27th Meeting of the Wiesbaden Group on Business Registers
- Online International Roundtable on Statistical Business Registers Management under the COVID-19 pandemic

Mexico, 20 – 24 September 2021

Iliyana Iskrenova, Isabelle Collet
Eurostat

Session No. 7
Quality and statistical coverage

Data-quality indicators for European statistical business registers

Abstract

The commitment to quality is at the heart of the work of Eurostat. As the statistical office of the European Union, Eurostat's mission is to provide high-quality statistics for Europe. Notably, in line with the general quality standards of the European Statistical System (ESS), Eurostat monitors the quality of the European Statistical Business Registers (SBRs) of Member states and EFTA countries.

The European SBRs include the National Statistical Business Registers (NSBRs) and the EuroGroups Register (EGR). The quality of NSBRs directly affects both national and European statistical aggregates. Moreover, because NSBRs feed into the EGR at micro-data level, the quality of the EGR is especially reliant on the quality of NSBRs.

The SBRs are core elements of business, macroeconomic and trade statistics. They are also used in a wider range of statistical areas. It is therefore vital to provide legible and straightforward information on the quality of SBRs to users. Measuring quality in statistics is a well-known field based on standard quality dimensions, among which 'relevance', 'accuracy and reliability', 'timeliness and punctuality', 'coherence and comparability' and 'completeness'. They can be used to create quality indicators. It is for this reason that Eurostat has built up a composite indicator to summarise the quality of SBRs.

Based on Eurostat's experience in assessing the quality of European SBRs, this paper will present the different components, weightings and settings of the indicators. Each component, matching a quality dimension, could be adapted according to information availability and specific interest of other countries. The weighting and setting for each component could also be adapted. This paper discusses the benefits and disadvantages of such a composite indicator.

Improving the quality of the SBRs will: (i) ensure the quality of official statistics; (ii) reduce the costs and burden of statistical production; (iii) improve the production efficiency of NSBRs; (iv) improve the coherence and consistency of statistical products; and (v) improve technical and statistical harmonization.

Keywords: Data quality, indicator, composite, quality dimensions, European framework for statistical business registers.

1. Introduction

The commitment to quality is a key principle for the European framework of SBRs, which covers both: (i) the NSBRs and the EGR of multinational enterprise groups in the EU and EFTA countries; and (ii) data exchanges between these registers. The quality of NSBRs directly affects national and European statistical aggregates. Moreover, because NSBRs feed into the EGR at micro-data level, the quality of the EGR is especially reliant on the quality of NSBRs.

The NSBRs and the EGR are the basic infrastructure for deriving high quality and harmonized statistical business-register populations to produce European statistics. NSBRs are used as the main source of information for: (i) statistical analysis of the business population and its demography; (ii) defining the survey population; and (iii) establishing the link to administrative data sources. The EGR as a statistical business register of multinational enterprise groups is used for statistical purposes to coordinate cross-border information on multinational enterprise groups in the EU.

To improve the quality of the European SBRs, a data-quality programme (DQP) for the NSBR and EGR was introduced in 2016. The DQP helps EU Member States to improve data-quality management and adapt the SBRs to new user requirements by applying the 'plan-do-check-act' cycle. The DQP is made up of four components that are followed cyclically each year: Component A - Quality reporting; Component B - Quality targets; Component C - Quality assessment; and Component D - Quality improvement.

The SBRs are core elements of business, macroeconomic and trade statistics, but they are also used in a wider range of other statistical areas. It is therefore vital to provide legible and straightforward information on the quality of SBRs to users. Measuring quality in statistics is a well-known field based on standard quality dimensions among which 'relevance', 'accuracy and reliability', 'timeliness and punctuality', 'coherence and comparability' and 'completeness'. They can be used to create quality indicators. It is for this reason that Eurostat has build up a composite indicator to summarise the quality of SBRs.

2. Data Quality Programme

The DQP for European SBRs was set up in 2016. The overall objective of the DQP is to improve substantially the quality of the European SBRs, making it possible to fully achieve the SBRs' 'backbone' role in producing European statistics.

In line with the general ESS quality standards covering the implementation of principles derived from the European Statistics Code of Practice (ES CoP) and the quality assurance framework of the European Statistical System (ESS QAF), the DQP for the European SBRs is being implemented to further strengthening the quality of European statistics. The DQP includes four main components that are followed cyclically each year. Quality is improved continuously by focusing on the SBR user's needs.

The quality components of the DQP for SBRs are set out in the following four bullet points.

- Component A – Quality reporting or an 'as-is' state that conducts standardized quality and metadata reporting for NSBRs and the EGR.

• Component B – Quality targets or defining a ‘to-be’ state that specifies the quality targets to enable assessment of: (i) NSBRs and EGR frames throughout the entire production processes; (ii) related exchange processes; and (iii) resulting outcomes. Currently, Eurostat evaluates the NSBRs by using a set of compliance and quality targets.

• Component C – Quality assessment or assessing the gap between ‘to-be’ and ‘as-is’ is conducted on NSBRs and the EGR. Quality is assessed annually as part of monitoring compliance with the European legal framework. The results are documented in the assessment reports that are provided by Eurostat to Member states and EFTA countries.

• Component D – Quality improvement or ‘reducing the gap’ is dedicated to the annual improvement procedure. This quality component is focused on continuously improving the quality of national SBRs and the EGR to reduce gap between the ‘to-be’ and the ‘as-is’ states defined by quality targets.

DQP is implemented within a harmonised compliance monitoring and reporting framework in the area of European Business and Trade Statistics developed by Eurostat. Therefore, the DQP provides the European SBRs with an overall quality-management tool to ensure high quality and compliance with the European legal basis.

3. SBR quality assessment

Eurostat annually assesses the quality of the NSBRs within DQP Component C – ‘Quality assessment’. Quality is assessed by using a set of compliance and quality targets. The compliance targets are derived from Regulation (EU) 2019/2152 on European Business Statistics. The quality targets also include additional targets based on user needs that are not required by the legislation, e.g. optional variables. The assessment results are documented in individual national reports and an EGR report. The reports are provided by Eurostat to Member states and EFTA countries as part of monitoring the compliance and quality of the SBRs.

Although the DQP targets have been specifically designed under the European framework for national SBRs, the quality assessment relies on an assessment of basic quality dimensions i.e. ‘punctuality’ or ‘completeness’. The list of quality dimensions is rather large, so for illustration purposes this paper will focus on four main dimensions considered as significant for SBRs: ‘punctuality’, ‘completeness’, ‘accuracy and reliability’ and ‘coherence and comparability’. In the following sections, we will present an evaluation and assessment of these dimensions. Although the approach presented is based on Eurostat’s experience, the method could be tailored are re-applied by other countries.

Evaluation of quality

The evaluation is the way a quality dimension is measured. The measurement could be either quantitative or qualitative.

- ‘Punctuality’

Eurostat evaluates the punctuality of national SBRs by measuring the time lag between the actual date that data is delivered and the planned date of data delivery. For instance, the DQP work plan includes a calendar with deadlines for the delivery of data and

metadata. The work plan is made available to the countries at an early stage. Eurostat assesses punctuality in working days and only considers late delivery (i.e. for countries delivering data before or at the deadline, punctuality is assigned a value of 0, for countries delivering data after the deadline, punctuality is assigned a value equal to the number of working days of delay). However, one could choose another time unit or consider a negative value for early delivery.

Punctuality: P = (delivery date) – (deadline date) if delivery after the deadline; and respectively 0 if delivery is on time.

Where P is the integer measured in working days.

- 'Completeness'

Under the DQP, the completeness of national SBRs is measured using a list of variables related to SBR units listed in Regulation (EU) 2019/2152 on European Business Statistics. Eurostat focuses on the variables that the Regulation says national SBRs 'shall' provide to users, also called mandatory variables. For instance, for the statistical unit 'enterprise', the national SBR 'shall' provide an identifier and a name. For each unit and mandatory variable, completeness is evaluated by comparing the number of units and the number of variables available for this unit. For the variable that shall be available for all the units, such as the name of the enterprise, completeness is measured as a percentage, which can range from 0% to 100%. However, some variables are only available if the unit has specific characteristics. For instance, the enterprise date of final cessation of activities is only available for ceased enterprises. For these variables, completeness is either 100% if the national SBR can provide this information for the involved units or 0% if not. The SBR's completeness is evaluated by summing up the completeness of each mandatory variable and comparing it with the number of mandatory variables.

Completeness: $C = \sum$ (completeness of mandatory variable) / $\sum$ (mandatory variables)

C is a percentage ranging from 0% to 100 %.

- 'Accuracy and reliability'

To evaluate the 'accuracy and reliability' of national SBRs, Eurostat asked the country to specify in the metadata report the overall accuracy and reliability of their national SBR and what actions they take to systematically assess data quality and validate data, to identify, tackle and reduce potential bias. Thus, the countries report: (i) how they handle under-coverage/over-coverage or misclassification of the units; and (ii) how they spot and correct the data. For instance, some countries report possible misclassifications in the main economic activity in newly created legal units. To tackle this issue, national statistical institutes reported that they had set up specific surveys or carried out crosscutting checks, etc. Depending on the information the country provides, the 'accuracy and reliability' is evaluated using a binary variable (yes/no) whether or not a country has set up a newly created legal unit to report possible misclassifications.

'Accuracy and reliability': A = Yes if national statistical institute takes efficient measures to monitor SBR 'accuracy and reliability' and tackle potential bias, No if it does not.

A is qualitative variable: Yes/No.

- ‘Coherence and comparability’

Eurostat considers two types of SBR coherence: internal and external. Currently, several data sources are used by the national SBR, and it is important to consider the consistency of the information recorded in the SBR. Furthermore, the SBR should also consider its consistency with other business domains. Under the DQP, the coherence of national SBRs takes into account: (i) the coherence of the data provided by the countries within the report itself; and (ii) the coherence between the data provided in the report and data received by Eurostat related to other business domains. Thus, internal coherence in the countries’ reports is evaluated by the gap between aggregates reported in different parts of the reports. For instance, the total number of legal units with an activity code is cross-checked with the total number of legal units distributed by activity; the total number of enterprise groups is cross-checked with the sum of residential and multinational enterprise groups, etc. For external coherence, Eurostat evaluates the gap between the aggregates provided in the countries’ reports and the information provided to the EuroGroup register (EGR), the structural business statistic (SBS) or the business demography domains. However, for comparison between other domains, it is important to consider the scope and the reference year. In other words, different aggregates should be consistent in the order of magnitude or trend but not necessarily identical. To evaluate coherence, the absolute figure for internal coherence and percentage classes for external coherence are usually used.

Coherence: C = (Internal coherence; External coherence)

Internal coherence: Ci = $\sum$ (gaps between units reported in the report)

Ci is evaluated in number of units.

External coherence: Ce = percentage class including the gap between SBR and external aggregates. The gap is evaluated as a percentage in absolute terms.

Ce is a [a;b] class evaluated in percentages.

Assessment of quality

The evaluation process provides a figure for each quality dimension. However, to assess quality it is necessary to have a scale for each quality dimension, i.e. a qualitative ordinal measure. The scale includes classes and their respective thresholds. The number of classes will directly affect the progressivity of the assessment, and the thresholds will directly affect the scale’s requirement level. The classes can be mapped using color shades.

The assessment process should go beyond a simple evaluation and answer the question ‘is the information of sufficient quality?’ Thus, the assessment scale should include clear levels showing when the information has – or has not – acceptable quality. One can also include additional levels with warnings on potentially deficient quality.

The following examples illustrate the assessment for the punctuality dimension.

Punctuality: P = (delivery date) – (deadline date) if delivery after the deadline; and respectively 0 if delivery is on time P is an integer measured in working days.

Table 1. Punctuality assessment using 2 classes

On time	Late/Not accepted
If P= 0	If P > 0

By using two classes, the assessment is a binary variable (with no progressivity)

Table 2. Punctuality assessment using 4 classes

On time	Acceptable	Late	Not accepted
If P = 0	If 0 < P <= 10	If 10 < P <= 30	If P > 30

By using four classes, the assessment is more progressive, by moving the threshold of acceptance level from 0 to 30 working days the requirement is lower.

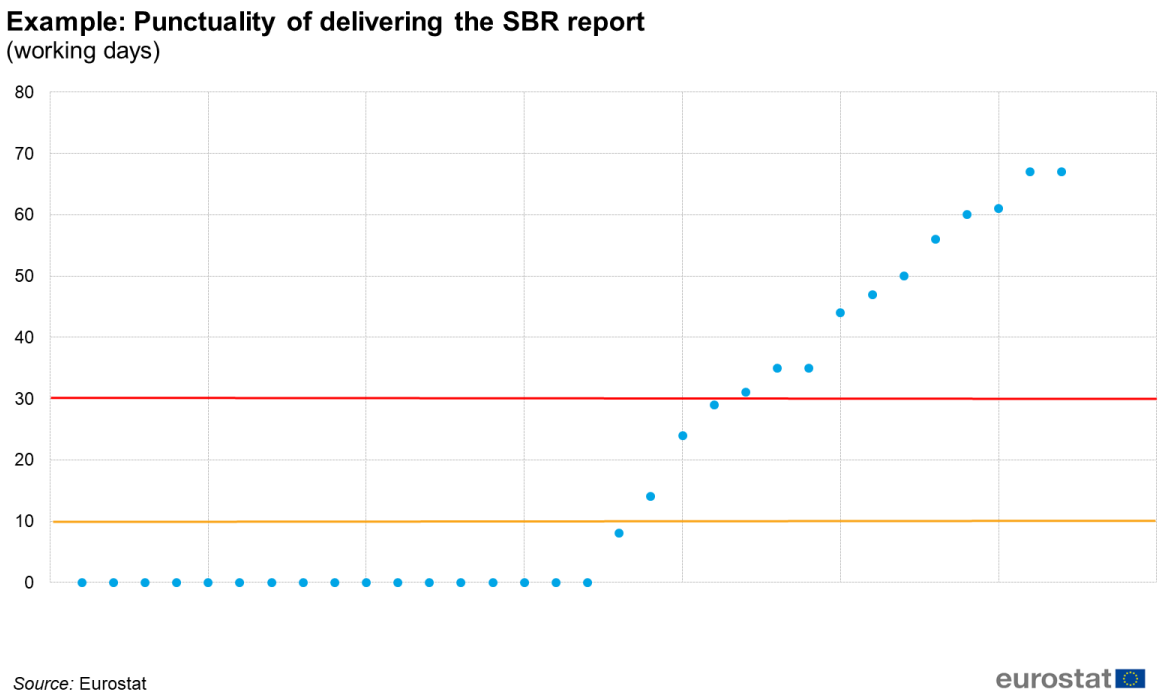

Figure 1. Distribution of the punctuality of delivering the SBR report by the NSI in working days highlighting the classes' thresholds.

There are no specific rules for setting the scale. However, from a pragmatic perspective, the assessment is rather unreadable with more than five classes. Furthermore, the classes should reflect the assessment's purpose. Here, for instance, Eurostat is specifically monitoring whether the countries deliver their data on time. There is therefore no need for an early-delivery class. The threshold settings should take into account: (i) the distribution, e.g. the former distribution of punctuality (as shown in the figure); and (ii) user needs. Thus, Eurostat considers both the feasibility of data delivery in the DQP but also the users' requirements. For instance, each country is legally required to deliver the data on time, because Eurostat is legally required to release EU aggregate according to schedule.

The assessment should also consider the type of measurement used in the evaluation. As previously discussed, the type of evaluation measurement can differ from one quality dimension to another. For instance: (i) 'punctuality' is measured by the number of working days; (ii) 'completeness' is measured as a percentage; (iii) 'accuracy and reliability' is measured as a binary variable; and (iv) 'coherence and comparability' is measured as both a number of units and a percentage class. The settings for numerical figures follow the same pattern as illustrated for punctuality, only the threshold has to fit with the evaluation units. For instance, because the evaluation unit for completeness is a percentage, the threshold should also be presented as a percentage. When the evaluation is a binary variable, as is the case for 'accuracy and reliability', the obvious assessment is a binary scale with two graduations. However, there is a way to introduce additional graduations as illustrated in the following example.

As previously presented, is the 'accuracy and reliability' evaluation a binary variable? = Yes / No

Table 3. 'Accuracy and reliability' assessment using 2 classes

SBR 'accuracy and reliability' is satisfactory	SBR 'accuracy and reliability' is not satisfactory
Yes	No

To get a more refined assessment, it is possible to split the evaluation by considering several accuracy areas. This means that instead of having an overall accuracy evaluation, accuracy can be evaluated by the three binary variables set out below.

1- Are the units' activity codes in SBR monitored and validated? Yes/No.

2- Is the under-coverage/over-coverage of recorded units in the SBR monitored and validated? Yes/No.

3- Are specific measures taken to tackle misclassification of the units? Yes/No.

Table 4. 'Accuracy and reliability' assessment using 4 classes

Satisfactory	Acceptable	Poor	Not satisfactory
Yes for the 3 variable	Yes for the 2 variable	Yes for 1 variable	Yes for 0 variable

4. Composite quality indicator

It is possible to monitor each quality dimension individually. However, we can also set up a composite quality indicator gathering the assessments of all the selected quality dimensions. The basic settings of such an indicator are the weightings of the quality dimensions and the selection of the indicator type. The indicator can also be used to consider quality as a whole. In this case, there is no need to individually assess each quality dimension. Thus, the indicator calculation will be based on the evaluation outputs.

However, by skipping the assessment of the quality dimensions, we will not be able to tackle specific issues.

Among the indicator settings, the weighting reflects the focus on the quality dimension needed. Although quality dimensions have equal importance, it may be necessary to monitor some quality dimensions more closely, for example to ensure compliance with a regulation or because they affect vital processes. One might also consider the users' needs and requirements when setting up the weighting. For instance, Eurostat launches regular national SBR users' surveys to monitor the use of – and any quality issues with – national SBRs.

The indicator type reflects the use, purpose and goal of the quality monitoring. There is a large range of indicators starting from the simple score to the more complex compound index. However, a basic rule is that all the stakeholders should be able to easily understand the indicator output, its settings, and its purpose.

The following example derived from Eurostat's experience illustrates a score indicator.

The example considers four quality dimensions: 'punctuality', 'completeness', 'accuracy and reliability' and 'coherence and comparability'. The indicator is a score ranging from 0 to 100, each dimension has a weight, and the sum of the weights is 100. Points are calculated for each quality dimension, taking into account the weight and the assessment. The sum of the country score is the sum of the points. The following example illustrates weightings and the distribution of the quality points according to the assessment.

Table 5. Quality dimensions' settings

Punctuality		
P = 0	0 < P <= 30	P > 30
On time	Late	Not accespted
Completeness		
C = 100%	95% <= C < 100%	C < 95%
Complete	Acceptable	Not accepted
Accuracy and reliability		
Yes		No
Internal coherence and comparability		
|Ci| = 0 unit	0 < |Ci| <= 10 units	|Ci| > 10 units
Satisfactory	Acceptable	Not satisfactory
External coherence and comparability		
|Ce| in [0%;5%[	|Ce| in [5%;25%[	|Ce| > 25%
Satisfactory	Acceptable	Not satisfactory

Table 6. Score settings

Quality dimension	Weights	Points		
Punctuality	20	20	20*(P/30)	0
Completeness	40	40	40*C	0
Accuracy and reliability	20	20		0

Internal coherence and comarability	10	10	10*(|Ci|/10)	0
External coherence and comparability	10	10	10*(|Ce|)/10	0
Total	100	100	∑ points	0
Score		Satisfactory	Acceptable	Not satisfactory

Key points

The examples above are based on Eurostat experience and are for illustration only. They cannot be considered as rules but instead should be considered as a toolbox. Thus, one can adapt and tailor the proposed assessment process to monitor more specific SBR quality issues or SBR data sources. Furthermore, the presented assessment process only considers four quality dimensions. However, other dimensions, such as timeliness, have a significant impact on the quality of the SBR and can be added to the monitoring.

Some threshold effects might negatively affect the assessment process. For example, one can discard some SBR information by setting high acceptance thresholds or accepting data even if they are of poor quality. To reduce this effect, one should test the assessment process using previous data, or create test samples by using previous data or bootstrap sample based on previous data. Furthermore, once the process has been tested, it should remain stable (i.e. it should not be changed) at least for some years. Thus, the assessment could monitor the quality of the SBR over a certain period. In addition, if we need to update the settings, we might consider back casting the quality indicator time-series.

As previously mentioned, a composite indicator is a valuable and user-friendly instrument to monitor overall quality. However, it can conceal or over-estimate specific quality issues. Composite indicators calculated as averages often overwrite the distribution and are affected by extreme values. For instance, it might be the case that the quality of the SBR as a whole is acceptable even if one of the SBR quality dimensions is not satisfactory. Conversely, it is possible that the overall quality of the SBR is not satisfactory even if one of the SBR quality dimensions is not satisfactory. This is why it is usually better to monitor each SBR quality dimension separately from the overall quality of the SBR.

5. Conclusion

In all EU Member States, the availability of good-quality SBRs is key to compiling business statistics based on survey frames derived from the SBRs. This is why the quality of these registers should be closely monitored. For instance, it will be very difficult to produce aggregates for the EU as a whole if the quality of data – either provided directly or through statistical surveys – varies between countries.

Improving the quality of the SBRs will: (i) ensure the quality of official statistics; (ii) reduce the costs and burden of statistical production; (iii) improve the production efficiency of

NSBRs; (iv) improve the coherence and consistency of statistical products; and (v) improve technical and statistical harmonization.

References

[1] European Statistical System Committee (2017), European Statistics Code of Practice, Publications Office of the European Union, Luxembourg (https://ec.europa.eu/eurostat/documents/4031688/8971242/KS-02-18-142-EN-N.pdf/e7f85f07-91db-4312-8118-f729c75878c7).

[2] European Statistical System Committee (2019), Quality Assurance Framework of the European Statistical System, Version 2.0, Publications Office of the European Union, Luxembourg (https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf/bbf5970c-1adf-46c8-afc3-58ce177a0646).

[3] European Union (2019), Regulation (EU) 2019/2152 of the European Parliament and of the Council of 27 November 2019 on European business statistics, repealing 10 legal acts in the field of business statistics (https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R2152&from=EN).

[4] Eurostat (2021). European business statistics methodological manual for statistical business registers — 2021 edition. Publications Office of the European Union, Luxembourg (https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-20-006)