Data quality and Indicators

Tübingen, March 2005
August Götzfried
Eurostat - Unit B5

Background

Eurostat works since more than 10 years on quality in statistics; the following results were achieved:

- The definition of quality
- Standard quality reports (plus handbook)
- Glossary of terms in quality
- Standard quality indicators
- Domain specific Regulations requiring quality reports

Organisational structure at Eurostat

The Eurostat Working Group on the Assessment of Quality in Statistics

- focuses on technical aspects
- co-ordinates the implementation of the quality reports in the various statistical domains

In addition, quality is treated in sectoral working parties such as the National Accounts Working Party, Structural Business Statistics, etc. related to the domains in question

The main criteria of quality

- **Relevance**..for users incl. c**ompleteness** (i.e. availability of information)
- **Accuracy:** closeness between the estimated value and the true (unknown) population value with the sampling and non sampling errors (e.g. non-response errors)
- **Timeliness and punctuality:** the punctuality of the statistics produced
- **Accessibility and clarity**: availability of the statistical data in the form users desire it

The main criteria of quality

- **Comparability:** over time, between geographical areas, between domains
- **Coherence**: between annual and infra-annual statistics, between provisional and final statistics, between social and business statistics, etc.
- In addition: the quality reports can also contain information on **costs** and **burden on respondents**

These criteria are normally treated in the quality reports.

The standard quality report

" describes the subjects that should be covered in a report on quality and list particular information that Eurostat expects Member States to provide"

The main reference documents are

- "How to make a quality report"
- "Glossary on quality terms"
- "Definition of Quality in Statistics"

These documents were prepared and are continuously revised by the Working Group on Assessment of Quality in Statistics. All documents available on CIRCA (dissemination on the web planned).

Quality indicators- producer oriented

Quality component	Indicator	1=Key 2=Supportive 3=Further development
Relevance	R1. User satisfaction index	3
	R2. Rate of available statistics	1
Accuracy	A1. Coefficient of variation	1
	A2. Unit response rate (un-weighted/weighted)	2
	A3. Item response rate (un-weighted/weighted)	2
	A4. Imputation rate and ratio	2
	A5. Over-coverage and misclassification rates	2
	A6. Geographical under-coverage ratio	1
	A7. Average size of revisions	1
Timeliness and Punctuality	T1. Punctuality of time schedule of effective publication	1
	T2. Time lag between the end of reference period and the date of first results	1
	T3. Time lag between the end of reference period and the date of the final results	1
Accessibility and clarity	AC1. Number of publications disseminated and/ or sold	1
	AC2. Number of accesses to databases	1
	AC3. Rate of released statistics with complete metadata information	3
Comparability	C1. Length of comparable time-series	1
	C2. Number of comparable time-series	1
	C3. Rate of differences in concepts and measurement from European norms	3
	C4. Asymmetries for statistics mirror flows	1
Coherence	CH1. Rate of statistics that satisfies the requirements for the main secondary use	3

3 Examples of sectoral quality approaches

- Quality in Structural Business Statistics
- Quality in the European Union Labour Force Survey
- Quality in the Community Innovation Survey

Quality in Structural Business Statistics

- Commission Regulation No 1618/1999 concerning the criteria for the evaluation of quality of structural business statistics adopted in 1999
- Quality reports and indicators defined for the main statistical results: overall co-efficient of variation, unit non-response rate and item non-response rate to be delivered to Eurostat for a number of main variables and breakdowns (e. g. by size class)
- Quality checks performed upon data reception in Eurostat
- Quality action plan developed and approved aiming to:
  - Closer follow-up of availability, timeliness, coverage and accuracy
  - Improved diagnosis: global quality report, SBS quality barometer
  - Better service to users: dissemination calendar, improved metadata
  - Specific actions on comparability and consistency

Quality in the European Union Labour Force Survey

- Every 3 year a report to the Parliament on the execution of the survey needed, including quality issues. (COM (2003) 760(01))
- Annual quality reports on voluntary basis since 2001, using the adapted standard reporting
- Quarterly quality reports on accuracy, non-response and major revisions since 2004
- Summary report for 2003 is being produced.
- Extensive documentation and meta-data available to users
- Constant review of methods/definitions/issues of comparability:
  - Task Force at work on atypical working times
  - Task Force will be organised in second half of 2005 for a total overhaul of the explanatory notes

Quality in the Community Innovation Survey

- In 2004, the survey questionnaire and the methodological guidelines of the Fourth Community Innovation Survey were agreed at European level
- The methodological guidelines also cover a section on data quality dealing with response rates, unit and item non response, imputation, weighting and calibration and the precision of results
- For key variables (such as the number of enterprises with innovation activity), certain precisions of results are to be reached which has implications for the sample size
- Eurostat will ask for the transmission of those quality measures which could then be disseminated as together with the meta-data
- Comprehensive standard quality reports will be prepared in the second step

Next work steps and conclusions

- Very good work progress made in the past years
- The statistical quality approach is more and more recognised as a standard tool within the European Statistical System
- Promote further the implementation of the statistical quality approach in the statistical domains which are not yet (sufficiently) active heron (e.g. R & D statistics), including legislation
- Work more on user oriented quality indicators
- Enlarge the quality approach to the New Member states
- Disseminate the quality documentation more widely