



Paper for the 84th DGINS conference in Stockholm 28-29 May 1998

Quality work and conflicting quality objectives

Tim Holt and Tim Jones
Office for National Statistics, UK

1. Introduction

As official statisticians, quality must be at the heart of all that we do. Today there is so much information, from so many sources and some of it – perhaps most of it – of dubious quality, that the output that official statisticians produce must be recognised as authoritative and of the highest quality. National Statistical Offices devote considerable effort and resource to this aim.

And yet ‘quality’ is not a well defined concept. It is multi-faceted and statistics that are regarded by users as having high quality for one purpose may be less adequate for another. Official statistics are often seen as general purpose tools, a single statistic being used for diverse purposes. It is more helpful to think of quality, not as an absolute property of a particular statistic but as a changing property depending on the use to which the statistic is put.

Just as quality is multi-faceted, so are the mechanisms that we, as official statisticians, use to achieve it. No single delivery mechanism or management process can produce statistics of adequate quality for all purposes. We use various inputs from high level office-wide initiatives, from the contributions of specialists for example through methodological programmes, and not least, from the contribution of individual staff members engaged in the ongoing operational work. Quality improvement can come as the result of large or small initiatives and all are important.

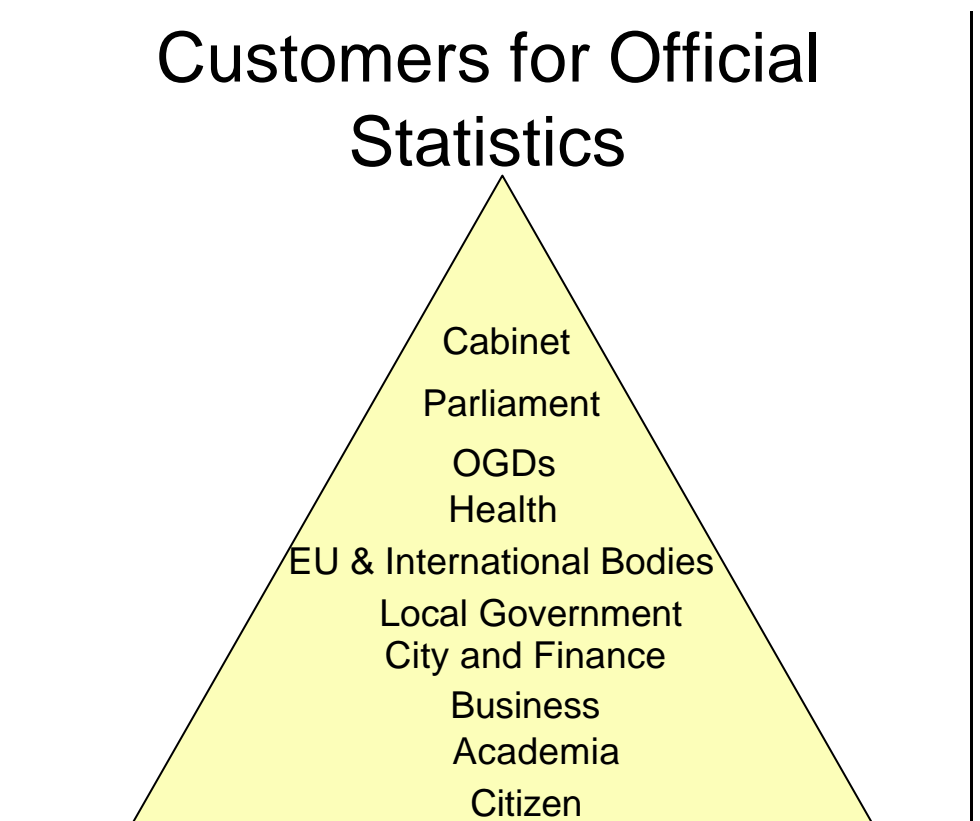
The fact that quality is multi-faceted carries with it the seeds of conflict. Inevitably there are trade-offs that must be made. These can cause a lack of confidence for public users and need to be treated by engaging users in the issues.

The paper is in three further sections. The next section sets out to describe the various facets of quality in official statistics, and the third sets out the ways in which the ONS tries to deliver it. The final section describes some of the conflicts that arise between different facets of quality.

2. Quality: what is it?

Years ago, quality in statistics might have been synonymous with accuracy. But nowadays it is a much wider concept. It is multi-faceted. It depends ultimately on what users consider to be important. The statistics we produce must address their needs. But there is a wide range of users. Understanding who they are and what aspects of quality matter most is crucial.

The following diagram illustrates the range of customers for official statistics in the United Kingdom.



At the top of the tree, in the UK, is government and Parliament. This includes HM Treasury, who is our principal paymaster and primary customer, especially for macro-economic statistics. It also includes the other major government departments (OGDs) who depend upon statistical information to monitor and develop policy and to allocate resources. With all of our Government customers, there has to be a lot of direct contact, of dialogue about needs and the same applies to our customers in the wider community for which we have established a wide range of user groups and advisory committees. In the UK in the 1980s, Government considered itself to be the sole customer. Others could benefit only incidentally. This has now changed, and we have made great strides in recent years in understanding the wider market and seeking to meet its needs.

In the UK the business community spends around 2 billion ecus per year on market information, most of it rapid up-to-date tactical and detailed market intelligence. Official statistics play only a small part in this market (less than 10 per cent). But businesses also require a range of more strategic background information. This market is less valuable but one which official statistics, given their current attributes, can more readily satisfy.

For high level aggregates and their main components, for example, population, the labour market, price movements, the balance of payments, the economy, official statistics are very hard to beat. But that is not to say there are no problems. The issue of public confidence in official statistics is important. In the UK for example, in parallel with a number of other countries we have suffered from a lack of public confidence in the unemployment statistics because of their dependence on politically determined rules for claiming unemployment benefit. This lack of confidence has persisted despite the fact that we have an excellent quarterly Labour Force Survey (using ILO definitions) and we have consistently published analyses of the differences between the benefit claim count and the ILO definition of unemployment. However while dedicated statistical surveys can provide good overall estimates, they do not satisfy all needs. Cost prohibits the sample size from being adequate for more detailed domains. Users may be in a position to check one element of a dataset and hence lose confidence in the whole set if it is subject to large standard errors.

So what are the features that users look for in our data?

2.1 Accuracy

As noted above, accuracy is fundamental to quality. There is perhaps little new to say about it. It is a relative concept, not an absolute one. Absolute accuracy is generally neither needed nor expected. But the relative size of the error attaching to a statistic and to the relationship of the statistic to others in the same context is crucial. It needs to be such as to minimise the risk of the user drawing a false conclusion and we need to educate users about the limitations for use. Too little emphasis has been given to this and we need to help prevent decision-makers from over-reacting.

The impact of inaccurate data on public policy can be disastrous. It is claimed that in the UK in the late eighties, signs that the economy was overheating were obscured by inaccuracies, which meant that suitable remedial action was not taken.

All this depends on having appropriate measures of error - in itself no easy task. Errors may come from many sources, and have been categorised in several ways including one developed by Eurostat¹ in the context of quality in business surveys. The sources of error may be subject to both variance and bias. The impact of

¹ Document Eurostat/D3/Quality/96/04 final

these on survey statistics can be reduced by high quality survey procedures. For example, appropriate interviewer training can help to minimise non-response and ensure greater consistency amongst interviewers; cognitive testing methods can be used to determine whether questions are working as intended.

In some cases it is possible to estimate the bias in survey estimates. In the UK, the Census of Population provides an opportunity to obtain information about non-responders to voluntary household surveys in progress at the time. The procedure has been to match addresses used in these samples with the corresponding census records. This information about non-responders to surveys can then be used to estimate non-response bias and to adjust survey estimates.

2.2 Relevance

The requirement for accuracy often implies that the quantity being measured is well defined. The number of voting papers in an election ballot box is one population that is probably better defined than most. The number of human live births in a given country and period is another, although there may be classification problems about what constitutes a live birth. There may also be one or two practical boundary problems (e.g. births that take place at close to midnight).

But whether someone is "unemployed" or not is a much less well defined concept. For this and most other elements in our statistical systems we have to rely on definitions and conventions that are to some degree arbitrary and open to interpretation at the boundaries. The difficulties that we have experienced over debt and deficit are other illustrations of this issue. Some constructs (for example index numbers of price or quantity changes) are not observable even in principle.

These conceptual problems are at the heart of the classification and estimation problems that official statisticians face. It is rare for the concepts that we strive to measure to be driven by a well defined theoretical construct. Often there is a degree of arbitrariness in the choice and the ideal concept will vary from one use to another. For example the definition of unemployment best suited to a labour economist analysing the downward pressure of unemployment on wage settlements will be different to that required by the sociologist investigating the impact of unemployment on family life. There is no perfect solution to this problem although we strive to present data as flexibly as possible to support different uses. Internationally agreed standards are essential – both as a means of striving for international comparability and as a source of guidance for national statisticians when faced with conflicting pressures.

Frequently statistics are available from an administrative source and hence are much less under the control of statisticians. The relationship between the ideal concept and the source can be tenuous if not actually misleading. An example is provided by the incidence of crime and of complaints of crime recorded by police. As mentioned above, a second example is the administrative benefit claim source that is

used as a measure of unemployment. But administrative systems can provide good quality data at little additional cost. What they actually measure, although not ideal, is usually well defined in its own terms. The figures for claimants of social benefit claims are clearly helpful in relation to public expenditure commitments. It is up to us to exploit these systems. Ideally we would have some influence over the design and coverage of the systems. We certainly have a role in interpreting the data; and analysing the difference between the source and the ideal concept. We can and should supplement them where necessary with appropriate additional information.

Some concepts may be relevant for macro-economic analysis, but meaningless in a business context. Value added as defined in the ESA (especially if adjusted for FISIM) might be one example. Businesses are more likely to be interested in transactions as measured in a commercial accounting framework.

The shift from manufacturing to services and the globalisation of business are not new. But the revolution in information technology and communications is adding impetus to them. It is contributing to the challenges we face in defining and measuring prices, outputs and productivity for services.

Although much has been achieved in the European and world context, the difficult task of bringing together national definitions and classifications is likely to remain for years to come, especially when we measure any aspect of the interface between the citizen and the state. Our social, cultural, legislative and linguistic diversity will inhibit true comparability.

As soon as one accepts that perfect accuracy is impossible even in conceptual terms, a host of other issues arise, such as coherence, consistency and continuity. These are discussed next.

2.3 Coherence and consistency

Ideally we would use the same definitions from all sources. This is not always possible, but there should be good reason for any differences. If the data are the responsibility of different agencies, it is likely that requirements will be different. There may be neither opportunity nor incentive for harmonisation. Coherence suffers. Our experience is that the greater centralisation of statistics in the ONS and the demand for cross-cutting analyses has provided incentives for harmonisation.

Frameworks such as the System of National Accounts are extremely valuable for achieving coherence. The introduction of the new ESA has provided a useful further incentive to consistency across Europe, although legislative, administrative, linguistic and cultural differences will all contribute to differences of interpretation and results. Even within one country, regional differences may affect coherence and consistency.

Even if the definitions are the same, collecting the "same" information about events from different sources will, with probability one, give different results. Estimates of

employment from employers gives one figure while those from households gives another. Intrastat exports exceed Intrastat imports.

In the UK in the past, measures of changes in productivity have not been reliable because of different sources of data about output and employment. Surveys of output were based on enterprises (which have a coarse industrial classification) and surveys of employment were based on establishments (having a finer industrial classification). A major initiative in recent years has been to merge these surveys; leading to productivity estimates of much improved quality.

2.4 Continuity

Comparisons over time are important in many situations. For macro-economic modelling and forecasting, lengthy time-series are required. As far as possible they must be free from discontinuities.

Every method of data collection and compilation has its limitations. The statistical system must change and develop over time, especially in the face of new requirements or methodological developments. Discontinuities will result. So it is important to be able to assess the extent of these breaks and to make appropriate adjustments to the back series.

2.5 Timeliness

For many decisions, especially in macro-economic management, up-to-date information is generally regarded as essential. For example, the Bank of England Monetary Policy Committee is a primary user of short-term indicators in determining whether or not to change interest rates.

In recent years, one of the most dramatic improvements in timeliness has been on the Labour Force Survey (LFS). When the current panel design for the survey was introduced in 1984, data was not available for analysis until four and a half weeks after the end of the data collection period. However following the advent of computer assisted interviewing (CAI) in 1990 and subsequent improvements to the processing systems, this has now reduced to four days. CAI has enabled most post-survey editing procedures either to be transferred to the field, or to be removed completely, as in the case of continuity checks. Hence data is clean and ready for analysis almost immediately it is returned to the office.

In addition this has resulted in improved quality of the data. In the days of paper questionnaires, data was checked once it had been returned to the office, and corrections could only be made on the basis of what was shown on the questionnaire. Under CAI any inconsistencies are identified at the time of the interview and can be clarified immediately with the respondent.

The need for "up to the minute" data is not so keenly felt in the case of other types of social data, given that social conditions are slower to change than economic ones. However, with the pace of change accelerating in the social sphere, as Linacre and Mahoney² point out, the greater concern is the gap between the time the need for data on a new issue is felt and the time the statistics become available. In the UK the "Omnibus" survey is designed to minimise this gap. This flexible, quarterly survey allows for question formulation to the delivery of results to be completed within six weeks.

2.6 Accessibility

It is not just the timeliness with which the data are collected and processed that is important. The speed at which relevant data can be located, retrieved, understood, and reused is a vital element of quality. An executive may often be required to produce a business case for something within a day or so. Unless the piece of information needed can be accessed at the touch of a button, by the time it has been found it may be too late. The figures must carry with them the appropriate metadata, so that the user can understand and interpret them correctly. In the UK, our initiative to develop an integrated database is designed to help meet this need.

2.7 Revisions

Very often our systems and resources make it impossible to obtain more than indicators of movement in the short term, so that when more reliable information becomes available it is often necessary to revise. Revisions provide some information on how accurate our initial estimates were, and have been used as performance measures for the organisation. Seasonal adjustment also leads to revisions within years because estimates of the seasonal factors towards the end of a series are bound to be provisional.

While revisions are a sign that we did not "get it right first time", performance indicators can send a message to staff that revisions are to be avoided. There would seem to be little point in moving from one inaccurate figure to another that is not significantly more accurate. Relative error must therefore play a key role here. Some users may see a sensible revisions process as a sign of quality control and attention to quality, and hence as a positive rather than a negative sign. A proper appreciation of the quality of our data and its liability to revision is something we need to foster among users.

2.8 Describing quality

Running through this categorisation has been a theme that we do too little to measure quality and to educate our users so as to improve their understanding and help them avoid over-interpretation. In the UK we prepared a booklet entitled *Statistical*

² Linacre and Mahoney (1997) Managing the trade-off between timeliness and other aspects of quality in the outputs of a statistical agency. ISI conference paper.

Quality Checklist GSS (1997) which invites statisticians to provide answers to forty-three questions about the statistics within five major groups: objectives, coverage, design, analysis and extensions. I am pleased to say that it has already been translated into Russian! We are also pleased to be taking part in the development of model quality reports on business statistics in partnership with Statistics Sweden, funded by Eurostat's SUP.COM programme.

However, whilst valuable this will not be enough. We need to get beyond supplementary reports and footnotes and build these issues into our primary outputs. The general question of how to use charts, graphs and tables to maximum effect so as to assist understanding but avoid over-interpretation is a much neglected issue.

3. Quality: how do we deliver it?

If the quality of our products is multi-faceted, so are the processes that we need to deliver it. No single mechanism can deliver quality. It depends to a significant extent on systems and processes that are designed to underpin quality. But equally it depends on people at all levels in the organisation identifying ways of doing things better, taking responsibility for the processes and exploiting their knowledge of them. All this needs to be done in the knowledge of what the customer needs. Some of the work going on in support of delivering quality is described in this section.

3.1 Meeting customers' needs

Understanding users' needs is a crucial step. To help clarify these needs within Government, 'service level agreements' or 'Concordats' have been established with other departments and are subject to regular review. Often the process of reaching agreement can be more illuminating than the agreement itself. Additionally we have established an extensive framework of user groups and advisory committees. We find that these are an invaluable way of learning about how our statistics are used and what aspects of quality matter most. A crucial skill for us is to determine to what extent and how we can meet those needs from the elements that are available to us: the data sources, appropriate methodology, processing capacity, etc.

3.2 The ONS framework for delivering Quality

ONS is still a new organisation, the result of a merger between two different Departments and part of a third. It operates in five locations. This has in itself brought about considerable change, affecting most people, as common systems have been adopted. Cultural and procedural differences were substantial although all of the component parts of ONS had a strong commitment to quality before the merger.

At the outset, the first priority was to get people working together and talking a common language. This was done mainly through a corporate management training programme under the umbrella of "ONS 2000". We set out a vision of where we wanted to be and what our values were. Quality is at the top of the list of values and the ONS2000 programme challenges staff to integrate this into their work programmes.

As a consequence we have chosen the ONS2000 programme as the main corporate mechanism for delivering quality, supported by a number of important initiatives which will be described below. ONS has chosen not to introduce an additional quality initiative such as a TQM programme. The ONS2000 programme has specific 'managing quality' modules. Our approach is to make quality more explicit in the ONS planning process and provide a broad framework to assist this. We will develop a more systematic framework for measuring, monitoring and evaluating quality as an integral part of the ongoing work programme. To ensure we are not collecting data unnecessarily, regular reviews of continuing surveys are carried out. These provide an opportunity to check that the survey is still relevant and meeting users' needs, while at the same time not over-burdening respondents. This set of initiatives, which is designed to engage all staff, is being underpinned with more structured support for communication and co-ordination.

However this approach is supported by a number of other initiatives. Two years before the creation of ONS we began to systematically strengthen the Methods and Quality Division, and when ONS was formed a methodology committee was created to underpin quality and disseminate best practice. This has been supported not only by increasing staff numbers and resources but also by entering into a substantial contract with survey statisticians at the University of Southampton. This has permitted a systematic programme of quality improvements. In addition staff from other universities have been engaged on specific research and development projects. Work has focused on areas that are related to specific business needs or on cross-cutting issues that will provide the most significant quality improvements across a range of work areas. Some of these are described below.

3.3 *Redesign of business surveys*

An extensive programme of methodological work related to business survey redesign has been undertaken. Almost all surveys have now been redesigned and migrated to common sample selection and processing software, yielding both efficiency and quality improvements. The redesigns have also allowed full integration of the surveys with the Business Register improving consistency and coherence. As a result of the merger responsibility for production surveys and employment surveys are now in ONS and the short term surveys have been integrated leading to quality improvements and more efficient designs. The redesign of the annual surveys is well advanced.

3.4 *Improving data supply*

An important element of quality lies in the methods of obtaining and validating data through surveys or from other sources. Exploiting modern technology to capture data offers improvements in both accuracy and timeliness. The use of CAI has already been mentioned. In business surveys and in the Census of Population, the use of imaging and character recognition techniques can speed the process and major research and development programmes are in progress.

Additionally the process of designing questions and their layout is another element in achieving quality results. The ONS has a unit that undertakes the analysis of questions and of respondents' understanding of what is required. One of its current tasks is testing forms of question for the next Census.

3.5 *Data compilation and exploitation*

Developing and applying soundly based statistical methods represent contributions to accuracy which can have a surprisingly significant effect. These methods include such processes as imputing missing values, allowing for non-response, and making the best use of auxiliary variables to derive estimators of population totals from samples. For example, the ONS has been working on optimal ways to identify, and to reduce the influence of, outliers in business surveys.

Ideally, the software to carry out these processes will be common to most surveys and incorporate standard routines. For some purposes, spreadsheet systems have offered great flexibility, but at a cost of reliability, difficult maintenance and a risk of error. There is a need to seek the optimal balance between the reliability of a closed system and the ability to respond to changing requirements and developing methodology.

To tackle these problems, the ONS has a central Methods and Quality Division and an ONS-wide Methods Committee, which have established a strategy for methodology in which important cross-cutting topics are identified.

3.6 *Small area estimation*

An example of a cross-cutting methodological issue that has the potential to lead to quality improvements in many parts of ONS is the problem of making small area estimates from survey data. This is a recurring issue encompassing employment, unemployment, health, business activity, Census undercount and many other areas. The underlying approach, using multi-level models, is in principle applicable to a wide variety of situations and has been identified as a methodological priority.

3.7 *Time series issues*

A similar example is the analysis of time series data. Issues surrounding seasonal adjustment and estimating current trends are broadly applicable within ONS. A

methodological programme with the purpose of identifying best practice and then raising all areas of application to this level is in progress.

3.8 Social classification

For many years ONS has been using two different but related social classifications, based on occupation, both of which have been subject to criticism. Currently major development work is under way, in conjunction with the British Economic and Social Research Council and involving university staff, to replace them with a single new classification. This will have a sounder conceptual basis than its predecessors and will be properly validated. The result should be a higher quality framework within which to provide relevant social information.

3.9 Examples of quality work in action

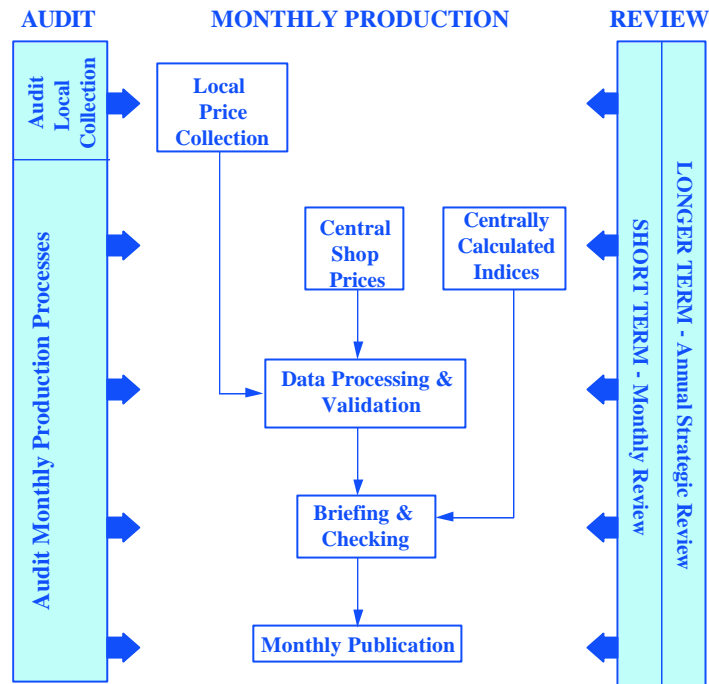
Two examples of initiatives currently in progress and related to specific business areas follow. The first concerns the monthly processing to produce the UK Retail Prices Index. (RPI). A formal system of "Quality Management" has been established for this. It is designed to meet the requirements of International Standard ISO 9002. There were several reasons for adopting this approach, including the costly nature of any mistakes and a need to enhance public confidence.

There are three main parts to the RPI Quality Management System: the production cycle itself, the auditing of the system and the review of the system. The system also includes assurances that all staff receive adequate training for their work. An outline of the structure of the System is shown in the chart below.

The basis of the Quality Management System is first of all that all aspects of the monthly production cycle are documented in three levels: policy, procedures and work instructions.

Each aspect of the monthly production cycle is audited at least once each year. The purpose of each audit is to verify that operational procedures and controls comply with the documented procedures and to determine their effectiveness in delivering products and services that are fit for purpose. The local price collection is also audited in two ways: an accompanied check of field collection and a random post-hoc check. The primary purpose of the accompanied check is to ensure that the collector is following the procedures and instructions laid down for price collection. The random post-hoc check involves randomly selecting each month a sample of locations and index items for which the prices are to be audited. The principal aim is to see if the rate of error is below the acceptable level, defined as 5% or less. The quality of item descriptions and the use of indicator codes (e.g. to indicate that an item is on sale) are also assessed.

RPI Quality Management System



There are two parts to the review system, designed to promote "continuous improvement": a short term monthly review and a long term strategic review. Towards the end of each monthly processing cycle, one day is allocated as a Quality Day. The day starts with a review of the previous month's work. It also includes a look forward to the next month. Any particular features of the coming cycle which need to be taken into account and planned for will be discussed. The day normally includes a presentation and discussion or group work on a specific topic. A high level review of the System is carried out each year, as part of the strategic planning cycle. The focus is on the quality of outputs, matched against an assessment of customer needs. The meeting reviews progress over the previous year, sets policies and strategies for the coming year, and decides on what improvement projects should take place.

Quality is also at the heart of planning for our next Census of Population and Housing in 2001. Having established the lessons to be learnt from the 1991 Census, we set up a research programme to evaluate the best methods of collecting, processing and disseminating census data in the new millennium.

A Census, by its nature, is designed to cover the whole population. As such, there are no sampling errors. In other respects, it is no different from a sample survey. For example, there is non-response bias - whole people and individual questions; measurement error - the public answering questions incorrectly; and system, processing or coding errors. Research is continuing to design census forms that are both easy to complete and enable new technology to be used to capture the data automatically on to the database. Editing and imputation systems will also be used to eradicate obvious errors and estimate answers where they are missing.

Non-response or underenumeration is the error which attracts the greatest attention and is the most important. While the estimated overall underenumeration in 1991 was small (2%), of more significance was its variation across population subgroups and different parts of the country. Census statistics are used in the formulae to calculate the resources to be allocated from central to local government. It is important therefore that the Census accurately identifies the demographic and socio-economic indicators which feed into this process.

The method of delivering and collecting forms to and from the public is being rethought, for example making more use of the postal system to enable enumerator resources to be targeted where they are needed. Publicity and liaison with community groups to 'sell' the Census will also be important in minimising the differential undercount.

It is inevitable, however that there will be non-response. Planning for this is essential. The *One Number Census* research project was therefore set up to do this. The primary objective is to determine how, following the Census, to obtain the highest quality population counts by age and sex for each local authority to which funds are allocated by central government. The second objective is the full adjustment of the Census database so that all counts derived from it add to the same 'one number' totals, but this will take longer.

4. Quality conflicts

In this section, some examples of conflicting quality objectives are given.

4.1 Accuracy and relevance

Many users require information at a more detailed level than can safely be provided. This may be in terms of geography, for example characteristics of the population in small areas, or products, where the size of the market for specific set of products is required.

Occasionally, there is a source of data for what is required with 100 per cent coverage - for example the Census of Population, Customs data on external trade. But even these may not be very reliable for particular small groups, because of undercoverage, mis-codings or other errors. Administrative sources such as the unemployed claimant count can also provide detail, provided the user is prepared to accept that the measure may not accord with the definition ideally required. It can be very difficult to manage this conflict.

Often detailed geographical data are required for the purpose of resource allocation. What we can do, and are doing, is to pool data in such a way as to provide better

estimates of *expected* values of the variable of interest in small areas, using multi-level modelling techniques. This process can provide better quality small area estimates than would be available from other approaches and may provide substantial gains in the effectiveness of the allocation.

4.2 Accuracy, timeliness and revisions

Users want statistics which are accurate and free from revisions. But the need for data which is up to date can conflict with this. In the UK, the first estimates of GDP are published three, eight and twelve weeks after the end of each quarter. The twelve week estimate is subsequently revised in future quarters and in the annual Blue Book. An analysis of revisions is published each year in Economic Trends. Revisions are caused by the receipt of later data, in particular annual data sources which are assessed as being more accurate than short term data sources.

The need to provide users with early estimates, while ensuring that account is taken of later, more accurate data sources, leads to revisions. We aim to minimise these by adjusting early estimates for known biases, and work within revision targets agreed with the Treasury.

4.3 Timeliness, relevance and accuracy

Another trade-off can come into play when new outputs of a composite nature are being developed. In the UK, the construction of a set of environmental satellite accounts provides an example. At the outset, the accuracy of information needed to achieve the desired result may be very patchy, with several gaps for which the roughest of estimates can be made. However, despite its shortcomings, such a prototype can be extremely valuable for demonstration purposes: the possibilities can be put before potential users in a practical and concrete way, with a view to stimulating interest and hence the support necessary to fill the gaps. Equally importantly, in cases of significant economic or social change, prototype outputs can stimulate users to contribute to the development of new concepts or new outputs so as to better meet user needs.

4.4 Consistency and timeliness

Frameworks such as the national accounts, or more particularly input-output tables can be used to ensure coherence, but it takes time to assemble the data and to carry through the adjustments that are needed to make the system balance. Having done that, there may be a problem of consistency between the adjusted data and the original sources.

The population estimates provide another example. Accurate figures are required for local areas. After a Census has been carried out, estimates of population could be provided based on the Census counts. However, if these are known to be subject to undercoverage (which may vary according to the area), they will not be

accepted. So it is essential to arrange for estimates of the undercount to be made. Adjusting the population estimates as a result of this work will take more time. It will also make the population estimates incompatible with the raw Census data,

Our contacts with users of Census data suggest this incompatibility often represents a lack of quality that users find difficult to cope with. In the *One Number Census* project mentioned earlier we are investigating the feasibility, both theoretical and practical, of adjusting the Census database itself, so that all Census counts are compatible with the adjusted estimates. Such full adjustment is likely to take longer, so timeliness may suffer. Many users may however be prepared to wait, rather than to have to cope with the incompatibilities themselves.

4.5 Continuity and change

There is clearly a tension between the changes needed to improve coherence and the desire of users for continuity.

A major initiative for the Business Statistics Group of the ONS is the integration of business surveys which in the past collected employment data and operating data separately. The designs of these surveys were quite different, and there were difficulties in comparing the results: productivity could not be measured in a satisfactory way. The employment surveys were directed at local units, while the other surveys were directed at enterprises. One obvious difference lay in the industrial classification. Classifying local units is not the same as classifying enterprises. Moreover, in some parts of the economy different auxiliary variables were used for stratification and estimation.

A second example of conflicting quality objectives in relation to continuity and change relates to the ONS social classifications mentioned above. Whilst more soundly based, a new classification will disrupt time series some of which go back to the early years of the century. Those responsible for the work are attempting to deal with this problem by producing a longer version of the classification which preserves the major distinctions of both the old and new. This can be then collapsed in different ways to produce reasonable approximations of either of the two older classifications or the new classification. Thus users will be able to use more than one classification for a time to compare differences and ease transition.

5. Conclusion

This paper takes as its theme the fact that quality, and the mechanisms for delivering it are multi-faceted. So too are the conflicts that arise when trade-offs are called for between different facets of quality. Quality is not so much an absolute property of a statistical estimate but is related to the purpose for which the estimate is used.

Conflicts between the different facets of quality inevitably arise and need to be resolved in consultation with users.

Quality is at the heart of public confidence in official statistics and evaluating it, and reporting on it, in a way that leads to an open discussion of quality issues is central to enhancing public confidence. But beyond this, official statisticians need to find ways to present statistics so as to convey quality, enhance understanding and avoid over-interpretation.

London
January 1998