

High quality official statistics – benchmarking as an integral part of a quality management system

David Fenwick, International Expert

Group of Experts on Consumer Price Indices – Geneva 2014

Introduction

Benchmarking against another national statistical institute provides a tool for the identification of potential issues relating to the compilation of official statistics and is one example of the processes that a statistics office can deploy to fulfil their public commitment to effective quality assurance of outputs to facilitate the cost-effective compilation of high quality statistics. Such benchmarking should operate within agreed operational guidelines and objectives and should focus on a pre-determined and mutually agreed set of performance indicators. The paper draws attention to a number of necessary conditions for a successful benchmarking exercise and makes a number of observations on how benchmarking can be an integral part of a quality management system.

Background

Benchmarking should be seen a fundamental part of the process of achieving cost-effective services in the public sector. It has had a long history. For instance, the Public Sector Benchmarking Service (PSBS) jointly launched by the Cabinet Office and HM Customs and Excise in late 2000 to facilitate knowledge sharing and collaborative approaches. Similarly, the European Commission spearheaded a programme of benchmarking where member states lead projects to facilitate knowledge sharing. These initiatives have generally focused on non-statistical areas but not always. For example, the UK Office for National Statistics (ONS) and the Australian Bureau of Statistics (ABS) undertook a benchmarking exercise of their respective Consumer Price Indices (CPIs) based on a shared view that there would be mutual advantage to be gained from comparing performance in the compilation of a high profile statistical output, such as the consumer price index, where the output is relatively well-defined internationally but where there can be significant differences in, for example, sample design and data collection practices. It was also seen as a way for addressing a number of specific issues to do with the compilation of the Australian CPI that the ABS thought warranted investigation, although care was taken to keep an open mind and not to pre-judge the direction of the exercise or the results.

Benchmarking can provide an external stimulus for change, to increase the effectiveness of delivery and improve performance. But success depends on objectivity in approach and no pre-determined agenda.

The Benchmarking process

There are a number of steps when undertaking a benchmarking exercise.

1. ***An initial face-to-face meeting*** to agree the basic scope of the benchmarking and the method of working for the benchmarking of consumer price indices. This meeting is critical as it determines the future direction of the project and lays down the foundations for co-operation. It will also secure an understanding of the mutual commitment involved which may be subsequently formalised in a memorandum of understanding. One of the main outputs of such a meeting should be to agree the objectives of the exercise and to produce a list of issues raised that both parties agreed would benefit from specific consideration¹. From these discussions can also emerge implicitly a number of principles that contribute to the determination of the scope of the audit. Criteria for the inclusion of issues might include, for example:
 - a. Aspects of data collection and index construction which are seen by at least one party as a potential issue i.e. there is prima facie case for believing that sub-optimal solutions are being pursued.
 - b. Potential concerns that are common issues to both parties, or where there is scope for sharing experience e.g. where the approaches taken by the two parties differ, thereby providing scope for comparisons of different solutions.
 - c. Issues with a broad focus that go beyond the bounds of methodology to issues relating to e.g. value for money where data collection and index compilation costs are compared with the quality of the statistical output.
 - d. Purely methodological reviews.

The success of the initial meeting is dependent on the effort put in by both parties into conducting sufficiently detailed preliminary background work before entering into bi-lateral discussions. Equally important is a realisation that pre-conceived ideas about the outcomes can seriously prejudice the objectivity of the benchmarking exercise. To guard against this great care need to be taken to ensure that the issues are expressed in a neutral way. For instance, on the use of handheld computers for price collection, the issue should be described in terms of “the relative merits of recording prices using handheld computer compared with pen and paper” rather than “the case for moving from pen and paper to handheld computers for the collection of prices”.

2. ***Collection and analysis of information.*** The process of collecting information for the benchmarking exercise, itself acts as a focus for inward-reflection and for generating a number of issues that have not been identified at the initial scoping exercise. The task of gathering and analysing this information is not a minor one and warrants a great deal of consideration to ensure that it generates material that is relevant and comprehensive. It usually involves an iterative process. For example:

¹ This might include, for instance, continuous versus point-in-time price collection, the use of handheld computers versus paper for the recording of prices, the organisation of fieldwork, the use of other data sources such as information from bar codes, the auditing of price.

- a. The drawing of process maps to give comparative overviews of the statistical processes.
- b. The design of a standard initial questionnaire for basic fact finding and its subsequent completion by both statistical offices, including footnotes with additional information as considered appropriate. This can be a major task but can be moderated by the use of readily available meta-data.
- c. An initial comparative analysis by followed by an email discussion and agreement on the main issues that are raised. This part of the process can also identify additional information that needs to be collected e.g. to follow-up the main issues raised and including financial information on costs.
- d. Repetitions of the above stage until issues had been satisfactorily resolved
- e. The drafting of the final report plus a post-hoc evaluation of the benchmarking process.

It should be noted that a point may be reached where the collection of additional information leads to diminishing returns e.g. with the extraction of more detailed financial information on costs e.g. on data editing and validation, where different financial recording systems may be in place and comparative information at a detailed level cannot be extracted easily and where approximate and high-level financial information may not be very useful.

3. ***The Report.*** One of the parties or an external facilitator will normally take the lead in writing the report. The expectation is that the report will go through a number of drafts as comments are exchanged by email or face-to-face meetings. But as all the key recommendations should have already been agreed during earlier discussion, drafting comments will normally relate to points of clarification and presentation.

Recommendations can cover a broad range of issues including: methodological issues (this may include e.g. the interpretation and adherence to regulations relating to index production); the use of handheld computers for collecting prices in shops; the editing and validation of primary data; general quality assurance procedures; longer-term index development. They can range from the very specific and concrete to a recommendation that an issue be investigated further. Some may go wider than day-to-day operational issues and address the perceived weaknesses in performance indicators and other management information such as the identification of costs.

4. ***Post-evaluation.*** Evaluation is an essential part of the benchmarking process and is best undertaken in two stages: an initial evaluation e.g. in the form of a workshop; a final evaluation at a later date when the delivery of benefits should be much clearer. These can be facilitated by a third-party. The discussion and conclusions of the initial evaluation should be summarised in a report and action points agreed. The latter may relate to general lessons learnt from the benchmarking exercise about the organisation and conduct and procedural improvements to be applied in future exercises. The matters raised might include, for example: the need for better guidance on the

collection of financial information and the need to focus more detailed cost data on areas where such information is critical; the need to use agreed standard terminology e.g. in the initial questionnaire; the value of qualitative performance indicators. The final evaluation should be more systematic and more focussed on determining whether the benchmark exercise achieves its objectives and represents good value for money. It should include a review of the follow-up action to the recommendations made in the benchmarking report and an assessment of the outcomes that have been delivered.

Benchmarking and Eurostat's Peer Review Process

5. The European Statistics Code of Practice sets the standard for developing, producing and disseminating European statistics. It builds upon a common European Statistical System definition of quality in statistics. The Code of Practice contains fifteen principles which address the *institutional environment* under which national and EU statistical authorities operate as well as the *production* and *dissemination* of European statistics.
6. Associated with the fifteen principles is set of indicators of good practice
7. The *institutional environment* relates to the observation that institutional and organisational factors have a significant influence on the effectiveness and credibility of a statistical authority developing, producing and disseminating European Statistics. The relevant issues are professional independence, mandate for data collection, adequate resources, quality commitment, statistical confidentiality, impartiality and objectivity.
8. For instance the first code relates to the *Institutional environment* and states that the “Professional independence of statistical authorities from other policy, regulatory or administrative departments and bodies, as well as from private sector operators, ensures the credibility of European Statistics”. It is assessed using the following eight indicators.
 - a. The independence of the National Statistical Institutes and Eurostat from political and other external interference in developing, producing and disseminating statistics is specified in law and assured for other statistical authorities.
 - b. The heads of the National Statistical Institutes and of Eurostat and, where appropriate, the heads of other statistical authorities have sufficiently high hierarchical standing to ensure senior level access to policy authorities and administrative public bodies. They are of the highest professional calibre.

- c. The heads of the National Statistical Institutes and of Eurostat and, where appropriate, the heads of other statistical authorities have responsibility for ensuring that statistics are developed, produced and disseminated in an independent manner.
 - d. The heads of the National Statistical Institutes and of Eurostat and, where appropriate, the heads of other statistical authorities have the sole responsibility for deciding on statistical methods, standards and procedures, and on the content and timing of statistical releases.
 - e. The statistical work programmes are published and periodic reports describe progress made.
 - f. Statistical releases are clearly distinguished and issued separately from political/policy statements.
 - g. The National Statistical Institute and Eurostat and, where appropriate, other statistical authorities, comment publicly on statistical issues, including criticisms and misuses of statistics as far as considered suitable.
 - h. The appointment of the heads of the National Statistical Institutes and Eurostat and, where appropriate, of other statistical authorities, is based on professional competence only. The reasons on the basis of which the incumbency can be terminated are specified in the legal framework. These cannot include reasons compromising professional or scientific independence.
9. Principles relating to *production* and *dissemination* deal with whether the available statistics meet users' needs and whether the statistics comply with European quality standards and serve the needs of European institutions, governments, research institutions, business concerns and the public generally. The important issues concern the extent to which the statistics are relevant, accurate and reliable, timely, coherent, comparable across regions and countries, and readily accessible by users.
10. For instance, Principle 12, covering outputs, relates to whether the European Statistics accurately and reliably portrays reality. It is assessed in the basis of three indicators.
- a. Source data, intermediate results and statistical outputs are regularly assessed and validated.
 - b. Sampling errors and non-sampling errors are measured and systematically documented according to the European standards.
 - c. Revisions are regularly analysed in order to improve statistical processes.
11. Compliance to the Code of Practice principles is assessed through a self-regulatory approach. Peer reviews focusing on the institutional environment and dissemination (principles 1 – 6 and 15 respectively) were conducted in 2006 – 2007 in the Member States and the EFTA countries. The exercise comprised a self-assessment by the National Statistical Institutes (NSIs), and peer reviews conducted by experts from other National Statistical Institutes and Eurostat. The peer reviews were organized during 3-days visits on-site during which the review team met with NSI representatives including NSI top and middle management and junior staff, as well as

with the main stakeholders, comprising other data producers, various user groups and respondents' representatives. The exercise resulted in a series of reports covering all reviewed countries², improvement actions by the NSIs³, and a set of good practices⁴. The Commission presented a final report on the exercise to the European Parliament and the Council in 2008. It was concluded that Eurostat's monitoring of the implementation of the Code of Practice based on peer reviews and country reporting was considered effective and proportionate.

12. A further round is due to start in mid-2014.

Conclusions

Benchmarking – the UK experience

13. A number of factors can contribute to successful benchmarking and a number of more general observations can be usefully be made for the benefit of those statistical offices who might wish to consider benchmarking:

- a. Initial discussions between benchmarking partners, prior to the process itself, provide a useful tool to identify issues through informal self-evaluation for the identification of potential.
- b. The value of benchmarking is not restricted by the limited range of performance indicators that may be available. Discussions can extend beyond the directly measurable to why two national statistical offices adopt different approaches to some aspects of index construction. For example, an exchange of views on some methodological issues including use of different formulae for elementary aggregation.
- c. Direct benefits more often than not continue to accrue beyond the benchmarking exercise. For example, from follow-up action points. Benchmarking can be seen as the beginning rather than the end of the process of improvement.
- d. Correspondingly, the further exploration in greater detail of issues raised during the course of the initial benchmarking can pay dividends. The general experience is that the work becomes more focused on specific issues as the benchmarking progresses and the issues of concern became more apparent.
- e. Longer-term benefits also include the subsequent opportunity for networking.
- f. Financial and management information compiled specifically for benchmarking can be useful management information in its own right. Performance indicators are a necessary ingredient of the process of continuous improvement and are not just short-term management tools.

² http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/code_of_practice/peer_reviews

³ http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/egreffe_adopted.pdf

⁴ http://epp.eurostat.ec.europa.eu/portal/page/portal/ver-1/quality/documents/PRS_ESSGOODPRACTICESJUNE08.pdf

- g. A stage can be reached where there are diminishing returns associated with the collection of additional data and greater benefits accumulate from further analysis and investigation of the data already accumulated. For example, the resources put into producing more detailed financial information may add limited value and may not influence recommendations. The challenge is in recognising when this stage is reached.

14. More generally, the success of the benchmarking project depended on a great deal of trust and mutual respect without which the discussions would have suffered from inhibitions over sharing detailed critical comment on performance with a third party.

Benchmarking – Eurostat’s peer reviews

15. Four characteristics have probably contributed to the success of Eurostat’s peer reviews.

- a. *Professionalism*. Competence & objectivity.
- b. *Helpfulness*. Making useful and beneficial suggestions. It is usually much harder to suggest how to fix problems than identifying them.
- c. *Realism*. Not setting standards that are too high, whilst not compromising compliance to Eurostat regulations.
- d. *Engagement*. Engaging in a dialog with the National Statistical Institute that is being reviewed.

16. In essence, the reviewer has been able to put themselves into the shoes of the reviewed.

The views expressed are those of the authors.