

OFFICIAL STATISTICS

A first interim assessment of the third round of peer review of the European statistical system

Agustín Cañada

Subdirección General de Difusión, Instituto Nacional de Estadística, agustin.canada.martinez@ine.es

Received: November 3, 2022. Accepted: December 26, 2022.

Abstract: Peer Reviews are exercises to assess compliance with the principles and indicators of the European Statistics Code of Practice by the members of the European Statistical System: Eurostat and the national statistical systems (composed of statistical offices and other institutions). Peer Reviews are carried out periodically (every 5/6 years), by agreement of the European Union. To date, three rounds have been carried out: in 2006-2008, in 2013-2015, and a third round is underway between 2021 and 2023. Although the third round is still ongoing at the time of writing (December 2022), based on the experience of a representative group (14) of the countries already reviewed, a first assessment can already be made of the degree of achievement of the objectives pursued. The aim of this document is to provide a first input for a future comprehensive "lessons learned exercise" and to contribute to the debate on aspects to be taken into account in future peer reviews.

Keywords: Code of Practices, Peer review, Quality of official statistics

MSC: 62-03, 62-04, 62P99

1 Introduction

Peer reviews (PR) are exercises to assess compliance with the principles and indicators of the European Statistics Code of Practices (CoP) by members of the European Statistical System: Eurostat and national statistical systems (NSI, other institutions...). The ultimate objective of these assessments is to strengthen the statistical systems at national and European level, thus increasing the reliability and credibility of users and other stakeholders in European official statistics (Cañada, 2019).

The PR are carried out periodically (every 5/6 years), by agreement of the European Union: There have been three exercises ("rounds"): The first "round" of PR took place between 2006 and 2008; the second, more far-reaching one, was conducted between 2013-2015 and a third round is underway between 2021 and 2023.

At the time of drafting this report (November 2022), the halfway point of the third round of peer reviews has been reached, both in terms of time and number of countries evaluated, with 19 of the 31

planned evaluation visits having been completed. Of these, Eurostat has published 14 final reports on its website. In the case of Spain, the experts' visit to INE took place from May 31 to June 3, 2022, and the final report is not yet available. However, based on published reports and Spain's experience, it is possible to make an initial assessment of the degree of achievement of the objectives pursued. The purpose of this document is to provide decision-makers with a first input for a future and more complete "lessons learned exercise" and to contribute to the debate by the countries on the aspects to be considered in future peer reviews.

2 Peer review of the European Statistical System: An overview

Table 1 summarizes the key features of the PR, highlighting the differences between the three rounds conducted to date.

The first "round" of reviews focused only on some CoP principles and was conducted under a proper peer review approach: statistical institutes in each EU country were assessed by experts from statistical institutes of other countries.

The second round, carried out between 2013-2015, was broader in scope: all the principles of the Code of Practice were assessed; (a sample of) other institutions producing European statistics in each country - the so-called Other National (Statistical) Authorities ONA (ministries, agencies, etc.)- was brought in; the process was applied not only to the EU countries, but also to the four "European Free Trade Association" (EFTA) countries; an attempt was made to move closer to an audit approach, using teams of expert reviewers from outside the NSIs, to avoid the impression of less independence and objectivity linked to the internal reviewers as in the first round.

From the point of view of its audit approach, conventional audit principles are followed: information on the national statistical system is prepared by the NSI, based on a self-assessment questionnaire (SAQ), supporting the answers with documents that serve as evidence; the team of reviewers analyses this information in depth; there is an "audit" visit to each NSI in which the reviewers complete their on-site knowledge of the system assessed; the last stage is the preparation by the reviewers of a report on the degree of compliance with the CoP by the NSI (countries) that includes "recommendations" on areas for improvement; and then, in response to these recommendations, the NSI would prepare a multi-year action plan for improvement. Implementation will be monitored annually by the European authorities.

The third round, which runs from 2021 to 2023, shares common features with the second round, as it covers the same scope of countries (all EU countries plus the four EFTA countries), the same content, as it covers all CoP principles, and the same reviewed institutions, Eurostat, the NSIs and the ONA. However, it presents differences that seek to improve or correct more controversial aspects of the second round, which can be grouped into six fundamental ones (Cañada and Muñoz, 2016):

1) Methodological approach. Although like the 2013 PR, the new phase is still close to the audit methods, it is nevertheless intended to be a combination of audit and PR approaches. This is perfectly illustrated by the change in the composition of the teams explained below.

There are also changes in some of the tools used: in particular, the NSI self-assessment questionnaire has been simplified so that, without losing its global character (it covers all the principles of the Code), it allows for greater agility and speed in the review process. The questionnaire is much simpler than that of the 2nd round: the questions are adapted to the 84 CoP indicators, which in fact

Topic	2006-2008	2013-2015	2021-2023
1. Territorial scope	EU countries	28 EU + 4 EFTA countries	27 EU + 4 EFTA countries
2. Conceptual scope	Principles: 1 to 6 and 15 of CoP (2005).	- All CoP principles (2011) + coordination + EU cooperation	- All CoP principles (2017). Emphasis on certain principles
3. Institutional Scope	NSI	NSI + ONA + Eurostat	NSI + ONA + Eurostat
4. Methodology	Peer review approach	Quasi-audit	Combination PR + audit
4.1. Reviewers	Internal to the ESS: NSIs active staff	Experts from outside the system (+ Eurostat observer)	Internal + External + Eurostat Member + Eurostat Harmonization Observer (in some PR)
4.2. Procedure	PR approach: self-assessment; review; NSI visit; reporting	"Audit-type" approach. self-assessment; review; NSI visit; report	"Audit-type" approach.
4.3. Self-assessment questionnaire(s)	CoP indicators only for the selected principles	a) NSI: 3 SAQ: - CoP SAQ: defined by principles/ indicators (broken down at QAF detail) - 2 Ancillary SAQ on: Coordination and Cooperation. [Total number of questions: 400] b) ONA: simplified questionnaire	a) NSI (only one SAQ) - 84 questions on each indicator of the CoP. - 60 SWOT questions. - 9 questions by area of the CoP. - 29 additional questions (innovations, COVID-19 pandemic) [Total number of questions: 182] b) ONA: simplified questionnaire
4.4. Evaluation report	Country-specific reports: NSIs	Country-specific report, focusing on NSIs (with some reference to ONA)	a) Specific report for each country, covering both the NSI and the ONA. b) Recommendations differentiated between Compliance-relevant and improvement-related
5. Others			Eurostat + countries communication campaign

Table 1: Peer review of the ESS: key features of the three rounds

reveal compliance with the 16 principles of the Code; some other questions are included (see Table 1) to reach a total of 182 questions. We can remain that in the 2nd round, the SAQ was structured according to QAF (quality assurance framework, Eurostat (2019)) methods linked to the CoP indicators, which involved completing a questionnaire with more than 400 questions.

2) Seek greater harmonization of reviews across countries. In the 2013 round, there was great heterogeneity and lack of harmonization in the country reviews and in the PR reports: there were no well-

established general criteria on what the most relevant and accessory points could be when assessing countries; the heterogeneity of the reports and of the recommendations to the countries also translated into heterogeneity of the improvement action plans.... This heterogeneity greatly conditioned one of the objectives of the PR, which is to contribute to the improvement of the European statistical system, by developing actions that would have led the countries to advance along common lines.

Therefore, more "harmonization of the review along the different countries" is a Priority goal of the current 3rd round. Several elements are used to achieve this harmonisation (Eurostat, 2021a):

- On the one hand, an effort towards greater standardisation of the methodologies applied by the different teams: Although the report and the recommendations will obviously depend on the outcome of the evaluation, and the review teams have autonomy of decision, priority themes have been recommended for the review (according to European agreements). A guide of suggested types and categories of recommendations, has been published in one of the annexes of the methodology (Eurostat, 2020a).

- Furthermore, a differentiation into two types of recommendations was introduced: (...) "The recommendations issued by the peer review team should be split into "Compliance-relevant" - recommendations fundamental/important to ensure compliance/alignment with the ES CoP - (...), and "Improvement-related" - less critical/technical supporting improvements recommendations-.

- An additional element is the role of Eurostat technicians in the process, which is discussed in a later section.

3) Composition of the review teams. In the third round, the composition of the teams of reviewers uses a mixed formula from previous experiences: each team of reviewers includes both "internal" evaluators from the statistical system (-PR approach-) and experts external to official statistics (-audit approach-). This is an attempt to solve a problem detected in the second round: having reviewers external to the system guarantees greater objectivity but has the counterpart that they may not have a sufficiently updated knowledge of the situation and trends of the statistical system.

For this reason, the third round has opted for this compromise between the two previous ones, combining the independent and objective vision provided by the external reviewers with the updated knowledge of the practices and criteria of European official statistics that can be provided by the NSI and Eurostat staff.

Finally, the teams of reviewers are made up of four members: a serving member of an NSI ("internal", "peer review" approach); an expert external to the ESS ("audit" approach); an expert from Eurostat (participating as a reviewer); and, as Chairman of each team of experts, a senior statistician, with experience in NSI management.

The first group, people who are currently working in the NSIs, is proposed by the countries. Within this group, some coordinators of the PR Process (and/or responsible of the Quality management) in their own countries, are simultaneously reviewers for other countries, which reveals the difficulty in finding people specialised in quality topics. An implicit issue is whether this direct involvement of quality managers as reviewers may give an impression of less objectivity of the process, as already raised in the first round of the PR.

4) Role of Eurostat. One of the aspects of the previous round questioned by the countries was the lack of active participation of Eurostat in the reviews (only a Eurostat technician participated as a "mere observer" in the country visits). In response to this criticism, Eurostat has taken a more active role in the 3rd round:

- On the one hand, a Eurostat technician participates as a member of the expert team. This active role, besides serving to improve the quality of the process, is guided by the objective of improving harmonization in the assessment and reporting of countries.

- Moreover, in pursuit of this harmonization objective, Eurostat proposed to incorporate as an ad-

ditional "observer" to the process an expert, specialized in quality issues, who would support the expert team during the country visit and in the drafting of the report and, especially, of the recommendations to the countries. "The role of Eurostat observers would be to support the expert team in formulating more harmonized recommendations, especially in terms of scope and magnitude, during the PR visit and, in particular, on the last day of the visit." (Eurostat 2020b). However, due to lack of resources, Eurostat observers are only involved in some of the reviews. It is Eurostat itself that chooses the visits in which it wishes to participate, with the prior approval of the country under review.

5) Strengthening the role of ONA. This round of the PR aims to strengthen the participation and role of the ONA. Although they had already been included in the previous round, their role in the process and in the reports was considered very marginal. Therefore, the new round aims to give a greater role to the ONA and, in short, to have a more complete view of the situation of the National Statistical Systems as a whole. To this end, a specific procedure and objective criteria agreed upon by Eurostat and the NSIs have been established for greater participation of the ONA in the process.

6) Improved communication on the PR to stakeholders. One of the novelties of the third round is the attention given to communication aspects, through the design and implementation of a communication campaign by Eurostat and the countries on the process, its objectives and results. This is a reaction to one of the most questioned aspects of the second round, which was the limited impact of the process outside the statistical world. There was a critical view of the (limited) dissemination of the process.

But communication is essential to achieve one of the fundamental objectives of the PR: to contribute to the external image of quality and credibility of official statistics. That is, to demonstrate to the institutions most closely linked to statistics (the "stakeholders": informants, policy makers, users), but also to society at large, that the European Statistical System "operates within a sound quality framework". In short, to contribute to the credibility and confidence of users in statistical institutions. At the same time, it will also promote that governmental institutions support the improvement actions derived from the PR.

To this end, a communication strategy to accompany the third round of ESS peer reviews was defined based on the design of common instruments and means of dissemination for all countries.

3 Some drawbacks of the third round: Lessons learned and alternatives for future PR

In November 2022, the halfway point of the third round of peer reviews has been reached, both in time and in the number of countries evaluated, with 19 of the 31 peer review visits having been completed.

In the case of Spain, the experts' visit to INE took place from May 31 to June 3, 2022. Part of the meetings were held with the statistical production units of INE (35 managers and technicians of the institution participated in the meetings), but also with 30 representatives of the main stakeholders of the Spanish Statistical System: other producing institutions, such as Ministries and Bank of Spain; qualified users -business federations, trade unions, non-governmental organizations-; managers of administrative registers; representatives of the scientific community: universities and researchers; media. At the time of writing (November 2022) only a provisional version of the report is known. Based on the data known so far from the reports published by Eurostat on its website (www), a first assessment can already be made of the degree to which the objectives pursued are being achieved.

1) Harmonization as a main challenge. A greater harmonisation of the countries' processes was

one of the basic objectives of the third round the ESS peer reviews. A challenge of this objective is how to achieve more harmonisation of the reports while there is flexibility for the reviewers' team in the choice of principles/indicators to review. Therefore, harmonization has been focused on the recommendations: "The aim of harmonisation is about the outcome/ the final results of the peer review meaning the scope, magnitude and number of recommendations" (Eurostat, 2021b).

Concerning scope and magnitude, to ensure greater harmonization, several elements already mentioned, were introduced: a guide of the types and categories of recommendations; the above-mentioned distinction between "Compliance-relevant" (CR) and "Improvement related" (IR); the inclusion of a Eurostat observer...

As a reference of the outcome of the process, we can summarise figures for the 14 countries: a total of 251 recommendations were made, most of them (221) belong to the "Improvement" category, and only 30 are "Compliance-relevant". This is undoubtedly a very positive result, as it reveals the high degree of compliance with the CoP by European countries. However, an in-depth analysis of the reports reveals some questionable aspects: on the one hand, there are some countries without any "CR" recommendations. Although this is possible, as it is reflecting a very high level of compliance with the CoP, in practice it is questionable, if we remember how the PR self-assessment questionnaire is designed: For those countries where no CR recommendations have been identified, this means that the level of compliance with the 84 CoP indicators should be almost total or perfect.

The second doubt that arises from the differences in interpretation among the different teams of reviewers as to what is included in one category or another, even with different evaluations for similar recommendations. Of course, the recommendations and their classification in one category or another is responsibility of the Expert teams; and they are autonomous to guarantee objectivity of the process. But more homogeneous criteria would be useful.

Another aspect linked to harmonization is the aforementioned issue of Eurostat observers, who, due to limited resources, only participate in some of the visits (chosen by Eurostat). Thus, this initiative would have more scope if observers could participate in all reviews.

2) Recommendations and the communication issue. One of the aims of the 3rd round is the effort to a more intensive dissemination and communication campaign of the process to reassure stakeholders about the quality of European Statistics. On the other hand, being one of the main objectives of the PR assessing whether NSI are fully compliant with the CoP, that means to identify aspects for improvements and/or where there is not compliance with the CoP (as in any auditing process). That is, the final and more evident outcome of the PR is the list of the recommendations stated to the countries. The obvious problem is that if the report places excessive emphasis on the recommendations (by their number, by their nature) and/or points to be improved by the country under evaluation, the final view portrayed to the stakeholders about the country and its statistical system is debatable. This may have the opposite effect of what was intended by the review process: by causing stakeholders to question the quality of official statistics.

Back to the figures for the 14 countries available, the most frequent number of recommendations is 22. This number of recommendations necessarily implies devoting a good part of the country report to their justification (moreover, the recommendations appear twice in the report: in the executive summary and in a specific chapter). Thus, for a non-expert reader or a reader outside the statistical world, it is uncertain what impression can be obtained of the statistical situation of a country from a report of these characteristics.

3) The scope of the PR. In terms of scope, the PR aims to assess the overall status of countries' statistical systems, through the level of achievement with the CoP as a whole. This objective is clear, but probably too broad or ambitious, given the current complexity of statistical systems and the obvious resource and time constraints faced by the reviews. In practice, it is a difficult challenge,

firstly for countries to summarize the status of the statistical system in a simple questionnaire and supporting documentation; secondly, it is also difficult for the team of reviewers, a small group of people working in a necessarily small amount of time, to analyze the documentation in detail and to adequately understand and assess a country's statistical system.

For example, the official self-assessment questionnaire (SAQ) of the third round contains a total of 182 questions: 84 for the CoP indicators; 60 for the typical SWOT questions; 9 for a general self-assessment by groups of principles (institutional environment, processes, and products); 29 for additional questions on other topics.

Moreover, in the case of Spain, the SAQ has been designed under a structure like that of the 2nd round questionnaire; that is, describing not only compliance with the 84 CoP indicators, but also compliance with the methods recommended by the QAF 2019. And, in support of this questionnaire, INE prepared more than ninety documents, in some cases written or updated (and translated into English) especially for the PR. The investment made by the Spanish NSIs in the preparation of the PR was very considerable.

After the analysis of all this documentation, the second step of the PR is the visit to the country assessed by the reviewers. Over the course of five days, the reviewers try to complement their vision of the statistical system through meetings with INE staff and different stakeholder groups.

Realistically, despite the undeniable effort and professionalism of the experts, it is still a system that faces obvious limits (in terms of time and resources) to adequately capture the complex reality of current European statistical systems. And the question remains whether the final reports could not reflect the great efforts made by the NSIs in this field.

Returning to one of the recurring questions throughout the different rounds, it is worth asking whether, instead of a global approach aiming to assess the full coverage of the Code, a more in-depth but narrower scope analysis, focused on groups of principles, would be more appropriate. In addition, a more concentrated and narrower scope of the PR could also contribute to the recurrent objective of harmonization.

4 Final comment

There is no doubt that the PR assessments are contributing to the improvement of the quality of the European statistical system. They constitute a balanced mechanism between simple internal evaluations and audit exercises, adapted to the peculiarities of official statistics. The third round represents a new step forward, overcoming some of the limitations of the previous rounds.

Although the third round is still in progress at the time of reviewing this document (December 2022), based on the final reports for 14 of the countries already reviewed, (and the practical experience of the author in the case of Spain) a first assessment can already be made of the degree of achievement of the objectives pursued. As a result of this analysis, some suggestions are made on areas where there is still room for further efforts: in relation to the scope of application, as in previous rounds, the question arises once again as to whether the PR can be approached from a perspective more focused on specific topics, rather than attempting to cover the entire CoP; it is also pointed out that further progress is needed in the harmonization of results between countries, with greater homogenization of criteria among the review teams; aspects that could contribute to the objective of better communication of the nature and value of these exercises aimed at the key players in the statistical system.

References

- Cañada, Agustín (2019). Peer reviews of the european statistics and the involvement of users in the quality assurance of official statistics. *Spanish Journal of Statistics* 1(1), 33–40.
- Cañada, Agustín and Luisa Muñoz (2016). Peer review 2013-2015. lessons learned, challenges and opportunities. In *European Conference in Official Statistics 2016, Madrid*.
- Eurostat (2019). Quality assurance framework. Technical report, Eurostat.
- Eurostat (2020a). Guides' annex vi: Formulation of issues and recommendations. Technical report, Eurostat.
- Eurostat (2020b). Overall methodology for the third round of peer reviews. Technical report, Eurostat.
- Eurostat (2021a). Annex on the harmonisation of results of the peer reviews. third round: 2021-2023. Technical report, Eurostat.
- Eurostat (2021b). Frequently asked questions on the 3rd round of ess peer reviews (last update: 3.6.2021). Technical report, Eurostat.