

Annex B: International approaches to audit arrangements and official statistics

Introduction

The National Statistician has prepared a guidance document¹⁵ around the use of administrative data. It describes the statistician's role in assuring and communicating the quality of administrative data for producing official statistics. It sets out the broader quality assurance context for determining the suitability of the administrative data, and emphasises the importance of: investigating the data source, establishing good relations and processes for the supply of the data, validation checks, and preparing clear explanations of the associated quality issues. It highlights the need to understand the audit arrangements of the administrative data within the critical questioning about the data source.

This report seeks to articulate the regulatory standard expected by the Authority on quality assurance and audit arrangements by, among other things, exploring international approaches. We reviewed some material published online from national statistical institutes (NSIs) and Eurostat to find examples of how audit features in their quality assurance approaches and how these are informed by the underlying quality frameworks.

ESS Quality Assurance Framework

Eurostat and its member states use the ESS Quality Assurance Framework¹⁶ (ESS QAF) with its six dimensions of quality: relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability, and coherence. Accuracy denotes the closeness of computations or estimates to the true values.

Eurostat emphasises the importance of providing quality reports to accompany official statistics, describing how each dimension is met. It has produced guidance¹⁷ to NSIs on the content of quality reports. It states that the accuracy section in quality reports should describe the methodology used, and identify the key variables and the main sources of random and systematic error of the statistics. An assessment of bias can be either described in quantitative or qualitative terms. It should set out the main sources of potential bias and reflect actions taken to reduce bias.

ESS QAF identifies some of the quality indicators that can be produced for administrative sources:

- Coverage, i.e. over- or under-coverage of the population that should have been included according to the definition or classification used

¹⁵ <https://gss.civilservice.gov.uk/wp-content/uploads/2012/12/Interim-Admin-Data-guidance.pdf>

¹⁶ <http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/introduction/>

¹⁷ http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/quality_reporting

- Errors in register variables – similar to measurement error in a survey (errors occurring during the collection)
- For event-reporting systems, an estimate of the rate of unreported events

ESS QAF recommends the preparation of a quality assurance plan in which statistical producers should:

- Assure the quality of data collection, including the use of administrative data
- Ensure the systematic examination of possible trade-offs within quality

These elements of the quality assurance plan should include the assessment of the collection arrangements within the data supplier organisation and encourage the proactive role of statisticians in influencing the system development to support the use of the data in producing official statistics.

ESS QAF identifies some specific activities that can be undertaken in relation to quality reporting to comply with Indicators in the ESS Code of Practice:

- Indicator 4.4 (the regular and thorough review of the key statistical outputs using external experts where appropriate) highlights the need for a plan for implementing quality reviews such as auditing and self-assessment for statistical outputs. This reference to auditing appears to be in relation to the conduct of peer reviews on the statistical outputs as opposed to the internal or external audit of the underlying data source. We found a number of examples of NSIs using the peer review approach for the auditing of statistical outputs, including Statistics Canada, Italy and Sweden. ONS has had a programme of National Statistics Quality Reviews that involved external experts and experts within the Government Statistical Service. The programme was recently reintroduced by the National Statistician, commencing with a review of the ONS Labour Force Survey¹⁸.
- ESS QAF also identified the need to carry out regular updating of business registers to ensure that the register reflects the changes in the population of businesses (under Indicator 7.3) and to conduct external evaluations of the methods used with external experts (under Indicator 7.7).
- In relation to administrative data the ESS QAF specifically highlights the distinction between statistical and administrative data processing and the use of appropriate validation rules and specific procedures for checking quality, in using the administrative data for statistical purposes (under Indicator 8.1).

Each of these elements of ESS QAF is relevant to the quality assurance of administrative data by statistical producers. However, these do not clearly set out the need for determining the assurance processes and robustness of measures taken by

¹⁸ <http://www.ons.gov.uk/ons/guide-method/method-quality/quality/quality-reviews/list-of-current-national-statistics-quality-reviews/nsqr-series--2--report-no--1/index.html>

data supplier organisations in the collection and processing of the administrative data.

Quality assurance arrangements in National Statistical Institutes (NSIs)

Our review of some material on the quality assurance approaches in various NSIs highlighted the importance of validation checks being conducted by statistical teams within the statistical producer organisations. The guidance appears in line with the ESS approach to reporting quality, identifying the steps to check and validate the data and statistics by the statistical producer. We found two NSIs – Statistics Netherlands and US Census Bureau – that specifically refer to the need for data supplier organisations to audit the administrative data and for the statistical producer bodies to assess the quality of the data identified as a result of the supplier's assurance activities.

Statistics Netherlands has developed a quality framework for administrative data sources¹⁹ for NSIs. It highlights three 'hyper-dimensions' of quality: Source, Metadata and Data.

- 'Source' reviews the quality aspects of the data source as a whole, the data collection organisation and the delivery of the data to the statistical producer. The detailed dimensions are: supplier, relevance, privacy and security, delivery and procedures
- 'Metadata' describes the quality aspects relating to the information required to understand and use the data: clarity, comparability, and data treatment by the data source keeper
- Data involves quality aspects that are mainly related to the accuracy of the administrative data such as over-coverage, under-coverage, unit non-response, item non-response and measurement – which includes activities such as 'external check: has an audit or parallel test been performed?'

It developed a checklist to assist NSIs in the assessment of the quality of the administrative data sources that covers the first two hyper-dimensions which both include some aspects of the assurance arrangements of data suppliers. The information is mainly captured through a structured study of data quality indicators under the 'Data' hyper-dimension.

The US Federal Committee on Statistical Methodology, the US Census Bureau, has also developed a tool²⁰ to assist in the assessment of the quality of administrative data. Its Data Quality Assessment Tool for Administrative Data provides a set of questions that can aid data providers in evaluating the quality of its administrative data, as well as assisting users (or the statistical producers) in determining the

¹⁹ <http://www.cbs.nl/NR/rdonlyres/0DBC2574-CD4E-4A6D-A68A-88458CF05FB2/0/200942x10pub.pdf>

²⁰ www.bls.gov/osmr/datatool.pdf

suitability of the data for an intended use. The tool comprises 43 questions to prompt the investigation of the data quality. It also uses a quality framework, with six dimensions: relevance, accessibility, coherence, interpretability, accuracy and institutional environment. 'Accuracy' has the same meaning as used in the ESS QAF. In relation to the assurance and audit of the underlying administrative data the US tool specifically asks:

- What investigations/analyses have been conducted that reveal data quality characteristics (such as Government Accountability Office reports, Office of Inspector General audits, internal agency reviews etc)
- Describe the checks the administrative agency performs to ensure the quality of the data and the typical results for your production processes
- Describe the principles, standards, or guidelines the agency uses as a basis to determine what is considered acceptable quality
- Describe the findings and corrective actions of studies, evaluations or audits to assess compliance with quality standards

The US tool is particularly helpful in highlighting the information that official statistical producers can obtain about the underpinning administrative data sources to better understand the quality implications for the statistical outputs. This information could be used in conjunction with the National Statistician's interim good practice guidance document on quality assuring and using administrative data (circulated to Heads of Profession for statistics in April 2014).

These tools have been influential among NSIs seeking to develop quality indicators for administrative data. In particular, Statistics Netherland's tool has informed the development of a quality indicator instrument as part of the ESS network ('ESSnet') project on administrative data²¹ focusing on business and trade statistics. It is intended to be used by NSIs in routinely verifying the quality of administrative data received from data suppliers, and includes indicators for each of the ESS quality dimensions. The quality indicator instrument will be useful for statistical producers conducting their own quality assurance of the data received from suppliers, particularly in highlighting possible errors for further investigation.

The ESSnet project also identified checks to be conducted when first considering the use of an administrative data source for producing official statistics²². The indicators highlight the need to be clear about what is needed from the statistical use of the data and to build a comprehensive understanding of the administrative source. They suggest that producers compare definitions between the need and the source, and determine the coverage, such as geographical area, reporting delays, and thresholds

²¹ <http://essnet.admindata.eu/WorkPackage?objectId=4257>

²² <http://essnet.admindata.eu/WorkPackage?objectId=4252>

in size or quantity. Producers can also contact data suppliers and seek expert opinion from those managing the collection process, to become acquainted with practical aspects of the collection. The project also emphasises the benefits of comparing with other data sources, to gain a more objective view, particularly with a trustworthy source. It may be possible to use methods such as capture-recapture procedures to estimate duplicates in the administrative data. Also visual inspection will enable outliers and/or processing errors to be identified, through charting the data and making comparisons of distributions.

Findings

- ESS QAF provides some broad guidance on determining the quality of administrative data
- Activity in the national statistical institutes (NSIs) for member states has tended to focus on peer review of statistical outputs rather than checking the auditing and assurance arrangements within data supplier organisations
- Statistics Netherlands and the US Federal Committee on Statistical Methodology have developed tools to assist statistical producer bodies in determining the quality of administrative data
- The US tool is a helpful steer for official statistical producers when gathering relevant quality information from data suppliers about administrative data and forming a judgment of their suitability
- The ESS network project on administrative data for business and trade statistics has produced guidance for producers on making the initial judgment about the suitability of the data and a set of quality indicators that can be incorporated into routine checking of input administrative data

Wider lessons

The need to gather information about the auditing and assurance arrangements of administrative data suppliers has not been widely promoted among NSIs. However, some guidance is available for statistical producers to understand these arrangements and inform the use of the administrative data. The US quality assurance tool provides a useful starting point for investigating audit arrangements.

We have drawn on the audit items from the US tool, together with elements of good practice identified in the six case studies (Annex C) and from the ESS network project, to form a model describing the practice areas to consider when reviewing the quality assurance and audit arrangements for administrative data. It has informed a QA maturity matrix (see Part 4) which is designed to assist statistical producers in critically evaluating the robustness of the data for use in producing official statistics.