

**STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

**COMMISSION OF THE
EUROPEAN COMMUNITIES**

CONFERENCE OF EUROPEAN STATISTICIANS

EUROSTAT

**Joint UNECE/Eurostat Work Session
on Statistical Metadata**
(6 - 8 March 2002, Luxembourg)

Working Paper No. 20
English only

Topic (iii): Metadata and quality

**ASSESSMENT OF DATA QUALITY FOR COMPARISONS ACROSS COUNTRIES:
EUROSTAT'S EXPERIENCES AND THE LEADERSHIP GROUP ON QUALITY (LEG)
RECOMMENDATIONS**

Submitted by Eurostat¹

Invited paper

Keywords: Quality reports, metadata quality, quality indicators and international comparability.

ABSTRACT

In 1994, Eurostat launched its approach to quality measurement based on the experience of (statistical) organisations around the world and with the explicit support of several national statistical offices of the European Union.

Much of the methodological work for measuring and reporting on the quality has been discussed and agreed upon with the Member States of the European Union. The main forum for this has been the Working Group on the Assessment of Quality in Statistics. In addition to general quality related methodological work Eurostat has set up ad-hoc task forces with the Member States and has started laying down quality provisions in legal acts. Special task forces have been created for the measurement of quality in National Accounts, Balance of Payments statistics, Labour Force Survey, Foreign Trade statistics and variance estimations.

The Leadership Expert Group on Quality (LEG)² was set-up in 1999 and the main purpose was to attain improved quality in the European Statistical System (ESS). The LEG provides a number of recommendations for the ESS regarding its quality work, and the final work of the LEG was approved by the Statistical Programme Committee (SPC) meeting in September this year. Some of the recommendations are highly linked to the Eurostat's activities on assessment of data quality. For example, the recommendations related to reporting of product quality, documentation, and dissemination of information.

The first part of the paper highlights the recommendations of the LEG that have bearing on quality reporting and metadata quality.

The second part of the paper gives an overview of Eurostat's experiences concerning quality reporting. Such work is going on for most of the areas producing statistics in Eurostat and the first experiences are now available. Special focus is on international comparability requirements for quality information that

¹ Prepared by Hakan Linden and Harald Sonnberger.

² The LEG was chaired by Statistics Sweden and the other members were Eurostat, INSEE (France), Statistics Netherlands, ONS (UK), Statistisches Bundesamt (Germany), INE (Portugal), ISTAT (Italy) and the National Statistical Service of Greece.

has to be documented and proposal for quality indicators that can be used for summarising information about the quality of the statistics produced.

1 INTRODUCTION

1. Many activities related to the assessment of data quality across countries have been undertaken in the European Statistical System during the last years.

2. In 1994, Eurostat launched its approach to quality measurement based on the experience of (statistical) organisations around the world and with the explicit support of several national statistical offices of the European Union.

3. Since then much of the methodological work for measuring and reporting on the quality has been discussed and agreed upon with the Member States of the European Union. The main forum for this has been the Working Group on the Assessment of Quality in Statistics. The results are summarised in documents such as Definition of Quality in Statistics [Eurostat, 2000a], Standard Quality Report [Eurostat, 2000b], or Glossary on Quality in Statistics [Eurostat, 2000c].

4. In addition to general quality related methodological work Eurostat has set up ad-hoc task forces with the Member States and has started laying down quality provisions in legal acts. Special task forces have been created for the measurement of quality in National Accounts, Balance of Payments statistics, Labour Force Survey, Foreign Trade statistics and variance estimations.

5. In September 1999 the Management Board of Eurostat decided that Internal Quality Reports (IQR) should be produced for all data available in Eurostat ideally during the year 2000. Exceptions were however accepted for aggregate data like National Accounts.

6. In parallel, a Leadership Expert Group on Quality (LEG) was set-up in 1999 and the main purpose was to attain improved quality in the European Statistical System (ESS). The LEG provides a number of recommendations for the ESS regarding its quality work, and the final work of the LEG [Lyberg et al., 2001] was approved by the Statistical Programme Committee (SPC) meeting in September 2001. Some of the recommendations are highly linked to the Eurostat's activities on assessment of data quality.

7. The first part of the paper highlights the recommendations of the LEG that have bearing on quality reporting and metadata quality.

8. The second part of the paper gives an overview of Eurostat's experiences concerning quality reporting. Special focus is on international comparability requirements for quality information that has to be documented and proposal for quality indicators that can be used for summarising information about the quality of the statistics produced.

II. THE LEG ON QUALITY RECOMMENDATIONS

9. The LEG on Quality has provided a number of recommendations related to total quality management in the ESS. In this chapter the focus is on a few recommendations directly linked to product quality and metadata information.

II.1 Product quality

10. Quality consists of a number of features reflecting user needs. In this setting, quality can be defined along a number of dimensions. All these dimensions constitute the product quality. The Eurostat quality vector has the following seven main components: relevance of statistical concept, accuracy of

estimates, timeliness and punctuality in disseminating results, accessibility and clarity of information, comparability, coherence and completeness.

11. The LEG on Quality recommends that each National Statistical Institute (NSI) reports product quality according to these quality dimensions and sub- dimensions.

12. However, the measurement of some of the quality dimensions or components can be very difficult in many respects. The Working Group on “Assessment of Quality in Statistics” has tackled many problems related to the measurability of the quality dimensions but there is currently a lack of adequate measures (e.g. coherence) while for other components (e.g. accuracy) measures do exist but are difficult to calculate on a continuing basis.

13. Therefore, the LEG on Quality also recommends that the measurability of each ESS quality dimension and sub- dimension should be improved.

II.2 Process quality

14. In order to be able to understand and improve the product quality the producer of statistics also need information about the underlying processes. It is unlikely that the product will have good quality if the underlying process is not up to par.

15. The process quality is improved by identifying key process variables (i.e. those variables with the greatest effect on product quality), measuring these variables, adjusting the process based on the measurements, and checking what happens to product quality. If improvements do not materialise, alternative adjustments are made or new key variables are identified and measured.

16. The LEG on Quality recommends that a handbook on the identification of key process variables, their measurement, and measurement analysis should be developed.

II.3 Documentation

17. The main purposes of documentation are to ensure and improve quality, and to facilitate the understanding and the use of data. The users are, in general, particularly interested in the information content of the statistical product. The producer also needs information about the production processes (involving all steps of the activity from the planning phase to the data dissemination phase).

18. With regard to the statistical product quality and the process quality the LEG writes that extensive documentation is required to satisfy the different levels of information needs since there are different kinds of users and even producers. The actual documentation should consist of metadata on the production process and the information content, quality measures and indicators concerning the product, and data on the producing organisation’s strategies, policies and user relationships.

19. The LEG on Quality recommends that the ESS members should analyse their documentation status in a report. The report should include an action plan with clear priorities for improvement and a timetable.

III. QUALITY REPORTING IN EUROSTAT

20. The basis for Eurostat's quality reports is a “standard quality report form” that follows the structure of the seven main components of the Eurostat quality concept. The form is very general and has to be tailor-made to the need of each statistical area.

21. These forms are to be filled in by the Eurostat units concerned and, even if some of the information asked for refer to user assessments, it is so far, a producer oriented way of assessing the

quality of statistics. The information in the quality reports should (and will) be supplemented by information from users such as results of user satisfaction surveys/assessments, user comments on databases and publications, users feedback in the press etc.

22. At the time of writing this paper, quality reports are available or preparatory work towards quality reports have begun for some 30 statistical areas in Eurostat. The general experiences so far are:

- ?? The current collection of information has to concentrate more on the availability of information rather than on the collection of the information itself.
- ?? Most of the reports follow the structure of Eurostat's definition of quality.
- ?? The information asked for is easy to collect for some of the quality components or sub-components, while others are more complicated to evaluate (mainly accuracy, comparability, and coherence). The evaluation of some of the errors requires a lot of resources and can therefore only be assessed on an intermittent basis and in co-operation with the Member States.
- ?? It is well known that links exist between the different components. Timeliness and accuracy are not independent; timely figures could be heavily revised. Revisions might influence accuracy, comparability, and coherence of statistics. Therefore, quality has to be judged in a global way taking all the components into account at the same time.

IV. SOME SPECIAL ASPECTS RELATED TO INTERNATIONAL COMPARABILITY

23. From the point of view of users of international statistics as produced by Eurostat, the geographical comparability is one of the most important aspects of quality as the statistics are based on - not necessarily fully harmonised - data from different countries. However, the assessment of comparability is still in its infancy, and a lot of further methodological research is needed for building up practical experiences as well as the theoretical framework for assessing comparability. Some problems and reflections directly linked to the assessment of comparability and its effects on accuracy and coherence of international statistics are given below:

IV.1 Coefficients of variation

24. In the Internal Quality Report (IQR) the coefficient of variation is asked for one important variable from each Member State that provides data collected by sample surveys. The experiences from the available quality reports are that firstly, CV's can be provided in most cases by the Member States, secondly CV's vary quite a lot between countries and, thirdly, little is known about the error components taken into account in the calculation of the CV's.

25. This example shows that detailed specifications need to be agreed with the Member States, for instance on the components to take into account when making aggregations/ estimations.

IV.2 Mirror statistics

26. Mirror statistics can be used for the evaluation of the quality of flow statistics. In particular the accuracy of variables, the comparability between countries and the coherence between statistics.

27. Several examples of the use of mirror statistics can be found in the quality reports. An example is Foreign Trade statistics where mirror exercises are regularly made in order to provide Eurostat with an estimate of the margin of error related to intra-EU flows (arrivals and dispatches). Another example is Tourism statistics where the number of inbound tourism nights according to supply statistics in a Member

State can be compared to the number of outbound tourism nights in the Member State according to demand statistics from other Member States.

IV.3 Complex Statistics

28. For complex statistics, i.e. statistics that are based on different types of data sources, it is not always possible (or meaningful) to calculate “objective” error margins. The experience from the quality reports shows that the evaluation of the quality of such statistics should focus more on the compilation process for the production of the statistics. An example is given in [Eurostat, 2001i] for National Accounts (NA), where three layers are separated: the basic figures, adjusted figures, and final NA estimates. The first layer aims at describing the different sources used, the second layer measures the various adjustments made due to differences in concepts and so on and their relative importance, and the third layer assesses the balancing adjustments of the final NA estimates.

IV.4 Time series and seasonal adjustment

29. Data in time-series are often presented in two different ways: original and seasonally adjusted data. As the methodology of seasonal adjustment is far from being harmonised at international level detailed methodological information is needed for international comparisons. An example can be picked from the quality report for Short Term Business Statistics [Eurostat 2001f] about the need for information for example on the methods applied or the software packages used.

30. This is an example of metadata that should be provided together with the data from the Member States, as well as of the needs for common rules in the area of seasonal adjustment.

IV.5 Revisions

31. Available IQRs show that first, final and sometimes even revised estimates are produced in many statistical areas depending on deadlines for publishing results or lack of basic data. Examples come from Short Term Business statistics where EU15/Eurozone figures are calculated by Eurostat if at least 60% of the information is available, from Infra Annual statistics where several updates are usual, or from Intra-Community Trade statistics. A study on the updating process of these trade data over the period 1994-1997 [Eurostat 2001a] shows that revisions can represent more than 7% of the definitive figures.

32. Information about the impact of the revisions on the estimates should always be reported.

V. FIRST PROPOSALS FOR QUALITY INDICATORS

33. Since the first IQRs are available they can now be used as an input to define a set of indicators for describing the data quality and its evolution, at an aggregated level, for the statistics produced by Eurostat. The objective is to identify (annual) indicators to measure and to follow over time the quality of the data and the quality of the information provided in the reports.

34. Different internal uses of such indicators have been identified:

- ?? Eurostat's management: to observe and analyse the quality of the statistics produced.
- ?? Product managers and the quality co-ordinators: to get quality related information, to monitor the production processes and to launch improvement actions.
- ?? Input to the identification of further needs for methods and tools for quality measurement.
- ?? Comparison of the different products with respect to the seven quality components.

?? Input to product documentation, e.g. metadata.

V.1 Performance and monitoring indicators

35. Two types of indicators can be identified: firstly, performance indicators trying to measure the level of the quality dimension concerned, and secondly monitoring indicators which are management indicators referring to the quality of the quality reports. Quite high numbers of indicators need to be developed looking at the range of aspects included in some of the seven quality dimensions. A list of such quality indicators can be found in annex I. It is worth noting that performance indicators are currently missing for relevance, accessibility and clarity, and coherence. Further investigations are currently undertaken to close these gaps.

V.2 Summary indicators

36. The previous sections have looked at the requirements for quality information from a producer perspective. These requirements imply quite a high number of indicators. Such information is not always of interest for the users of statistics. Without going into detailed descriptions of different users needs for quality information we can state that there are, internal and external, needs for indicators summarising the quality of the statistics produced. Such summarising indicators have to be based on the indicators described in section 5.1. However, it is not always possible from a theoretical point of view, or meaningful from a practical point of view, to define such overall quality indicators.

37. Therefore, these indicators have to be based, to some extent, on subjective assessments taking into account information from different quality aspects accompanied by descriptive information. A list of such more user oriented quality indicators is given in annex II.

VI. CONCLUSIONS

38. From a conceptual viewpoint, a common agreement has been achieved on how to define quality in statistics. Despite some minor remaining controversies, it is now generally agreed, in Eurostat and even within the ESS, that quality in statistics can be reported according to the seven main quality components. Additionally, useful models of quality reports and operational documentation have been prepared.

39. The use of Eurostat's Internal Quality Reports has proven successful though sometimes burdensome. First experiences show that useful information can be reported in all seven components of quality. These reports are also a useful data source for quality indicators enabling a description of the individual statistics for quite different uses.

40. Eurostat's Management Board has recently acknowledged the basic idea of developing a set of indicators covering all dimensions of the underlying quality definition through which the producer and user is able to get a comprehensive picture of the quality of the data presented.

41. The Eurostat Unit for Research and Development is undertaking methodological work in order to fill the existing gaps of indicators. As soon as the lists are complete they will be tested on the basis of existing quality reports. Once the methodology is established, it is proposed to use the indicators for all tables produced by Eurostat.

42. An Implementation Group is currently being set-up with the task of collecting information and co-ordinating the recommendation activities from the LEG on Quality. The Implementation Group can be seen as a Quality Advisory Group to the SPC and will be chaired by Eurostat. The Group will lead and co-ordinate the work, and it shall also cooperate with the Working Group on the "Assessment of Quality in Statistics".

REFERENCES

- Bocci, L., Brancato, G., and Signore, M. (2000), The SIDI system: a generalised tool for the quality control of ISTAT surveys, Italy.
- Carson, C. S. (2001), Toward a Framework for Assessing Data Quality, IMF Working Paper, Statistics Department, International Monetary Fund, February 2001.
- Grünwald, W., and Linden, H. (2001), Quality measurement- Eurostat experiences. Paper presented at Statistics Canada Symposium 2001, October 17-19, 2001.
- Eurostat (1998), ECHP Data Quality, December 1998.
- Eurostat (2000a), Definition of Quality in Statistics, Document n° Eurostat/A4/ Quality/00/General/ Definition.
- Eurostat (2000b), Standard Quality Report, Document n° Eurostat/A4/Quality/ 00/Standard/Report.
- Eurostat (2000c), Glossary on Quality in Statistics, Document n° Eurostat/A4/Quality/ 00/Standard/Glossary.
- Eurostat (2000c), Report from the Commission to the Council On Harmonization of Consumer Price Indices in the European Union. COM(2000) 742 final, Brussels, 2000-11-21.
- Eurostat (2001a), Foreign Trade Statistics Quality Report, Eurostat C4, February 2001.
- Eurostat (2001b), Farm Structure Survey 2000. National Methodological Report, 2001-05-17.
- Eurostat (2001c), FSS Quality Assessment System. Annex 1: Accuracy Indicators, Document n° ASA/SB/580- Annex 1, 2001-07-10.
- Eurostat (2001d), FSS Quality Assessment System: A proposal. Document n° ASA/SB/580, 2001-06-19.
- Eurostat (2001e), Implementation of a standard quality report in Structural Business Statistics, 2001-03-26.
- Eurostat (2001f), Quality Report for Short-term Statistics. Draft version 1, 2001-06-15.
- Eurostat (2001g), Quality Reporting Labour Force Survey, Task Force on quality report LFS, 2001-07-03.
- Eurostat (2001h), Quality Reporting within Eurostat, Draft 2001-08-02.
- Eurostat (2001i), Task Force on Accuracy Assessment of National Accounts Statistics. Final Report, June 2001.
- João Zilhão, M. (2001), Quality Indicators Panel, INE-PT, Document n° Eurostat/ A4/Quality/01/Quality Indicators Panel.
- Ladiray, D. (2001), More Refined Quality Indicators for Euro-SICS, Eurostat, Draft 2001-01-18.
- Ladiray, D., Nugnes, F., and Picciolo, M. (2001), Assessing the quality of the Euro-Sics database, Eurostat, 2001-04-13.
- Linden, H. (2001), Some Comments on the Assessment of the Data Quality Reports at Eurostat, Draft 2001-08-20.
- Lyberg, L. et al. (2001), Summary Report from the Leadership Group (LEG) on Quality, Eurostat and Statistics Sweden, 31 July 2001.
- Petterson, H. (2001), Quality Indicators at Statistics Sweden, Statistics Sweden, 2001-05-27.
- Viseur, J. (2001), Quality Reporting within Eurostat, Draft, Eurostat, 2001-08-02.

Annex I. Current list of (producer oriented) quality indicators (December 2001)

| | Indicator from the IQRs | Quality component | Type of indicator | Comment |
|----|--|-------------------|----------------------------|--|
| 1 | Number and proportion of statistical products measuring user satisfaction | Relevance | Monitoring | In a first phase user satisfaction can be gathered from other sources like rolling reviews and various pieces of information from the datashop network |
| 2 | Number and proportion of sample survey based products that calculate CV | Accuracy | Monitoring | |
| 3 | CV for one important variable for statistics based on sample surveys. | Accuracy | Performance | In a first phase only the sampling variability is requested to be taken into account. |
| 4 | Number and proportion of statistical products that provide information about unit non-response | Accuracy | Monitoring | |
| 5 | The unit non-response rate | Accuracy | Performance | Unweighted and weighted |
| 6 | The item non-response rate for one important variable | Accuracy | Performance | Item non-response before imputation |
| 7 | Number and type of different sources that the statistics are based on. | Accuracy | Performance/ Monitoring | A "simple" way to measure the sources of errors that might occur depending on the complexity of the statistics produced. |
| 8 | The difference between first results and revised (final) results. | Accuracy | Performance | To measure the reliability of the statistics produced. |
| 9 | Number and proportion of statistical products that provide information about other non-sampling errors | Accuracy | Monitoring | |
| 10 | Production lead times in weeks between first and final results | Timeliness | Performance | |
| 11 | Average time between the end of the reference period and the date of first publication | Timeliness | Performance | |
| 12 | Proportion of publications released in time | Punctuality | Performance | |

| | | | | |
|----|---|---|-------------|---|
| 13 | Number and type of means used for disseminating statistics | Accessibility | Monitoring | The following classes could be used: Reference databases, Statistical Yearbooks, Monthly Bulletins, Newsletters, Statistics in Focus, Press releases, Other thematic publications, CD Rom and floppy discs. |
| 14 | Proportion of statistical products that publish metadata or methodological references | Clarity | Monitoring | |
| 15 | Number and proportion of the statistical products, that have indicated changes in concepts and/or measurement over time, and provide quantitative estimates of the effects | Comparability (over time) | Monitoring | |
| 16 | Number and proportion of the statistical products, that have indicated differences in concepts and/or measurement from the European norms, that provide quantitative estimates of the effects | Comparability (geographical, and between domains) | Monitoring | |
| 17 | Net and gross differences for the comparability over time | Comparability (over time) | Performance | The non-comparability due to deviations from the European concepts. The gross difference is the absolute value of all deviations when making the comparisons. |
| 18 | Net and gross differences for the geographical comparability | Comparability (geographical) | Performance | |
| 19 | Net and gross differences for the comparability between domains | Comparability between domains | Performance | |
| 20 | Number and proportion of statistical products that provide descriptive or quantitative information about the coherence with statistics from other domains | Coherence | Monitoring | |
| 21 | The rate of available statistics | Completeness | Performance | Number of statistics provided divided by the number of statistics requested |

**Annex II. Current list of (user oriented) quality indicators
(December 2001)**

| | Summary indicator | Quality component | Type of indicator | Reference to annex 1 |
|---|---|---------------------------------|--------------------------|---|
| 1 | <i>Qualitative assessment of the overall accuracy</i> | Accuracy | Performance | Given that information is provided for indicators 2 to 9 of annex I an overall qualitative judgement of the accuracy should be provided |
| 2 | <i>a) Average production lead times between first and final results b) Average time between the end of the reference period and the date of first publication</i> | Timeliness | Performance | Identical to indicators n° 10 and 11 of annex I |
| 3 | <i>Proportion of publications released in time</i> | Punctuality | Performance | Identical to indicator n° 12 of annex I |
| 4 | <i>Qualitative assessment of the comparability over time</i> | Comparability (over time) | Performance | Identical to indicator n° 17 of annex I |
| 5 | <i>Qualitative assessment of the geographical comparability</i> | Comparability (geographical) | Performance | Identical to indicator n° 18 of annex I |
| 6 | <i>Qualitative assessment of the comparability between domains (other than geographical)</i> | Comparability (between domains) | Performance | Identical to indicator n° 19 of annex I |
| 7 | <i>Rate of available statistics</i> | Completeness | Performance | Identical to indicator n° 21 of annex I |