TRUSTWORTHINESS, QUALITY, VALUE: WHAT ASSURANCE DO INDEPENDENT ASSESSMENTS OF CODE-COMPLIANCE GIVE?

Richard Laux

UK Statistics Authority, London, UK; Richard.laux@statistics.gov.uk

Abstract

This paper explores the role that independent assessments of statistical systems – including the ESS peer reviews - have in enhancing trustworthiness. It argues that such reviews are necessary but not sufficient to make the most of investment in official statistics – other steps are also needed (at the European and national level); and that while a focus on trust is important, the ESS should be building on the foundations of trust and quality, and concentrating increasingly on ways to enhance the value of European Statistics.

Keywords: Trust, peer-review, assurance, public-value.

1. Introduction

This paper has both a national (UK) and an international (European) focus; it draws on emerging thinking in the UK which appears relevant to the European Statistical System (ESS) and which might resonate across the wider international statistical system. The paper refers to statistical codes of practice, and the way in which compliance with such codes is evaluated. In the UK, this means the Code of Practice for Official Statistics[1], and the statutory assessment function[2]; in the ESS the equivalents are the ES Code of Practice[3] (ESCOP) and Peer Review[4].

---

[1] https://www.statisticsauthority.gov.uk/monitoring-and-assessment/code-of-practice/
[2] https://www.statisticsauthority.gov.uk/monitoring-and-assessment/assessment/

Section 2 distinguishes between the concepts of *trust* (as applying to an organisation) and *quality* (as applying to a set of statistics). Section 3 introduces the concept of the *value* of sets of statistics, setting out two strands – *importance* (relative benefits against cost: how valuable is a particular set of statistics?), and *utility* (how can the public value of a particular set of statistics be enhanced?). Section 4 reviews the extent to which levels of trust, quality and value are evaluated in independent assessments of statistical codes of practice. Section 5 concludes that:

- The concept of value should be researched, defined, codified and operationalised. It should be articulated in ways that embrace official statisticians' innovations (such as the use of Big Data and 'digital' approaches to collection and presentation/dissemination) and that distinguish constructively between the roles of official statisticians and others in the numerical information business.
- Assessments of Code-compliance should evolve to provide assurance that, at an organisational level, statistical producers are ensuring that their statistics are inherently valuable and are managed in ways that make them increasingly useful.
- Assessments of Code-compliance should be aligned with and complemented by reviews of sets of statistics which explore both their quality and their value (utility).
- There may be merit in clear communication about what compliance with the ES Code of Practice means and, as part of this, the development of a 'brand' that supports the label of 'European statistics'. This should involve exploring whether the current assurance/review/evaluation mechanisms are sufficiently aligned with our vision for European statistics.

---

[3] http://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice
[4] http://ec.europa.eu/eurostat/web/quality/peer-reviews

2. Trust and quality: how do we conceptualise Code-compliance?

2.1 Trust (and trustworthiness)

Both the UK and European Codes include requirements – referred to as practices in the UK, and indicators in ESCOP - that apply at the organisational level, and requirements that relate to particular sets of statistics. For example, at the organisational level:

- "Ensure that those producing statistical reports are protected from any political pressures that might influence the production or presentation of statistics" (Principle 3 practice 2 of the UK Code);
- "Statistical work programmes are published and periodic reports describe progress made" (Principle 1 indicator 5 of ESCOP).

In relation to particular sets of statistics:

- "Provide information on the quality and reliability of statistics in relation to the range of potential uses …" (Principle 8 practice 1 of the UK Code);
- "Statistics are comparable over a reasonable period of time" (Principle 14 indicator 2 of ESCOP).

Official statisticians tend to equate Code-compliance with a rather ill-defined notion of 'trustworthiness' – so, one might argue that if a code of practice is being complied with, then the statistics to which it relates are 'trustworthy'.

This is helpful in so far as it addresses the adage that "if statistics aren't trusted, then they won't be used". But what does it mean to trust a set of statistics? Does it mean that you trust them to answer a particular question? That you trust them to be correct? That you trust them to have been compiled in a cost-effective manner? Deep down these questions have little meaning, because 'trust' and 'trustworthiness' are concepts that apply to organisations and individuals: what statements about 'trustworthiness' and 'statistics' are really saying is that the statistics are the product of an organisation that is trustworthy.

2.2 Quality

'Quality' is the analogous concept relating to sets of statistics. Official statisticians understand quality as a multi-dimensional concept which includes, for example: accuracy, precision, timeliness, relevance and accessibility (although in common parlance 'quality' tends to be equated with 'accuracy'). And statisticians understand that quality is a relative concept; that a set of statistics can be of high quality for one type of use, but inadequate for another use: fitness-for-purpose.

2.3 Trust and Quality

A trustworthy producer of statistics should be expected to produce statistics of a given level of quality and to describe that level of quality meaningfully.

An organisation whose chief statistician was politically appointed may not be inherently trustworthy, but the organisation might still produce high quality statistics. Conversely, an organisation that is highly trustworthy might produce some statistics that are of inherently low quality.

Levels of trust and quality have different types of drivers. An organisation might be trusted because it is well known, with an identifiable leader, a culture of transparency, strong statutory underpinnings and so on. A set of statistics might be high quality because the main uses to which it is put are well-understood and it has been invested in to meet the needs of most users.

3. The concept of value

3.1 Value and importance

It seems a reasonable assertion that not all statistics are equally valuable (important) in terms of the nature of the decisions that they inform. Most people would probably accept that GDP

statistics are inherently more valuable than statistics on collection rates for local authority taxes – although it is reasonable to note that for an individual whose role is to maximise local authority tax revenue, the latter set of statistics might well be more directly valuable.

In the absence of an agreed way of attributing value[5] to different sets of statistics through, for example, a cost-benefit analysis, we tend to use proxy indicators – such as whether statistics are required under legislation, or depending on the nature of the user (an approach which is always likely to favour major institutional users – such as Central Banks and Finance Ministries – over, for example, academics). Determining the relative value (importance) of different sets of statistics is more of an art than a science, and judgements about importance are likely to change over time and differ between countries.

Why does it matter? Other things being equal, we might want our most valuable (important) statistics to be of highest quality (so, value should influence resource allocation) and produced by the most trustworthy organisations.

Arguably this conceptualization of 'value' (importance) is closely related to one of the dimensions of quality, that of 'relevance'.

3.2   Value and utility

What distinguishes value in the current context is that we are also looking at ways in which an individual set of statistics, and the associated statistical service provided to users, might be made more valuable – more useful. The notion of public value is central to this. Public value[6] is akin to shareholder value in the private sector. It emphasises improvements to the quality of decision making by "calling for public managers to engage with service users and the wider

---

[5] Though see http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2016/mtg/CES_11_-ENG_G1602756.pdf
[6] https://www.statisticsauthority.gov.uk/wp-content/uploads/tempdocs/uk-official-statistics---from-new-public-management-to-public-value.pdf

public; it seeks to promote greater trust in public institutions and meet head on the challenge of rising expectations of service delivery". Public value can derive both by meeting the needs of government directly - with assumed benefits to society generally - and by explicitly meeting the needs of groups in society (this distinction may be referred to as *government value* and *societal value*).

The public value of official statistics can be enhanced in several ways, such as:

- An emphasis on analysis, explanation and interpretation – drawing out the main messages in a set or series of official statistics, supported by appropriate graphics - as opposed to simply producing a table of data.
- Conversely, the release of some sets of statistics - whose value derives from extreme timeliness – with very minimal commentary.
- Ensuring that detailed data are available and supported by sufficient metadata to enable designers of APIs to readily link different sources of information.
- The provision of a user insight (local intelligence) service, to improve producers' understanding of users' needs.
- An emphasis on coherence[7] within and between statistical domains, as opposed to the separate release of statistics on related topics with no attempt to read-across.
- An investment in informing significant areas of public debate, rather than in simply 'collecting and counting'.
- Improved data sharing and linkage, together with the analytical exploitation of the linked dataset.
- Working in partnership with organisations whose core skills complement those of official statisticians.

---

[7] Coherence is regarded as part of the ESS Quality Framework although, unlike other dimensions (such as relevance or accuracy), it is not an attribute of a single set of statistics.

This points to a conceptually simple hierarchical framework to characterise official statistics:

- Trustworthiness. All official statistics should be produced by organisations that are trustworthy.
- Quality. There should ideally be a positive correlation between the quality of a set of statistics and the importance of the decisions that the statistics inform – the more important the type of decision, then the more important the statistics are, and the higher their quality should be.
- Value. The inherent public value of any set of statistics can be enhanced by improving different aspects of their production, presentation and interpretation; and by improving the statistical service associated with the statistics.

4. Independent assessments of Code compliance

4.1 Peer Review

The stated objectives[8] of the latest (2013-15) round of Peer Reviews (PRs) were to:

- Enhance the credibility of the ESS
- Strengthen its capacity to produce high quality European statistics
- Further reassure stakeholders about the quality of European statistics and the trustworthiness of the ESS
- Assess progress made in compliance with the principles of the ESCOP
- Assess progress made in the development of the ESS

---

[8] http://ec.europa.eu/eurostat/web/quality/peer-reviews

Put another way, the PRs were focused on trustworthiness and, at an institutional level, quality. This latter point is noteworthy. Peer Reviews against the ESCOP are not intended to explore the quality of individual sets of statistics. Other tools – such as Eurostat's deep methodological reviews of National Accounts – are intended to serve this purpose. In a wider international context, "Global Assessments" attempt to explore both compliance with the standards set out in ESCOP and also to evaluate the quality of the major statistical domains (such as macroeconomic statistics, business statistics, and population and demographic statistics).

The concept of European statistics is generally understood to refer to those official statistics supplied by EU Member States to Eurostat, and subsequently published at an aggregate (eg EU28) level. However, despite the role and objectives of the PRs there does not appear to be a description of the characteristics – the "brand" – of European statistics such as, for example:

- "credible statistics for European decision makers"; or
- "high quality statistics about Europe produced by trustworthy organisations".

The focus of peer reviews is on European statistics – essentially, those sets of statistics which are subject to EU Regulations. As noted above, the legislative backing for the production and publication of these statistics is an indicator of their importance – in other words, one aspect of their value. But peer reviews have not tended to focus on public value/utility - the ways in which National Statistical Institutes and other (national) statistical authorities produce and disseminate their contributions to European statistics.

4.2   Assessment

In the UK model there is a subset of official statistics, assessed against the Code and, once found to be Code-compliant, labelled as "National Statistics". In the first four years of the Authority's Assessment function (2009-2012), during which all 'legacy' National Statistics – that is, those previously selected by the producers as meriting the label – were independently assessed against the Code, the "National Statistics" label was taken to mean 'fully compliant with the Code'. More recently the Authority has developed its thinking about the National

Statistics brand, which it is now referring to as signifying sets of statistics whose producers meet the highest standards of trustworthiness, and which themselves are of high quality and of high public value.

As part of a stocktake of the UK's Code of Practice[9], the Statistics Authority is exploring whether the formulation of the National Statistics brand might be better supported by the Code and the assessment function. Two aspects are relevant here:

- The UK's Code and assessment function enable an evaluation of trustworthiness, of quality (as managed at an institutional level) and, to a certain extent, of the quality of different sets of statistics. The Statistics Authority's assessments over the last few years in particular have increasingly focused on this latter question, and the recommendations of the Bean Review[10] will stimulate further work to deepen the way in which assessment explores the quality of individual sets of statistics.
- It seems likely that there is a significant element of the stock of National Statistics that are trustworthy, and of high quality - but which are not as valuable, either in the sense of importance or of public value (utility). One strand of thinking in the Code stocktake is whether it might be possible to better align the National Statistics brand with the stock of National Statistics – for example, by identifying those statistics which are not important or are not of high public value and enabling their voluntary de-designation as National Statistics.

5. Conclusions

This paper has outlined a conceptual hierarchy of trustworthiness, quality and public value. While it has set out some aspects of value, further work is needed. The concept of the public

---

[9] https://www.statisticsauthority.gov.uk/monitoring-and-assessment/code-of-practice/code-of-practice-stock-take-latest-news/
[10] https://www.gov.uk/government/publications/independent-review-of-uk-economic-statistics-terms-of-reference/independent-review-of-uk-economic-statistics-terms-of-reference

value of official statistics should be researched, defined, codified and operationalised – perhaps learning from the way that 'coordination' has been approached in the context of the latest round of ESS peer reviews, with the development of an initial set of principles and subsequent 'soft' evaluation as an add-on to peer review. Such an approach to *value* may go beyond the current Code, and should allow for national specificities. It should be articulated in ways that embrace official statisticians' innovations (such as the use of Big Data and 'digital' approaches to collection and presentation/dissemination) and that distinguish constructively between the roles of official statisticians and others in the numerical information business.

Assessments of Code-compliance are undoubtedly important in providing specific types of assurance to stakeholders – funders, and users. However, such assessments should evolve to provide assurance that, at an organisational level, statistical producers are ensuring that their statistics are inherently valuable and are managed in ways that make them increasingly useful.

Assessments of Code-compliance have tended not to systematically explore the characteristics of particular sets of statistics in detail. Assessments could usefully evolve by aligning them with, and complementing them by, reviews of sets of statistics which evaluate both their quality and their value (utility).

There may be merit in clear communication about what compliance with the ES Code of Practice means and, as part of this, the development of a 'brand' that supports the label of 'European statistics'. At a practical level this might help to distinguish between statistics produced by Eurostat and those produced by other parts of the Commission. But perhaps there is even more to gain by deciding what we think are the desirable characteristics of the 'brand' of European statistics, and then exploring whether the current assurance/review/evaluation mechanisms are sufficiently aligned with our vision for European statistics.