DESTATIS
wissen.nutzen.

METHODS – APPROACHES – DEVELOPMENTS

Information of the German Federal Statistical Office

Edition 2/2018

German Federal Statistical Office, D-65180 Wiesbaden

Published by: German Federal Statistical Office, Wiesbaden

Subject-related information
on this publication:

Division B 2,
tel: +49 (0) 611 / 75 26 53
fax: +49 (0) 611 / 75 39 50
institut@destatis.de

General information
on the data supply:

Information service,
tel: +49 (0) 611 / 75 24 05
fax: +49 (0) 611 / 75 33 30
https://www.destatis.de/.../Contact

Periodicity: (generally) twice a year

An archive of all editions as from 1/2000 is available at www.destatis.de/Methods

Published in December 2018

The catchword

Confidentiality in the 2021 Census

Background

Confidentiality in official statistics is defined in Section 16 of the Federal Statistics Act (BStatG). In accordance with these provisions, individual data on personal and material circumstances provided for federal statistics must not be disclosed by those entrusted with the production of these federal statistics unless otherwise stipulated. As a consequence, it must be guaranteed that no information can be derived on individual cases from any 2021 Census output released.

Traditionally, cell suppression methods for blocking specific information are used in official statistics to ensure data confidentiality. When the 2011 Census evaluation programme was prepared it turned out that complete and consistent confidentiality could not be achieved by using suppression methods. This is mainly due to the fact that there is no predefined "final" set of tables but a flexible online evaluation system for users. Therefore, the SAFE procedure for the secure anonymisation of individual data was used in the 2011 Census to ensure the statistical confidentiality of tables of counts based on enumeration[1]. SAFE prevents the disclosure of individual data by slightly modifying the microdata themselves and then creating the evaluation tables from these modified data.

A method used in the Australian Census since 2011 was regarded as better suited for the 2021 Census. That method[2] does not modify the microdata themselves but makes modifications at the time the results are generated: a random procedure called cell key method permanently assigns a small perturbation amount to each result (or cell). The total of each original result and the assigned perturbation amount is released instead of the original result; this is called stochastic perturbation of the original results.

The relevant perturbation specifications (e.g. maximum perturbation) will be defined once, as a standard, by the statistical offices before the results of the 2021 Census are published. Anyway, as the perturbation amounts added to the original results average zero and mostly range from -2 to +2, they will have no effect, virtually, on larger results.

To specify the stochastic features, a probability distribution is defined once for the perturbations. It is a conditional probability distribution recorded in the form of a so-called perturbation matrix. A conditional probability distribution is needed as the meaningful target frequencies *j* may deviate, depending on the original frequency *i*. Original frequencies of zero are to remain unchanged in the 2021 Census, and results published are to contain neither negative values nor values of 1 or 2 (whether original or perturbed). The perturbation matrix is used to specify the probability of an original frequency *i* being perturbed to the target frequency *j*. Relevant parameters for creating the perturbation matrix are:

- the maximum deviation, i.e. the absolute value of the maximum deviation between original frequency and target frequency,
- the variance, i.e. the dispersion measure for the distribution of the deviations.

[1] In contrast to tables of counts based on extrapolation where statistical confidentiality is guaranteed anyway by the fact that the result is uncertain due to sampling (standard error).

[2] Fraser, B., Wooton, J. (2006): A Proposed Method for Confidentialising Tabular Output to Protect against Differencing, in Monographs of Official Statistics. Work session on Statistical Data Confidentiality, Eurostat - Office for Official Publications of the European Communities, Luxembourg, 2006, pp. 299-302, and Thompson, G., Broadfoot, S., Elazar, D. (2013): Methodology for the Automatic Confidentialisation of Statistical Outputs from Remote Servers at the Australian Bureau of Statistics, paper presented at the Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality (Ottawa, 28-30 October 2013) available at http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2013/Topic_1_ABS.pdf

In addition, a probability will be defined for the 2021 Census,
- the probability that an original frequency will remain unchanged.

Stochastic perturbation using the cell key method

First, the microdata records are extended for the purposes of the cell key method. A random number distributed uniformly in [0,1] (called record key or seed) is assigned to each record or statistical unit of the data stock (persons, buildings, dwellings, households and families).[1]

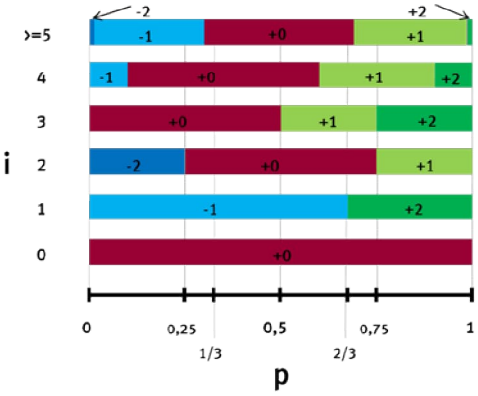
When the frequency counts are made (i.e. number of statistical units with cross combinations of variables that are of interest), the record keys are also added up for all these counted combinations of values of variables in order to produce the results to be published. The record keys which have been added up (and retransformed to the original interval [0,1]) are called cell keys (or seed sums). In addition to its frequency value, each evaluated combination of values of variables is thus also assigned a cell key. Since record keys are added up, logically identical combinations of values of variables are always automatically assigned the same consistent cell key.

In the last procedural step called lookup, the small perturbation amount of each cell is looked up as illustrated in Chart 1, using the original value and the cell key. The perturbation amount added to the original result is the difference between the target frequency and the original frequency. As the cell keys are consistent, the perturbation amount is always the same for logically identical combinations of values of variables. This is how the method provides consistent tables and therefore has to deal with all results in the same manner - also with marginal totals and sub-totals.

Perturbation matrix

i (Originalhfkt.)	j (Zielhäufigkeit)								
	0	1	2	3	4	5	6	7	8
0	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1	0,67	0,00	0,00	0,33	0,00	0,00	0,00	0,00	0,00
2	0,25	0,00	0,00	0,50	0,25	0,00	0,00	0,00	0,00
3	0,00	0,00	0,00	0,50	0,25	0,25	0,00	0,00	0,00
4	0,00	0,00	0,00	0,10	0,50	0,30	0,10	0,00	0,00
5	0,00	0,00	0,00	0,01	0,29	0,40	0,29	0,01	0,00
6	0,00	0,00	0,00	0,00	0,01	0,29	0,40	0,29	0,01

Perturbation chart



Perturbation legend: -2, -1, 0 (no perturbation), +1, +2

Chart 1: Perturbation matrix and graphical representation in the form of a perturbation chart[2]

The perturbation amount of a result which is shown as a marginal total in a table is not equal to the total perturbation amounts of the individual interior table cells. This procedure makes sense because it prevents a greater loss of accuracy in table margin values. It is similar to commercial rounding where marginal totals of tables are first calculated precisely and rounded afterwards instead of the rounded values of inner table cells being added up. The familiar sentence "as a consequence, there can be minor differences between totals and the sum of their components" also applies in stochastic perturbation. Therefore, tables subjected to this procedure will as a rule not be exactly additive.

[1] Statistical units representing the same items should be assigned the same values of the random variable in order to avoid paradoxes (e.g. five people living in three one-person households).

[2] The perturbation matrix can be presented graphically as a perturbation chart. Each bar represents an original value, different colours stand for different perturbations and the length of the coloured parts of a bar for the probability laid down in the perturbation matrix that the respective perturbation will be applied to the original value concerned. At the lookup stage, the perturbation is looked up in the perturbation chart in the row determined by the original value i at the point p=cell key.

Non-additivity is accepted, however, as the method has two important advantages.

1.) Cross-tabular consistency: whatever the table that contains a specific result (e.g. "number of under 7-year-olds"), the added perturbation value, and thus the result shown, will always be identical, whether it is, for example, the marginal total of the two inner cells "under 7-year-olds, male" and "under 7-year-olds, female" in one table or the total of two age groups (e.g. "0 to 3-year-olds" and "4 to 6-year-olds") in another.

2.) Accuracy: it is ensured that several perturbations which happen to have the same direction (either positive or negative) do not accumulate when they are added up and thus do not lead to somewhat larger differences between the original values and the confidential values in some cases. This can be illustrated by our example: if the "under 7-year-olds" are shown in a table as a total of the respective 7 individual age years (0, 1, 2, ...) and, by chance, the perturbation amount is negative for all seven age years, the perturbation value resulting from the sum of all seven individual perturbation amounts would be not that small and unnecessarily reduce the overall result strongly.

Conclusion and outlook

The statistical offices of the Federation and the Länder have decided to apply the cell key method in the 2021 Census. In view of the advantages, i.e. the high accuracy of the procedure and the cross-tabular consistency of cells with identical content, non-additivity seems not much more than a cosmetic flaw.

The confidentiality procedure now has to be planned and specified in detail. In this context, the issue of whether additivity can in part be restored subsequently and how ratios should be handled still has to be studied. It also needs to be clarified how IT tools can be supplied to municipalities with isolated statistics units and to the research data centres of the statistical offices of the Federation and the Länder which enable them to create official statistical results from confidential statistical results based on the original individual data provided, in an uncomplicated and consistent manner.

Birgit Kleber, tel: +49 (0) 611 / 75 42 85, e-mail: Birgit.Kleber@destatis.de

Sarah Gießing, tel: +49 (0) 611 / 75 27 01, e-mail: Sarah.Giessing@destatis.de

Methods of federal statistics - Further development

Use of mobile phone data for official statistics

As part of the generally increasing use of digital technology, official statistics agencies are facing the challenge to explore and employ new data sources and to organise their processes and procedures accordingly. For that reason, the Federal Statistical Office (Destatis) is involved in several European projects as part of which first feasibility studies are carried out to determine the usefulness of new digital data, such as mobile phone data, for official statistics. The use of such data is seen as potential for a quicker, more precise and more cost-effective production of official statistics and may help to reduce the burden on respondents.[1]

A number of scientific studies has already been conducted concerning the usability of mobile phone data for statistical purposes. De Meersmann et al. (2016), for instance, assessed the quality of mobile phone data as a source of official statistics. Deville et al. (2014) used mobile phone data for dynamic population mapping. Makita et al. (2013) investigated whether mobile phone network data could be used to estimate small area population, and Schmid et al. (2017) combined socio-demographic indicators with mobile phone data to present these indicators for small areas as well.

In order to research the use of mobile phone data for official statistics, Destatis entered into cooperation with T-Systems International GmbH and Motionlogic GmbH (both wholly-owned subsidiaries of Deutsche Telekom AG) in September 2017. The conceptual designs of the planned feasibility studies were developed in coordination with the Federal Network Agency, the Federal Commissioner for Data Protection and Freedom of Information, and T-Systems. The medium and long-term objective is to use mobile phone data to provide a valid picture and estimate of the daytime and resident population, of commuting flows and of tourist distribution in the whole of Germany. The population figures of the 2011 census as well as the Commuter Atlas published by IT.NRW[2] are used as a benchmark to check the representativity of the data.

However, basic research is needed before new digital data can become an integral part of the production of official statistics. For that purpose, Destatis has started first feasibility studies to examine the suitability of mobile phone data for official statistics in the fields mentioned above. At European level, such studies were already conducted as part of 'ESSnet Big Data 2016 – 2018'.[3] In this context, Destatis participated, for example, in the WP5 working group which looked into the usability of mobile phone data for official statistics at European level. Follow-up projects on this issue are planned for 'ESSnet Big Data 2018 – 2020'.

The project implementation at national level and first feasibility studies are currently being carried out by Destatis in cooperation with IT.NRW and are still limited to the Land of Nordrhein-Westfalen. Due to data protection rules, Destatis only receives anonymised, aggregated data on mobile phone activities from T-Systems. The aggregated data provided make it possible to map the daytime and resident population.

The data record currently available to Destatis contains mobile phone activities[4] of Telekom customers in Nordrhein-Westfalen (NRW) for a statistical week that consists of 24-hour days which were selected from the months of April, May and September 2017.[5] The mobile phone

[1] See also Wiengarten et al. (2017).

[2] For more information on the Commuter Atlas see: https://www.pendleratlas.nrw.de/.

[3] See: https://ec.europa.eu/eurostat/cros/content/essnet-big-data_en.

[4] A mobile phone activity is defined as a length of stay at a location without movement, with all signalling data being evaluated, i.e. phone calls, text messages and data connections. Furthermore, signalling data are produced automatically and only register the location of the cell tower to which a mobile device is connected at a specific time.

[5] School holidays and public holidays are excluded here.

activities comprise the average activities on the weekdays selected. The weekdays are categorised according to five types of day, with the days from Tuesday to Thursday being grouped together. In addition, mobile phone data contain information, for example, on socio-demographic characteristics of mobile phone users, such as age group and sex[1]. In compliance with data protection rules, the mobile phone activities were anonymised[2] and aggregated. Only values based on a minimum of 30 activities per grid cell[3] were transmitted to Destatis.[4] The grid cells conform with INSPIRE and correspond to the census grid cells of the 2011 Census Atlas.[5,6]

The number of mobile phone activities depends on the location and number of cell towers in the various grid cells. Depending on the cell towers' location (rural or urban), their frequencies differ and, as a result, they are sometimes distributed unevenly across the regions. Consequently, an existing geometry may contain 5 to 20 cell towers. For that reason, some geometries are combined to ensure a minimum of 30 activities per grid cell. Since the number of mobile phone activities depends on the dwell time of mobile devices, long mobile phone activities corresponding to the length of the dwell time are counted and included in the data record, while short mobile phone activities are left out of account. The dwell time is defined as the length of stay of a mobile device at a location or in a grid cell without any movement. The dwell time in the data record available is two hours in order to filter out short mobile phone activities which result, for instance, from quick movements between the grid cells.

The aim of the first analyses was to provide a valid picture of the resident, daytime and working population. To this end, the total number of active SIM cards per grid cell was identified in the aggregated data studied. First of all, the figures were used to determine the correlation between mobile phone activities and census values by type of day and time of day for NRW, as shown in Figure 1 below. Overall, the values reveal a high correlation of 0.8 between mobile phone activities and census values throughout Saturday and Sunday. On weekdays, the correlation declines to less than 0.7 between 5 a.m. and 4 p.m., which indicates significant differences in the resident population according to the 2011 census and according to the location of mobile phone activities within the given period.

[1] This includes contract, prepay, congstar and business customers. The characteristics, however, are only available for contract customers.

[2] Telekom AG uses a procedure agreed with the Federal Commissioner for Data Protection and Freedom of Information (BfDI) to anonymise the data.

[3] A grid cell is a square-shaped geographical unit of varying or uniform grid width, with clear cell and spatial reference. Grid cells do not follow national administrative boundaries but represent a suitable territorial delineation. Several cells together form a grid as a large-scale reference system.

[4] The data provider Motionlogic has no access to individual or raw data, either.

[5] For more information on the Census Atlas please refer to: https://atlas.zensus2011.de/.

[6] In addition, the mobile phone activities are supplemented by three socio-demographic characteristics indicating the gender, age group and nationality proportions of the SIM cards per grid cell.

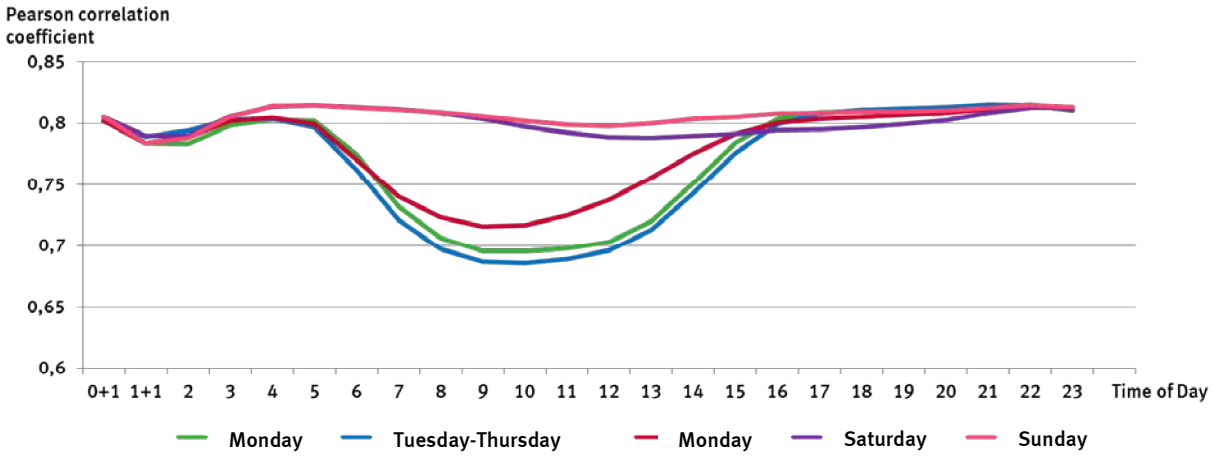


Figure 1: Pearson correlation coefficients referring to the census values and the mobile phone data, by statistical day and period.

The results show that, to some extent, the mobile phone data available could provide a good picture of the population. The differences observed between the population figures based on mobile phone data and those based on census values may partly be explained by the time difference between the mobile phone data from summer/autumn 2017 and the census data from 2011, but they may also result from the extrapolation method used by Motionlogic. The extrapolation is based on Deutsche Telekom's regional market shares across the mobile communications market in the whole of Germany. The regional market shares are determined using postal codes. The mobile phone activities are weighted based on the location or postal code of a mobile device's first signal at the very beginning of its activity chain. This means that the extrapolation of all the activities of a mobile device counted throughout a day depends only on the market share of the postal code of the first activity counted. In addition, the figures are only extrapolated to the total number of mobile phone users. At present, roughly 80 percent of the German population own a mobile phone.[1] Consequently, 20 percent of the population are not taken into account in the extrapolation. The issue of biases and selectivities will be discussed in future papers.

The data studied allow to distinguish between daytime and nighttime population. Using the data currently available, it is however not yet possible to describe the commuting patterns as such, i.e. the movement in space. Nevertheless, the results allow to deduce commuter regions[2]. The forthcoming *'Pendler Mobil'* project, which will be carried out in cooperation with IT.NRW, will have the objective of identifying the domains where mobile phone data may contribute to complementing the commuter accounts. Using origin-destination matrices, it is possible to employ mobile phone data for mapping commuting flows during the day. If the SIM card nationality is taken into account, it is even possible to record cross-border in-commuters.

In addition, the SIM card nationality facilitates analysis for the purpose of tourism statistics. Further cooperation is planned to conduct feasibility studies on the usability of mobile phone data for tourism statistics, such as accommodation statistics. One objective is to analyse the

[1] Cf. https://de.statista.com/statistik/daten/studie/585883/umfrage/anteil-der-smartphone-nutzer-in-deutschland/, accessed on 8 June 2017.

[2] They can be identified based on the territorial units with above and below average population densities during the day, especially in the middle of the day compared with the morning and evening hours.

influence of fairs, festivals or special events on tourism in a region under study and on the accommodation units located therein.[1]

At European level, there are plans to use an equivalent data record for the ESSnet project *'City data from LFS and big data'*. This project will examine whether and to what extent various Labour Force Survey (LFS) indicators can be estimated at the level of functional urban areas. According to the model described by Schmid et al. (2017) and as part of collaboration between Destatis and Freie Universität Berlin, mobile phone data are intended to be linked to LFS indicators and to be estimated for small areas using a small area estimation method. The basic question is where and to what extent small area estimation could be used in combinations of mobile phone data and official statistical data.

The first step has been taken with the start of the feasibility studies on the use of mobile phone data as described here. The cooperation initiated with T-Systems/Motionlogic is aimed at implementing existing procedures in Germany and at refining them, again in close collaboration with the Federal Network Agency and the Federal Commissioner for Data Protection and Freedom of Information. The representativity of the data across Germany will continue to be of essential relevance. To ensure this, further steps will have to be taken to obtain data from additional mobile network providers in Germany. The cooperation with T-Systems is just the start of the process of integrating this new source of digital data into official statistics.

References

N. Makita, M. Kimura, M. Terada, M. Kobayashi & Y. Oyabu (2013). Can mobile phone network data be used to estimate small area population? A comparison from Japan

F. De Meersman, G. Seynaeve, M. Debusschere, P. Lusyne, P. Dewitte, Y. Baeyens, A. Wirthmann, C. Demunter, F. Reis, H.I. Reuter (2016). Assessing the Quality of Mobile Phone Data as a Source of Statistics (mirror site), Q2016

P. Deville, C. Linarde, S. Martine, M. Gilbert, F.R. Stevens, A.E. Gaughan, V.D. Blondela & A.J. Tatem (2014). Dynamic population mapping using mobile phone data, PNAS 2014 111 (45) 15888-15893

T. Schmid, F. Bruckschen, N. Salvati and T. Zbiranski (2017). Constructing socio demographic indicators for National Statistical Institutes using mobile phone data: estimating literacy rates in Senegal. Journal of the Royal Statistical Society: Series A (Statistics in Social Sciences). doi:10.1111/rssa.12305

L. Wiengarten und M. Zwick (2017). Neue digitale Daten in der amtlichen Statistik, Wirtschaft und Statistik, issue 5.

Sandra Hadam, tel: +49 (0) 611 / 75 34 52, e-mail: Sandra.Hadam@destatis.de

[1] It must be noted here that the mobile activities are counted separately for each SIM card nationality, without including any socio-demographic characteristics and with no possibility of linking the data.

Resource management expenditure

Background

How much money does a society spend on the management of water, forest, energy and mineral resources? Who finances the expenditure and what are the effects? What is the demand for resource management goods? What are the turnover and employment figures related with the production of such goods?

This is only part of the questions the resource management expenditure accounts (ReMEA) can answer. They are to show society's expenditure on maintaining the stock of economic resources and are outlined in the System of Integrated Environmental and Economic Accounting (SEEA Central Framework 2012), the international environmental-economic accounting standard. ReMEA may serve, among other things, to analyse the impact of resource management policies and determine the need to finance specific activities. Resource management activities include:

- slowing the depletion of natural resources, e.g. by means of recovery, re-use and recycling,
- restoring stocks of natural resources,
- managing natural resources in a general sense, e.g. supervising, monitoring and controlling,
- producing goods and services needed to manage or maintain natural resources.

In 2016 and 2017 a partly EU-funded feasibility study was carried out in the Federal Statistical Office which, among other things, assessed ways to cover resource management expenditure in Germany. The concept and the methodological framework of the environmental protection expenditure accounts already implemented in environmental-economic accounting can be adapted to suit ReMEA. ReMEA, too, is largely compatible with the European System of Accounts (ESA 2010). Monetary transactions, most of which are already considered in the national accounts, are shown separately and broken down further. ReMEA is based on the Classification of Resource Management Activities (CReMA):

- 10: Management of waters
- 11: Management of forest resources
  - 11A: Management of forest areas
  - 11B: Minimisation of the intake of forest resources
- 12: Management of wild flora and fauna
- 13: Management of energy resources
  - 13A: Production of energy from renewable sources
  - 13B: Heat/energy saving and management
  - 13C: Minimisation of the use of fossil energy as raw materials
- 14: Management of minerals
- 15: Research and development activities for resource management
- 16: Other resource management activities

As CReMA has not been agreed upon at international level yet, its definitions and the structure of the individual classes have no legal basis. Efforts are being made at EU level to have a working group revise the environment-related classifications.

Methodological approach

With the help of experts of the various resource areas, recommendations were first made on how to adjust the CReMA classes. The aim was to describe them exhaustively and without overlap and make them more concrete by fitting examples of goods (commodities and services).

This made it possible to identify such goods in existing statistical classifications. As part of the feasibility study, a functional delimitation was made based on groups of goods. The approach assumes that goods can be distinguished by their resource management support function and identified by the official goods statistics classification. The most detailed breakdown is provided by the Product Classification for Production Statistics (GP 2009). The functional delimitation of the resource management economy was a challenge in that the definitions of many goods are too general in nature to reveal their purpose. As a consequence, only a very small part of the goods could be clearly classified to resource management (e.g. solar cells), while the majority of the goods (e.g. pumps or measuring instruments) may also serve other purposes ("dual use" problem).

Based on the GP 2009, a list of potential resource management goods was compiled for each CReMA class. As the GP 2009 serves as a descriptive classification for the individual classes of the Classification of Economic Activities (WZ 2008), the list could be used to establish links to the WZ. The aim was to cover resource management services as well. However, this aggravated the problem already encountered with the much more detailed GP classification: the degree of detail of the WZ classification does not suffice to clearly identify resource management activities. The same applies to other official classifications such as the Classification of the Functions of Government (COFOG 1999). That classification is not detailed enough, either. COFOG 04.3 "Energy", for instance, does not distinguish between conventional and regenerative energy sources.

In addition to the lack of potential data sources there is another problem which should not be disregarded: many resource management activities such as the use of paper instead of energy-based plastics have both a positive effect on the stock of one resource (CReMA 13C) and a negative effect on another resource (CReMA 11). It still is not entirely clear how such cases can be handled. As a rule, the assessment of resource-efficient products requires an analysis of their complete life cycle.

Data availability

The degree of detail in official statistics is not sufficient yet to present expenditure on resource management activities. It was possible in some cases to estimate the share of resource management on the basis of non-official data, e.g. for the turnover of paper made from waste paper as a proportion of the turnover of total paper production. This was used as a basis for estimates or for the further differentiation of aggregated official data. On the one hand, there was considerable effort and uncertainties, mostly with regard to individual components of the expenditure accounts such as persons employed or exports because the data situation was more critical here. On the other, it was not possible to guarantee compatibility with the national accounts.

Conclusions

When it comes to resource management activities, goods cannot be clearly identified in existing statistical classifications. It is therefore not possible at the time being to calculate individual items of the expenditure accounts such as "turnover", "persons employed" or "exports".

Katharina Schlesag, tel: +49 (0) 228 / 99 643 - 81 92, e-mail: katharina.schlesag@destatis.de

Foreign trade statistics: change to the monthly press release

The major variables collected for the publication of foreign trade statistics include not only the commodity code and the value and quantity of goods, but also the country of origin and of consignment (imports) and the country of destination (exports). While the country of origin is the country where the goods were entirely obtained or produced, the country of consignment is the country from which the goods were directly moved to the statistical territory. An example is banana imports from Ecuador (country of origin) via the Netherlands (country of consignment).

As a rule, the country of origin of a commodity is indicated as the partner country in the publications of the Federal Statistical Office. Regarding imports from non-EU countries, the relevant information is included in the data derived from the customs declarations and transmitted electronically for statistical purposes. In intra-EU trade, the country is to be indicated in the monthly Intrastat declarations by the businesses obliged to provide the Federal Statistical Office with information on their imports.

Since the results were published for reference month January 2018, the initial monthly press release of foreign trade data has also been compiled in accordance with the country of origin principle, which governs the other publications, too. As a result of the further development of the supplementary estimation procedure, the presentation of results for the various groups of countries in the press release (EU Member States, euro area, non-euro area, third countries) now corresponds to that of other publications of the Federal Statistical Office. Another advantage is that, as of reference month January 2018, it has been possible to publish the list of Germany's trading partners compiled in accordance with the country of origin principle and an overview according to chapters of the Commodity Classification for Foreign Trade Statistics at the same time as the press release.

Compared with the previous results, changing over to the country of origin concept in the monthly press release has led to an increase in imports from third countries and a decline in imports from the EU Member States. This also affects the foreign trade balance. Germany's export surplus regarding third countries has declined, while the surplus of exports to the EU Member States has grown. However, this change only refers to the results of the monthly press release. A table showing results for the different groups of countries is available for download on the internet.

Unlike Germany, Eurostat publishes results by country of origin only for the trade with third countries. The data on the trade between EU Member States are rather based on the country of consignment principle. Therefore the EU database Comext for instance indicates the Netherlands as Germany's trading partner for imports of goods from Ecuador (country of origin) which are imported via the Netherlands (country of consignment). In accordance with the country of origin principle, however, the German statistics assign the same imports to Ecuador.

As the European Union sees it, indicating the country of origin is only useful for imports via the EU external borders because citing the country of origin several times (Ecuador for imports to the Netherlands and again Ecuador for imports to Germany) would inflate the foreign trade figures. However, information on the "origin" of a commodity, that is, the country where it was entirely obtained or produced, is requested by users to obtain a picture of Germany's trade relations in the context of globalisation.

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

Events

EU-wide seminar on major aspects of Intrastat modernisation, Lisbon, 17-18 April 2018

Foreign trade statistics will in future be governed by the Framework Regulation Integrating Business Statistics (FRIBS). An essential milestone is the modernisation of Intrastat in accordance with the relevant decision taken by the European Statistical System Committee (ESSC) in May 2016. The aim is to reduce the burden on the parties responsible for providing information by using innovative and flexible compilation methods while at the same time maintaining a high level of quality. The approaches to achieve this include an integrated statistics production using various data sources and the exchange of microdata on intra-EU exports between the Member States involved. However, the integrated production approach requires largely harmonised intra-EU trade statistics. To meet this requirement, Eurostat has set up a multiannual work plan: preparations for the microdata exchange take place in 2017 and 2018, and the changeover to the new system is scheduled for 2019. Initially, however, the beginning of implementation was scheduled for 2020.

In the context of this multiannual work plan the Member States, together with Eurostat, are currently working on a total of nine work packages on various topics. The two-day seminar in Lisbon focused on three work packages: confidentiality, globalisation and compilation.

Confidentiality work package

Regarding foreign trade statistics, the relevant legal provisions at European level[1] provide for "passive confidentiality". Under this principle, the microdata of importers or exporters are kept confidential if they request to do so. The statistical office of the Member State which collects the data considers whether the request for confidentiality is justified. Confidentiality cases are recorded under the collective "Confidentiality cases in foreign trade statistics" code or the "Non-disclosed countries" country code.

The issue of passive confidentiality is a central topic regarding the future exchange of microdata (MDE). Germany presented its national practices of primary and secondary confidentiality in relation to this topic. The experience gained with confidentiality of Land data was considered a valuable contribution regarding a confidentiality practice across the network of statistical offices. As the future microdata exchange also requires harmonisation of confidentiality approaches, Eurostat will not only conduct a survey of national confidentiality practices. It has rather also considered Germany's proposal to examine the potential of centrally determining secondary confidentiality with a view to minimising the 'domino effect'.

Globalisation work package

The "nature of the transaction" is one of the most important variables published in foreign trade statistics. The variable provides information on whether a given cross-border trade in goods transaction is a purchase or sale, return of goods, delivery of goods under aid programmes or operation with a view to or following processing under contract not involving transfer of ownership. This information is used, for instance, in national accounting or by Deutsche Bundesbank to compile Germany's current account of the balance of payments.

Based on a previous survey among the Member States, the topics of the seminar centred on doing without a separate labelling of operations under joint defence projects or other joint intergovernmental production programmes and on expansion to cover globalisation aspects such as re-exports, quasi-transit, intra-company trade and mail order sales. The proposal to include in

[1] Art. 10 of Regulation (EC) No 471/2009, Art. 11 of Regulation (EC) No 638/2004

the discussion the major users in the area of national accounts and the balance of payments was welcomed by Eurostat.

Compilation work package

a) Country of origin

Some countries consider it a burden for businesses to indicate the country of origin in the intra-EU export declarations and have therefore conducted projects to identify possible ways of estimating the country of origin. For this purpose, the usability of both customs data and the European production survey Prodcom was analysed. However, the results presented at the seminar were basically not regarded as satisfactory so that nearly no Member State considers replacing the relevant data collection by estimates.

b) The future Intrastat data collection system

Ideas regarding the future system of Intrastat data collection were also presented at the seminar. As the exchange of microdata is currently under preparation, concrete simulations could not yet be carried out. Therefore theoretical aspects were considered which would have to be taken into account in future modelling. Furthermore, first experiences gained by several Member States in collecting variables on a voluntary basis whose collection will be compulsory in the future (VAT identification number of the recipient of exports and country of origin) were also presented to the participants.

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

Ilda Duarte Fernandes, tel: +49 (0) 611 / 75 34 27, e-mail: ilda.duarte-fernandes@destatis.de