Proceedings of Q2006
European Conference on Quality in Survey Statistics

Using Quality Measures to Manage Statistical Risks in Business Surveys

Paul Schubert, Tenniel Guiver, Robyn MacDonald and Frank Yu[1]

Abstract

Identification, measurement and monitoring of key quality measures at different stages of the survey process aid the survey manager in assuring quality or in identifying potential quality problems. In the first part of this paper, we describe recent developments in the Australian Bureau of Statistics extending the use of quality measures through the concept of quality gates, which can be seen as validation points in the end to end statistical process. The key quality measures are assessed and measured against acceptable limits or benchmarked against set targets. Throughout the statistical process, there will be a number of validation points or gates which should be 'passed' before moving on to the next stage. This information is also useful for investigating quality issues which may emerge during the output analysis and clearance phase before data release.

The second part of the paper describes the way our systems are being enhanced and integrated in order that the quality measures used as inputs to the gates are produced as a by-product of the various (e.g. editing, estimation) systems, and piped automatically to a central Quality Infrastructure System repository where they are easily accessible to users. Experience in registering standard measures and implementing the system will be discussed.

Key words: quality measures, business surveys, statistical risks

1. Introduction

Estimates or other statistical outputs from sample surveys are the culmination of a series of often-complex processes. Once the requirements for a survey are known, the frame from which the sample has to be constructed; the sample, collection instruments and processes designed; the data collected and processed; the raw data transformed into aggregates or other outputs; the outputs analysed, scrutinised and adjusted where necessary; and the final outputs published and disseminated. In each of these

[1] Paul Schubert, Australian Bureau of Statistics, Locked Bag 10, Belconnen ACT Australia 2616, paul.schubert@abs.gov.au; Tenniel Guiver, Robyn MacDonald, Frank Yu all same address, tenniel.guiver@abs.gov.au, robyn.macdonald@abs.gov.au, frank.yu@abs.gov.au

processes, there is potential for errors to be introduced which could affect the quality of the final outputs and lead to incorrect analyses, conclusions and decisions by users.

Various quality measures or indicators have long since been used to assist survey managers in assessing and assuring quality of key components of the survey process. For example, the number (or estimated number) of units that are alive on the survey frame (especially compared with previous time points) tells us something about the quality of the frame; monitoring the response rate during data collection is useful for initiating when action may be required if it is too low in some parts of the sample; the contribution of imputed values or outliers to the final sample estimates are an indicator of potential bias.

As do other statistical agencies, the Australian Bureau of Statistics (ABS) has a formal output clearance process prior to the release and dissemination of publications and other statistical output. A number of the quality measures collected during the collection process are useful in this clearance process for assuring managers that there are no unresolved issues affecting the integrity of the statistics, or for identifying any problem areas that can be addressed before they are publicly released. Our experience is that it is not uncommon for quality issues or potential quality issues to be identified at this stage. Occasionally, errors are only discovered after the data have been released, leading to the need for the outputs to be corrected and reissued, and potentially damaging users' confidence and trust in the reliability and integrity of the statistical agency.

Many of the quality issues could be identified much earlier in the collection process. There are a number of reasons why they may not be: the relevant information which would identify the problem may not be available, or not easily accessible; there may not be sufficient understanding within the relevant area that the quality measures are indicating a problem; or the relevant information may not be looked at or assessed until later in the process.

The ABS has been working on a number of initiatives to help us better manage what we term *statistical risk* - the risk that something untoward in the collection process affects the integrity of the statistical outputs. Much of this work has focused on business surveys, so the examples given in this paper will be focused on economic collections; however, the same principles could easily be applied to population and household surveys. In section 2 of this paper, we will describe our development of *quality gates*, which are specific validation points within the collection process. In section 3, we will describe the Quality Infrastructure System (QIS) that is being developed to help capture, store and disseminate quality measures as a by-product of other statistical processes, in order to feed into quality gates or other quality assurance activities.

2. Quality Gates

2.1 The concept of quality gates

The concept of quality gates is used in other fields. For example, it has been applied in the automotive industry and forms part of the quality standards QS-9000, a derivative of the ISO-9000 quality series standards (International Automotive Sector Group, 2003). It has also been applied in the Information Technology field, particularly in IT development projects where quality gates have been applied at the end of each of the planning, design, development and deployment phases of the projects (Charvat, 2003).

Generally, a quality gate is a checkpoint consisting of a set of predefined quality criteria that a project must meet in order to proceed from one stage of its lifecycle to the next (Workshop on Cross-Lingual Information Retrieval, 2004). We are applying the concept a little more loosely in the ABS, as the ‘acceptance’ may be a subjective judgment based on a mixture of quantitative and qualitative information; nevertheless, a quality gate can be seen as a validation point in the end to end statistical process. Throughout the statistical process, there will be a number of validation points which must be 'passed' before moving on to the next stage.

Quality gates can be used to improve the visibility of quality in the statistical processes as well as being used to measure and monitor quality in real time at various points in the end to end statistical process.

2.2 Developing quality gates

As stated above, the concept of quality gates is akin to an acceptance criteria imposed at predetermined points in the end to end statistical process. The quality gate consists of a set of indicators to sign off on. These indicators may be qualitative or quantitative and need to be agreed on by all stakeholders; that is, those who are signing off on the quality and those who are receiving the quality product.

2.2.1 Using quality measures in quality gates

The range of quantitative quality measures utilised in quality gates are used to quality assure surveys both within the collection cycle and at the end-of-cycle. These are useful tools in monitoring the quality and progress of the survey to sound the alert should a quality incident be looming.

Quality measure categories include frame quality, response rates, adjustments to data, number of revisions, estimates, and data capture. A core set of quality measures has been developed (Guiver, 2005) to assist in the real time monitoring of the quality of the collection. There will also be a more comprehensive list to cover more extensively the quality measures needed that weren't readily available to be captured in the initial phase. This core set of quality measures will be developed in negotiation with a range of stakeholders to ensure relevancy. These quality measures include within cycle and end

of cycle measures.

The judgment and experience of the survey manager is required in deciding which quality measures to monitor. Deciding on the suite of quality measures for a particular survey should be done in consultation with survey methodologists and the senior managers who will be signing off on the quality of the collection as fit for purpose. It also comes into play when deciding whether the data is signalling a quality incident or is moving within an acceptable control limit.

The measures are often more useful when presented as a time series e.g. in graphs. This enables the measures for the current period to be assessed in context of what has happened in previous periods e.g. identifying whether a response rate in the current period is unusually low for a particular collection.

2.2.2 Placement of quality gates

In determining which part of the process should be monitored by using quality gates, the survey manager and stakeholders should base the decision on a top down assessment of the statistical risks associated with the process. They need to be positioned to take into account critical issues or decisions. Processes which are subject to higher risk of quality degradation, or known ‘quality hotspots’ should be monitored more intensely than lower risk ones.

When positioning the quality gates, the survey manager needs to remain cognisant of the time lines around the survey balanced with good quality management. The focus should be on a small number of key measures for each process. The introduction of too many gates, or gates at inappropriate junctures will only serve to slow the process down and may ultimately devalue all quality gates.

The placement of a quality gate may be different for each survey or collection. There will also be different gates at different levels. For example, at the overall collection level there may be a quality gate between each of the broad phases of the survey cycle. At the same time, each of the phases may have quality gates within lower level activities or tasks, to be monitored and managed by local operational managers. Within the process of acquiring data, for example, there is frame preparation and sample preparation. It may be that there will be a quality gate signifying that frame preparation is completed and signed-off and sample selection can begin. For other surveys, the quality gate may only be after the sample preparation.

Business analysis and process mapping is a useful tool to map out the processes at the various phases of the statistical process. Mapping processes will help understand how a system works, and identify how a system interacts with other systems and processes. It will provide a simple conceptual framework to identify logical and key areas where quality gates can be used to monitor quality.

2.3 Examples of quality gates

2.3.1 Example 1: Sample frame and selection

A quality gate could be inserted at the point in the collection process where a sample has been selected from a frame for a particular survey, and the survey is about to enter the enumeration phase. In many statistical agencies, there may be a different area responsible for the data collection than for the construction of the frame and drawing of the sample, so it would be appropriate that the inputs of frame and sample to the next part of the statistical process are quality assured and signed off by the key stakeholders.

In this case, possible quality measures to be used as part of the quality gate could include:

- original frame size in number of units (perhaps broken down by key characteristics, such as industry, region, or size of business) and percentage changes from previous frame (if for a repeated survey)
- benchmark totals for frame (again, could be broken down by key characteristics) e.g. if administrative data on the frame for sample selection or estimation are to be utilised
- number of units selected in sample (by key characteristics)
- units common to previous sample (count and rate)
- units rotated into sample for first time (count and rate)
- units rotated out of sample (count and rate)
- units in previous sample but now not on frame (count and rate).

Other qualitative information used at the quality gate could help to inform the survey managers and stakeholders involved in the quality gate sign-off of particular issues or reasons why some of the measures have changed markedly from their usual levels. This could include information about particular events that may have influenced the current frame size such as strong economic growth or decline or changes in processes in source agencies for administrative data used in the frame, whether (and how) the sample design has changed for the current period, etc.

2.3.2 Example 2: Response rates

At a strategic level, there may simply be a quality gate at the end of the enumeration period which utilises quality measures such as response rates, which could be assessed against a target for acceptability (this target may vary from survey to survey).

However, in the area responsible for the data collection, the local manager may have a series of quality gates - for example, at the end of each week - to monitor the response rate. It may be assessed against a particular target, or an acceptable range may be prespecified. If the response rate is falling below the target or outside the

acceptable range, it would be appropriate to investigate the possible causes and initiate remedial actions.

3. Quality Infrastructure System

3.1 Introduction

For a quality gate to be effective, the information must be readily available to the survey manager. This information needs to be accessible within the statistical cycle and at the end of the cycle as well as being flexible enough to allow the survey manager to manipulate the data and to set appropriate targets or limits. Lack of availability of measures, or the fact that it often takes extra resources and effort to produce them has been a barrier in the past to regular use of quality measures by survey managers.

In the ABS, the Quality Infrastructure System (QIS) which is intended to deliver quality measures to the desktop. QIS is an integrated infrastructure that allows the capture, storage and use of quality measures over the end-to-end business processes of all ABS data collections. QIS is a key enabler of making the quality of statistical processes and products visible within and between collections and collection cycles, and will facilitate the use of quality gates and other quality assurance activities. QIS will allow quality measures to be readily accessible so users can understand and improve processes, report on quality and make quality declarations about data.

The goals of the QIS are to embed the production of quality measures into the whole collection process infrastructure so that they are produced and captured as by-products of the detailed statistical processes, and to automate (or make available on demand) delivery of these measures to survey managers and other stakeholders.

3.2 The QIS Quality Measure Repository

Central to the development of QIS is the production a repository for the storage of quality measures. The repository is built around a ‘star schema’ data warehouse. This model allows for a variety of different types and levels of quality measures to be stored for both economic and household collections. A star schema is composed of a central ‘fact table’, a database table that contains the observed values of all quality measures and series of ‘keys’ that link each quality observation or ‘fact’ to a number of different ‘dimensions’. These dimensions are themselves database tables, containing the keys and classificatory details. Dimensions used in the QIS repository include: quality measure (e.g. response rate, standard error), geography (e.g. state, part of state) and industry. Wherever possible standard classifications are used; for example the industry dimension uses the Australian and New Zealand Standard Industry Classification (ANZSIC).

Centralised storage of authoritatively defined and accessible quality measures allows for investigations of quality within collections, between collection cycles, and between

collections using relevant and understandable measures. It also provides for the efficient and timely identification and resolution of quality related issues. Another benefit of a centralised store, like QIS, is that quality measures are unambiguously and authoritatively defined, which is the process of registration. This ensures the quality measures are fit for purpose and sound methodologically.

The QIS project has had a strong metadata focus. Standard data item definitions and classifications are used to define the dimensions in the QIS repository. While quality measures are a class of metadata in themselves, they also require their own definitional and operational metadata.

Around the central QIS quality measures repository are a number of ‘services’ that enable quality measures to be sourced from existing systems (e.g. provider management systems, estimation systems), and to be disseminated to users either within the statistical agency or (potentially) to external users via publications or quality statements e.g. on the agency website.

3.3 Quality Measure Storage Services

Two different service types have been developed to store quality measures. The PutQM service enables existing systems to send quality measures directly to the QIS Repository. Existing donor systems undergo minor modification to enable them to send quality measure mail messages to the ABS business process management infrastructure. A ‘subscription service’ monitors this process management infrastructure and automatically loads quality measures to the repository as they arrive.

The LoadQM services are the services that source quality measures from existing systems and load them to the QIS Repository. LoadQM services, independently from existing systems, access data stores these existing systems create and load these measures to the QIS Repository. These services do not require any re-engineering of donor systems.

The PutQM services are preferred as they are based on a single integrated system. It is envisaged that in the future, requirements to provide information to QIS will form part of the specifications of any new statistical systems built so that the PutQM service can easily be incorporated as minimal marginal cost.

3.4 Quality Measure Access Services

QIS will deliver information to the desktop. The information will allow the survey manager to make informed decisions about the quality of the collection at a point in time, at end of cycle and between cycles. The survey manager will be able to set parameters around the quality indicating limits or targets.

It will be up to the survey manager, in consultation with survey methodologists and key stakeholders, to decide what quality measures to monitor and have accessible from the

desktop. A minimum will be a core set of quality measures that has been agreed as standard across all ABS economic collections, as well as any other measures as identified by the survey manager with a focus on a small number of key measures.

As with the storage services, there are two generic access service types in the QIS that have been generated to allow standardised access to the repository. The StoreQM service allows quality measures to be automatically copied from QIS to other parts of the ABS Corporate Metadata Repository. For example, the Collection Management Systems is a centralised store that records details on the concepts, sources, methods, timing, collection procedures and output of each ABS collection. The StoreQM service allows for the relevant quality measures to be included in this documentation.

The other generic access service type is the output services, utilising business intelligence software such as Oracle Discover and SAS, to deliver regular and automated reports on quality e.g. response rates over time, production of tables and data for clearance documentation. These can be used by business areas as part of the signing off the output of a data collection (e.g. in quality gates). The same software can be used for more sophisticated analyses of the quality data stored in the QIS repository. For example, these analytical systems allow the user to explore the quality measures, drilling down to finer levels of details or making comparisons across data collections and collection cycles.

3.5 Experiences in implementing QIS

At the time of writing, QIS is still in the early stages of implementation, being trialled for a small number of collections. However, the many of the services being developed will then be able to be easily implemented for other surveys e.g. the PutQM services being developed to enable existing corporate systems to send quality measures to the QIS repository will also be able to be used for other collections using the same corporate systems. Nevertheless, there are already some lessons emerging from the development and implementation.

The first is that some access control issues have been identified. Particularly when pre-published data are under embargo, there is a need to be able to restrict access to the quality measures to only those areas involved in the collection. More generally, some areas have been concerned that wider (internal) access to the quality measures without an understanding of the particular issues associated with a collection may lead to uninformed statements about survey quality. This issue is also important for external dissemination; it is important to provide explanatory material along with any quality measures presented so that users understand the implications.

A major challenge in the development has been in developing the necessary linkages between different systems, which may be operating on different platforms and utilising different software, and QIS. The cost of connecting these systems is another driver for collections to utilise generic systems rather than systems developed uniquely for their own collections. Nevertheless, if there is a need for a new special-purpose system to be

built in the future, the need for production of the quality measures can be done cost-effectively if considered in the design of the new system.

The format and software utilised for the output services has been another area that has generated considerable discussion in the QIS development. It may take a little time to develop and refine ‘best practice’ in this area.

A key issue to address in any similar development is to obtain ‘buy-in’ from collection areas. Our experience is that they often need some convincing that the initial investment in defining quality measures appropriate for their survey and doing any system-related work to enable to quality measures to be captured by QIS will be more than repaid over time by the ability to monitor quality measures in real-time and their usefulness in operational decisions. Our strategy has been to gain corporate support from the executive level, demonstrate the usefulness through implementation in a small number of key collections, and working with each area to define measures and output which will be useful in each specific case.

Finally, there is a need for centralisation of the quality measures registration process. Many areas currently monitor various measures and management information which cover similar concepts, but there are differences in the definitions used. Our methodology area is acting as the registration authority. The registration process includes a consultation period where proposed measures are circulated to stakeholders and comments considered, before they are ratified and tabled at a committee overseeing survey integration issues. The registration authority will also have a role in strongly enforcing that the measures, when used, conform to the agreed standards. This is important in making the quality of statistical processes and products visible within and between collections and collection cycles

References

Charvat, Jason P. (2003), “How to use quality gates to guide IT projects”, available from http://builder.com.com/5100-6315-1061893.html
Guiver, Tenniel (2005), “Core quality measures”, unpublished paper, Canberra, Australia: Australian Bureau of Statistics
International Automotive Sector Group (2003), “QS-9000”, available from http://www.qs-9000.org
Workshop on Cross-Lingual Information Retrieval (2004), “About Quality Gates”, available from http://www.uni-hildesheim.de/~lecliq/aboutQualityGates.php