

The system aspect of statistical quality

Norbert Rainer, Statistics Austria (norbert.rainer@statistik.gv.at)

Josef Richter, University of Innsbruck (josef.richter@aon.at)

Abstract

In order to guarantee the quality of European statistics according Regulation No 223/2009 Article 12 seven criteria are to be met. The first five of them refer to the quality of a statistical project seen in an isolated way. The last two criteria ‘comparability’ and ‘coherence’ address the relationship between statistical projects; they deal with system wide aspects. In many cases the assessment of quality concentrates on the first five criteria and in particular on the criterion of accuracy. The interdependencies between statistical projects within a system such as business statistics are too often neglected. Based on the findings of the “ESSnet on Consistency” the paper concentrates on the role of consistency and coherence seen from the perspective of an integrated system of business statistics. Special emphasis is put on the need to arrive at a certain hierarchy of objectives taking into account that the criteria of statistical quality as laid down in Regulation No 223/2009 are often conflicting goals.

1. Introduction

In order to guarantee the quality of European statistics according Regulation No 223/2009 Article 12 seven criteria are to be met: relevance, accuracy, timeliness, punctuality, accessibility, comparability and coherence. The first five of them refer to the quality of a statistical project seen in an isolated way. Only the last two criteria ‘comparability’ and ‘coherence’ address the relationship between statistical domains; they deal with system wide aspects:

- Comparability refers to the measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas, sectoral domains or over time;

- Coherence refers to the adequacy of the data to be reliably combined in different ways and for various uses.

In many cases the assessment of quality concentrates on the first five criteria and in particular on the criterion of accuracy. The interdependencies between statistical projects within a system such as business statistics are too often neglected.

However, in the last years more and more emphasis was put on ‘comparability’ and ‘coherence’. This re-orientation is also caused by the fact that users increasingly ask for coherent data, as the phenomena to be studied become more complex and interrelated. As a consequence the so called “Vision paper”[1] argues in favour of an integrated system of statistics replacing the present stove-pipe model in which each statistical domain is viewed in an isolated way.

Within the objectives of the MEETS Programme [2] one of the main projects was devoted to the consistency of concepts and methods of business and trade related statistics. European business statistics are compiled in accordance with EU legislation, which has developed over the years and is not necessarily consistent in terms of concepts, scope, definitions, methodology, etc. and furthermore, may not uniformly be applied over the Member States. Based on a the “Consistency Study”[3] which identified and analysed the inconsistencies found in the legal framework of European business statistics, the objective of the “ESSnet on Consistency” was to prepare an inventory of the implementation of the concepts in the Member States and to derive appropriate proposal how to reduce inconsistencies that were identified in the legal basis and as well as in the implementation in the Member States. The “ESSnet on Consistency” analysed all relevant determinants of a statistical project:

- Statistical units
- Target population
- Frames
- Reference period
- Classifications
- Characteristics and variables

[1]Communication from the Commission to the European Parliament and the Council on the production method of EU statistics: a vision for the next decade; COM(2009) 404 final

[2] Programme for the Modernization of European Enterprise and Trade Statistics (MEETS) (Decision No 1297/2008/EC)

[3] Eurostat (2010); External Study on Detail Evaluation of the Legal Acts in the Areas of Statistics which were Identified by Member States as Areas to Revision, Luxembourg 2010

Even if the ESSnet presented proposals and recommendations for all the determinants for consistent statistics, the focus was on conceptual issues, and how these could be adapted or developed to achieve consistent data. However, it was also recognised that achieving consistency needs also to harmonise the statistical processes. This was especially emphasised concerning the use of the statistical business registers which should serve as the backbone for all business statistical domains. To achieve the vision of integrated statistics necessitates integrated processes.

Consistency of concepts is a necessary, but not a sufficient precondition for consistent statistics. The transformation of inconsistent stove-pipe procedures to an integrated system needs a new paradigm on how the statistical data are to be produced: harmonised concepts – which are yet not at all achieved in European business statistics – have to be applied by harmonised statistical processes.

2. Characteristics of business statistics of special relevance for consistency and coherence

The standard quality criteria and the methods to quantify ‘quality’ are not always adequate for business statistics. In this context a number of specific characteristics of business statistics such as the complexity of units, the skewness of distributions and the type of variables [4] used need to be taken into account.

Most of the key variables in business statistics have their basis in the accounting system and are thus very dependent on the legal and institutional background. Unfortunately, the information basis of the respondents is by no means homogeneous. It is not even homogeneous within a country because it differs by size classes and by type of economic activities. In a number of countries there are no mandatory standards for accounting. Certain standards only are compulsory for units with a certain legal form and/or beyond a certain size. The information basis available to the respondents and the characteristics of the administrative data are even less comparable between countries. Accounting rules, tax laws (with the exception of VAT) and social security regulations are not standardised within the EU. To make matters worse, accounting rules and other administrative standards are changing over time.

[4] As regards these characteristics see also Kloek (2011) and the Memobust handbook (2014)

In this difficult environment business statistics aim at transforming incomparable basic information into statistical results comparable among branches and countries as much as possible. Users should be aware that this fundamental challenge is much more relevant for business statistics than in other fields of statistics, such as e.g. social statistics.

A second feature which is specific for business statistics and very relevant for ‘consistency’ and ‘coherence’ is that the results are the outcome of a process characterised by a very high degree of division of labour, already at the level of the respondents, especially in the case of bigger units. In a very simplified manner five consecutive steps in the data generating process can be distinguished [5].

Step 1 - In this step the management information and accounting systems are designed and the basic observations are made. The statistical reporting obligations only play a very limited role in the design of the information systems. Usually institutional factors such as tax laws, social security regulations, reporting obligations to headquarters in the case of multinationals and the like are considered more important.

Step 2 - In this step the basic information available in enterprises is condensed according the concepts laid down in the management information and accounting systems. The result of classifying and aggregating very detailed pieces of information is the (aggregated) information to be found in the accounting and related information systems in the enterprises.

Step 3 - It is not before this step that statistical criteria enter the picture. Managers are confronted with questionnaires. The definitions, concepts and classifications behind these questionnaires are determined with specific analytical goals in mind. Because of the high degree of division of labour in the previous steps the respondent will have to rely on information from very different sources within the own reporting unit. As regards employment and wages and salaries, the information will be delivered by the personnel department. As far as output and intermediary inputs in monetary terms are concerned information is derived from the accountants’ department. Colleagues from the research department will provide information on R&D.

Statisticians have to cope with the fact, that the accounting conventions and valuation methods used typically differ from those required by business statistics and national accounts.

[5] This distinction is to a good deal inspired by Blackburn (1996).

Therefore the definitions used do not always correspond to the ones used in the previous steps of organising information on the level of the enterprise.

Step 4 - In this step the individual micro data sets provided by the respondents are edited in the Statistical Offices. Statisticians have to deal with item- and unit-non-response and other data editing issues. After all these processes edited micro data sets are available.

Step 5 - In this step - also performed in the Statistical Office - classification and aggregation lead to data on the meso level such as for example to information by industries in structural business statistics.

3. Sources of inconsistencies

Two main sources of inconsistencies in business statistics can be distinguished and need to be treated separately. On the one hand side, inconsistencies - in particular between statistical domains - resulting from differences in concepts. Most of these concepts are laid down in European legal acts. To help to eliminate these inconsistencies was the task of the “ESSnet on Consistency”. On the other hand inconsistencies occur in the data generating process. In this process the approaches and methods used in the reporting units and in the Statistical Offices are of decisive importance for consistency and thus for statistical quality in a system wide context. These methods are not necessarily connected with legal provisions or recommendations laid down in Manuals.

3.1 Inconsistencies on the level of single respondents

Two sources of inconsistencies can be distinguished on the micro level. On the one hand present business statistics use slightly different definitions for variables such as for example turnover or number of persons employed. Such problem areas have been identified and the work done by the “ESSnet on Consistency” will contribute to arrive at changes in the legal basis.

The second source of inconsistencies already on the level of the reporting unit is due to the fact that different people and different processes (such as coding) are involved in the preparation of the micro data. The persons responsible for filling in the PRODCOM questionnaire might classify a certain good in a different way than the persons responsible for INTRASTAT, although the classifications used in the two domains are to a considerable

degree compatible. This division of labour and the danger of a lack of internal communication may result in a different ‘translation’ of internal codes or verbal descriptions into statistical classification systems and thus for example to exports of goods by a unit which were (according to the information delivered to the Statistical Office) not produced in this unit. Other factors at the micro level are the correct understanding of the reporting obligation concepts and the possibility and willingness to transform the micro data accordingly.

3.2 Inconsistencies on the micro level of micro data between units

Special problems occur in situations when the statistical unit is not identical with the unit for which the aggregated information after Step 2 is available. Under such circumstances respondents have to be motivated to supplement information directly available from their accounts by additional information or by estimates. In the case of KAUs and LKAUs this includes splitting up of information and modifying data available for directly observable units.

Some respondents will organise their internal data generating process according Step 2 in such a way that starting from the basic records at least some of the requirements of the statistical system can be met. Other respondents might not be willing to devote resources in such a process and will base their reports on estimates. Usually information on the hypotheses behind these estimates is lacking or not accessible to Statistical Offices. It is by no means guaranteed that the model calculations performed by respondents are consistent estimates even within a single report. Inconsistencies which occur in this step are not caused by inconsistencies in concepts.

3.3 Inconsistencies on the level of Statistical Offices - edited micro data sets (Step 4)

At present in most countries the editing process of incomplete and implausible micro data sets is performed for each domain of business statistics separately. Different methods of coding, imputation and editing in general are applied. Because of this stove-pipe principle information on the micro level which was perhaps incomplete but consistent might become inconsistent. The data will to a certain extent reflect the differences in methods applied in the different domains. Inconsistencies which occur in this step are also not caused by inconsistencies in concepts.

3.4 Inconsistencies on the level of Statistical Offices – meso data (Step 5)

Inconsistencies between the results of different statistical projects on this level result from differences in the underlying statistical units, ties to different versions of the register, missing ties to the statistical business registers, the reference periods chosen, the scope, the breakdown of results by activities and size classes, periodicity and the like.

4. Towards quality of a system of business statistics

Business statistics have to meet quite dissimilar demands by different groups of users. As a consequence the relative importance of the various dimensions of statistical quality also differs by the three main functions of business statistics:

Monitoring

This function is closely related to meeting policy needs, the aim is to control the progress made towards politically defined targets. The focus of monitoring is often on one (or a small set of) variable(s). Accuracy plays a major role and as regards consistency comparability of the results over time and between Member States is considered extremely important.

Basis of analyses

Results of business statistics provide the empirical basis of analyses of all kinds, both on the level of aggregated meso data and on the basis of microdata. All complex analyses have to use various results from business statistics jointly. Consistency across statistical domains is thus more relevant than accuracy of a single variable. Comparability over time is essential for analyses on the basis of time series. Reduced coherence will lead to biased analysis and will pose restrictions to complex modelling exercises.

Empirical foundation of national accounts

For the EU and its Member States, the figures from the European System of national and regional accounts in the European Union ESA play a major role in formulating and monitoring their social and economic policies (ESA 2010, 1.19). For these functions consistency across domains and vertical consistency are of utmost relevance. It is an illusion that consistent ESA figures can be compiled on the basis of inconsistent business statistics.

If the needs of all three user groups are to be met ('relevance' in the terminology of Regulation No 223/ 2009) a fully integrated system of business statistics has to be established.

As a starting point of such a system the hierarchy of objectives (quality criteria) needs to be reconsidered. In the present system (too much) emphasis is laid on accuracy of single variables. The criteria for the evaluation of the quality of structural business statistics as laid down in Regulation No 275/2010 can serve as a good example. This Regulation mentions all quality criteria, but only few quantitative indicators are to be provided by Member States: If sample surveys or a combination of sample surveys and administrative data are used coefficients of variation for the key variables have to be included in the quality reports (Section IV, 1.). If sample surveys, a combination of sample surveys and administrative data or administrative data only are used Member States shall report the weighted degree of unit non-response by industries (Section IV, 2.).

If the focus is on the quality of the entire system of business statistics it is essential to arrive in a first step at consistent concepts and a consistent way to implement the concepts. A harmonisation of definitions would eliminate the inconsistencies already occurring on the level of the single respondent. Many of the inconsistencies occurring in Step 5 could be eliminated if the strategic role of the statistical business registers would be fully acknowledged. All related projects need to be based on the same version of the statistical business registers. Furthermore, the breakdown of results by activities and by size classes should be harmonised.

It is important but not sufficient to arrive at consistent concepts. A high degree of harmonisation of processes is indispensable if the goal is system wide consistency. Measures on two levels of the data generating process should have priority.

Micro level – reporting units

In order to arrive at consistency on the micro level respondents need to be motivated to provide reports that are consistent across different statistical domains. It will become necessary to inform respondents about the logical relationships between different statistical domains for which they have to fill in questionnaires. These relationships should play a prominent role in the explanation of terms and concepts.

Edited micro data

A specific strategy of coordinating the processes mentioned in Step 4 and applied for coding, imputation and editing across domains will be required to arrive at consistent micro data sets.

The plausibility checks carried out by Statistical Offices should never be limited to a single statistical domain. Already on the micro level they have to be performed across all statistical domains. At the moment this is not an easy task because the various surveys are carried out at different dates and have to meet different deadlines.

In particular there is an urgent need for a specific strategy to guarantee consistency on the micro level for big units. A number of countries have already successfully launched projects in this direction. Given the skewness of distribution inconsistencies on the micro level of big and complex units may do more harm to the overall consistency than for example small differences in definitions.

Providing consistent solutions would be much easier both for respondents and for Statistical Offices in an integrated system not split up into (too many) domains. The Memobust handbook could serve as a tool for arriving at a higher degree of standardisation of methods and thus for a higher degree of international comparability.

There is a price to be paid for higher quality seen from the viewpoint of the entire system of business statistics. If the emphasis is put on consistency less flexibility in particular as the fundamentals of business statistics are concerned can be granted to Member States. The output orientation dominating the present system needs to be redefined. The output orientation at present seen for a single statistical domain has to be replaced by an output orientation of a fully integrated system of business statistics, which implies a certain degree of input harmonisation.

In an integrated system of business statistics the quality assessment of statistics will become much more complex than in the present system characterised by the 'stove-pipe' principle. Traditional quality measures (e.g. sampling error) will become less relevant as many results will be based on different sources and not just on the use of sampling techniques. A new quality assessment methodology will therefore have to be developed. In particular it needs to be taken into account to which extent the data editing processes applied in the Statistical Offices pay due attention to consistency across domains.

5. Conclusions

In a system wide view of statistical quality the criterion ‘relevance’ is not just the first one mentioned in Regulation No 223/2009. It should be considered on the top of the hierarchy of criteria. When assessing the quality of business statistics it has to be taken into account that it is the adequacy of the data to be reliably combined what makes them ‘relevant’ for most exercises. The criteria ‘consistency’ and ‘coherence’ are necessary conditions to make the results ‘relevant’ to users who are not just interested in a single variable. Therefore, ‘consistency’ and ‘coherence’ must have a high priority seen from a users’ perspective.

Consistency is dependent on consistent concepts but at the same time also on methods applied by respondents and in the Statistical Offices. The harmonisation of concepts is a necessary step but needs to be supplemented by a certain re-orientation in the data generating processes which are carried out by Statistical Offices.

The obvious trade-off between the ‘best solution’ for a statistical project seen in an isolated way and a ‘consistent and coherent solution’ seen from the perspective of the system of business statistics must be taken seriously. The ‘accuracy’ of an estimate for a specific variable might be (a little bit) higher if the methods of editing can concentrate on this variable without the need to take the relationship to other variables into account. Checks of the consistency relative to other variables take some time; therefore the ‘timeliness’, of the availability of information of a single variable might also somewhat be reduced in the case of an integrated system. No major implications must be expected for ‘punctuality’ and ‘accessibility’ and ‘clarity’.

The present way of quality assessment is adequate to the function of ‘monitoring’ and meeting the needs of users interested in a single aspect of economic reality. If the functions ‘basis of analyses’ and ‘empirical foundation of national accounts’ are also considered important ‘relevance’ will only be given in the case of a consistent system of business statistics. Only such an integrated system will offer the degree of flexibility to meet “potential needs of the users” as requested by Regulation No 223/2009 in defining ‘relevance’.

If phenomena that cover multiple dimensions are to be analysed a fully integrated system of business statistics such as it is envisaged by the FRIBS initiative has to be established. For such a system a clear priority has to be put on the criteria ‘relevance’, ‘consistency’ and ‘coherence’ thus stressing the system aspect of statistical quality.

6. References

Legal acts

Communication from the Commission to the European Parliament and the Council on the production method of EU statistics: a vision for the next decade; COM(2009) 404 final; referred to as “Vision paper”

Commission Regulation (EU) No 275/2010 of 30 March 2010 implementing Regulation (EC) No 295/2008 of the European Parliament and of the Council, as regards the criteria for the evaluation of the quality of structural business statistics

Decision No 1297/2008/EC of the European Parliament and of the Council of 16 December 2008 on a Programme for the Modernisation of European Enterprise and Trade Statistics (MEETS)

Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics and repealing Regulation (EC, Euratom) No 1101/2008 of the European Parliament and of the Council on the transmission of data subject to statistical confidentiality to the Statistical Office of the European Communities, Council Regulation (EC) No 322/97 on Community Statistics, and Council Decision 89/382/EEC, Euratom establishing a Committee on the Statistical Programmes of the European Communities

Regulation (EU) No 549/2013 of the European Parliament and of the Council of 21 May 2013 on the European system of national and regional accounts in the European Union; referred to as ESA 2010

Other publications

Blackburn K. (1996), After frameworks; what then? Paper to the 24th General IARIW Conference, Lillehammer 1996

Deliverables and documents of the ESSnet on consistency of concepts and methods of European business and trade-related statistics, <http://www.cros-portal.eu/content/consistency-0>

Eurostat (2010), External Study on Detail Evaluation of the Legal Acts in the Areas of Statistics which were Identified by Member States as Areas to Revision, Luxembourg 2010, <http://www.cros-portal.eu/content/external-study-detail-evaluation-legal-acts>

Kloek W. (2011), What Makes Business Statistics Different? Paper to the European Establishment Statistics Workshop, Neuchatel 2011

Memobust handbook (2014), The Memobust Handbook on Methodology for Modern Business Statistics, www.cros-portal.eu/content/memobust