Journal of Official Statistics, Vol. 17, No. 1, 2001, pp. 107–112

Comment

Dennis Trewin[1]

1. Introduction

The answer to the specific question raised by the article (Platek and Särndal) has to be **yes.** I would not like to take a defeatist attitude. But it is important to look at whom we are delivering to. In my view, it must be the end users of the statistics. Their needs are different of course but, in terms of the question posed in the article, I think we should concentrate on the more sophisticated users, who are both more knowledgeable and challenging on quality issues.

The article provides many messages. The ones I have chosen to comment on are as follows:

- Survey methodology is a set of practices with no unifying theory - such a theory is highly desirable.
- There should be greater use of Total Survey Error models.
- There are different views of quality – it is difficult for an individual statistician to feel motivated by all of the components.
- There is too much emphasis on sampling error in the presentation of information about quality. Nevertheless, a lot of the other quality information that can be provided is of little meaning to users.
- Trust in the national statistical agency is the criterion by which most people judge the quality of its product.
- A lot of new technical work (e.g., small area estimation) is underutilised.

2. No Unifying Theory for Survey Methodology

There is no unifying theory but does it matter. I think not. There are other ways of judging the appropriateness of different methods. I do not support the view that ''one cannot challenge any particular way of operating, because no firm theory dictates the choice.'' There are methods which are not reliant on a unifying theory, but more on experience and judgement.

For example, benchmarking is one effective way of developing best practice. We have started using this extensively in internal comparisons of our collections. National statistical agencies run many business and household collections so there is a lot of scope for internal comparisons. Benchmarking and ''best practice workshops'' are two means we have used to identify effective survey methods (both design and operations).

The Australian Bureau of Statistics (ABS) has also been facilitator of international benchmarking comparisons for consumer prices, household expenditure surveys and

[1] Australian Bureau of Statistics, Belconnen, ATC, Australia.

manufacturing surveys. A second round is planned. The strongest interest is in the differences between the methods used by the national statistical agencies. Some may be justified but discussions on the reasons for the differences are often revealing. We will be making important changes to the three collections mentioned above as a result of the first round of the international benchmarking studies.

For some functions, it is possible to make benchmarking comparisons with organisations other than international statistical agencies. For example, we participate in Gartner benchmarking studies of information technology costs. This is an international study. We also make comparisons of our corporate costs with other public sector agencies.

There are some functions where standard ''best practice'' procedures should be followed, based on common sense and research rather than a unifying theory. For example, in the ABS:

- We use a common framework for all our quarterly surveys; one of these four quarterly frameworks is also used for annual surveys.
- The same ''Common Survey Frame and Maintenance Procedures'' are used across all surveys with the support of a standardised software package.
- As a consequence of using a common framework, we use consistent approaches to estimating new business provisions.

These procedures have led to greater coherence between our business surveys to the benefit of national accounts and other users.

3. Total Survey Error Models

I agree that there should be more use of Total Survey Error Models. In the ABS, we made an attempt to model the errors of a Construction Industry Survey to help optimise the design across all types of errors (see Linacre and Trewin 1994). The design, optimised using Total Survey Error models, was radically different to that of the previous survey, which was based on sampling error considerations only.

However, use of the Total Survey Design approach is unusual. Why? First, I think many (survey methodologists and subject matter statisticians) find Mean Squared Error a difficult concept to interpret when explaining the accuracy of a point estimate. Second, and most importantly, the models are very difficult to specify.

It may be possible to take a Total Survey Design approach without having an underlying mathematical model. Judgement would be involved on the likely error contributions (possibly in ranges) for each component of the survey design and how these might vary with expenditure on that component. Of course, more precise estimates of accuracy will be available for some components (e.g., sample errors with an approximate relationship with sample size). The cost of each component, and how accuracy varies with cost, also needs to be known.

This approach was used in a National Account Measurement Error (NAME) study undertaken in the ABS. Experts were used to assess the likely error range for each component of the national accounts. The object of the exercise was to determine the relative effort that should be placed into the source data series that contribute to the national accounts. Unfortunately, this interesting piece of work was not published.

4. Different Views of Quality

The authors describe the different categories of statisticians found in a large statistical agency (theoretical statistician, survey methodologist, subject matter specialist, etc.). They argue that they have different views or perspectives on quality and it is difficult for an individual statistician to feel motivated by all the components. They see project teams, with representation from most of the categories, as the main means of communication between the different categories of statisticians.

For ongoing work, project teams are usually not created. It is far more important to create a culture of collaboration between the different categories of statisticians. We believe we have done this at the ABS. Things that have helped include:

- Strong recognition and support for the mathematical statisticians and survey methodologists. The Chief Methodologist is a member of the Executive Board and has a strong remit to provide leadership on quality issues.
- Methodologists are encouraged to transfer to other parts of the organisation, in part to ''spread the gospel'' but also because their skills are valued. A very high proportion of senior staff in the ABS have come from the Methodology Division.
- In recognition of this, mathematical statisticians are accepted as one of main graduate recruitment streams.
- All the categories are represented in the cross-cutting fora that are responsible for strategy and coordination activities associated with particular fields of statistics.
- There is an expectation that concepts, sources and methods manuals will be made publicly available for the different fields of statistics, increasingly in electronic form.
- Most of the categories of statisticians are represented at the clearance meetings before the release of the most important statistics. The mathematical statisticians have a significant influence on the way key statistics are presented and explained together with accompanying material on quality.
- More generally, a collegiate culture exists now and is reinforced whenever possible.

Through these initiatives the different views of quality are harmonised and a unified view provided to our users.

5. Too Much Emphasis on Sampling Error

This is undoubtedly true. They are relatively easy to calculate and will be easier to present in electronic form than on paper publications. There is a risk that there will be even more emphasis on sampling errors as cleverer ways of presenting them become available.

It does not have to be that way if all the different categories of statisticians are involved in presentations on quality of data. The quality of official statistics can be judged primarily by asking how well the data fit the uses to which they are put. Within the ABS, statistical quality encompasses the dimensions of relevance, accuracy, timeliness, accessibility and coherence. Australian Bureau of Statistics (1996) is a good example of the presentation of a range of information about quality, including many of the dimensions mentioned in the previous sentence. For example, there is a great deal of discussion about the behaviour of revisions over time.

The material presented on quality should be largely influenced by what the users of the statistics want. As the authors state, a lot of the quality information is of little meaning to users. They recognise that fitness for use and user orientations are important themes. Quite frankly, sampling errors would be among those components of quality that are not very meaningful to many users even though this is traditionally where there has been most emphasis. Their judgements on quality are judged to a far greater extent by:

(a) the general reputation of the national statistical agency,
(b) its openness in providing prior information (''no surprises'') on factors that might affect statistical series, e.g., methodology changes,
(c) sound explanations of unusual movements in statistical series,
(d) the history of revisions,
(e) the volatility of the series (in part influenced by sampling errors), and
(f) a willingness to listen to, and act on, informed criticism.

It is not difficult to provide useful information on (b), (d) and (e). We have also found the publication of trend series to be useful in explaining statistics particularly where they are volatile or subject to unusual movements. They are subject to less sampling error than original and seasonally adjusted series (considerably less sampling error for most statistics at the national level). They also provide a context for interpreting seasonally adjusted movements. Specifically, we have found that the interpretation of seasonally adjusted movements has improved considerably since we started presenting trend series as a standard output for sub-annual series.

6. Trust in the National Statistical Agency

Fellegi (1996) provides a strong argument that the trust in the national statistical agency is how most users judge the quality of its statistical products.

''Credibility plays a basic role in determining the value to users of the special commodity called statistical information. Indeed, few users can validate directly the data released by statistical offices. They must rely on the reputation of the provider of the information. Since information that is not believed is useless, it follows that the intrinsic value and usability of information depends directly on the credibility of the statistical system. That credibility could be challenged at any time on two primary grounds; because the statistics are based on inappropriate methodology, or because the office is suspected of political biases.''

The ABS places great reliance on adherence to its core values, namely:

- Relevance – good statistical planning, which requires a keen understanding of the current and future needs for statistics, is essential, as is the need for statistics to be timely and relatable to other statistics.
- Integrity – our data, analysis and interpretation should always be objective and we should publish statistics from all collections. Our statistical system is open to scrutiny, based on sound statistical principles and practices.
- Access for all – our statistics are for the benefit of all Australians and we ensure that equal opportunity of access to statistics is enjoyed by all users.

- Professionalism – the integrity of our statistics is built on our professional and ethical standards. We exercise the highest professional standards in all aspects of ABS statistics.
- Trust of providers – we have a compact with respondents; they are to encourage us to provide accurate information and we ensure that the confidentiality of the data provided is strictly protected. We keep the load and intrusion on respondents to a minimum, consistent with meeting justified statistical requirements.

All these values are important for trust in the national statistical agency and therefore the ability of a statistician to deliver.

It is also necessary to work hard to obtain strong recognition and support amongst decision makers and the community. It does not happen automatically. One pre-requisite is continuing high performance and credibility achieved by managing ourselves well and ensuring that our work is of high integrity. We use the media to promote awareness of the ABS and its performance, mainly through statistical releases. This requires an open relationship with the media, making ABS staff available to the media and ensuring they have the necessary training. We also encourage the active participation of ABS staff in external fora. All these activities are necessary if the ABS is to maintain a profile as a highly respected and trusted institution.

7. Nonuse of Technical Work

A major frustration of methodologists is that well thought through research is underutilised, even though years of intellectual effort might be involved. I believe small area estimation is one such field. Over the last 25 years, it would have been one of the most researched fields of survey statistics yet the utilisation of these methods is still not extensive. Why is this so?

I believe that one of the main reasons is that the mathematical statisticians have not been able to explain these sophisticated models in terms that the potential users of the models understand (or even some fellow mathematical statisticians). Some might argue that it is lack of reliable measures of quality but I doubt whether this is the case. Subject matter statisticians largely get their own feel for quality of statistics by the performance of statistics and their reception by users. After all, many important statistical measures such as the national accounts and the CPI do not normally have formal measures of quality.

In my view, this reinforces the need to have a close relationship between the different categories of statisticians, particularly the theoretical statisticians and the subject matter specialists.

8. Conclusion

Platek and Särndal should be congratulated for raising an important and provocative question – can a statistician deliver in terms of the quality expected by users of statistics. The answer to this question has to be in the positive for the long-term visability of national statistical agencies. However, ''delivery'' requires much more than good quality assurance procedures or a theory of statistical production.

This is the main point I am trying to make in this response. ''Delivery'' means that national statistical agencies address some of the so-called softer issues such as:

- community trust, particularly on the part of those who can influence public opinion,
- organisational values that are understood and upheld by staff,
- a collaborative relationship between the different categories of statisticians, where each category understands and respects the roles of other categories, and
- a preparedness to react to informed criticism.

9. References

Australian Bureau of Statistics (1996), Information Paper: Quality of Australian Balance of Payments Statistics (Catalogue 5342.0).

Fellegi, I.P. (1996). Characteristics of an Effective Statistical System. International Statistical Review, 64, 2.

Linacre, S. and Trewin, D. (1993). Total Survey Design – Application to a Collection of the Construction Industry. Journal of Official Statistics, 9, 3, 611–621.

Received December 2000