and Definitions.

### **Major changes**

The cross-border movement as the decisive criterion in foreign trade statistics is now defined more clearly and distinguished more precisely from the change of ownership principle. The latter, which is the main criterion of both the balance of payments and national accounts, should only be used if the cross-border movement principle cannot be applied (e.g. acquisition/sales of vessels and aircraft).

Explanations are now much more detailed for the recording of transactions involving processing under contract, intra-firm trade, incl. territorial enclaves and exclaves, satellites and their launchers, electronic commerce, trade in used goods, and trade in electricity, crude oil and natural gas. However, other movements of goods not to be recorded, such as transit trade, are also defined more clearly.

New chapters include those on the mode of transport, data compilation strategies (using customs and non-customs data sources), data quality and metadata (measuring and assessing data quality and metadata requirements) and additional topics such as foreign trade indices and seasonal adjustment.

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

---

<sup>1</sup> Critics consider the recommendation to indicate the country of consignment in addition to the country of origin/destination for imports and exports a problem.

## **Simplification of intra-Community trade statistics: current situation and trends**

When the European single market was created in 1993, secondary surveys of intra-Community trade operations via the customs authorities were replaced by the Intrastat system. Since that time, businesses have been required to submit information on their monthly flows of goods into and out of other EU Member States directly to the Federal Statistical Office.

Last year, about 244,000 businesses liable to pay taxes exported goods from Germany to other countries in the EU territory, while some 476,000 businesses imported goods from other EU countries to Germany.

Not all of those businesses were responsible for providing information. As a result of the political initiatives aimed at reducing bureaucracy at both the national<sup>1</sup> and the European<sup>2</sup> level, measures were taken in the past to lower the burden on respondents. Furthermore, those initiatives have promoted the discussion about a basic simplification of intra-Community trade statistics.

### **Simplification measures enshrined in law**

Several EU legal acts regarding foreign trade statistics provide the legal framework for simplification measures. The Member States are required to satisfy the demand for statistical information without placing an excessive burden on respondents.

Simplifications permitted are largely laid down in Article 10 of the Intrastat basic regulation (Regulation (EC) No 638/2004)<sup>3</sup> and in Articles 8 and 13 of the Intrastat implementing regulation (Regulation (EC) No 1982/2004)<sup>4</sup>.

Businesses liable to pay taxes which are engaged in dispatches or arrivals of goods within the European single market profit from simplifications if the value of their transactions in the previous and the current year does not exceed the specified threshold. Actually, thresholds can be defined separately for the two directions of trade and be updated at annual intervals.

As laid down in Article 10, Para. 3 of the current Intrastat basic regulation, the Member States may set the thresholds at a level that ensures that the value of at least 97% of the total dispatches and at least 95% of the total arrivals is covered. If the value of its intra-Community goods transactions of the previous/current year is below the threshold, a business is fully exempted from providing an Intrastat declaration for the relevant direction of trade.

In addition, the regulations provide for simplification measures that can reduce the scope of statistical information to be submitted.

Pursuant to Article 10 Para. 4 a of the Intrastat basic regulation in conjunction with Article 9 of the Intrastat implementing regulation, net mass does not have to be indicated for commodity codes requiring the indication of a supplementary unit.

---

1 As a result of the discussion about the burden caused by official statistics, the previous government formulated bureaucracy reduction as a political goal (2005 Coalition agreement between the SPD, CDU and CSU "Gemeinsam für Deutschland. Mit Mut und Menschlichkeit", pp. 74ff.).

2 Based on the SLIM initiative (Simpler Legislation for the Single Market) of the European Commission, efforts have been made at the European level to simplify the Intrastat data collection system since May 1996.

3 Regulation (EC) No 638/2004 of the European Parliament and of the Council of 31 March 2004 on Community statistics relating to the trading of goods between Member States (OJ No L 102 of 7 April 2004, p. 1) amended by Regulation (EC) No 222/2009 of the European Parliament and of the Council of 11 March 2009 (OJ No L 87 of 31 March 2009, p. 160).

4 Commission Regulation (EC) No 1982/2004 of 18 November 2004 implementing Regulation (EC) No 638/2004 of the European Parliament and of the Council on Community statistics relating to the trading of goods between Member States amended by Commission Regulation (EC) No 1915/2005 of 24 November 2005 (OJ No L 307, p. 8).

Furthermore, Member States may define other thresholds below which the benefiting parties may be exempted from providing information about the nature of the transaction (Art. 10 Para. 4 b of the Intrastat basic regulation) or be given the opportunity to report a maximum of 10 of the commodity codes that are the most used codes in terms of value (Art. 10 Para. 4 c of the Intrastat basic regulation).

Pursuant to Article 10 Para. 5 of the Intrastat basic regulation in conjunction with Article 13 Para. 3 and 4 of the Intrastat implementing regulation, the parties responsible for providing information may report simplified information for 'small individual transactions' whose value is less than EUR 200 under Commodity code 9950 00 00.

In accordance with Article 8 of the Intrastat implementing regulation, a calculation of the statistical value<sup>1</sup> is not required for purchases or sales and for supplies for consignment or with the intermediation of a commission agent, if the relevant thresholds are not exceeded.

### **Simplification measures in place**

Currently, the declaration threshold is EUR 400,000 for each direction of trade in Germany. As a result, some 82% of all businesses dispatching goods to other EU Member States and about 91% of all businesses receiving goods from other EU Member States are exempted from the obligation to provide monthly information about their intra-Community trade operations. All in all, about 66,000 businesses are responsible for providing information. In terms of the commercial values declared by those parties required to give information, 98.2% of all dispatches and 97.0% of all arrivals of goods are covered.

Additional simplifications refer to the indication of both net mass and statistical value. Net mass does not have to be identified for some 750 commodity codes requiring the indication of a supplementary unit. Also, declaring the statistical value for the direction of trade concerned is not required for purchases or sales and for supplies for consignment or with the intermediation of a commission agent, if the value of the goods traded amounted to a maximum of EUR 42 million (dispatches) or EUR 30 million (arrivals) in the previous year. Thus not more than about 1,900 businesses with intra-Community dispatches and some 1,700 businesses with intra-Community arrivals are obliged to declare the statistical value.

Further legal simplification measures such as reporting the 10 commodity codes that are the most used codes or combining small individual transactions have not been implemented. The reason is that the reduction in the burden on the parties responsible for providing information can be expected to be rather small and thus would not justify the quality losses which would result from an implementation of those measures.

Technical simplifications are not laid down by law. However, they have developed as a result of automating the accountancy systems. At the beginning of collecting data in the context of the Intrastat system, the monthly goods transactions were declared in writing using special forms. Today, only 4% of all declarations are submitted in paper form. About 91% of all business declarations are transmitted online and 5% on magnetic data media.

### **Trends in the development of simplification measures**

In addition to the simplification measures described above, the total of eight-digit commodity codes has been considerably reduced since the introduction of the Intrastat system in 1993. Furthermore, repairs and maintenance were excluded from the list of movements of goods to be declared. From the perspective of the businesses concerned, however, intra-Community trade statistics continue to be one of the most time-consuming surveys. For this reason, much more substantial measures are under discussion with the aim to clearly reduce the relevant burden.

---

<sup>1</sup> Commodity value at the German border.

Due to their great potential for lowering the burden, the following two options are in the focus of the current discussion: a further increase in the declaration threshold and the application of what is called a 'single flow system' whereby data would be collected for one direction of trade only (namely dispatches). Information on arrivals would be derived from the dispatch data of the relevant EU partner country. While the declaration threshold could be raised at relatively short notice, introducing the single flow system would certainly be a medium to long-term measure which would require a great deal of preliminary work.

By introducing different minimum coverage rates for the two directions of trade, a step has been taken towards applying the single flow system in the current Intrastat basic regulation. In the future, more of the businesses receiving goods from other countries in the EU territory will be exempted from the obligation to submit an Intrastat declaration due to the reduction in the coverage rate. Nonetheless, the higher coverage rate of 97% for dispatches will be kept for quality reasons.

As a result of reducing the minimum coverage rate to 95% for arrivals, the frame specified by the current declaration threshold would no longer be fully required. *Theoretically*, the uniform declaration threshold for both directions of trade could be raised to a maximum of EUR 800,000 in Germany by amending the relevant stipulations in the national Foreign trade statistics implementing regulation (AHStatDV). This would mean that the number of parties responsible for providing information would decline to about 46,000 or 8% of all businesses that are liable to pay taxes and are engaged in intra-Community trade operations.

As a result of introducing the single flow method, some 44,000 businesses which receive goods from other EU countries and are therefore obliged to provide information on arrivals of goods would be exempted from the obligation to submit information. A big advantage would be the elimination of discrepancies between the data of the EU countries due to the application of this system. However, the single flow method carries some risks regarding data quality and timeliness which should not be underestimated. As the data on arrivals depend on the dispatch data of the partner countries, the chances for checking and correcting them would be limited. And the date of publishing the relevant data would depend on the 'slowest' country. In addition, data would no longer be collected for certain arrival-related variables defined at the national level if the single flow method were implemented under the conditions prevailing.

Eurostat has approached the tasks which have to be fulfilled before the single flow method can be implemented. Presently, a comprehensive quality concept is being prepared which specifies various quality goals and describes appropriate methods and procedures to be implemented in the Member States. The results of the different countries will be reviewed in the context of regular assessment processes. The concept as a whole is aimed at harmonising different approaches and reducing discrepancies in the results obtained. It is only after the successful adjustment of the individual systems that the application of the single flow method can be supported from a data quality perspective, too.

Susanne Meyer, tel: +49 (0) 611 / 75 20 72, e-mail: [susanne.meyer@destatis.de](mailto:susanne.meyer@destatis.de)

## Quality assurance in public finance statistics

### 1 The function of public finance statistics

Within the federal structure of the Federal Republic of Germany, the function of public finance statistics is to provide and ensure a comprehensive and accurate picture of the state's financial situation<sup>1</sup>. Data from the largely independent levels of public administration are combined in public

<sup>1</sup> Explanatory memorandum for the first Finance Statistics Law of 1960, German Bundestag, Bundestag printed paper 1367, Annex 1, p. 2.

finance statistics to form a consistent overall picture of public finance. The aim is to facilitate a precise assessment of whether the state's overall finances are stable and to determine the purposes for which financial resources are used by the state. Public finance statistics are an important data basis for political decisions at federal and Land level, for the presentation of the general government sector in national accounts and, hence, for the international comparability of public finance within the scope of the European Stability Pact.

## 2 The quality currently achieved in the accomplishment of tasks

The data for public finance statistics are directly collected from the public budget and accounting systems and therefore provide a very detailed picture of public finance<sup>1</sup>. The ways and means of how the functions of public finance statistics are fulfilled essentially depend on the budgets of the central, regional and local authorities which are at the centre of these statistics. The statistical reporting on finance focuses on public expenditure (on a cash basis) and on the revenue required to meet that expenditure. Well co-ordinated classifications for the federal, Land and municipal budgets have so far ensured transparency and a high quality of the data. The cameralistic budget management of the central, regional and local authorities does not include a balance sheet of financial assets and liabilities. Public finance statistics, however, record not only revenue and expenditure but also public debts, so that estimates can be made of future budgetary pressures resulting from interest and repayments. In 2005, an EU regulation came into force which made it mandatory to compile statistics on public financial assets (currency, accounts receivable, shares and other equity). However, these data are supplied by the public administrative authorities and agencies from diverse and hardly standardised sources, which means that they are less robust than the data on revenue and expenditure.

## 3 Difficulties arising from the growing flexibilisation of public budget management

Since the end of the last century, the comparability of public budgets has been increasingly affected by various actions to give more flexibility to budget management. This includes the separation of specific accounts from the core budgets, the introduction of global budgets, and deviations from the agreed framework of classifications. Adjustments to the group of respondents have made it possible to statistically reintegrate separate units and global budgets with the core budgets of the central, regional and local authorities in an overall picture of public finance<sup>2</sup>. Following severe criticism by the Federal Constitutional Court<sup>3</sup> of the comparability of public budgets, the Ministers of Finance of the Federation and the Länder agreed to adhere more closely to the common framework of budget classifications.

## 4 Reform of public budgets – Opportunities and risks concerning statistical quality

The transition to new double-entry accounting systems has enabled public finance statistics to represent public financial management in a more comprehensive way, which includes a balance sheet of all the state's assets and liabilities and gives a broader view of the generation and use of resources. However, the manner in which the transition to the new systems proceeds is a serious detriment to the quality of public finance statistics in Germany. The Interior Ministries of the Länder, which are responsible for the municipal budgets, and the Finance Ministries of the Federation and the Länder have chosen different paths of reform, which has a negative impact on public finance reporting. The municipalities are successively changing over to double-entry accounting, but at varying paces from Land to Land. The new Budgetary Principles Act (HGrG) gives complete leeway in how new accounting systems are applied to the federal and Länder budgets. Public finance statistics will therefore be affected by highly different public budget and accounting systems for a long time.

---

1 See Rehm, H.: Statistiken der öffentlichen Finanzen - aussagekräftiger und aktueller, *Wirtschaft und Statistik* 3/2006, pp. 282 ff.

2 Ibidem, p.281

3 Federal Constitutional Court judgment concerning Berlin of 19 October 2006 (2 BvF 3/03).

Under the amendments made to the Budgetary Principles Act in July 2009, the **Federation and the Länder** are entitled to introduce double-entry accounting as their sole budget and accounting system. The Länder of Hessen, Nordrhein-Westfalen, Bremen and Hamburg have opted for double-entry accounting. The majority of the Länder and the Federation, however, continue using cameralistic accounting (with the Federation and some of the Länder using a cameralistic accounting system that has been extended to include a balance sheet of assets and liabilities). Thus, cameralistic accounting still dominates public finance reporting. To ensure comparability across the Länder and the Federation and to enable an overall picture of the entirety of public finance to be gathered, the HGrG has made it mandatory for the Länder using double-entry accounting to provide consistent financial statistical reports on their revenue and expenditure on the basis of the cameralistic classifications by object and function. The HGrG expressly provides that “the requirements of public finance statistics and of national accounting shall be taken into consideration<sup>1</sup>” when deciding on the classifications. While this is a formal guarantee for consistent financial statistical reports by the Federation and all the Länder, the separation of the budgetary records (double-entry accounting) from the statistical records (revenue and expenditure) leads to a significant loss of transparency. What is most regrettable from a statistical point of view is that the classifications for double-entry accounting by the Länder have not been aligned to the municipal system of double-entry accounting. In contrast to municipal double-entry accounting, the (administrative) framework of accounts does not include an integrated financial account which would have ensured the transparency and traceability of the statistical records on revenue and expenditure.

In 2003, the Interior Ministers of the Länder jointly agreed that their **municipalities** should change over to a new double-entry accounting system, which offers the advantage of combining the use-of-resources approach with the liquidity approach. In addition to the two main components of commercial accounting (income account, asset and liability account), the municipalities have to compile an – integrated – financial account to provide an accurate record of the generation and use of liquidity (inpayments and outpayments). As a result, the municipal accounting system with its three components is more informative than traditional commercial bookkeeping. Also, the integrated financial account provides an interface with the former cameralistic system and thus helps to ensure consistent statistical reporting during the transition phase. However, serious difficulties arise from the fact that the Interior Ministers of the Länder do no longer take into consideration the requirements of public finance statistics and national accounts in implementing the reform of municipal budget management. While the present budget classifications of revenue and expenditure (structural classification, classification by object) are well adapted to statistical requirements and therefore guarantee a rather good quality, the Interior Ministries of the Länder have issued entirely different classification guidelines to the municipalities for double-entry budgeting and accounting. For that reason, the municipalities’ financial reporting is not only affected by transition difficulties but also by permanent system-induced disturbances of comparability.

To regain robust statistical data on municipal finance it is essential that – as for the Federation and the Länder – uniform standards of budget management and, above all, uniform classifications should be established for the municipal budgets throughout Germany, which make allowance for the requirements of public finance statistics and national accounting. The statistical offices of the Federation and the Länder have drawn up a step-by-step plan which offers important starting points. This step-by-step process should be started immediately together with the Interior Ministries of the Länder in order to recreate appropriate framework conditions for reliable statistical reporting on municipal finance in the medium term and to meet the request by the Federal Constitutional Court for comparable budgetary principles.

*Hans Rehm, tel: +49 (0) 611 / 75 41 74, e-mail: [hans.rehm@destatis.de](mailto:hans.rehm@destatis.de)*

---

1 Act on Modernisation of Budgetary Principles (HGrGMoG), Article 1 no. 22, Federal Law Gazette 2009 Part I No. 51, pp. 2580 ff.

## **Agreed earnings online: The agreed earnings database is a new basis for analyses and activities in the statistics of agreed earnings**

The agreed earnings database, which was installed in November 2009, is going to bring about far-reaching changes for the specialised unit dealing with the statistics of agreed earnings. The application software specially developed for Destatis is intended to improve both the internal work flow and the possibilities of external dissemination of information on agreed earnings. To be able to assess the innovations and opportunities involved in the introduction of the agreed earnings database, it must be examined how the work has been done so far.

The agreed earnings statistics provides comprehensive and detailed information on agreed earnings and their development. To achieve that, several hundred collective agreements, which are received on paper, are inspected and presorted every month. From the collective agreements that are relevant for the agreed earnings statistics, the required information is extracted manually for each sub-area. Due to a lack of suitable technology, information on agreed earnings and provisions contained in collective agreements have so far been published only as extracts and only in a static form in online publications. In the current situation, it has not been possible to meet the increasing user wishes regarding timeliness, extent of agreed earnings information and additional possibilities of analysis.

Those gaps will now gradually be bridged through the new agreed earnings database. In a first step, the data gaps will have to be filled manually once the old data stock has been migrated. As only one fifth of the required data are available as data files, the introduction of the agreed earnings database at first involves considerable additional work. In a second step, collective agreements recently received and new collective agreements will also be entered into the agreed earnings database.

So this will be the first time that all information required for the statistics of agreed earnings is captured in a centralised and standardised way and, through a web application, provided on the intranet for further internal processing. For every collective agreement up to eighty variables will be stored, ranging from general information on the area and economic branch covered by the collective agreement as well as data on weekly working hours to the earnings range of the salary brackets. In addition, special facts are included such as performance-related pay or opening clauses.

Introducing the agreed earnings database has a major impact on the work organisation in the agreed earnings statistics unit: While collective agreements have so far been analysed for particular purposes, all collective agreements will now, in the early phase, almost completely be captured prior to the other work steps. This includes not only preparing the *Fachserien* on agreed earnings but also providing agreed earnings data for the calculation of the index of agreed earnings. In addition, among other things, it will be possible for the first time to do analyses on current topics – such as lower wage limits, agreed working hours and extra payments – across all collective agreements.

For users of the statistics of agreed earnings, too, the agreed earnings database will provide a new range of data. Once the database has been launched, which is scheduled for summer 2010, any users will be able to view agreed earnings data through the internet according to their own specifications. Access to the agreed earnings database will be provided through a search form where information can be entered, for example, on the area covered by the collective agreement and on how specific collective agreements are classified in terms of subject area. Users who do not have such special information can start a search by using information on the occupation, Land or economic branch. The desired collective agreement can then be retrieved direct from the hit list. For the collective agreements available, data can be retrieved not only on earnings and percentage increases but also, among other things, on lump-sum and one-off payments, working hours, vacation and opening clauses. All results can be exported as pdf or csv files.

Optimal use of the agreed earnings database requires, however, critically reviewing the range of data offered so far. Once the web application has been launched, which will be available free of

charge and without registration to any user, it will be necessary to reorient the statistics of agreed earnings towards a user-oriented and theme-oriented presentation of results. For example, the highly comprehensive *Fachserien* offered so far could be replaced by more reader-friendly publications focusing on collective bargaining issues such as opening clauses or extra payments. In addition, studies dealing with one specific economic branch at a time might be published. Apart from that, it is intended to do analyses across all collective agreements. For that purpose, the agreed earnings database already offers pre-programmed searches, for example, the listing of agreed bottom or top earnings or the calculation of the average length of vacation in all collective agreements covered. To permit additional ad-hoc analyses, statistical analysis tools will be used to directly access the data stock of the agreed earnings database. This will open up a new field of activity for the agreed earnings statistics unit, including the opportunity to supply well-processed information to all users. What is more, this creates the preconditions for replacing the ADABAS/NATURAL system currently used to calculate the index of agreed earnings, which is scheduled for the end of 2012 at the latest.

The technical implementation of the agreed earnings database is based on internet and java technologies/PHP and on the MySQL database management system. The internet system consists of the same systems and software components as the intranet system. The difference is that in the internet system only selected data are stored. For example, the policy not to publish enterprise agreements will be maintained. The systems MySQL database, JBoss, Tomcat and Apache web server run only in the DMZ, so that it is not possible to access the intranet system from the internet system. As a matter of principle, no calls aimed at the internal network are initiated from the internet system. Data exchange between the intranet system and the internet system is done by mirroring the data stock. For that purpose an SSH server has been configured on the internet system. Data storage is done only in the internal MySQL database.

Jörg Decker, tel: +49 (0) 611/75 24 42, e-mail: [joerg.decker@destatis.de](mailto:joerg.decker@destatis.de)

## Events

### **18th Scientific Colloquium: “Information visualisation – Graphical processing and statistical data analysis”**

In co-operation with the German Statistical Society, the Federal Statistical Office hosted the 18th Scientific Colloquium on the topic of “Information visualisation – Graphical processing and statistical data analysis” at the Federal Statistical Office's premises in Wiesbaden on 19 and 20 November 2009.

The colloquium was moderated by Prof. Dr. Claus Weihs, Chairholder of the Chair for Computational Statistics at the Faculty of Statistics of TU Dortmund University, who also provided an introduction to the subject. In his contribution, Professor Weihs gave an overview of the results of the Milestones Project, which describes the development of the visualisation of statistical data up to the year 2000. In addition, Professor Weihs presented new developments and further possibilities of arranging graphical techniques.

The specialist discussion was opened by Prof. Dr. Heidrun Schumann, Professor of Computer Graphics at the Institute for Informatics of Rostock University, who gave a speech on the “Visualisation of information: methods and perspectives”. Professor Schumann outlined the latest developments and explained them, using specific methods and applications, especially from the Rostock area, as examples.

Prof. Dr. Stefan Wrobel, Director of the Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS and Professor of Computer Science at Bonn University, presented a talk on the “Presentation of spatial-temporal data”, which gave an insight into the field of visual analytics and its bases with regard to spatial-temporal data. Using a number of case studies, Professor Wrobel illustrated how spatial-temporal data can be analysed by means of appropriate methods and systems.

A paper on “Visual web communication” was presented by Sibylle von Oppeln-Bronikowski, Head of the Department “Questions of Principle of Federal and International Statistics, Dissemination of Information” of the Federal Statistical Office, and Tanja Raschke, Member of Academic Staff in the Federal Statistical Office's section for Online Services. They described the various types and uses of interactive applications at the Federal Statistical Office, such as animated age pyramids, the personal inflation calculator, the business cycle monitor and the price kaleidoscope/Voronoi diagram. These interactive applications can be accessed via the Federal Statistical Office's website.

Afterwards, Dr. Martin Theus, Senior Project Manager of Telefónica o2 Germany in Munich, explained the basic functions of interactive statistical graphs and their efficient use in data analysis in his contribution on “Data analysis using interactive graphs”.

The last talk of the conference day was given by Dr. Maximilian Stempfhuber, Head of the Communication and Information Technology Department of the Library of RWTH Aachen University. His contribution on the “Integration of texts and facts in information systems” described the overall model of an informational architecture for the integration of texts and facts in information systems. This informational architecture can be used both at a scientist's workplace and as an integral part of virtual research environments.

At the end of the first day, the Gerhard Fürst Award was presented for the eleventh time as part of the scientific colloquium. This year, an outstanding scientific study received the Gerhard Fürst Award, the scientific award of the Federal Statistical Office. Two other papers were awarded Promotion Prizes for Young Researchers. The 2009 Award and Promotion Prizes were presented by Vice-President Peter Weigl. For more information on the Gerhard Fürst Award please refer to the

relevant article in this publication or the "[Scientific forum](#)" heading on the Federal Statistical Office's website at [www.destatis.de](http://www.destatis.de).

At the beginning of the second day, Thomas Heumann, Head of the Editorial Department for Informational Graphics of the Frankfurter Allgemeine Zeitung (F.A.Z), presented a contribution on "Informational graphics and maps in newspapers" which explained by way of examples the differences between informational graphs in newspapers and comparable representations in scientific or technical publications. In addition, Mr. Heumann described how informational graphs and maps are developed and designed at the F.A.Z.

The paper on "Visualisation of missing values" by Dr. Matthias Templ, Project Leader at Vienna University of Technology and researcher at the Methodology Department of Statistics Austria, examined the central problem of missing values when processing or analysing data and discussed possibilities of visualisation.

In his contribution entitled "The RapidMiner data mining system (open source data mining)", Dr. Ingo Mierswa, Founder and Managing Partner of Rapid-I GmbH in Dortmund, presented the RapidMiner open source software for data analyses.

The 18th Scientific Colloquium ended with the paper on "Visualisation and statistical modelling" by Prof. Dr. Adalbert Wilhelm, Professor of Statistics at the Jacobs University Bremen and Holder of the Commerzbank Chair of Information Management. Professor Wilhelm's contribution focused on combining the graphical representation of data and statistical modelling. By way of selected examples, it illustrated how visualisation and modelling can supplement each other.

Further information on the 18th Scientific Colloquium can be found under the "[Scientific forum](#)" heading on the Federal Statistical Office's website at [www.destatis.de](http://www.destatis.de). From early in 2010 onwards, all the contributions will be available for free download from the website.

*Birgit Baptistella, tel: +49 (0) 611 / 75 26 03, e-mail: [birgit.baptistella@destatis.de](mailto:birgit.baptistella@destatis.de)*

## **Gerhard Fürst Award 2009**

On 19 November 2009, the Gerhard Fürst Award 2009 was presented to the award winners on the occasion of the 18th Scientific Colloquium in Wiesbaden. Upon recommendation of the panel of independent jurors, Vice-President Peter Weigl granted the Gerhard Fürst Award, the scientific award of the Federal Statistical Office, to an outstanding scientific paper. Two other papers were awarded Promotion Prizes for Young Researchers.

In the category of diploma/master's dissertations, the jury regarded the diploma dissertation on the topic "The influence of body height on wage level and choice of occupation: Current state of research and recent results based on the microcensus" as an outstanding achievement. The dissertation, which was written by economist Fabian Spanhel at Ludwig-Maximilians-Universität München, won the Gerhard Fürst Award 2009 and a prize money of EUR 2 500.

In the category of doctoral theses, two papers were awarded Promotion Prizes for Young Researchers. One was the doctoral thesis on "The topicality of the labour theory of value. Theoretical and empirical aspects", written by Dr. Nils Fröhlich at Chemnitz University of Technology. The other doctoral thesis on the topic "Collective bargaining coverage, wage structure and potential for flexibilisation in collective agreements" was written by Dr. Wolf Dieter Heinbach at Universität Hohenheim. The jury considered both papers as an enrichment to the scientific discussion within the official statistics community. They were awarded a promotion prize and a prize money of EUR 2 000 each.

By offering the Gerhard Fürst Award every year, the Federal Statistical Office seeks to strengthen the existing links between the work of the official statistics agencies and that of institutions of higher education. At the same time, the award is intended to encourage junior scientists to make extensive use in their empirical research of the wide variety of data offered by official statistics.

An appropriate setting for the Gerhard Fürst Award, which was granted in 2009 for the eleventh time, was provided by the 18<sup>th</sup> Scientific Colloquium on the subject “Information visualisation – Graphical processing and statistical data analysis”. The Colloquium was held jointly by the Federal Statistical Office and the German Statistical Society in Wiesbaden on 19 and 20 November 2009.

For more information on the 18th Scientific Colloquium please refer to the relevant article in this publication or the "[Scientific Forum](#)" heading on the Federal Statistical Office's website at [www.destatis.de](http://www.destatis.de).

Details on the Gerhard Fürst Award are also available under the "[Scientific forum](#)" heading. There you will find abstracts of the award-winning papers and information on the conditions of participation for the Gerhard Fürst Award 2010, whose deadline for submissions is 31 March 2010.

Professor Dr. Hans Wolfgang Brachinger (University of Fribourg/Switzerland), the chairman of the expert jury, delivered the laudatory speeches at the award-giving ceremony on 19 November 2009. These speeches were published in Volume No. 12/2009 of the journal “Wirtschaft und Statistik”. Early in 2010, the award-winning authors published detailed articles on their scientific papers in this journal.

*Birgit Baptistella, tel: +49 (0) 611 / 75 26 03, e-mail: [birgit.baptistella@destatis.de](mailto:birgit.baptistella@destatis.de)*