

Comment

Eva Elvers¹ and Lennart Nordberg²

Richard Platek and Carl-Erik Särndal have written an interesting and challenging article on an important topic. We, who are two mathematical statisticians and methodologists at Statistics Sweden since about twenty years, agree with much of what they say, but overall we have a somewhat more optimistic view.

Quality of Statistics

Although quality of official statistics can be defined in many ways, it seems – as pointed out by Platek and Särndal – as if many national statistical agencies have adopted a concept with roughly the following five main components: (1) Contents, (2) Accuracy, (3) Timeliness, (4) Coherence especially Comparability, and (5) Availability and Clarity.

The characteristics are often the same, even though somewhat different structures and different terms are used. The term *Relevance* is much used instead of *Contents*, but it has the disadvantage of carrying a value. A further aspect is *Objectivity*. It is fundamental, but also difficult to discuss and assess, and mostly left out.

The approach taken emphasizes the users and their needs, both existing and potential users.

The Users

Platek and Särndal state that the user will ask questions such as: “Can I trust the measurement procedure, the nonresponse adjustments and the estimation procedure behind the results ?” We believe that only a minority of the users will ask such insightful and technical questions. There are many kinds of users with very different preferences, and they emphasize the different quality components in very different ways. For some timeliness is very much at the top, for others comparability over time, for still others a specific domain of estimation, etc.

Some users are subject matter specialists. Other users are ordinary people; perhaps school children surfing on the Internet. Some users are producers of statistics; perhaps colleagues in other countries. Some users are beginners. Others have used official statistics for years. Some users are looking for something about a certain subject matter. Some other users are interested in long time series and a wide perspective. Still others have a particular interest and are very focused. Some users ask for economic indicators at the very moment of publication.

¹ Statistics Sweden, Stockholm, Sweden

² Statistics Sweden, Örebro, Sweden

Many users will, more or less consciously, trust the figure they find, because it is there, from the national statistical agency. It is convenient to have just one figure and to use it without questions. An uncertainty interval or some other measure of inaccuracy would be worrying and troublesome for many users.

Even though most users do not like working explicitly with accuracy, many of them are aware of its importance. At present, several important users of Swedish economic statistics express wishes like: “The statistics should be published earlier, but without deteriorating the accuracy.”

Quality Assessment, Quality Measurement, and a Quality Declaration

In accordance with the user-oriented quality concept, it is the user who makes the quality assessment. This implies that the statistics must be accompanied by such information – a quality declaration – that the statistics can be interpreted and used in a correct way. This is necessary in order to say that the statistics have been delivered fully, using the terminology of Platek and Särndal. Since the users are so heterogeneous, they need quality declarations written in different ways and with different amounts of information. This makes the task for the producer difficult. To provide separate pieces of information is a step in the right direction; both general descriptions and technical reports.

We think that a user-oriented approach to quality is desirable and useful. Nevertheless we agree with the authors that this approach may have some side effects. Accuracy is very difficult or even impossible for the user to assess without proper assistance from the producer. Since measuring accuracy can be difficult and expensive, this may shift the focus of the agency towards those quality components that the average user can most easily perceive and react to and towards monitoring customer satisfaction in just these respects. For example, there may be a temptation to emphasize the quality components availability and clarity. There is always a pressure on timeliness. The communication within the agency and with the users should be more open and clear about the balance between all quality components, among them accuracy. There is a risk that accuracy is not valued as long as the statistics seem reasonable.

However, giving low priority to accuracy is short-sighted and dangerous. We agree with Platek and Särndal that the primary source of user confidence derives from a feeling of general trust in the statistical agency. Hence a good reputation is a crucial asset for a statistical agency. To build good reputation takes much time and effort. Just one or a few mistakes may be enough to demolish this laboriously gained reputation.

There will always be some “demanding customer,” such as the user mentioned by Platek and Särndal, who will ask questions such as: “Can I trust the measurement procedure, the nonresponse adjustments and the estimation procedure behind the results ?” When this customer arrives, the agency must be able to “deliver,” i.e., to provide a good answer.

Measurement of accuracy is indeed important also for internal use within the agency. This is the only proper way to acquire knowledge of the production process – its weak and its strong points – and hence to provide a basis for improvements. The following example illustrates this point. A few years ago we were both involved – together with other colleagues from Statistics Sweden, the Office for National Statistics (ONS) in the

U.K., and the Universities of Southampton and Bath – in a project called *Model Quality Report in Business Statistics*. This project was initiated and funded by Eurostat under SUP-COM 1997, lot 6. As part of the project four surveys, two British and two Swedish, were analysed and attempts were made to measure and report, as well as possible, accuracy and its components: sampling errors, frame errors, measurement errors, processing errors, non-response errors, and model assumption errors.

For one of the two Swedish surveys it turned out that the major error components were not the ones previously assumed. This was an interesting outcome, and the report provided a solid basis for reallocation of resources.

Design, Process, and Product

Platek and Särndal distinguish between design and process, indicating that the former is more theoretical and preparatory, and that the latter is more applied and operational.

From 1993 onwards, there has been an increasing focus in the working language at Statistics Sweden on TQM (Total Quality Management) and processes. Here, the distinction is not made between design and process, but rather between process and product. The product is, of course, the end result of the process. Hence, the quality of the product depends on the quality of the process.

There is an emphasis on the process, on different subprocesses, and on related processes that assist the production of statistics. The very rapid IT development enables new possibilities for these processes, but it also requires much attention. All staff categories are involved. There is an increasing pressure to spend time on IT issues at the expense of core areas such as subject matter and methodology.

For this and other reasons, measurement of product quality seems presently to be a bit neglected, even though the customer should be in focus. As an example, accuracy is important, but the quantitative product facts are often in practice limited to an approximate sampling variance. It is much easier to measure nonresponse rates in the process than to analyse the bias in the product due to differences between respondents and nonrespondents. Results from a bias analysis are, however, much more informative. At least some indications may be obtained by comparisons with other surveys and administrative data. This information may be important not only for the current statistics, but also for the next round of the survey and its follow-up of nonresponse.

We feel that the tools of statistical methodology are presently under-utilized at Statistics Sweden, not only in connection with quality measurement but also in design and processes. More methodological work should be used to improve processes, for example a more systematic statistical approach to editing.

Is a (More) Firm Theory Needed?

Platek and Särndal seem a bit pessimistic, for example, when they talk about a desire for a more satisfactory theoretical framework. We do not share their pessimism on this point.

We would like to – and we should – know more about the quality of the statistics we work with, in particular about the accuracy. We do not primarily need more theory to this end. The main problem, in our view, is that it is costly to measure accuracy, in particular nonsampling errors.

As an example, we fear that measurement errors are substantial, but we rarely have the money for such studies. The willingness to respond has decreased lately, but again, we have little knowledge in terms of hard facts about the consequences. What is the loss in terms of reduced accuracy? What could we do to reduce the burden and to increase the willingness to respond? How could we better define and explain what we ask for and what unit the response should refer to? These questions and the answers to them are essential for the statistical agency in order to be able to reduce the errors.

To put some money into such very applied studies would be a wise investment in our view. We would not call this a new or firm theory, though.

A statistical agency should foster a culture, which gives high priority to statistical methodology. Applied methodological studies should be a natural part of work, with this work leading to such knowledge that important error sources can be identified and eliminated or reduced. It is fascinating to think about a total survey error model, which can explain all survey errors, and what this model should contain. However, we do not expect to find such a model, and we do not find it necessary for our work. We believe that it is possible to come a very long way towards “full delivery” in the terminology by Platek and Särndal by a systematic use of presently available methodology. This does not mean, of course, that we are opposed to statistical research, but only that we find the main obstacles elsewhere. A strong statistical culture should be a self-evident and manifest goal of a statistical agency, and if it is not a top priority, we should be worried.

Conclusion

To conclude, our answer would be “yes” to the question posed in the title by the authors: “Can a statistician deliver?”. We do not see any fundamental scientific reason why quality could not be appropriately measured and analysed. In this respect we are optimistic.

However, according to our experience, the answer would be “no” to the question whether national statistical agencies actually deliver fully. In this case we find the main obstacles to be nonscientific ones, in particular the high costs of measuring nonsampling errors, and perhaps also the “culture” and choice of priorities within agencies. On this point we are not quite so optimistic. Still, even if the obstacles are there, they should not be insurmountable. We see many challenges for statisticians and statistical agencies.

Received December 2000³

³ Note: Due to a mistake on the part of the Chief Editor this Comment did not reach Platek and Särndal in time and could therefore not be taken into account in their Rejoinder.