

## **QUALITY MEASUREMENT- EUROSTAT EXPERIENCES**

Werner Grünewald and Håkan Linden<sup>1</sup>

### **ABSTRACT**

The first round on quality reporting of the statistics produced in Eurostat has almost been completed. This paper presents the experiences so far and in particular some of the methodological problems encountered when measuring the quality of the statistics that are produced for international comparisons. A proposal is also given for indicators that summarise the quite detailed information provided in these quality reports. Two sets of indicator are discussed. The first set is more producer oriented and the second set is more user oriented.

KEY WORDS: Quality reports, data quality, international comparability, and quality indicators.

### **1. INTRODUCTION**

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. Today, Eurostat has a basic structure for the management of quality in statistics and a single reference framework for the definition of quality in statistics and for the reporting on quality ("quality reports"). In September 1999 the Management Board of Eurostat decided that such quality reports should be produced for all data available in Eurostat ideally during the year 2000. Exceptions were however accepted for aggregate data like National Accounts.

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; relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability over time and space, coherence, and completeness.

The activities on how to measure and report on the data quality have begun in most of the statistical domains in Eurostat, and first versions of Internal Quality Reports (IQR) are available for a number of statistical products.

This paper summaries the first experiences on how to report on quality of statistics (in a supranational environment), and how to summarise the information that can be used for analysis of the quality of different steps in the production processes and for assessing the quality of the final statistics produced.

### **2. WORKING ON QUALITY WITH THE MEMBER STATES**

Much of the methodological work for the reporting on quality has been discussed in detail and agreed upon with the Member States of the European Union. The results are summarised in documents such as

---

<sup>1</sup> Werner Grünewald, Håkan Linden, European Commission, L-2920 Luxembourg  
email: [werner.grunewald@cec.eu.int](mailto:werner.grunewald@cec.eu.int), and [hakan.linden@cec.eu.int](mailto:hakan.linden@cec.eu.int)

Definition of Quality in Statistics [Eurostat, 2000a], Standard Quality Report [Eurostat, 2000b], or Glossary on Quality in Statistics [Eurostat, 2000c].<sup>2</sup>

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.

There is no general legal act on quality in Statistics in the European Union at the moment. Nevertheless, quality issues have been progressively introduced in specific legal acts during the recent years, and it is interesting to observe how quality provisions have steadily expanded their scope over time. In the 1998 legal act on the Labour Force Survey<sup>3</sup>, the focus is only on the representativeness of the sample; the word “quality” is not even used. The act simply requires that Member States shall ensure that sampling errors, measured in terms of coefficients of variation (CV), are kept below a certain level, that appropriate imputation is made for missing values and that weighting factors are correctly calculated. The 1999 legal act on Structural Business Statistics<sup>4</sup> is the first one that includes a precise, although limited, requirement for quality indicators. It partially addresses the accuracy component asking for sampling errors with the calculation of CV and for non-response errors with the calculation of non-response rates. The 2000 legal act for Labour Costs Statistics<sup>5</sup> is the first one that contains the complete and precise quality provision. It reviews each of the seven quality components and recommends the collection of appropriate indicators for each of these components. For more information see [Viseur, 2001].

### **3. THE COMPILATION OF THE INTERNAL QUALITY REPORTS (IQR) IN EUROSTAT**

The internal quality report of Eurostat is a questionnaire type summary form of relevant aspects of all different quality components as fixed in Eurostat's quality definition (for details see [Eurostat, 2000b]). The form is very general and has to be tailor-made to the need of each statistical area.

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.

#### **3.1 The process**

Each producing unit in Eurostat is responsible for the preparation of its quality report. Eurostat's unit for Research and Development in Statistics (A4) supports them, amongst others through the production of an electronic Quality Report Form, and documents on How to Fill in a Quality Report, the Production Process of a Quality Report, and Frequently asked Questions. Training is also provided whenever asked for.

It was clearly stated from Eurostat's management board that the “first round” of internal quality reports should only request information already available inside Eurostat, in order to limit the burden on the units and the Member States. However, several units preferred to address the quality reporting in an interactive way with the Member States.

---

<sup>2</sup> These documents are available on the internet under <http://forum.europa.eu.int/Public/irc/dsis/qis/home>.

<sup>3</sup> Council Regulation (EC) N° 577/98

<sup>4</sup> Commission Regulation (EC) N° 1618/1999

<sup>5</sup> Commission Regulation (EC) N° 452/2000

### **3.2 Some general comments on the reporting**

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 on a global way taking all the components into account at the same time.

## **4. THE EXPERIENCES FROM THE FIRST ROUND OF INTERNAL QUALITY REPORTS (IQR)**

The following general problems have been identified for the individual quality components:

### **4.1 Relevance**

This quality component addresses the aspects related to the description of the users, description of their needs and to which extent the users needs are fulfilled. It requires the identification and the classification of the users, and the measurement of their satisfaction for example through user satisfaction surveys. Such requirements are very costly and time-consuming and therefore very few units, if any, can afford such an operation at the moment. Instead, relevance in most of the IQR are measured through feedback from the most important users, i.e. often other services of the European Commission or the European Central Bank.

However, the methodology for carrying out user satisfaction surveys should certainly be developed in order to capitalise on some useful common experience and know-how, and prepare for general investments, such as standard questionnaires and sample designs.

### **4.2 Accuracy**

This is probably the most complicated quality component to assess. Many sub-components have to be addressed and the evaluation of the various types of error often requires a lot of human and financial resources. So far, Eurostat units have only covered some sub-components in their IQR, most commonly the sampling errors. Hence, some IQR provide variances or even CV, e.g. the IQR on structural business statistics.

With respect to non-sampling errors, non-response is the most commonly addressed sub-component, as it can often be simply obtained as a by-product of the statistical production process. Other non-sampling errors, i.e. coverage error, measurement error, process error, model error and therefore the total error are generally not given in the available IQR.

Attempts to measure some non-sampling errors, particularly the measurement errors, should be encouraged as, in some situations, this type of error could be far more significant than sampling errors.

### **4.3 Timeliness and punctuality**

These aspects of quality mainly refer to pre-established and factual reference periods of data and publication dates. With the evolving practice of providing users with a timetable for the release of statistics, punctuality has become easy to observe and is usually measured. Timeliness requires more attention but is still a fairly simple component. Information on both is often available in the IQR.

### **4.4 Accessibility and clarity**

These aspects of quality refer to how accessible the statistical data are, if they are available in the form the users desire and if they are well documented. These aspects are assessed from a producer perspective in the IQR. Even if built upon objective criteria such as size of publication, existence of an information service etc. the measurement of this component is not fully satisfactory in the IQR. Particularly, clarity indicators and the methodology for their assessment, i.e. the quality of the associated metadata, should be further elaborated.

### **4.5 Comparability**

The comparability component refers to comparability over time, geographical comparability and comparability between domains (other than geographical). The experiences from the IQR shows that these aspects of quality, together with the accuracy component, are the most difficult to assess. Normally, descriptive information is given about the differences in concepts or changes over time, but quantitative information about the effects of the “lack of comparability” is rare.

Further methodological research is needed for the quantitative assessment of comparability.

### **4.6 Coherence**

For many of the statistical products it is relevant to compare their data with the results of other domains or sources. These domains/sources are specified in most of the quality reports. However, due to the considerable resources and time needed for assessing the coherence it is rare that checks of the coherence are done. With the future development of Eurostat's quality programme and the gradual existence of quality reports it is expected that such checks will be easier in the future.

### **4.7 Completeness**

In the ideal case, completeness should be assessed by the users of statistics. In the quality report this component is a bit superficially addressed. The information asked for is limited to the number of statistics requested and the number of statistics provided. However, this information is available in most of the IQR.

## **5. SOME SPECIAL ASPECTS RELATED TO INTERNATIONAL COMPARABILITY**

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.

## **5.1 Coefficients of variation**

In the 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.

## **5.2 Mirror statistics**

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.

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.

## **5.3 Complex Statistics**

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.

## **5.4 Time series and seasonal adjustment**

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.

This is an example of meta-data 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.

## **5.5 Revisions**

Available IQR 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.

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

## **6. FIRST PROPOSALS FOR QUALITY INDICATORS**

Since the first IQR 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.

### **6.1 Internal uses of the quality indicators**

Different Eurostat 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. in meta-data.

### **6.2 Performance and monitoring indicators**

Two types of indicators can be identified: firstly, performance indicators<sup>6</sup> 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 necessary to close these gaps.

### **6.3 Summary indicators**

The previous two 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 6.2. 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. 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.

---

<sup>6</sup> In order to have some input for defining performance indicators the experience of a few countries has been used as benchmark reference [Linden, 2001].

## 7. CONCLUSIONS

Work on quality issues in statistics has considerably evolved in Eurostat over the last decade. It began slowly in the early 1990s and has seen a growing interest in recent years. Several concrete initiatives have now paved the way for further progress.

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 so far identified<sup>7</sup>. Additionally, useful models of quality reports and operational documentation have been prepared.

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. Though the list of indicators is far from being complete it shows the direction in which the development of such indicators should go.

Eurostat's Internal Quality Reports are therefore a useful tool to improve the quality of European data. They are one important element for the implementation of the principle on product quality commitment of the "Quality Declaration of the European Statistical System" approved and signed by the heads of the National Statistical Institutes at their meeting in Luxembourg in September 2001. Other principles concern the focus on users, continuous improvement, accessibility of information, partnership within and beyond the European Statistical System, respect for needs of data suppliers, commitment of leadership, systematic quality management, effective and efficient processes and staff satisfaction and development. All these principles will finally help the European Statistical System to fulfil its mission, i.e.

to provide the European Union and the world with high quality information on the economy and society at the European, national and regional levels and to make the information available to everyone for decision-making purposes, research and debate.

---

<sup>7</sup> Some other agencies use a definition in 5 or 6 components, but it is often a matter of breakdown and wording rather than divergences in contents. This, however, does not ease the situation....

**ANNEX I CURRENT LIST OF (PRODUCER ORIENTED)  
QUALITY INDICATORS (DECEMBER 2001)**

	<b>Indicator from the IQRs</b>	<b>Quality component</b>	<b>Type of indicator</b>	<b>Comment</b>
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

### 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.
- 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.

Pettersson, H. (2001), Quality Indicators at Statistics Sweden, Statistics Sweden, 2001-05-27.

Viseur, J. (2001), Quality Reporting within Eurostat, Draft, Eurostat, 2001-08-02.