Quality Work
and Quality Assurance
within Statistics
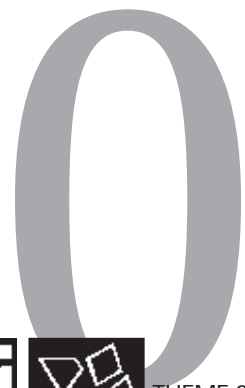
Statistics & quality go hand-in-hand

Quality of statistics was the theme of the annual conference of presidents and directors-general of the national statistical institutes (NSIs) of EU and EEA countries, organised in Stockholm on 28-29 May 1998 by Statistics Sweden in collaboration with Eurostat.

Quality has always been one of the obvious requirements of statistics, although the notion of 'quality' has changed over the years. Nowadays a statistical 'product' has to exhibit reliability, relevance of concept, promptness, ease of access, clarity, comparability, consistency and exhaustiveness. While all these features form part of the whole product, individual users will attach more or less importance to each one. When statistics - gross domestic product or inflation, for example - have a financial impact, accuracy and comparability are vital. But if the same data are being used by someone interested in short-term trends, then the speed with which they are made available is the key feature.

It is for users to decide. They are the people who determine quality criteria. Statisticians are no longer 'number freaks' in a world of their own, but have become managers of statistics, in constant touch with those who make decisions.

Such a transformation is possible only if the whole production process is ready for change, because, as a rule, scientists such as statisticians tend not to pay much attention to the needs of people outside their own world. Scientists prefer to talk to other scientists.

Now that the need for change is understood, how to bring it about? This question is being addressed by most of those in charge of national statistical institutes. The discussions about various aspects of quality and the importance of 'quality schemes' that took place during this conference enabled the NSI presidents and directors-general to demonstrate that statistics and quality are one and the same thing.

I hope this publication will be a source of inspiration and encouragement!

Yves Franchet
Director-General Eurostat

TABLE OF CONTENTS

5

Theme 2:
Quality work - some national experiences

Theme 3:
International lessons and prospects

0.1 Erik Åsbrink, Ministry of finance, Sweden

Independent and high quality statistics are essential for a democratic society.

I would like to begin by thanking Eurostat and Statistics Sweden for organising this conference. These annual meetings of Director Generals of the National Statistical Institutes are important meetings since they provide the opportunity for exchanging experiences and learning about the best statistical practices and improving the comparability of statistics across national borders.

The ability to compare statistics at a European level is becoming more and more important, especially owing to developments in the European Union. Current examples illustrating the increased need for comparability are the creation of the Economic and Monetary Union and the signing of the Stability and Growth Pact, which is a result of EMU . The need to determine the extent to which Member States meet the convergence criteria as the requirement for sound, sustainable public finances set out in the Stability and Growth Pact has intensified the demand for reliable and comparable statistics.

It is interesting to note that the convergence criteria and the efforts to meet these have not only had an impact on the 15 current Member States of the EU. I have noticed similar interest and efforts in a number of other countries. There is an unprecedented vigour in formulating policy objectives, if one assumes that the goals are reasonably formulated and that they steer developments in the desired direction. I am convinced that we will observe more examples of similar needs for increased comparability in the future both in the EU and in other international forums. Statistics on employment and unemployment in our countries are undoubtedly one such area of importance.

In addition a large number of countries wish to become members of the European Union. It is gratifying that among the participants at this conference there are representatives of the candidate countries for the first time. This is a concrete expression of growing European co-operation. We want you to feel especially welcome.

Reliable and high quality statistics that are largely independent are a feature of democratic society. Without independent statistics, there is considerable risk that politicians and other decision-makers may manipulate figures in a manner that distorts reality in order to achieve their particular aims. It is no coincidence that only democratic societies have been able to produce high quality and independent statistics, nor is it a coincidence that statistics in undemocratic societies are of lower quality and often reflect a selective and warped view of reality.

Openness and transparency are important prerequisites of a democratic society. Independent and high quality statistics are, in turn, an important constituent of openness and transparency. That these statistics can be developed and improved even more by extensive international co-operation is in fact a way of further strengthening democracy in our societies.

Sweden has a long tradition of statistics. As far back as 1747 the 'Tabellverket' -literally the Table Bureau--the predecessor of today's Statistics Sweden, was founded by a decision of the Swedish Riksdag. The aim was to compile better demographic information. The actual reason was a wish to promote population growth. Even if population data naturally had shortcomings, we have received from early on a relatively accurate picture of the size and occupations of the population as well as the population changes through births, deaths and migration both within Sweden and between Sweden and the rest of the world.

This not only signifies that we in Sweden have uniquely extensive population statistics. It also signifies that we had concrete practice relatively early in applying and developing statistical methods in

the areas considered important to society. The social debate and proposals for reform could largely draw on knowledge of actual conditions. The existence of this statistical information in itself constituted a stimulant to debate and reforming zeal and in this way training in building a democratic society.

In recent decades Sweden's statistical production has undergone considerable change. During a large part of the post war period Statistics Sweden expanded. The need for statistics grew and increased resources were allotted to this area. A centralisation of the responsibility for statistics also took place. A great many of the statistics that earlier had been produced by various sectoral authorities were gradually transferred to Statistics Sweden. The idea behind this was both to take advantage of co-ordination and synergy gains, but also to ensure a uniformly high quality. A part of the statistics that earlier had been produced by the sectoral authorities was not of sufficiently high quality from a statistical point of view.

In recent years statistical production, like most other expenditure areas in the state budget, has experienced some reductions. The reason for this was that a few years ago Sweden had one of the largest budget deficits of all the OECD countries. The deficit at its highest amounted to over 12 per cent of GDP. The expenditure reductions were unavoidable but many times were difficult or even painful to implement. These naturally also affected the area of statistics. In truth, it should be added that there initially was a certain margin or slack in the public expenditure system, where the reductions up to a certain limit were both warranted and easy to carry out.

One particular factor that has naturally been important for Statistics Sweden has been the increasing demand for statistics as a consequence of Swedish membership in the EU. Tighter budget restrictions simultaneous with increased demand for statistics is of course not an ease situation to handle. I can only certify, as the minister responsible for statistical matters, that Statistics Sweden has handled the new situation in an exemplary fashion and has been successful both in satisfying the many demands for pertinent statistics and at the same time has safeguard-

ed the high quality of statistical production.

I would like to add that a very extensive budget consolidation programme has been carried out and that this consolidation has already shown a surplus this year and in the coming years this surplus will grow. This should certainly not be interpreted as a promise of increased expenditures in the future either for the production of statistics or for other purposes. But I can probably go so far as to say that the toughest and most restrictive period should now likely be over.

One important change in the production of statistics that has been made in recent years means that a part of the financing has been shifted to different sectoral authorities. About half of Statistics Sweden's revenues now come from commissions, of which most orders placed by such sectoral authorities, which must also answer for corresponding financing. The reason for this change is to be better able to monitor how great the demand is in certain statistical areas both with respect to extent and design. The need for statistics in a certain sector can thus also be weighed against other measures in the same area.

As to the other half of the production of statistics, the financing is still made through direct appropriations to Statistics Sweden. These appropriations apply to comprehensive and important areas of statistics such as economic statistics, labour market statistics, demographic statistics and substantial parts of the welfare statistics. In addition Statistics Sweden also has a co-ordinating role and responsibility for the development of methodology, et cetera in the area of statistics.

I am aware that this arrangement has raised questions and even misgivings that the independence of Statistics Sweden would be circumscribed. I do not believe that these concerns are warranted.

It is important to understand that the Swedish government authorities--including Statistics Sweden--have an independent status that is unique from an international perspective. In most countries the authorities form parts of large ministries. Sweden, like Finland, has developed a tradition of gover-

nance that is different from that in other countries. In Sweden the ministries are small and are responsible for the overall policy formulation. The day-to-day management of matters in the public sector based on the laws and general guidelines decided by the Government and the Riksdag is handled by independent government authorities that are separate from the ministries to a large extent. This independence is guaranteed by the constitution. This system also guarantees a high degree of independence and integrity for the production of statistics in Sweden. I cannot see any signs that this independence is jeopardised.

This is not to say that the present system for the production of statistics does not have to be followed up and analysed continually. Some time ago I appointed a government committee with the task of reviewing the present system and to consider the need for any changes in the future. Irrespective of the views that may be put forward I want to declare that the high demands made on statistical production with respect to integrity, independence, quality and reliability will also have to be met in the future.

Let me mention a few other points.

One area that is important to me as Minister of Finance--and to the finance ministers of all other countries--is the quality of the national accounts. Here there is a problem because of the revisions that are being made--in some cases many times and sometimes with a time lag of several years. This is not just a problem for finance ministers endeavouring to meet the convergence criteria and the like. It also creates a certain confusion in the broad economic policy debate.

When I point out this problem to the Swedish statistical authorities, I usually get the answer that the revisions and the time lag are even worse in many other countries. This is possibly true but it does not do much to alleviate my problem.

I am of course aware how difficult and complicated it is to improve the quality of the national accounts and to reduce the time lag until the final figures can be ascertained. A reform of this area is being con-

ducted and there is no doubt that international co-operation is badly needed. But I would still like to stress that at least in the field of economic statistics, improvements in the quality of the national accounts must be a top priority.

We live in a time when it may be necessary to make major social changes within a short time span. This necessity accentuates he demands for follow up and evaluation. In this respect statistics have an important role to play.

At the beginning of the 1990s we carried out a very extensive tax reform in Sweden. Practically all parts of the tax regime were changed. The intention was to get away from an old, very complex system with high tax rates, generous possibilities for deductions and a number of special provisions and exceptions. In its place we imtroduced a system with broad tax bases and lower tax rates and tried as much as possible to eliminate all distortions and exceptions.

By the time that the reform had been agreed, we decided to evaluate it thoroughly. This was done both during the implementation phase and after the reform had been fully implemented and had been in force for some time. The evaluation work was done by economists, behavioural scientists, political scientists and statisticians. It was a very difficult task since there were so many changes at one time and since the behaviour of various participants was changed--in some cases in an intended manner and in other cases more surprisingly.

With the evaluation we gained valuable experience and were able, among other things, to study the extent to which the effects strived for were achieved.

We will need to make similar studies of other areas of society.

The work to improve quality is something that is playing a bigger and bigger role in modern society and I note with pleasure that this is one of the matters to be discussed at this conference. I personally find this very important. It concerns working to improve quality in the production of goods and services to a greater extent than previously. In order to make sure that this actually happens, our instruments of measure must be further refined. Methods

of measurement have to be developed in those areas where measurement of the activity has traditionally been viewed as difficult or even impossible, such as the production of services.

It is my impression that in Sweden the business sector, especially industry, has made more progress in this area than the public sector. Therefore it is important that increased emphasis be put on strengthening work on improving quality in the public sector. The ultimate objective is to give taxpayers better value for their money. Without doubt statisticians have important tasks to perform in this area.

I have given a few examples in the area of statistics that I think are important for continued development and expanded international co-operation. More examples could certainly be given. But I would like to use these examples to emphasise that you confront important tasks. Several main areas will be featured in conference discussions, according to the interesting agenda that has been presented.

I would like to wish you all luck in your important work, both at this conference and in the future. You are making a valuable contribution within your discipline to strengthen co-operation and solidarity in Europe.

 Thank you for your attention.

0.2 Jan Carling, Statistical Office, Sweden

Quality and statistics.

We have before us two days of - hopefully - lively discussions about matters that are important to all of us.

We have chosen different aspects on quality as a theme for this meeting. There are several reasons that we made that choice.

One is that high quality meaning high reliability is a fundamental prerequisite for a statistical institute. We are supposed to produce and publish facts as a base for debate, decision making and research. Our data must be accurate. Otherwise they might lead to the wrong conclusions. And our data must be regarded as reliable, otherwise they will not be used.

Consequently our task is not only to produce and disseminate reliable statistics but also to convince our potential users that our output can be relied on. We have to be able to demonstrate how we handle the quality issues and to provide error-measurements related to our figures.

But accuracy is only one aspect on quality. The modern quality philosophy - regardless of which of the many varying concepts you think about - is based on customer or user needs. And user needs goes far beyond accuracy. We are all aware of the pressure from users for timely data. Only a few weeks ago we at Statistics Sweden were criticised by one of our customers for being late with the presentation of our CPI in the form of "data shooting". We had been 25 seconds late, and the customer found that unacceptable.

And we all have experienced demands for less costly products. These - and many other examples could be given - are also examples of aspects of quality in the user perspective.

Within SCB we work since about five years to develop our operations with the help of a quality improvement system of the kind that has been used in the manufacturing industry for many years. Our work is based on what is called TQM - total quality management, but it is our opinion that each organisation must choose its own mix of ideas, priorities and tools. All kinds of comprehensive quality pro-

grams contain the same basic ideas: customer focus, the participation of all the staff, concentration on working processes, systematic and sustainable work, continuous improvement.

Of course there are differences between a public sector organisation and a private one. Defining the customer is often more difficult in a public sector organisation. The final objectives are more easily described and evaluated in a private company. And there are of course lots of other differences. Still, I believe that representatives of the public sector often tend to overestimate the differences between the public and the private sector in these respects.

 You can certainly implement a quality improvement program in your organisation in a low profile way with different measures introduced successively. That is one way of doing it - it has its advantages. You can for instance avoid some of the - sometimes very understandable - resistance to new concepts from parts of the staff. If - on the other hand - you choose a higher profile when introducing the program and send loud signals out into the organisation that from now on we are going to work in a new way you may be sure that many of your staff will use up a lot of energy to oppose the program. We chose the high profile way , and we certainly have had our share of opposition.

Still, I'm quite convinced that we would not be where we are today if we had chosen another way of implementing our quality improvement program. Our approach has generated not only opposition but also a lot of positive response. Many improvement projects have been carried through, a considerable part of our staff has got concrete of working with TQM tools and are very enthusiastic about the concept. There is an impatience within the organisation to go on with the development of our systematic quality improvement work

Now, this is one way of looking at the quality issues concerning a statistical agency. There are others and we have tried to present different angles at this seminar.

We will start this morning by discussing. conflicting quality objectives - something that is of a great practical importance for all of us: How do we strike a balance between conflicting goals in a situation where different users have different expectations.

We will discuss metadata - an issue of increasing importance when the technical development creates new possibilities for users who like to process the data themselves.

We will hear about quality assessment criteria from a general point of view - not the statistical perspective.

After lunch we will discuss two different approaches to quality improvement work: The Finnish TQM-approach and the Dutch way - statistical auditing, described as a pragmatic approach to quality management in a statistical office. How different are the approaches? And how pragmatic are the Finns? We will talk about that.

After that we turn to a specific quality issue: The treatment of non response. At least for us in Sweden this is a very important topic and a matter of growing concern.

Tomorrow we will go over to the international perspective and hear about a comparative study made a few years ago on quality improvement in some different countries. And finally we will discuss harmonisation of quality work in the European Statistical System - a major issue for all of us.

THEME 1: QUALITY IN STATISTICS - MAIN TRENDS AND CONCEPTS

1.1 Tim Holt and Tim Jones, Statistical Office, UK

Quality work and conflicting quality objectives

A. 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.

B. 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.

CUSTOMERS FOR OFFICIAL STATISTICS

Cabinet
***
Parliament
***
OGDs
***
Health
***
EU & International Bodies
***
Local Government
***
City and Finance
***
Business
***
Academia
***
Citizen

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?

B.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 Eurostat[1] in the context of quality in business surveys. The sources of error may be subject to both variance and bias. The impact of 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

---

[1] *Document Eurostat/D3/Quality/96/04 final*

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.

B.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.

B.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

B.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.

B.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 Mahoney[2] 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 for-

mulation to the delivery of results to be completed within six weeks.

B.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.

B.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.

B.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 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.

C. 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.

C.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

---

[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.*

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.

C.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-burden-

ing 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.

C.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.

C.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 recogni-

tion 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.

C.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.

C.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.

C.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.

C.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.

C.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

AUDIT   MONTHLY PRODUCTION   REVIEW

Audit Local Collection

Audit Monthly Production Processes

Local Price Collection

Central Shop Prices

Centrally Calculated Indices

Data Processing & Validation

Briefing & Checking

Monthly Publication

LONGER TERM - Annual Strategic Review

SHORT TERM - Monthly Review
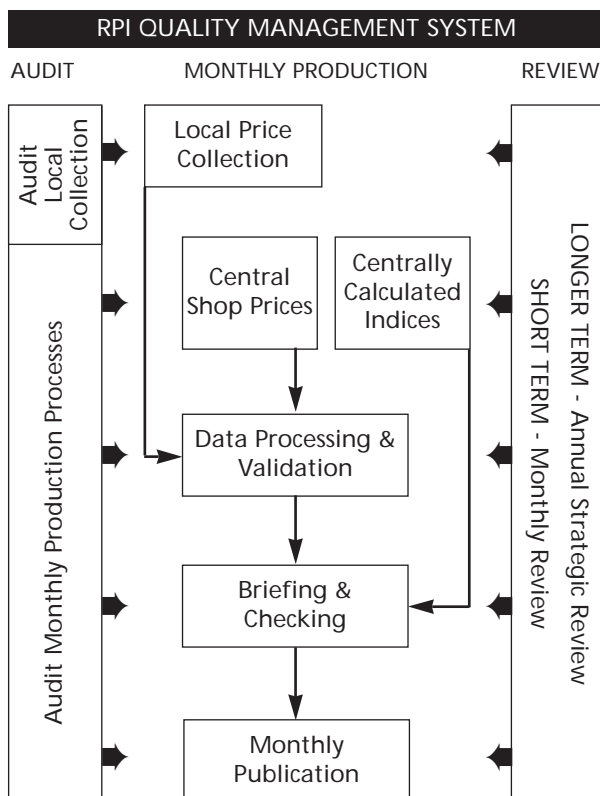
There are two parts to the review system, designed to promote "continuous improvement": a short term monthly review and a long term strategic review. Towards the end of each monthly processing cycle, one day is allocated as a Quality Day. The day starts with a review of the previous month's work. It 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 iden-

tifies 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.

D.   QUALITY CONFLICTS

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

D.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.

D.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.

D.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.

D.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 under-coverage (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.

D.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 compar-

ing 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 classification for a time to compare differences and ease transition.

E. 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.

1.1.1 Discussion: Erich Bader- Statistical Office, Austria

Standards determine the culture of statistics

The paper of our British colleagues deals with the complex topic "Quality in Statistics" in such a systematic, comprehensive, balanced and practical way that I confine myself to illustrate a small number of additional items.

First of all, it is good to hear that the topic of safeguarding quality has become more and more important not only with regard to economic enterprises but also in statistic during the last few years. A number of reasons are responsible for this development

◗ The period of quantitative expansion of statistics is coming to an end for financial reasons as well as due to a growing resistance against the conduct of surveys. More and more importance will adhere to our efforts to render the existing data more expressive. This is part of the duties of safeguarding quality.

◗ A growing number of institutions, companies, and persons depend in their decision taking upon objectives, plans and estimates produced by means of statistical material. Therefore, it is of basic importance to know the limitations of data quality and to reduce the inadequacies of data.

◗ In particular, many decisions of the EU committees are based upon statistics. This is why it is very important to the European Union, apart from having high-quality statistics on individual countries, to be able to guarantee a minimum quality standard for all member states. Moreover, international data require special demands on quality, namely, to be comparable in content and in fact, or, in the terminology of Holt/Jones, the demand of coherence and of consistency, which might be completed by the term of international comparability.

◗ Increasingly, quality safeguarding is not merely the task of individual statistical experts but also a management duty. I can only stress what Holt/Jones have said with regard to the term of quality, and particularly that it exceeds the term of accuracy by far, and that it is finally defined by the requirements of the users of the data.

This last item, however, is relative: Mr Holt says that quality is defined by the users of statistics and what they consider important. This, of course, implies the question: which kind of users? As different users have different views of what is important. Thus, a civil servant, a scientist, a journalist, a politician, a representative of a particular interest group and an interested citizen normally differ greatly in their expectations of statistics. They probably all have in common that each of them requires data on a current topic as soon as possible. With regard to the accuracy of data or the relevance of the concepts used respecting the topic in question, the above mentioned user groups tend towards very different attitudes. To exaggerate, one might say that our experiences may be summarized in the following way: quality (in the sense of accuracy and relevance) is important, above all, to those users who usually receive the statistical data free of charge or at very favourable conditions and in those cases, respectively, where costs have to be born by the tax payer instead of by the immediate receiver.

On the other hand, not everything that is important to the user depends upon quality. A buyer of statistical data - as a buyer of cars - has different aspects of the product in his or her mind, and tries to find the right mixture for him or her. When you buy a car, a certain quality standard is a necessary precondition, for sure, but sometimes even more important are things like the design, the colour, the extra equipment, advertisement, and much more. Without transferring these individual components of car selection to the buyer of statistical data we may, nevertheless, state from our experience that the consumers of official statistics prefer many other aspects to the quality aspect. In this connexion, I would like to refer to two examples which I would like to summarize under the slogan "Fascination of the last digit".

When, in 1974, the Austrian data bank ISIS was

opened to the public, data were presented up to their third or fourth valent digits, only. One reason was the economy of the memory, another one was the accuracy of official data: in all probability, there is no set of official statistics, whose accuracy transcends the area of one per thousand, which makes the rest of the presented digits a "white nothing", that could equally be decided by throwing the dice. Many users of the data bank have protested against this form of presentation involving troubles: ISIS data did not correspond to the data published in books up to their last unit. What was more, due to the problems based on rounding tables stopped being additive. In consequence of the persistent resistance of the users and, of course, the permanent cost reduction in storage media OSTAT has published case figures up to the units place for a long time now, even though results have not become more accurate than before.

Another example: as it is usual to calculate indices up to one decimal digit some non-European! - countries with former hyperinflation still presented one digit **after** the decimal point, even in cases of already five-digit price indices and although the smallest monthly change had occurred in the second place in **front** of the decimal point.

Seeming accuracies are fascinating to the expert statistician, too. For years, the "Human Development Index" has gone round with the claim to enable a comparison of different societies by summarizing different statistics to form a kind of superstructure. Everyone of us certainly knows that the "grip for reality" often produces fuzzy data, and have this in the back of our minds when we are digging in basic data. On the other hand, we all have this dream of images of reality free from contradictions. Moreover, policy, media, and the population should confide in official statistics. This is why it is that difficult to explain to everybody why the required accuracy of every statistics must depend on its most important uses, and every extra effort is useless.

Therefore: part of statistical accuracy is its **psychological component** - with respect to the clients, the consumers and - let's admit it! - even with regard to the producers of statistics. There is nothing I should like to add to the therapeutic propositions made by Mr Tim Holt and Mr Tim Jones.

Safeguarding quality as a cross-sectional duty

To begin with safeguarding quality, there must be a system according to which we can act. In this, everyone of us will agree with Mr Tim Holt and Mr Tim Jones. This turns the issue into a management task, it requires guidelines, documentation, permanent monitoring, discussions between clients and users, and it makes part of the distribution of duties in a statistical office, in the development of personnel and in many other directing mechanisms. During the process of becoming a member of the European Union, we were busy in Austria to prepare the new individual statistics with their quality improved. In order not to neglect the cross-sectional duty we used the data standards of the International Monetary Funds, which were a great help, as well as the relevant Eurostat initiatives, as for example the foundation of a special unit for the safeguarding of quality. This signifies that, during the last few years, aiso Eurostat has increased its efforts with regard to quality work, and has defined this area as a central management task in the sense of a TQM (total quality management): ISO Standard Specification 8042 on the definition of quality at statistical offices was translated (producing a similar list of quality criteria as contained in the Holt/Jones paper), moreover, a pattern was worked out of total quality management at Eurostat's, a number of projects was started in this field and, finally, a quality manager was appointed responsible for the whole area of safeguarding quality, in the framework of a special unit working directly with the director general of Eurostat. This topic was already treated at length at the Mondorf Seminar on "The Future of European Social Statistics" in the paper of Mr Bernard Grais on "Statistical Harmonisation and Quality: The Case of Social Statistics", a seminar which took place in late March of this year.

The safeguarding of quality, however, is usually not free of charge. In some individual cases, it may result in economizing resources (if the analysis of a set of statistics opens up restrictional opportunities), in general, the expenses will exceed the savings. However: quality has its price, and guaranteed quality costs even more.

Little can be added to the individual components of extensive quality definition of Holt/Jones. It will, perhaps, be useful to compare their criteria of statistical quality with those worked out by Eurostat:

QUALITY CRITERIA HOLT/JONES	QUALITY CRITERIA EUROSTAT
Accuracy	Accuracy
Relevance	Relevance
-	Completeness
Coherence	Coherence
Consistency	-
Continuity	Comparability (with regard to time and place)
Timeliness	Actuality
Accessibility (of data)	Accessibility and transparency of information
Revisions	-
Quality description	-

If you compare these lists of quality criteria you realize that Eurostat explicitly mentions the criterion of local comparability, which, however, does not come as a surprise. The criterion of temporal comparability corresponds to the term of continuity used by Holt/Jones. Eurostat also uses the term of completeness as quality component meaning that the area where statistics are available reflects the requirements and priorities mentioned by the users of the European Statistical System. This term, however, might be integrated into the notion of relevance.

On the other hand, Eurostat does not explicitly state some of the quality components mentioned by Holt/Jones: consistency, infrequent revision, and quality description. it would certainly be interesting to know the relevant considerations of Eurostat.

In both lists, the existence of a sufficient documentation system of statistics seems to be lacking, as the best of statistics cannot be appreciated and interpreted accordingly if they are not described sufficiently for the user. The term of quality description used by Holt/Jones seems to cover only a small part of this documentation job, and, presumably, this is also true for the expression used in the Eurostat list of "Accessibility and transparency of information".

Special attention is paid to timeliness in the Eurostat paper, apart from the topicality mentioned in both lists. It refers to the publication at fixed dates announced in advance. In this connexion, I would like to point out the Special Data Dissemination Standards (SDDS) of the International Monetary Funds focussing this aspect of quality. In this way, all users of certain official data shall be guaranteed the same conditions of access to prevent starting advantages with regard to stock-market prices and other economic reactions. Under the heading of "Quality", these standards, by the way, comprise a short description of the methods and data sources used.

The safeguarding of quality is expensive, and therefore we should be careful with it. In the field of statistics which must be supplied to Eurostat, a division of labour between the national statistical offices and Eurostat seems advisable.

Another indispensable precondition of a systematic quality improvement to me seems the **better documentation of meta data**. In our data banks, the real data of statistics are well described. A documentation system of the same standard for meta data - that is, all information you need to be able to handle the real data - is lacking. First steps have been made, in the printing as well as in the data bank area: concepts, terminology, definitions, footnotes; sufficient in case of many individual sets of statistics, convincing and easy to use. With regard to official statistics, which considers itself to be a consistent system of duties, methods, and solutions making "the figures speak", this is not enough. A systematic documentation of meta data will be able to compensate for the "loss of language" produced by the computer.

Quality as a question of costs

All statistical offices are working under conditions of restraint. Quality today is no academic topic but a question of costs. I would like to stress this point because it seems to be just a peripheral phenomenon in the Holt/Jones paper.

I suppose that we all agree: quality improvement may lead towards economizing, for example, when samples are dimensioned and stratified. in general, however, they are associated with expenses, and that, in all stages of the statistical production process. And this expenditure is not always easy to explain to clients, media, little informed users, and, above all, to those who finance our work.

This is why I consider the working out of **minimum standards** very important representing supranationally agreed quality guidelines. These minima were first defined in an economic statistical connexion and are intended to be transferred to social statistics, as well. I once more refer to the paper produced by Mr B. Grais on the occasion of the recent Mondorf Seminar of Social Statistics ("Statistical Harmonisation and Quality: The Case of Social Statistics").

These standards must be produced in co-operation with those institutions which nationally and internationally - work as reservoirs of users' requests. The problem, of course, is the fact that many users have no conception at all of the expenditure involved. Statistical experts, therefore, must often plead not for more but for less but properly used accuracy. Nevertheless, we should encourage each other mutually to strive for quality even in the face of exaggerated users' requests, and to create this sort of minimum standards as tool.

Statistical quality as a characteristic of "political culture"

The whole text of Holt/Jones mentions the term "harmonisation" just a single time. Why is this magic word for compromise missing? Maybe because the safeguarding of quality is also part of our professional ethics, and because our professional ethics demands high standards with regard to the handling of reality.

Every conversation about questions of quality in statistics seems important to me as it confronts us once more with the "why" and "what for" of our work. Official statistics is an indispensable component of the "informational environment" of democratically constituted societies. It reminds of open questions, of problems which happen, of size and proportion. It is oriented towards certain obligations with regard to society and the public, to clients, respondents, and consumers. These standards determine the "culture of statistics" and, at the same time, the "political culture" of a country.

We owe to Mr Tim Holt and Mr Tim Jones that - apart from many practical examples - they have once again reminded us of this value of our work.

An examination of the quality of the whole process.

Quality - as we are all well aware of - is a very elusive concept, not least in statistics. It is a phenomenon that does not lend itself easily to precise definition - let alone precise measurement. How do we determine what part quality changes play in price developments? How do we break down an observed and measured nominal change in value into a pure price component, a pure volume component and a quality component? What do we mean by quality of life? How do we measure it? And I could go on.

Here, however, we are not concerned with quality as a term in statistics but with the quality of the statistics themselves, of the products and services of our institutions, even the quality of our overall operations. But what do we mean by quality is this respect? Even the most cursory study of the papers that have been submitted for our discussion, reveals that the term quality in the context of statistical production and services is no less elusive than its counterpart we have to grapple with in our statistical analysis. It seems to be very difficult to define and it is clear that we attach a number of meanings or interpretations to that concept. Or to use the expression so often used in the paper by Holt and Jones, it is multifaceted.

The British paper emphasizes this aspect. Gone are the days when quality equalled accuracy. Come have the days when we can more or less compare quality to satisfying user needs, as I understand the paper to say. More precisely; quality can be evaluated in terms of several criteria; accuracy, relevance, consistency, continuity, timeliness, accessibility. Or to use the tree criteria of Paul Champsaur in this paper on the metadata and quality work; the evaluation of the quality of statistical data can be made in terms of the criteria of their pertinence (relevance), their transparency and their independence. Such an evaluation, Champsaur points out, depends on metadata, its generation and availability.

Whenever one studies papers and participates in a discussion of this kind the underlying thinking is invariably - what do I get out of this? What lessons can I learn? What experience of others or models can I bring back home and apply in my own yard? From the British paper we learn about many different connotations of the quality concept and how these can be addressed. Moreover, the paper indicates that conflicts may arise between different objectives or facets of quality. Interpreting it rather freely, the paper can be said to emphasize a drastic change in both our statistical thinking and services. In the old days we probably assumed that there was a one to one correspondence between a certain data collection and a certain product, that a given collection of data would lead to one specific product or service. Nowadays, however, we may rather want to discuss needs and processes. We have to satisfy this specific need for information, it has to be done speedily, with a specified degree of accuracy, the information has to be accessible, understandable and documented by metadata. Looking at it in another way, a certain survey or collection of data can give rise to a whole host of products and services depending on the type of need we are addressing. To this I would like to add that it might be fruitful to differentiate between products or services, not only with respect to their content but no less on basis of the mode of dissemination.

In the analysis of the quality of our work we can not allow ourselves to focus on any one aspect of product or end result; we have to study the whole process in terms of quality and efficiency and evaluate the end result in terms of the quality criteria mentioned before.

But how can this be achieved in practice? Obviously, there are several approaches. But these may not all be suitable in all circumstance or not equally suitable. Here, I am mainly thinking about size, both the size of the operation or quality management project in question and the size of the institution where it is to be applied. I would for instance think that an overall quality project of the scope of the ones in Statistics Sweden, Statistics Finland and Statistics Netherlands would hardly be applicable in the very small setting of

Statistics Iceland. The reason is simply that it would be neigh impossible to absorb such large protects, they would be much too expensive and could neither be managed nor manned from within the organization nor without disrupting the regular operation and services of the institution. Hence, a partial and a more piecemeal approach may be called for. It is interesting to note that the ONS management seems to have opted for some sort of such an approach instead of the more overall one, such as the TQM approach of Statistics Sweden. I would be interested in hearing Tim Holt discuss why this is so, why such an option has been taken.

Another issue is that bringing the quality message to and convincing staff members in a small, close-knit institution may require an approach different from that in a larger institute. The basic task is, however, the same; to achieve a degree of cooperation and con-

tinuous consultation with staff members on the relevance of their products, the timeliness of the output, the accuracy, consistency and so on as well as on the documentation or metadata that has to be an integral part of the process. In short: the quality of the process. This can also be seen as consulting them on priority setting; in order to ensure that we do indeed give priority to products and services that are in visible demand or where a real need may be observed and that fulfill our quality criteria in other respects. In other words perhaps, there is a problem of ensuring that the priorities of the services and overall operation are viewed in the same way by staff and management. And if conflicts arise between management and staff with respect to the quality of the process and priority setting, how should this be treated? I leave on this note or this question hoping that it may be taken up in the discussion.

1.2 Paul Champsaur - Statistical Office, France

Metadata and quality work

Introduction

I would like to start this talk, in which I will be focusing on a domain in which I am not a specialist, with a clarification of the vocabulary I intend to use including, in particular, a statement of the meaning I attach to the concept of metadata.

Metadata and Metainformation

In the strict sense, metadata are data which provide us with information on the data themselves. For example, the monthly date of publication of the balance of trade is a metadata attached to that statistical series.

But this term is generally used in a broader sense, and refers us to every element of information regarding data, whenever it gives an insight into the production process and enables us to use the data correctly. Hence, for example, the description of the method of seasonal adjustment of the external trade figures will be treated as a metadata, although it should preferably be referred to as metainformation. It is generally treated as a metadata in practice, however, and I will not differentiate between these two concepts.

A. QUALITY OF STATISTICAL INFORMATION AND METADATA

The "quality" of statistical data can be evaluated in terms of three criteria: their pertinence, their transparency and their independence. But the evaluation of these criteria depends, precisely, on the production of metadata. Let us look at the facts:

Pertinence

The pertinence of an item of statistical information is determined by its capacity to observe the object it is intended to measure. Hence the need to know the concept on which the item of information is based in order to judge its capacity to give a valid account of the phenomenon one wishes to analyse. But it is equally important to know the degree of precision of the statistical observation. The evaluation of this cri-

terion therefore makes it necessary to take a very close interest in the concepts involved, the measurement of error in the estimation process, the rates of non-response, the representativeness of samples, and the analysis of revisions of the data.

Transparency

The transparency of an item of statistical information can be measured in terms of the evidence of its availability (the delivery of an item of information on a date which has been announced is a response to the expectations of the economic actors), its status (confidential, public, provisional result, definitive result), and its comparability across time and space. To evaluate this criterion, one therefore needs not only timetables for the publication of statistical data but also details of the methodologies used for collecting and processing the information, in order to analyse the production process and decide how the data can be used. It will also be evaluated via the quality of the documentation supplied to users, relative to their needs.

Independence

The independence of an item of statistical information is obviously an essential guarantee of its quality. It can be evaluated in terms of the way in which it is produced. The more publicly the choice of the objects to be measured and the corresponding concepts have been discussed, the stronger the guarantee that the information produced will actually be used. In France, that is one of the missions of the CINS (Conseil National de l'Information Statistique). The parties involved in the discussion of these choices are the statisticians and the social partners (i.e. representatives of the trade unions, employers, the administration and the universities).

Measurements of the precision of statistical data, definitions, and information on data are metadata. They are destined to provide a basis for measuring the quality of work carried out in the realm of statistical infor-

mation. An adequate supply of information on statistical data will enable the evaluation of their reliability and pertinence and the conditions of their use by deciders or other users.

B. THE GROWING NEED FOR METAINFORMATION

The internationalisation and transformation of the conditions of dissemination and utilisation of information are changing the patterns of supply and demand in the domain of statistical information.

On the demand side

The development of internationalisation in response to the globalisation of markets implies that various actors, and not only the traditional users of information, are keen to acquire a more detailed knowledge of economic and social conditions in different countries, in order to improve their decisionmaking capacity.

The development of special uses of statistical data by institutional decisionmakers (e.g. in the domains of jurisdiction in the European Union and the standards of the International Monetary Fund) makes it essential to provide them with a corpus of information on the data which they use, with the aim of enabling them to come to properly informed decisions.

On the supply side

In parallel with this growth of demand, the opportunities which have been opened up by technological progress are bringing about changes in the conditions of dissemination of information which are reshaping the contours of the supply of statistical data.

Just like other types of information, statistical information is now available much more quickly and much more easily than in the past, and consequently not only tends to circulate in larger quantities (inter alia because different statistics may well be intended to measure the same phenomenon) but are also of interest to a much wider public. On the other hand, the connection between the producers and the consumers of statistical information has become more tenuous, from the geographical, institutional and cultural points of view. The dissemination of more and more statistical information to more and more users

implies a desire, on their part, for access to information on the characteristics of the data, in order to be in a position to choose the statistical information that best meets their needs.

Statistical information has thus been plunged into an unprecedentedly competitive climate in which it is confronted (like other kinds of products) with the need to improve its quality. Every operator wants to have access to information which can be weighed against other information in terms of reliability and comparability. This can be likened, moreover, with what happens in the domain of industrial production. Knowing how to produce is not enough: one has to be able to do so under quality conditions that will optimise the satisfaction of one's "clients", whoever they may be, while at the same time providing them with the elements of information that will enable them to take informed decisions.

The needs of the "clients" in the domain of statistical information

In the domain of statistical information, it is possible to distinguish at least two main types of "clients", namely the "specialised" users and the "general" users. This is obviously a rough distinction, and the real situation can be more precisely described as a sort of continuum. The same user may have a general need for one type of information and specialised needs for other types of information. This kind of shift between the two types of need can also be a temporal phenomenon. The two kinds of clients have a wide and shifting range of information requirements. Let us now consider a few examples of two kinds of information which are most frequently needed.

The financing needs of public authorities

The specialised users, and especially the deciders who are responsible for the application of the Treaty on European Union, need a large volume of information for their evaluation of the pertinence of this indicator (information on concepts, definition of the sector, main components, timespan between provisional and definitive, dates of availability, jurisprudence of Eurostat with regard to the composition of the sector or the treatment of certain specific operations). This

information is available from the producers, but it needs to be processed to make it user-friendly.

The generalists among the users need more summary information consisting, for example, in the presentation of the objectives and the concepts employed, which is different from the information used by their specialist counterparts (a case in point could be the explanation of the difference between the public sector borrowing requirement and the budget deficit, or the presentation of the adjustments of the national accounts to ensure their conformity with the Decisions delivered by Eurostat).

Multiple indicators, such as consumer price indices

The specialist users of statistical information need to know why and how the results of a national index differ from those for the same index in its harmonised version (field of the index, technical treatment of the detailed elements of the index, weightings).

The generalist users need above all to know there are two types of index and to be informed of the reasons for this distinction, the contexts in which they are used, and what they measure, in order to be able to decide which of the two types they are going to use in the light of their specific problems at a national level or when working on an international comparison (involving, for example, the use of a homogeneous deflator).

Monthly trade statistics

The specialist users of these statistics may need a detailed item of information, which may be important for reasons with a more than strictly conceptual dimension. They will certainly not be satisfied with the balance of trade alone. They will at least want to know the fields covered by imports and exports. They may wish to use data in real terms and therefore to know the procedures used for splitting volumes and prices. They may want to analyse the trends across time and space, and therefore to know the method on which the delimitation of geographical (territorial) units is based, the product nomenclature that is used in this domain and the methods of calculation of the different coefficients of seasonal variation. For the

correct interpretation of short-term trends, they may even frequently need an outline statement of the criteria on which a commentary is based.

The generalist users need a more general insight into the precise definition of foreign trade, the ways in which it is connected with and differs from the balance of payments, and the accounting methods and their consequences as far as the interpretation of the results is concerned.

These few examples show that the scale of users' needs is potentially enormous and may be reflected in the need to mobilise a correspondingly large volume of information which will have to be processed on the basis of the information which the producers are in a position to provide. Failure to present it to the users in a form which meets their needs will greatly reduce the utility of this information.

C. ORGANISATION OF A SYSTEM OF METAINFORMATION AND ACTION TO ENSURE THE IMPROVEMENT OF QUALITY

Producers will have to respond to this general developmental process by stepping up their efforts to master the conditions of their production processes, with all due regard, in particular, for the existing financial constraints. It is also essential, at this level, to improve the efficiency of the NSIs by rationalising their production processes, in other words by organising the elements of information which are the sine qua non for success on this front.

Enrichment of the flow of information to users can help to raise the level of quality

Metadata are also of interest not only to the data producers themselves, irrespective of whether they are presently producers (as the managers of statistical operations) or future producers (as their successors) or more remote producers (colleagues in another NSI or an international organisation). This collection of metadata contains, a priori, every element of the information needed for the data production process. If the concepts or the nomenclature of the domain are standardised by an international refer-

ence, the standard is obviously a component of the metadata needed by the producer. But the statistical methods must also be defined; and the exchange of information on these methods can then enable the identification of best practices and promote the emergence of international standards. But the establishment of this kind of system is a major investment. It can sometimes be particularly difficult to differentiate, in the wide range of methods employed in several different countries, between those which reflect the characteristics of their statistical systems and those which reflect their choice of methodology.

The contribution of European harmonisation to the improvement of quality

The constraint of harmonisation, when it exists, is a positive factor. In the European Union, for example, the obligation to establish IPCs has been the essential precondition for the convergence of practices in this domain. But this had all been preceded by a long phase of reflection on the methodological aspects: the objective of harmonisation wasn't invented yesterday! In a sense, the European statisticians specialising in the production of price statistics had already established a large part of the corpus of metainformation they required. It also needs to be emphasised that the extent to which harmonisation can bring about a general improvement in quality depends on the extent to which it leads to alignment with best practices. I have already had occasion to say that aiming for exhaustiveness at any price would have a double impact on quality and the comparability of the harmonised indices of consumer prices.

Organisation of a system of metainformation

Methodology for the organisation of a system of metainformation

Statistical data are managed by information systems. Likewise, metadata must be managed by metainformation systems. I can only report here on the situation in France. Metadata have long been taken into account in our statistical applications, even if the term "metadata" has not been explicitly used. The two reasons for which the INSEE started to take a special interest in the documentation of its statistical

operations some 15 years ago were its concern to streamline the transmission of its know-how throughout the institution and its determination to improve its consistency on the data processing front. Hence its creation of a methodological framework, known as the DDS ("Dictionnaire de Données Statistiques"). In its operational form, this framework enables the electronic storage and management of the documentation relating to its statistical operations. In its early days, priority was given to the producers' point of view, but we now pay more attention to the requirements of the disseminators and users of statistical data, and it is presently possible to extract from such a reservoir of metadata the elements we need for the production of user-friendly documentation which can be made available in various forms. This general methodology for producing (i.e. collecting, processing, disseminating) metainformation has admittedly not yet been generally adopted. We have so far tended to leave the initiative to the persons responsible for each statistical domain. The cost of collecting and formatting all this documentation is frequently regarded as prohibitive by the managers of statistical operations, who have to work to strict production deadlines. We nevertheless intend, in due course, to define a minimal documentation standard for the operations carried out by the INSEE.

Striking the right balance between quality and cost

The costs of collecting, storing, formatting and providing access to metadata is far from negligible. The need to meet all these requirements raises the question of striking a balance between the cost of "production" of all this information and the level of quality we can reasonably expect to attain. This arbitration process will probably require a phase of reflection on the priorities to be established and the definition of the "quality standards" put into practice at the European statistical level. Hence the conclusion that Eurostat's efforts in this context, in the form of the project for the establishment of the European Reference Environment, can only help us in our endeavours to complete a new phase of development on this front..

1.2.1 Discussion: Jan Plovsing- Statistical Office, Denmark

Better availability of metadata

I would like to thank Mr. Paul Champsaur for an interesting paper on the subject of metadata and quality work. I believe you have highlighted some very important aspects about metadata as part of delivering quality to our users.

I agree that metadata should be seen in a broad perspective, which is also in line with the recommendations made by UN that statistical metadata are the data needed for the production and usage of statistical data.

You start by defining quality as relevance, transparency and independence and argue that metadata are needed to evaluate these criteria. I agree that these are important factors for evaluating the quality of statistical data. I would like to add an important distinction: 1) between the metadata needed to ensure proper production leading to statistics of a reasonable quality 2) and the metadata produced by the production process and used to evaluate the quality of the result. The first group would contain definitions of concepts, methods for data collection and processing and the production timetable. The second group is data about response rates, number of missing values, estimates of errors, etc.

You continue by giving some very useful arguments for the growing need of metadata among the users of statistical information. I would like to stress the word "growing", because the need for metadata is not a new phenomenon, it existed long before the word metadata was invented. My reason for stressing this is that the word metadata still seems to scare some people, because they don't realise that producing metadata has, in fact, been a major activity in their everyday work for years.

I would agree with you that globalisation and the building of the European Union increase the need for metadata to make more and more data collected by the national statistical institutes comparable. But also the increased use of statistical information by a growing number of analysts and researchers at the national level increases the need for metadata.

It is perhaps more precise to talk about a need for better availability of metadata. We may consider that the large corpus of metadata exists in order to conduct our production of statistical data in a proper way. The problem is perhaps more the availability of the metadata.

You talk about the general user and the specialised user. As you say, this is a rough distinction, but I find that it is useful for the debate. However, as I read your paper, the general user is the user needing and using metadata about concepts and definitions and broad methodological issues. We should, in my opinion, not forget the large group of users who are not using any metadata besides what is disseminated as titles, headings, labels and footnotes in direct connection to the statistical data themselves.

You suggest that, as the number of users grows and the number of potential sources of information increases, the users will demand more metadata to make the best choice of the statistics they want to use. But even if the number of people asking for metadata at the general level is growing, it will still be a group of modest size.

It raises the question as to what the real demand for metadata on quality is, at the "general" and expert level? How many users are really going to seek this type of information regarding any single survey ?

I am very pleased with your examples which clearly demonstrate the problem that what is relevant for one user may not be relevant for another user. What may be stressed more is that we should find a way to serve them as one group. This may be linked with the idea you propose later: to define a minimal standard for documentation, alias metadata.

Next you talk about the organisation of a metainformation system and actions to improve quality. The first step, as you mention, must be to make sure that metadata needed for the production process are available. If our basic production is not properly based on

metadata, we can't produce data at reasonable quality and declare that quality. However, is the problem lack of metadata, or is it lack of visibility? The fact may be that it is a mixture. What I miss in this part of your paper is some distinctions between different types of statistical systems. The problems related to a questionnaire based survey are clearly different from the problems of a survey based on reuse of administrative data like income-tax forms.

What is the degree of control we have of the way data are collected and the definition of the concepts? The way we have to document a survey may reflect these differences.

Another important distinction is between the basic surveys, whether based on direct data collection or reuse of data, and "secondary" statistical information systems like the national accounts. I point to this, because it raises the question if we should strive for one single system to handle all this information. We have to admit that we have not, so far, been very successful in achieving this. The metainformation may well be there, but it is scattered in an not very well organised set of documents. It has a low visibility, even to the producers themselves. And worst of all, it may only exist in our minds.

You also raise the question of promoting best practices and use of international standards. In my opinion, and as I read your paper, one of the basic problems is visibility. Try to locate information about NACE, or the Combined Nomenclature on the Internet. You will not find much. We may define any number of standard classifications or general concepts, but if they are to have any effect, it must be easy to access them. This is true at national level as well as international level.

Harmonisation as you suggest may be a road leading to improved quality. But we also know that harmonisation takes much effort and long time to achieve. There are also a number of areas, where harmonisation is in conflict with the fact, that the national concepts are based on national legislation. You may take our discussion yesterday about the concept of unemployment as an example. We are in need of ways to deal with the issue of international comparability

that can be implemented with less effort and in shorter time and on a broader scope of subjects.

I may suggest that in the short term we try to establish a reference system (or an international standard of contents) for different areas of statistics. Then the actual systems at national level should declare their difference to that reference system. So rather than harmonising, give the user a hint on how to compare.

I should like to comment on your last item, the organisation of metainformation system. Your experience in collecting metadata I think is shared by many organisations. The people in production and subject-matter departments claim that they don't have the time to make metadata. They have to meet deadlines of productions, they are short of staff etc. etc. But can we do anything to remedy this? If we had unlimited resources, then of course there would be no problem. However, that's not the fact. But I have argued along the line that metadata very often are already being produced, do already exist. Is part of the problems that we don't collect metadata when they are produced? Are we asking the producers to deliver metadata in specific formats, using specific forms? Are we blinded by the wish of the ideal system? If that is true, should we consider to make less ambitious but still useful metadata demands? Should we just make a simple document store, taking metadata in whatever format is has? Could that be a way to meet the high demands of specialised users? Do we need a shift from a highly structured database-oriented organisation to a more loosely structured organisation using a technology like the Internet and Intranet? Is that a way to reduce the costs of collecting and storing metadata ?

This does not, of course, solve the problems of establishing a set of minimal metadata for the general user or a set of metadata to describe comparability (at the international level). Here I would rather point to the Swedish system of product descriptions, a system we are going to emulate at Statistics Denmark and that is a central part in the UN recommendations.

I would like to conclude as you do that there must be a balance between quality and cost (of metadata, as elsewhere). This may lead to a need for reconsidera-

tion. Are the goals we set reasonable? Can we achieve a decent level of service to the market at lower costs? Can we afford to build very ambitious projects like the European Reference Environment or should we strive for less perfect but more afford-able solutions?

Questions.

1. Is the existence or the availability of metadata the greatest problem?

2. What is the real demand for metadata on quality, at "general" and "expert" levels?

- How many users will seek this type of information?

3. Should we strive for one single database-oriented system to handle all metadata information?

- Is a less structured datastore a way to reduce the costs of handling metadata?

4. Are the goals we set reasonable?
- How ambitious projects can we afford?

Johny Lindström- Swedish Institute for Quality

Quality assessment criteria

It's an honour for me to be able to address you here.

Let me first introduce myself. I am the president of the Swedish Institute for Quality which is a non profit foundation of 7 years.

We are quite a small organisation of 20 people and we are located in Gothenberg. Our mission is to stimulate and to contribute to positive development in all aspects of the Swedish society in a specific area. We will accomplish this by generating, compiling and distributing up to date knowledge in customary and distance development or quality management.

If one looks back into industrial history there are a lot of organisations working with inspection and in such companies the management style was very much about direct control.

Going on through the history of quality the self-inspection level appears. This self-inspection level was some kind of a delegated power for inspection.

But after that it was felt better to avoid the control and instead to do the right thing from the beginning. The management style in these kind of organisations was for example management by objectives. The customers were looked upon as a resource and you could work together with this resource to reach your goals.

One could use abbreviations such as TQM to describe it. What is very important in these type of organisations is that you focus on the customer who is the equal partner.

Then the modern way of looking on quality would be how to work in a service company.

The first challenge is customer as a concept. To understand who is the customer and to solve this problem you need to understand with whom you are working. The customers are those for whom

your organisation exists. Let me look at a statistics bureau then. Perhaps the main process is to deliver information to customers that they can use for their decisions. So what you are delivering, your product then, is something to be used for decisions based on data collections, statistical methods and so on. Further you have to define the customer and you may find that you have several customers. You may have the government, other agencies, buyers of specific studies, international organisations, internal customers, and the customers' customer is very often the public.

It is very important to continuously work on quality and to see to a follow-up. This brings me to the "what" and "how" concept. What do I mean by the "what" concept? Well traditional quality work at controller based quality work is to have some kind of input - output measurements. Or an output related measurement. You are focussing on input - output relations and you use words like profitability, efficiency, effectiveness, activities etc. This is necessary but to achieve real quality you must change the focus to the "how" concept. The "how" concept can be described like this: you have to ask questions on how are we working. The first question here is how are we working for example to identify the needs of the customer. The most common answer to this kind of question is that not so much is done, there is no systematic approach.

The next question then. You conclude that you do have an approach but to what extent is it used.

Then you must go back to the "what" question again and the question here is, what are the results achieved from the deployment of the approach. Of course those results should be achieved in relation to target values.

The last question is a very relevant one. How do you evaluate and improve your approach and deployment? To this kind of question it is often said

that there is no systematic approach or activity for doing evaluation and improvement of work.

To sum up, I think it is very necessary to change the approach in the quality thinking from the "control approach" to the "management approach".

To do an evaluation a set of core values are necessary to evaluate against.

These core values are described on the slides:- customer orientation, committed leadership, participation by everyone, competence development, long-range perspective, public responsibility, process orientation, prevention, and so on…

To conclude, organisations that have been successful on the world market or successful in the public sector changed their view on quality from control orientated to this management orientated view and the leadership managers are working very hard on this kind of learning perspective with questions like this "how" concept I just described.

Slide 1

THE SWEDISH INSTITUTE FOR QUALITY - **SIQ**

- Non profit foundation
- Founded 1990 (1986)
- Founders: The Swedish Government & SIQ Sponsoring Association
- SIQ Sponsoring Association..
- 130 members from all sectors of the Swedish society

SIQ MISSION

The Swedish Institute for Quality (SIQ) shall stimulate and contribute to a positive development in all aspects of the Swedish society.

We will accomplish this by generating, compiling and distributing up to date knowledge in customer oriented business development, as well as methods for implementation. We will also provide new impulses through international co-operation.

Slide 2

Yes but…   Job = to be occupied

This is none of My business   Anonymous customers

Management by
Direct control

Talk about each other   Quality = inspection

Doesn't know the employees

The Inspector

Group feeling   Interest for the job   Customers = target-groups
                                                        with certain characteristics

Management by Budget

Talk to each other
Informs   Quality through self inspection

Knows a bit about
the employees

The Delegator

Personal Responsability	Engaged in the work	Customers = resource to be used to reach goals
Speak with each other Two way communication	Management by objectives	
The Coach	Knows the employees	Quality through Quality System (QA)

Full responsibility in the group	Devoted to the mission	Customer = Equal Partner
Communicates together	Management by holistic	Total Quality Management (TQM)
The Leader	Feels for the employees	

THE QUALITY PHILOSOPHY OF SIQ

◗ Focusing the customer
◗ Management's responsibility.........
.... everyone's concern
◗ Continuous improvement
◗ Following up

FOCUSING THE CUSTOMER

The biggest challenge
in service quality

FOCUSING THE CUSTOMER

Challenge #1: Customer as an concept
Challenge #2: Not focusing the employees
Challenge #3: Not focusing the internal organisation

CUSTOMERS AS A CONCEPT

The customers =
Those for whom your organisation exists

FOLLOWING UP

The WHAT concept
The HOW concept

THE **WHAT** CONCEPT

Input/output or output related measurements reflecting
- profitability
- efficiency
- effectiveness
- activities

To what extent?	What result will be achieved in relation to the target values?
How are you working?	How do you evaluate and improve your work?

THE **HOW** CONCEPT

Q1: HOW ARE YOU WORKING?
Level 0:   No activity / no systematic approach
Level 1:   Some activity more reactive than proactive
          limited systematic approach
Level 2:   ....
....
Level x: Systematic, proactive and unanimous …

THE **HOW** CONCEPT

Q2: TO WHAT EXTENT?
Level 0:   No activity
Level 1:   Some activity shown in some areas with some degree of regularity
Level 2:   ....
....
Level x: …… everywhere ….. always

THE **HOW** CONCEPT

Q3: WHAT ARE THE RESULTS ACHIEVED (IN RELATION TO THE TARGET VALUES)?
Level 0:   No results available
Level 1:   Some positive trends can be shown
Level 2:   Many positive trends can be shown as well as some good levels
Level 3:   ………
....
Level x: World class resistant results derived from approach and deployment

THE **HOW** CONCEPT

Q4: HOW DO YOU EVALUATE AND IMPROVE YOUR APPROACH & DEPLOYMENT?

Level 0:   No activity / no systematic approach

Level 1:   Some activity but limited systematic approach

Level 2:   ………

…………

Level x: Systematic, preventative and unanimous
                 evaluations and improvement activities

	Approach	Deployment	Results
100%	Systematic, constantly being improved	Always in all relevant processes	First-class, consistent
0%	Anecdotal	Not deployed	Not reported

CORE VALUES

◗ Customer Orientation

◗ Committee Leadership

◗ Participation by Everyone

◗ Competence Development

◗ Long-range Perspective

◗ Public Responsibility

◗ Process Orientation

◗ Prevention

◗ Continuous Improvement

◗ Learning from Others

◗ Faster Response

◗ Management by Facts

◗ Partnership

SWEDISH INSTITUTE FORQUALITY

Fabriksgatan 10

S-412 50 Göteborg, Sweden

Phone:+ 46 31-335 17 00

Telefax:+46 31-773 06 45

E-mail:siq@siq.se

Serving the past and the future

Firstly I want to thank Mr Lindström on the interesting philosophy of the Swedish Institute for Quality. I must also say I'm grateful to Mr Carling for allowing Mr Lindstrom to give his presentation first, if not I wouldn't have anything to comment on since I only got his slides yesterday.

In listening to Mr Lindström I found that it was a bit difficult to translate the rather general framework of the institute into the actual operation of a statistical institute. However the main message as far as I can read is nicely summarised in one of his slides where he states that the focus should be on the customers, quality is the management's responsibility but everyone's concern. There should also be continuous improvement and of course most importantly how to follow-up - it's not a once and for all exercise. He makes this statement more concrete by defining three or maybe four dimensions in the actual approach firstly of all the approach to quality, then the deployment that is the application of the approach and then results of course and then again evaluation of both the results and the approach and the deployment and so on.

He has provided us with a general framework of what we will have presented to us in papers late this day and tomorrow but I have some question marks to some of his statements. The main one is the fact that their customer orientation or as we may say it user orientation, forms the basis of his model. And closely related to this is the "what" and "how" concepts. However, it seems to me that these concepts and frameworks are basically based on the production and selling of goods and services in the market situation. This framework is then reinterpreted to be also applicable in the public sector setting. But as spelled out for instance by Tim Holt earlier this morning and by Timo Relander later on this is not an easy exercise and I will here try to explain why.

In my view the problem is both with the customer concept and with the nature of the project we are provided; namely official statistics.

But first I will refer to the definition of quality according to the ISO 9000, which contains both stated as well as implied needs. I will say that it is easier to have some ideas about stated needs, and very often stated in a market place, but that it is much more difficult to have some ideas of this when referring to implied needs. These implied needs predominate in the demand for statistics and these are the most difficult to ascertain. The usefulness of official statistics does not decrease but increases the more it is used. This means that official statistics has to be at least partly tax financed to be provided at an optimal scale. Just reading Tim Holts paper and his listing of who are the customers, is a clear evidence of that you can't have anyone to pay the full price of something that is available to everyone. Even if we are talking about the policy makers they have a double role - both as user and provider of the funds necessary for the production of official statistics, and they will have to extend their views: or to pay for it or fully finance the production of official statistics.

What does this really mean to us? First of all it must be an important part of quality work in a NSI to determine the scope and the composition of the statistical system. Only part of that system can be determined through the market place. We get of course a lot of information from the market place on the products we are selling but that is only on the margin, not the product as such because that basically had to be fully financed or nearly fully financed by our tax payers.

In addition there are various methods for the measurement of user needs and user satisfaction and these methods are spelled out in some papers we will listen to later this afternoon, so I will not go into that. But I would add that our quality considerations should go beyond what we can find in the market place and what we can find from various surveys of user needs and expectations. Just let me mention a few items. First of all we are producing something that is part of the infrastructure of a nation and you can not find out much about the need for such an

infrastructure by just looking around. You have to go much deeper into it and I am referring now to what Paolo Garonna said earlier this morning about listening to singles, but what kind of singles and from whom. Secondly we are in the business with long leads in the production of statistics because if we are thinking about publishing new statistics it takes years.

First we are going though discussions, then formulating it, then we have to collect the data and then we are publishing. It may take 5 years from starting until we have the final product on board. The user needs there may have changed from when we started so we have to think ourselves for the future to be inventive and have imagination and think about what will be the future needs. Another effect is that these products last forever, more or less. For example, it is now 50 years or more since the establishment of the system on national accounts and the basic principles are still with us and are still very influential in both the policy formulation, in the research and so on. So we are not only producing something for the user today, we are really forming their ideas of how the society works in the future. Because these concepts of national accounts are used everywhere and those who invented this system 50 or more years ago, I don't think they realised that this exercise whould have such a long lasting effect on thinking and policy formulation for instance.

Secondly we have the responsibility of the memory of the nation. We have to produce statistics that are maybe not useful today but we have to do that to prepare for future generations studying the situation today. This is a very important task of a national institute.

So to conclude I think that the national statistical offices have to take into account other aspects and other purposes of official statistics than the preferences of our present users. It goes far beyond that and to me this seems to be the most difficult and yet not well structured part of our quality work. To satisfy user needs and requirements is only part of the work on quality within national statistical institutes and we have to be very aware of that and not concentrating only on serving the present. We have to serve the past and the future.

Thank you.

THEME 2 : QUALITY WORK- SOME NATIONAL EXPERIENCES

2.1 Timo Relander - Statistical Office, Finland

Total Quality Management and statistics

A. INTRODUCTION

Rapidly changing information societies are formulating new roles and ways of working for both Public Administrations and for statistical agencies. Total Quality Management (TQM) and related quality management approaches have proven to be successful in creating performance excellence in the private sector. In some countries TQM-thinking has also strongly penetrated Public Administrations. A number of statistical agencies, including Statistics Finland, have likewise started their journey towards TQM.

Today actually, it is impossible to find an organisation, e.g. a statistical agency, which is not aiming at the improvement of quality and continuous development of its operations. As a matter of fact, statistical agencies have long traditions of quality thinking. The statistical quality concepts belong as an essential part to our daily work. Standardization and documentation form a basis for statistics. It seems to be relevant to ask what will then be the benefits of TQM work to us? Is TQM only a new tool of skilled business and marketing consultants to expand their markets to the public sector and to statistical agencies? Or, could TQM, after all, be of some use to statistical agencies in their continuous efforts of improving quality? And if so, are the TQM-concepts applicable to public administrations and statistical agencies?

This paper discusses these questions by using our experience on total quality management. Chapter 2 deals with the general ideas of Total Quality Management as presented in the current management literature. Chapter 3 discusses the applicability of TQM to the public sector in the light of the Finnish experience. Chapter 4 focuses on the relations between TQM and official statistics and Chapter 5 describes some activities of the current TQM processes at Statistics Finland. Finally, concluding remarks are presented in Chapter 6.

B. TOTAL QUALITY MANAGEMENT - WHAT DOES IT MEAN?

During recent decades the concept of quality has changed from a very narrow definition to a very comprehensive one. The road towards TQM started initially with quality inspections, followed in the '30s by quality control where the focus shifted from product quality towards process quality. In the '50s the concept of quality assurance was introduced where the quality of the organisation came into view. In the period thereafter the concept of TQM was developed, where continuous improvement, empowering people, caring for people and involvement were guiding values, instead of compliance to specifications and allocating blame. This development was reinforced by the process of emancipation of the customer. See Appendix 1 (Dale 1994).

Like the term "quality" TQM also has many definitions. Mainly in any TQM definition there is some mention of management practices designed to continuously improve the performance of organisational processes to profitably satisfy customers. Talking about quality may be very confusing and misleading, because it is hard, even impossible, to define exactly what quality means. Adding the words total and management, does not make it any easier. The following definitions are based on work done by the American Society for Quality (1997).

Quality management can be defined as "That aspect of the overall management function that determines and implements the quality policy". Quality policy is defined as "The overall quality intentions and directions of an organisation in regard to quality, as formally expressed by top management". A dominant way of applying quality management is known as TQM.

One of the definitions describes TQM as "a systems approach to management that aims continuously to increase value to customers by designing and continuously improving organisational processes and systems". As the word "total" implies, TQM is concerned with managing the entire system, and not only subsystems, processes or functional departments.

Two other examples of TQM definitions (ASQ 1997) are the following:

1. TQM is "a management approach of an organisation, centered on quality, based on the participation of all its members and aiming at long-term success through customer satisfaction, and benefits to all members of the organisation and to society".

2. TQM is "a business philosophy that seeks to improve the results, including the financial performance of an organisation's management system, to guarantee its long term survival through a consistent focus on improving customer satisfaction and meet the needs of all of its stakeholders".

One of the trends to define TQM is to use the Quality Award criteria as a definition. The Malcolm Baldrige Quality Award criteria (Appendix 2) draws attention to the following issues:

◗ customer-driven quality
◗ leadership
◗ continuous improvement and learning
◗ employee participation and development
◗ fast response
◗ design quality and prevention
◗ llong-range view of the future
◗ management by fact
◗ partnership development
◗ company responsibility and citizenship.

Factors which indicate that TQM has started to change the culture in organisations should be identified. If the following factors are found, one is driving on a good TQM-road (Dale 1994):

◗ People see for themselves the need for quality

management tools and techniques

◗ Motivators and champions start to emerge from various parts of the organisation

◗ People talk about processes and not of functions

◗ Changes to procedures and systems are easier to make

◗ People show a positive response to recognition

◗ Employees are viewed by senior management as an asset and not a cost

◗ People volunteer to take on tasks which previously would have involved considerable negotiations between management and unions

◗ Ideas and suggestions start to flow from the shopfloor

◗ The distinction between the 'manager' and the 'managed' becomes hazy

◗ Quality improvement continues in the face of organisational instability.

In the current literature (e.g. Dale 1994), attempts have been made in order to determine what is the organisations' development phase on their way towards TQM. According to some researchers, it is possible to classify organisations according to six different levels to which companies have adopted and committed themselves to TQM: uncommited, drifters, tool-pushers, improvers, award winners and finally world class level organisations. There are very few companies that have reached the two top levels. But anyhow it is said to be possible even for statistical agencies to reach these levels. Some survey organisations have reached an ISO 9000 certification level and one survey organisation has reached level 5 (award winners) by winning the Malcolm Baldrige Quality Award in the USA in 1996. This company is Custom Research Inc. which is a very small market research firm conducting market and customer satisfaction surveys employing 100 people.

C. QUALITY MANAGEMENT IN PUBLIC ADMINISTRATIONS

At the moment, enhancing productivity and quality of service is the all-embracing theme of development throughout the public sector in many countries. Financial problems in the public economy have made continual reforms and cost reduction essential. At the same time, citizens expect a high quality of service and a proper return for their taxes and service payments. In seeking ways to solve these problems, Total Quality Management and various methods of quality control have been adopted. At the same time, the discussion of the applicability of TQM to the public sector has grown. All this also reflects to quality work in statistical agencies as a part of the public sector.

At the meeting in May 1997, the directors-general for the public service in the EU described the dual problem of the public administrations concerning the quality management. On the one hand, the public administrations have a key role in modern society to provide the necessary conditions for supporting competitiveness to face the extremely high and global competition environment. Consequently they (PA) should not be seen as only regulators (in some cases as a necessary burden) but as suppliers of the basic and collective infrastructures and services at the service of the economic operators and the citizens and as facilitators to solve collective problems of the communities.

On the other hand, in order to face these changes, Public Administrations should be able to transform themselves through the introduction of a new mindset and the use of new management strategies. In fact, it is not only question of spending more or less money, of creating new public services or privatising the existing ones. On the contrary, it is a question of making the existing public sector organisations work more efficiently, of giving an entrepreneurial spirit to the Public Administrations.

In 1996, the Ministry of Finance and the Association of Finnish Local and Regional Authorities established a project to define the common quality strategy for the Finnish public sector.

The basic principles of TQM were taken as a starting point, but at the same time, the applicability of these principles to the public sector was analyzed. Many public sector characteristics seem to require re-thinking of TQM-concepts, particularly the concept of customer, and specifying of quality criteria.

The dual role of public sector, both as a representative of public services' users and as a producer and organizer of these services, makes it more difficult to consider the quality of services from a point of view of the customer. The production of public services is mainly financed by tax revenues. Some public services are obligatory for citizens, e.g. public defence and taxation services. Production of the majority of public services are obliged by law, e. g. a major part of official statistics.

As customers of the public sector can be individuals, separate groups of individulas, businesses, media and even the whole of society. The same customer can serve as a selector, user, payer, decision-maker and even as a part of the service producing chain like businesses in their role of data providers.

The users and the payers of public services are not necessarily the same persons or institutions. In the case of collective services, customer needs can even be conflicting. It is not always possible to take individual needs into account due to reasons of equality. The interests of users are not necessarily the same as those of payers.

In developing quality of public services, the different concept of customer and some basic principles of public administrations should be taken into account. In the Finnish discussion about the quality criteria of public services, the following issues have been taken into account:

◗ custom orientation and availability of services

◗ reliability and security

◗ justice, fairness and legal protection

◗ equal and impartial treatment of customers

◗ confidentiality

- accuracy and timeliness of public sector information

- customers' possibilities to influence and participate

- objectivity and kindness in service actions

- professional skills and training of personnel

- cost efficiency.

Many of these criteria need to have special attention while discussing quality of public services and also quality work of statistical agencies.

The outlined quality recommendations for the Finnish public sector emphasize that the work should be customer-driven. The customers and their needs have to be known. The service processes should be transparent and they should be continuously assessed and developed.

The development of assessing public services forms a fundamental part of quality work. The assesment has to be systematic and the related professional skills have to be developed in administrations. In our experience, the statistical offices can play an important role in this assessment process. It offers possibilities to develop new services of official statistics because a part of key figures and statistical numbers needed for assessment are produced by statistical agencies.

The quality recommendations for the public sector emphasize the role of personnel in quality improvement processes. There is a need for a common, clear and challenging objectives of quality work and a continuous development of professional skills and knowledge, expertise as well as willingness and ability towards customer-oriented service.

Among the tools recommended for quality management are e.g. self-assessment process by using quality award criteria. The annual meeting of the European directors-general for the public administrations last year saw the use of quality award criteria as a good mean and starting point for quality improvement process. Starting a self-assessment of the organisation by using the European Quality

Award model is probably one of the most frequent ways to launch a quality management initiative. This has a certain number of advantages, in particular, it allows the organisation to know its exact starting point and at the same time provides a good tool for management motivation.

D. TQM AND OFFICIAL STATISTICS

Emphasizing that process quality generates product quality, many national statistical agencies have in recent years shifted their focus to more general quality approaches and tools such as TQM. Published papers on experiences of specific quality applications have however been rare. The collection of conference papers published last year by Lars Lyberg and others, "Survey Measurement and Process Quality", is very welcome to fill this gap in its own part. Experiences of quality work in several national agencies are presented, including, among others, Statistics Canada, Australian Bureau of Statistics, Office of National Statistics (UK), Statistics Netherlands and Statistics Sweden. I would like to mention especially the article of Cathryn Dippo and the article of two Davids, namely Morganstein and Marker. They stress the importance of a broad, process-oriented quality framework to support the efforts of improving quality of products and services in a statistical agency.

It is not straightforward to implement successfully TQM principles for official statistics. In their article, "Quality management: Development of a Framework for a Statistical Agency", Michael Colledge and Mary March from Statistics Canada have summarized certain aspects specific to statistical agencies when compared to private companies with respect to the application of quality management principles. It might be advisable to refer to their observations.

For a statistical agency, the position in the marketplace, and the nature of products and processes, are obviously different when compared to a private company. Products of a statistical agency are generally priced to cover only the additional costs due to publication and dissemination. This can provide a position of a monopoly for such activities. Thus customers cannot exercise control on quality by

changing suppliers. Moreover, when compared to a private company, a statistical agency is often restricted in its choice of products. Certain products can be legislated by national laws or EU directives. This can reduce the scope of TQM by preventing fundamental changes in products. As a third issue, the diversity of customers and products is a characteristic feature of statistical agencies. This makes it difficult to identify some customer types, and therefore customer requirements are sometimes hard to determine.

On a more technical issue, Colledge and March argue that statistical estimates are likely to contain errors such as sampling and nonsampling errors that commonly occur in surveys. Thus, a TQM goal of zero defects (i.e. errors) cannot be literally applied. Secondly, customers are not in a position to detect errors by observing the data outputs alone. Even if the errors could be precisely measured, their impact in terms of customer dissatisfaction depends on the diverse and often unknown uses to which customers put the data.

Colledge and March conclude that these points do not invalidate the application of QM principles. They present a set of activities that make a useful basis for quality improvement in a statistical agency. I summarize briefly their suggestions.

Expand customer focus

Customer focus can be expanded by employing more and better co-ordinated contacts with customers, and by better identification of customer problems and the corresponding data needs. It is important, and perhaps most difficult, to define quality in customer terms. In this respect, using a standard definition, the quality of products of a statistical agency can be determined in terms of four components, namely relevance, accuracy, timeliness, and cost. Relevance, timeliness and minimum cost are features associated with customer satisfaction. Lack of accuracy relates to defects and customer dissatisfaction. Although these four factors do not lead immediately to a quantifiable definition of quality of a statistical product, they do give some perspective.

Promote the concept of internal customer-supplier partnerships

The concept of internal customer-supplier partnerships can be promoted by identifying internal customers and defining quality in customer terms, and by the evaluation of customer satisfaction.

Make suppliers part of the process

Suppliers can be made as part of processes by ensuring that quality standards are well defined to paid suppliers, and by reducing and improving measurement of respondent burden and error.

Expand continuous improvement activities

Continuous improvement can be expanded by initiating more quality-improvement teams, by using generic methods and systems for re-*engineering, and by using more data integration.*

Promote total employee involvement

Total employee involvement can be promoted by involvement of staff in continuous improvement through examination of processes and empowerment to make changes.

Encourage development of quality/performance measurement

Development of quality/performance measurement can be encouraged by identifying the data elements needed for performance assessment and decision making and working towards their routine production.

Reinforce quality-management organizational structure

Quality-management organizational structure can be reinforced by establishing a quality council, by developing and promulgating a QM strategy blended with the agency's strategic plan, and by establishing a quality support unit.

Promote quality management by training and recognition

Quality management can be promoted by training

staff in quality awareness, teamwork, problem solving and quality measurement. Team and personal quality achievements can be incorporated into reward and recognition systems. Workshops and seminars can be conducted focused on QM experiences and success stories.

Enhance communication

Communication can be enhanced by cross-functional and multilevel membership of quality improvement teams, with discussions with both internal and external customers and suppliers, and by setting up advisory groups to provide external feedback on the relevance, accuracy and timeliness of products.

E. TQM AND STATISTICS FINLAND

E.1. Background

At Statistics Finland, previous emphasis in quality work has been in the quality of products, development of production processes, and in the quality of surveys. A more systematic orientation to total quality management approaches and tools has begun quite recently. A short-term development programme was initiated in 1996. An overseas support organization was appointed to help our efforts in the implementation of certain parts of the TQM programme. The TQM concept used by the consultants is described in Morganstein and Marker (1997).

Statistics Finland started its TQM efforts in order to shift focus from product-oriented organisation to a more process-oriented organization. A key point is that the agency itself is viewed as a system of processes that produce a final statistical product. Our view is that the diverse and weakly coordinated quality activities in management, process planning, and production processes can be analyzed and developed in a more systematic way under the TQM framework. In connection with this, serious attempts are made to avoid an isolated "quality world" within the agency.

E.2. Ongoing work

Present quality activities at Statistics Finland consists among others of such items as basic training in TQM, self-assessment of the agency with respect to quality award criteria, development of concrete measures for monitoring quality, coordination of quality work by a quality guidance team, documentation and standardization of systems and processes, standardization of statistical methods, identification of core processes, and promotion of scientific research to support quality improvement.

Our TQM consulting partner has provided quality training for the management, quality pilots and the leaders of quality projects. During past two years, over 120 people have participated in the basic quality training of which about 20 are quality pilots. Training has focused for example on statistical process control, current best methods approach, production and use of checking-lists and other quality tools, and team work.

As a learning phase, training has been practised in about ten quality projects so far. On grounds of the experience on these projects, a similar quality approach will be applied to a wider amount of new projects and, step by step, to the whole organisation as a natural way to operate.

In 1996 Statistics Finland made a first attempt to assess the overall performance of the agency according to the Finnish Quality Award criteria. This analysis was a basis for the first quality programme. During this year the Quality Award criteria will be used more systematically by the board of directors in the identification, selection, development and monitoring the most important areas of quality improvement. This will constitute a basis for a more comprehensive QM strategy.

Also in 1996 we started work to generate concrete measures for a systematic follow-up process. It was noted that in addition to the traditional statistical quality measures, more comprehensive data and new indicators were needed for assessing the success of the TQM work. Among the data needed are e.g. results from customer and employee satisfaction surveys and more information on the competence of the personnel and the functioning of processes. So far, the following surveys have been implemented :

- Self-assessment studies to evaluate the quality of statistical systems were initially made in 1991, 1993 and 1995. A new study is currently in progress.

- A customer satisfaction survey was carried out for the first time in 1992. Since then, the survey has been made annually by a private company on a contract basis. The survey covers some 20-30 different measures and it is under continuous development.

- Employee satisfaction surveys were undertaken in 1993, 1994 and 1996. The next survey will be made this year.

In order to support the board of directors in quality work, a quality guidance team was launched last year to cover the main TQM activities. The team consists of employees from different departments and different levels of the organisation. The main tasks of the guidance team are the following:

- promotion of quality thinking

- updating goals in the quality policy

- standardization of quality efforts and co-ordination of separate quality activities

- increasing communication in quality work

- measurement of quality improvements.

Better documentation and standardization of statistical systems and processes are some of the key issues in quality management. So far, there are some examples of the progress made in this area. We have continuously and systematically developed tools and facilities in the area of metadata and data architecture. A manual including the basic rules and checking-lists for press releases and other publicity processes has been completed. Internal guidelines regarding e.g. cost calculation and contract and pricing practices for marketing activities are also completed. However, there is still a lot of work to be done in the area of standardization and documentation of systems and processes. A further step will be the production of a Quality Manual in order to collect together the separate pieces of quality guidelines.

Successful quality improvement programs should incorporate statistical methods as well. In this area we have concentrated on the standardization and development of the survey methodology. This covers for example the use of up-to-date estimation methods and the accompanying software products, and an extensive use of the available register data to improve the estimation. Computer-aided techniques have been implemented in most surveys using CATI and CAPI tools. Focus is also in pre-survey operations. A new cognitive laboratory for questionnaire design and testing will be established this year.

Identification, definition and measurement of the core processes is currently under progress. The aim is to complement management by results with process management. According to the experiences obtained from quality work, adopting process thinking is the most difficult and time-consuming "change operation". This "paradigm change" however can most effectively improve the performance of an organisation. It helps the organisation to change its focus from product orientation to service orientation and from partial (unit/department) service to total service. Improving cross-departmental co-operation is also one goal as well as lowering borders between departments.

The core processes of Statistics Finland are identified as follows:

- maintaining and developing the data capital (production/survey process)

- customer commitment process

- developing the products and services

- management process.

As a first step in this project, the customer core process was analysed and certain proposals for improvement were made. The evaluation and implementation of the proposals is now in progress. The next analysis to be started this year deals with the production/survey process or the process of

"maintaining and developing the data capital" as we also call it.

The quality declarations form a basis of information about the quality of statistics to the users. These declarations are usually based on the quality criteria of official statistics. So far, we have added product presentations to our annual publications. In the future, our aim is to define the quality criteria of official statistics more properly and add more comprehensive and standardized quality declarations not only to publications but to all means of dessemination. This will likely be done as separate metadata information attached to the dissemination databases. A tentative list of quality criteria for the Finnish official statistics is presented in Appendix 3.

We believe that approaches, concepts and methods commonly used in scientific research are important for quality improvement in a statistical agency. To promote scientific research last year we launched a second scientific programme, "Main Lines of Research and Development in 1997-1999". Many of the specific research projects in e.g. survey methodology, index techniques, population projections, measurement and reduction of respondent burden, etc. serve directly the goal of quality improvement.

One of the characteristics of the quality organisation is company responsibility and citizenship. How does an organisation promote quality in its environment? Statistics Finland has been an active member in planning the quality policy for Public Administrations including municipalities in Finland. The project is very comprehensive and final results are expected in the near future. Statistics Finland is also a member of a guidance team to promote quality efforts and share information in Public Administrations.

F.   CONCLUDING REMARKS

The key issue is to see TQM as a quality of management on different levels of organisations. The transformation of the organisation from a product-oriented organisation into one which focuses more on customers and data markets demands consider-

able efforts. Such an organisation will have to know more about itself, its customers, data markets and future trends and start to work in a new way and to create a different management system and working climate. It is really a new mind-set, which directors and managers and people in the organisation have to adapt to. TQM thinking can help organisations to improve their performance giving proper tools to meet this change. TQM for example gives an opportunity for organisations to work in a more proactive way. This will promote an organisations ability to adapt faster to new and rapidly changing circumstances, where a new role for public organisations, globalization, new ways to collect, produce and share information and increasing demands from customers are challenges for every organisation. These challenges must be used as opportunities.

The most important lesson to be learned from TQM is perhaps its basic idea of shifting the emphasis of quality thinking from a single product (that is a single statistics) to the whole statistical process and wider to the whole service and management process of statistical agencies. The development of activities should be seen as a process where the customer needs on the one hand and the competence of the personnel on the other hand play the key role. It is also important in the TQM-process to find a proper way to identify and prioritize the development objectives, to plan and schedule the right measures and to assess the results systematically. Actually, TQM should be seen as a new management model while, at the same time, forming a systematic framework and offering a number of good tools for quality improvement and for reforming work methods.

As to the statistical offices, TQM can be seen as a complementary and more comprehensive way of improving quality rather than a substitute of a limited and traditional thinking of statistical quality. Both are needed in order to answer the increasing demands from customers in circumstances of restricted financial means.

BIBLIOGRAPHY

Arie de Geus (1997), The Living Company

ASQ (1997), American Society for Quality, certified quality manager

Dale Barrie ed. (1994), Managing Quality

Dippo Cathryn (1997), in Lyberg Lars ed. (1997)

Colledge Michael & March Mary (1993), ASA Journal of Business & Economic Statistics Vol 11, No. 2

EFQM (1996), European Foundation for Quality Management,The European Way to Excellence

Gibson Rowan (1997), Rethinking the Future

Hannus Jouko (1994), Prosessijohtaminen

HM&V (1997), Tilastokeskuksen asiakasprosessin kehittäminen

Lyberg Lars ed. (1997), Survey Measurement and Process Quality

Mintzberg Henry (1997), Yritystalous-magazine 5/1997

Morganstein David & Marker David (1997), in Lyberg Lars ed. (1997)

Norton&Kaplan (1996), The balanced scorecard

Meeting of Directors -General for the Public Service in the EU (1997)

The Quest for Excellence (1997), The Official Conference of the Malcolm Baldrige National Quality Award

Silen Timo (1997), Kansallista laatustrategiaa koskeva selvitys

Stitt John (1990), Managing for Excellence 1990

Statistics Finland (1997), Main lines of research and development in 1997-1999, Reviews 1997/6.

APPENDIX 1.

Total Quality Management: an overview

APPENDIX 2.

BALDRIGE AWARD CRITERIA FRAMEWORK
A Systems Perspective
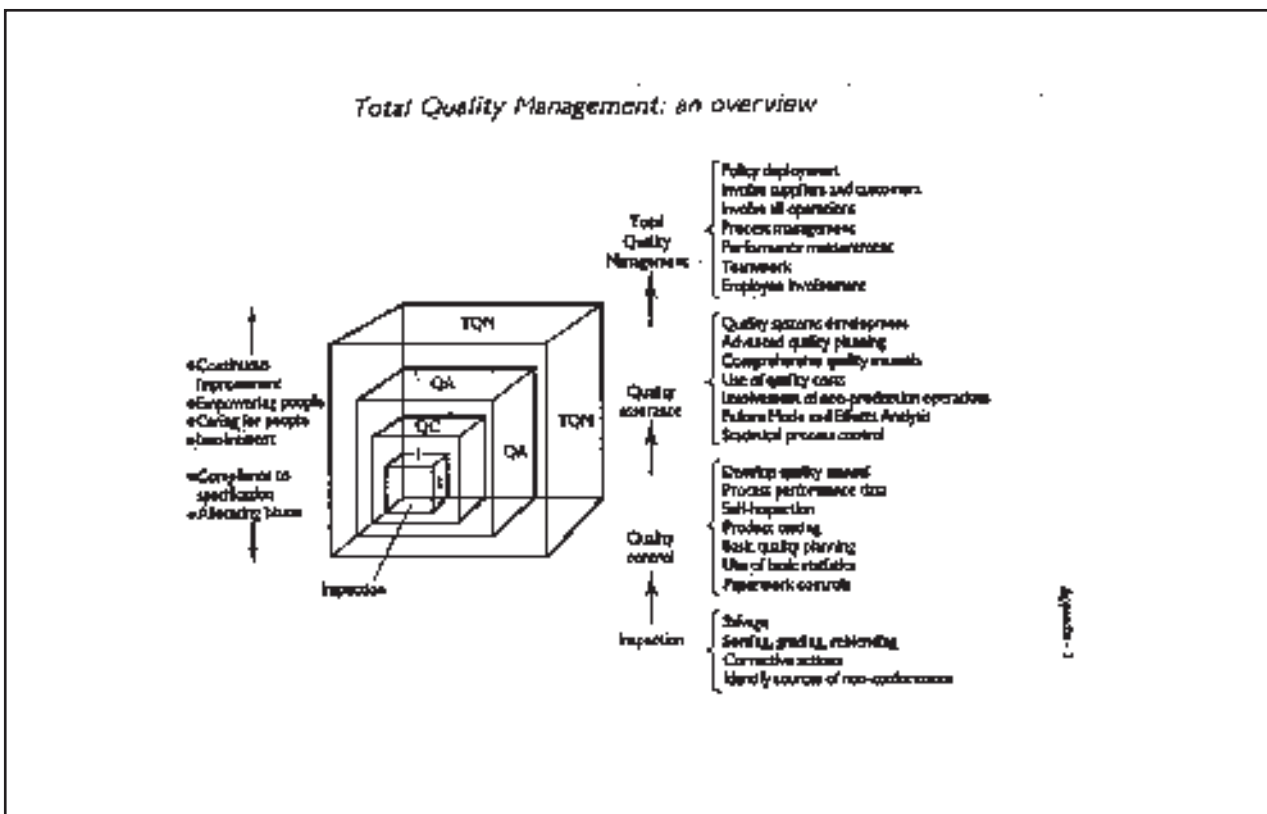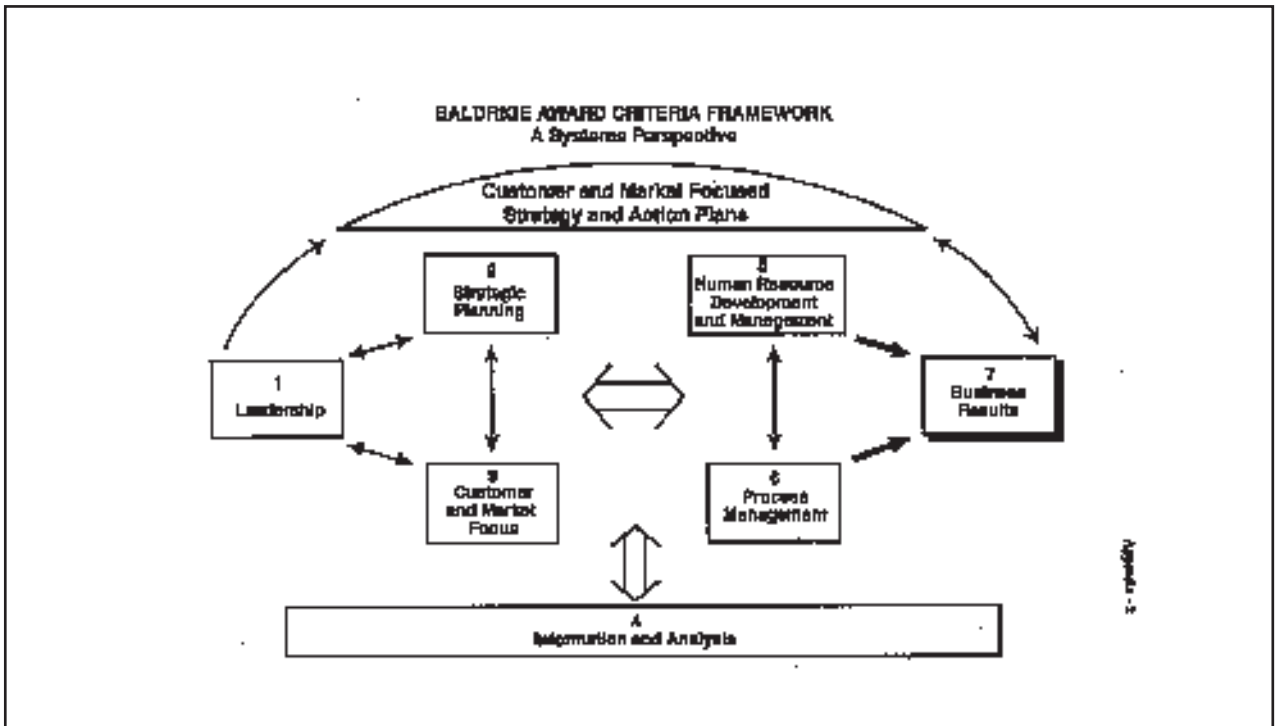
APPENDIX 3.

1.   The Quality Criteria of Official Statistics

◗ Relevance
- delineation, population, content, degree level of detail, extent and level of analysis, recurrence

◗ Reliability
- accuracy, errors in measurement and processing, non-response, coverage errors

◗ Timeliness
- topicality, up-to-dateness, punctuality

◗ Accessibility
- circulation, familiarity, technical availability, user-friendliness

◗ Comparability across the time

◗ Coherence
- definitions, classifications
- methodology

◗ Documentation
- product/process descriptions
- quality data

◗ Information potential

◗ Cost efficiency

2.1.1 Discussion Donald Murphy - Statistical Office, Ireland

Success depends on good management and a focus on quality at all levels

A. INTRODUCTION

On everybody's behalf I wish to thank Timo Relander for a most stimulating and informative paper. In it he provides us with::

◗ a good description of the *Total Quality Management (TQM)* process;

◗ an outline of some of the difficulties in transferring this private sector process to the public sector; and

◗ details on how he is implementing the approach in Statistics Finland (SF) based on a benchmarking against the Finnish quality award criteria.

The thoroughness of the approach being followed and the commitment of management to the initiative augurs well for its success. Particularly helpful in my view is the fact that the initiative is being undertaken in a Finnish public-sector wide TQM context with presumably strong political support. All of us will be keeping a careful eye on developments so that we can benefit and learn from the experience.

This leads me to my **first Question** - what special particular difficulties did Statistics Finland encounter in introducing the TQM process, what mistakes were made, what missteps were taken? The paper is silent on these points. The answers would help the rest of us to avoid these problems in any similar initiatives.

B. QUALITY FOCUS IN NSIS

Quality, as Tim Holt said in his ONS paper, is at the heart of all our work. It characterises the business we are in. I have spent my life in statistical quality work - ranging from methodological design, sample selection, training and supervising office and field personnel and so on through the full gamut of the various other phases of statistical activity. Most of the NSI mission statements that I have seen highlight the

quality aspect. In fact, the first High Level Goal in the Irish CSO's 1998-2000 Corporate Plan is the *"continued improvement of the quality of our statistics".*

The implementation of the quality philosophy is the challenge for management in all enterprises, both public and private. TQM is one, and many would argue, the best, approach. Timo mentions different definitions of TQM. A definition I came across at one stage, and personally like because it best conveys to me the essence of the approach, runs as follows:

> *TQM is a well-planned, organisation-wide process, integrated into the enterprise's business plan, that achieves the goal of continuous improvement of all business process in order to satisfy customer requirements, internal and external.*

All definitions convey the central point that TQM is "total" in the sense that it aims to involve everyone in the organisation in an enterprise-wide integrated effort toward improving performance at every level. It is not intended to be a "quick-fix" solution, but the culture an organisation should have and the way it should be doing its business. As one author I have read aptly put it - *"it is a process and a journey, not a destination".*

C. IRISH PUBLIC SECTOR APPROACH

We have not formally adopted the TQM approach in the Irish CSO. However, we are travelling on a similar road in the context of a *Strategic Management Initiative (SMI)* initiated by the Government in 1993 to modernise the Irish civil service and to provide a quality service to all its customers.

In this SMI approach, planning is linked to implementation by involving staff at all levels in a participative framework. The emphasis in this process, which is being very actively driven by Government and senior Civil Service management, is on customer service, results focus, human resource development,

staff empowerment and performance management. As part of the programme the CSO, in common with other Government Departments, was required to publish a two-year *Quality Service Plan* at the end of last year.

In parallel with the SMI, the functions of our Comptroller and Auditor General have been legislatively extended to include *Value for Money* dimension in addition to his traditional financial auditing role. This has increased the focus on the economy, efficiency and effectiveness (EEE) of all public sector operations.

The new management environment being developed in the Irish Civil Service has many of the characteristics of the TQM approach. We in the CSO may well take on board some of its technical constructs as we progress the implementation of our customer service plan and the goal of improving the quality of our statistical outputs.

D. THREE FURTHER QUESTIONS

My interest in TQM was kindled during a two-year management programme at the end of the 1980's. At that stage I saw it as a very effective and attractive approach but felt that it was mainly applicable to the manufacture and marketing of physical products in the competitive market environment. I saw then, and still do, difficulty in transferring the process to the public sector context of delivering services in a monopolistic non-market context. I, therefore, welcomed the call to speak to Timo's paper as it prompted me to catch up on the recent TQM literature.

I intend to focus on three particular concerns highlighted in the literature and to pose the issues they raise for the application of TQM in the public sector as ***three further questions*** which Timo might respond to in the light of Statistics Finland's experience. These concerns, and the questions they raise, are:

◗ TQM failures in the private sector - are there lessons to be learned by NSIs?

◗ TQM transfer to the public sector - are there difficulties NSIs should be particularly wary of?

◗ TQM's human dimension - is there a possible downside?

E. PRIVATE SECTOR TQM FAILURES - LESSONS TO BE LEARNED BY NSIS?

Despite the TQM fanfare I found in the more recent literature a surprising degree of ambivalence about how successful the process has really been in the private sector. There is, of course, a danger that isolated failures may have been over-emphasised. Equally well, I have a feeling that the positive messages of the gurus and management consultants may have conveyed too favourable picture of TQM.

In surfing the WWW (the first place to start nowadays) one of the early references I came across related to the article with the provocative title *"Is TQM Dead"*[1]. First emphasising that Mark Twain, the American writer, once remarked that reports of his death were greatly exaggerated, the author highlighted as follows, what he termed, the barrage of US media attacks on TQM as the latest of a long string of failed management fads and some hard evidence:

Media attacks

◗ "Totaled Quality Management"

(*Washington Post*)

◗ "Quality Programs Show Shoddy Results"

(*Wall Street Journal*)

◗ "Quality? Schmality"

(*USA Today*)

Hard Evidence

◗ McDonnell Douglas failed TQM process;

◗ A Baldridge-winner bankruptcy;

◗ the fall-off in Baldridge applications from 106 in 1991 to 76 in 1993 (71 in 1994);

◗ a 1992 *Rath & String Survey* reporting that only 20% of Fortune 2000 companies were satisfied with their TQM process;

◗ an American Electronic Association survey show-

---

[1] *Paton, Scott Madison. "Is TQM Dead?", Quality Digest, April 1994.*

ing that 73% of a sample of electronic companies had quality programmes in place, but 63% of these said they had failed to improve quality by even as much as 10%.

The author concluded that *"No, TQM isn't dead - TQM failures just prove that bad management is still alive and kicking"*. He also advised that the following six rules should be followed to ensure the success of TQM:

1) Measure costs

◗ *quality is not cheap despite what the guru Philip Crosby said*

2) Integrate TQM into your mission

◗ *TQM must be driven by business plans*

3) Use the right tools

◗ *TQM tools do little good if inappropriate or used incorrectly*

4) Balance people, processes and technology

◗ *too much focus on one can adversely affect to other;*

5) Don't loose people to teams

◗ *teams are critical to TQM but individual excellence is essential*

6) Allow time for the organisational culture to change

◗ *TQM takes time to become ingrained in an organisation's culture*

In the light of his experience to date Timo's views on these rules would be interesting.

For those NSIs that are not formally travelling the TQM road a 1996 article[2] reporting on a detailed empirical comparison of quality management in TQM versus non-TQM firms gives interesting results. While concluding that:

◗ in general TQM firms tend to do better than non-TQM firms;

◗ the more rigorous execution of TQM constructs the better the quality performance of a TQM firm;

◗ the quality performance of a non-TQM firm is determined by the extent to which it incorporates TQM constructs in its quality strategy;

it also showed that:

◗ A higher quality performance can be achieved by a firm through a "good management" philosophy without formally christening itself a TQM firm and it is possible for such a firm to exceed the performance of the TQM firm.

This means that TQM is not the only road to travel to achieve quality in an organisation.

F. TQM TRANSFER TO THE PUBLIC SECTOR - DIFFICULTIES NSIS SHOULD BE PARTICULARLY WARY OF?

Timo usefully summarises some of the difficulties in transferring the TQM process to the public sector. This is, in fact, a relatively neglected area in the literature - a recent article highlighted that of the 6,500 TQM articles catalogued in *ABI-Inform* (an US database covering business, management and related articles) only some 250 were concerned with public sector applications!

The problems involved are, however, analysed in detail in the 1994 publication[3] that looks at the problems involved from an international perspective. The authors consider that the difficulties fall under 5 headings:

1) The nature of TQM itself inhibits its application to the public sector

The authors reject the argument that the original manufacturing nature of the TQM body of ideas and practices, designed for physical products in a competitive marketplace, don't readily transfer to the delivery of public services on the grounds:

---

[2] *Ahire, S.L, Waller, M.A., and Golhar, D.Y. "Quality management in TQM versus non-TQM firms: an empirical investigation " International Journal of Quality & Reliability Management, Vol. 13, No.8, 1996, pages 8-27*

[3] *Morgan, Colin and Murgatroyd, Stephen, Total Quality Management in the Public Sector: An International Perspective, Open University Press, 1994.*

- the reasoning confuses "produce" and "process";

- it is the elimination of quality-impairing variations in the *processes* of activity that is all important, irrespective of whether these are *product processes* or *service processes*;

A basic problem, of course, is the fact the cutting edge of TQM is missing in the public sector context because of the absence of the competitive market.

2) The nature of the public sector itself is inimical to the reception of TQM applications

The issue here is that it is not the nature of TQM that limits its application outside the commercial sector, but rather certain essential features of the public sector itself, namely:

- it is more resistant to change;

- its resourcing is disconnected from performance;

- its managers are not rewarded for performance or quality attainment of quality;

- its managers are not able to manage as freely as their commercial counterparts;

- increasing service quality without increasing costs is intrinsically more difficult.

The extent of these limitations vary between countries due to differences in public sector institutional structures and human resource cultures.

3) The work culture of the professionals in the public sector are inimical to TQM

The difficulties highlighted here are:

- The multiplicity of professional specialisms in the public sector;

- The primacy of the individual professional transaction;

- The authority and autocracy of seniority and status hierarchies.

4) The customer is a more problematic concept

Timo has highlighted the problems with the customer concept in the public sector. These stem mainly from the fact that:

- the ultimate test of quality control exercised by customers by their purchases in a competitive market situation is missing; and

- public agencies must serve a wide variety of customers who generally have widely divergent and even contradictory demands.

There is also the issue of the appropriateness of using the term "customer" in the public sector context. At a recent seminar I heard the management guru *Henry Mintzberg* state that it did not want to be called a "customer" in his dealings with Government. His reasons were the term's business connotation of the arms length relationship and the hawking of products and services in the sole interest in profits. He considered that "client" was the appropriate term! Interestingly, many NSIs use the term "statistical user".

The conflicts of differing user requirements in terms of statistical quality criteria is highlighted in Tim Holt's paper. The TQM principle of satisfying all customers begs questions and can result in uneasy compromises. However, this does not deflect from the validity of the TQM concept in the case of internal customers.

5) The public sector environment is more complicated than the private sector

This complexity arises from political, public and legal accountability; extensive regulations and controls; openness of the decision making process; multiple values and goals; multiple stakeholders; annual budget battles, etc.

The authors' conclusion is that TQM is a management framework that can span all organisation settings both commercial and public, product or service or any particular mix. They feel it has a very significant role to play in the public sector, but only if it is modified to fit the unique characteristics of the particular public body. They saw

the reality of this fit as one where it is necessary to identify, from the range of TQM constructs, those having for the particular public sector organisation in question:

◗ **core relevance** *(i.e. generic constructs that can be applied to significant advantage);*

◗ **adaptive relevance** *(i.e. other constructs essentially belonging to TQM's manufacturing origins that can be adapted to fit the particular situation);*

◗ **problematic relevance** *(i.e. other constructs that are either not applicable or almost impossible to adapt).*

I wonder if this framework reflects Statistics Finland's experience?

G. TQM'S HUMAN DIMENSION - IS THERE A POSSIBLE DOWNSIDE?

All variants of TQM hold that people are critical to its success. However, a recent article[4] argues that an exclusive emphasis on quality service can have negative consequences for the people who work in public sector organisations. The author considers that the following four aspects merit careful consideration:

Middle Managers' Egos and Fears

◗ The TQM principle of self-management teams and empowerment for front-line employees reduces the decision-make power and role of middle managers and their opposition can be expected;

◗ Middle managers are the repositories of organisational experience and knowledge; delayering can result in a kind of organisational amnesia that can have negative repercussions.

Employee Motivation

◗ Despite Deming's call to the contrary TQM is remarkably silent on the people aspect of quality; its principle focus is on training in a limited technical sense;

◗ Staff can resent the imposition of a TQM bureaucracy and view it as just another "fad" being imposed by management;

◗ The time spent at a lot of meetings could be considered by some of preventing doing their "real" work (i.e. participation fatigue can set in);

◗ The TQM approach might appear cold/analytical and thereby take the fun and human emotion dimensions out of work which often are the catalysts for intrinsic motivation and commitment.

Trivial Employee Participation

◗ Empowerment can be considered by some employees to be a form of vertical job dumping of tedious administrative work; participation means involvement in meaningful decision making (TQM may be more talk than action in this regard);

◗ Not every employee values participation (some want just to do their job and leave it behind when they go home);

◗ Resistance to participation on the basis that it means more work, higher responsibility and accountability (because of empowerment) with no pay adjustment.

Coercive Teams

◗ TQM is a "totalising" concept; its central focus (i.e. the word "total") can have a sinister meaning and reality in some circumstances;.

◗ Self-managed teams can be coercive in the achievement of self-imposed targets (e.g. unpaid overtime);

◗ Work team commitment can transcend one's obligations to family, etc.

These concerns emphasise the need to take care to avoid any human downside in adopting the TQM philosophy. Two simple questions highlight the issues involved:

◗ These concerns emphasise the need to take care to avoid any human downside in adopting the TQM

---

[4] Connor, Patrick E., "Total Quality Management: A Selective Commentary on its Human Dimension" Public Administration Review, Volume 57, No. 6, November/December 1997.

philosophy. Two simple questions highlight the issues involved:

◗ TQM enthusiasts care about customers, process improvement, performance, etc. - *do they care about staff?*

Trade union attitude is another important human dimension. TQM initiatives can encounter resistance if public sector unions adopt a negative attitude because of concerns about such matters as:

◗ a decline in their influence and power according as the extent of empowerment and self-managed work teams increase;

◗ a view that TQM is only a "warmed-up" exploitation of labour;

◗ increased efficiency resulting in a reduction in the number of jobs and the promotion prospects for their members;

◗ workers not being rewarded for increases in productivity and responsibility.

These concerns are not helped in current circumstances where the focus of most Governments is on controlling or downsizing public sector employment levels in the context of public expenditure constraints.

Careful attention must be paid to human aspects in introducing a TQM approach. A partnership approach in which management, unions and staff work together would appear to be the best way of avoiding problems. Was this a feature of the Statistics Finland initiative?

H.   CONCLUDING REMARKS

Clearly, whether under the umbrella of a formal TQM campaign or otherwise, success in any organisation critically depends on good management and a focus on quality performance at all levels of its operations. I wish Statistics Finland well in its TQM initiative and hope it will be very successful. Based on its experience to date Timo's comments on some of the issues that I have raised would be interesting.

2.2 Willem de Vries and Richard van Brakel- Statistical Office, Netherlands[1]

Quality systems and statistical auditing :
A pragmatic approach to statistical quality management

Summary

This paper is about quality management in Statistics Netherlands (SN), and in particular about statistical auditing. In 1996, an overall quality program was adopted: the business plan 'SN 2000'. One of the objectives of this business plan is to introduce quality systems in all statistical departments. A standardized model for such a system will be finalized by the end of 1998. Provisional guidelines for quality systems were issued in 1997. Simultaneously, a system of 'statistical auditing' was set up. The aim of this system is to check how quality management in statistical departments is functioning and how the quality of statistical products and procedures may be improved.

Introduction

As providers of essential information for public debate, and for decision making at various levels of society, national statistical offices (NSI's) have to take the quality of their products and services very seriously. This has always been so, and quality issues have been around as long as NSI's have existed. However, in particular since the performance of NSI's and indeed of government services in general have come under closer scrutiny in many countries, quality management has lately become a focal point for many NSI's. Obviously, there are many sides to the 'quality' of official statistics. To mention some of the most important aspects: official statistics must be

◗ relevant

◗ timely and

◗ accurate,

 but they should also be

◗ produced in a cost-effective manner, and

◗ without too much of a burden for data providers.

Each of these major quality aspects of official statistics requires its own quality management approach.

Quality systems

In national statistical institutes (NSI's) there appear to be various approaches to quality management. Some NSI's have opted for a system of Total Quality Management (TQM), others aim at certification along the lines of the ISO-9000 system.

In 1996, Statistics Netherlands (SN) adopted a kind of comprehensive quality program of its own, laid down in the form of a 'business plan' for the next decade[2]. Apart from some general background information about this 'business plan' this paper in particular highlights two specific component of this overall quality program. On the one hand it is about the *quality guidelines* we are introducing. On the other it is about a system of what we have named *'statistical auditing'*. The focus of statistical auditing in this sense is on the quality of the *statistical production process*. This implies that it relates *primarily* to the quality elements 'timely', 'accurate', 'produced in a cost-effective manner' and 'without too much of a burden for data providers'. 'Relevance', though an important part of the quality guidelines, is usually not covered *in depth* by the statistical audits. To measure user satisfaction with the SN work program in general and with individual sets of statistics in particular, other mechanisms are in place. These too are set out in the 'business plan'.

We certainly do not claim that statistical auditing has been 'invented' by Statistics Netherlands. In fact, the approach we have taken was partly inspired by similar activities that have been going on in Statistics Canada for a number of years[3].

What is quality?

Essentially, quality is a *subjective* measure of the properties of a product or service. Quality assessments by users of products and services depend to a large extent on their specific needs and also: their expectations. This implies that what one user considers to be insufficient, may at the same time be excellent for another. Another useful definition of quality is therefore: 'fit for purpose' (i.e. the purpose of a specific user). To illustrate this with an example in the area of official statistics: macro-economic policy makers will generally be satisfied with fairly quick, reasonably accurate, highly aggregated statistics about international trade, while these same statistics, in that particular format, will be virtually worthless for a user who needs numbers for market research in the area of cosmetics or Scotch whisky. Therefore, it is rather difficult to assess the quality of statistics in simple 'objective' terms. In addition, there are, as mentioned before, quite a few aspects to the 'quality of statistics'. Statistics are indeed a fairly complex product. If McDonalds need many pages of text to properly define the quality of a Big Mac, which is, with all due respect, just a bun with a hamburger inside, it is no wonder that it takes a whole book to properly describe, say, the quality of the Consumer Price Index, let alone the quality of the National accounts.

ISO or not ISO

Quality systems encompass the organizational structures, responsibilities, procedures, processes and infrastructure for the implementation of quality care. There are several types of quality systems, different in philosophy, degree of 'regulatory ambitions' and applicability. Nearly all activities and instruments of an organization affect the quality of its products. Therefore, quality care is strongly linked to organization and management issues. This is not the place to go in great depth and detail into these different systems, but some remarks about the so-called ISO-

norms are appropriate. The general philosophy of ISO is that, in order to ensure a certain minimal quality level of final products, an organization must be able to demonstrate that all details of the production process are in some way formalized and are thus in principle kept 'under control'. Under ISO, quality audits are used to monitor whether the system is actually in place, and whether it is respected by staff and maintained by managers.

Some authors, at least in The Netherlands, are critical of the ISO system[4]. Indeed, ISO seems to be rather bureaucratic and costly. Other authors think that certain preconditions have to be met before it is useful to introduce any quality system whatsoever[5]. They think that some of the necessary conditions to be met are: a shared strategic vision of top management, effective communication, effective management at all levels, in particular also the lower management levels, clear targets and objectives, and finally: a organizational climate and culture in which success and good performance are systematically recognized and rewarded.

After careful consideration, SN has decide not to go for ISO, but to adopt a more modest, pragmatic approach. One particular reason for this decision was, that the statistical process is in such a phase of dramatic changes (redesign of many statistical processes, introduction of EDI, combining registrations with SN data collections, organizational changes), that investing in a system that would primarily describe the present situation and procedures, was deemed to be inefficient. We did expect from local managers, however, that they had some kind of quality system in place. To promote the introduction of such systems, we have issued provisional guidelines. These will be discussed later. Ultimately, we

[3] The authors thank Henk van Tuinen, Director for Statistical Policy and Dick Kroeze of the Audit Secretariat for their comments, as well as René Huigen and Josée Nollen, who laid the groundwork for more systematic quality thinking in Statistics Netherlands.

[4] Swinkels, G.J.P., J.G. Verrijdt en G.J. van der Pijl, 1995, Kwaliteitszorg, kwaliteitssysteem en certificering, Maandblad Accountancy en Bedrijfseconomie, nr. 4, blz. 253-261.

[5] Spoelstra, G.J.P. en J. De Vries, 1993, Met de lijn de vloer op: het realiseren van integrale kwaliteitszorg, Personeelsbeleid 29, nr. 1, blz. 26-31.

want to establish more final guidelines for quality systems. Meanwhile, one of the main purposes of statistical auditing as we see it, is precisely to find out what quality systems are in place to guarantee a certain level of quality of the statistical product. On the basis of the best practices applied in our office, the final guidelines will be developed.

SN Business Plan

As mentioned before, SN adopted a general quality program (business plan) in 1996. The SN Business Plan (SN 2000 for short) sets out six major objectives:

A relevant work program

This objective has to do with all mechanisms to ensure that the work program of Statistics Netherlands meets the needs of the users. Decisions about the work program are made by the Central Commission for Statistics. To assess user satisfaction, regular 'evaluation rounds' will be held among all major user groups: ministries, government research and planning institutions, organizations representing employers and employees, academia etc. In addition, and this has to do in particular with the aim of Statistics Netherlands to flexibly approach new user needs, proposals to exchange 10% (in budgetary terms) of 'old' statistics for 'new' statistics will in the future be made in each four-year work program (starting with the work program for 2001-2004) that is presented to the Central Commission for Statistics, enabling the Commission to make real choices and to set priorities.

A substantially reduced response burden.

This objective will be achieved by a mix of different instruments, such as increased use of registrations kept by other government institutions, more cooperation with other data-collecting agencies, better 'speaking the language of the respondents', introduction of EDI-techniques and the systematic promotion of advanced statistical methods. To monitor the response burden, a 'response burden meter' was introduced in 1996. It shows the development of the statistical response burden as an index (1994=100). Each year, parliament is informed about how and to

what extent the response burden has been reduced, with the target to achieve a reduction of 12,5% in 1998. In addition, despite the fact that the Dutch system of official statistics is highly centralized, a Center for Government Surveys has been created. This Center has two basic functions: a) to detect where statistical activities (in particular: surveys) are taking place or are being planned elsewhere within the central government and b) to try and help the people who are having such plans to get the data through SN, which may either adapt or expand existing data collections or apply special analyses on data already collected.

Effective statistical information.

To measure the effectiveness of the statistical information it produces, Statistics Netherlands uses some specific targets and indicators. For example: the sales of printed publications in the year 2000 must be 20% higher than in 1994. 60% of SN's press releases, on average, must be printed by the seven main national newspapers. More importantly, by the end of 1997, **all** important SN data, including all necessary meta-information, must be available through the user data base Statline. Finally, by the year 2000 a broad range of 'client satisfaction monitoring instruments' will be in place.

Comprehensive quality management system.

The aim is to have a comprehensive 'quality management manual' by the end of 1998. In the year 2000, more than half of all statistical projects will comply with this manual. In addition, and this is what this paper is specifically about, each statistical project will be submitted to an auditing exercise every five years, including a follow-up to see whether deficiencies have been corrected. Some other specific targets are: a) the response rates in household surveys will in the year 2000 be 8 percentage points higher than in 1996 and b) SN will publish at least 100 research papers a year externally.

Adequately trained and motivated staff.

Due to dramatic changes in the statistical process, we foresee substantial changes of our staff as well, over the next five years or so. Increased mobility will

be required, both internally (the target is 10% per year) and externally. Quite a number of jobs will become redundant. Some people may be re-trained for new positions, but others will have to look for employment elsewhere. A major program of '*empowerment*' and training (at present, 2% of SN's budget is earmarked for training) is in place to promote internal and external mobility. There are also some financial incentives available. On the other hand, we wish to recruit a substantial number of young, highly trained professionals. Another, more specific target is that we want to reduce sick leave to an average of 5%. Finally, the general job satisfaction of our staff will be monitored systematically.

An efficient, well managed, flexible organization.

The present system of management contracts will be further developed. This also requires further improvement of our accounting structure and management information systems.

The ultimate aim of SN 2000 is to create a vital organization with a manageable budget.

Provisional quality guidelines

As it was said before, there are many aspects to 'statistical quality'. In order to measure and monitor quality effectively, in whatever systematic way, it is therefore necessary to define statistical quality more concretely and precisely. We think the quality criteria listed below are the most relevant and important in the area of the production process of official statistics. Therefore, we would expect that quality systems at all levels of the organization cover most of these aspects. In the provisional quality guidelines, a number of points to be taken into account is listed for each of five major aspects. The listing here is merely meant as an illustration and is therefore not exhaustive.

The aspects covered in the list are:

◗ 1. purpose of the statistical collection

◗ 2. the survey design

◗ 3. data input

◗ 4. data throughput

◗ 5. data output

Explicitly *not yet* covered by the points of the list are: marketing, the appearance of publications and management systems (other than statistical management in a narrow sense).

1. The purpose of statistical collections

◗ Who are the most important internal and external users of the statistics?

◗ When have they last been consulted about their needs?

◗ What are their needs as to: detail of variables, levels of aggregation, periodicity, coverage, comparability with other statistics, timeliness, accuracy etc.

◗ What is the real value of the statistics in relation to what the users would want?

◗ Which user needs cannot be met?

◗ What may be done to improve this situation?

2. The survey design

◗ Is the survey design documented?

◗ Are the statistics based on data collection 'in the field' or on integration of existing data sources?

About data collections:

◗ What type of sampling (if any) is applied and why?

◗ Which sampling frames are used?

◗ To what extent is the sampling frame complete and up-to-date?

◗ Does the frame contain the right kind of statistical units?

◗ How do you cope with imperfections in these respects?

◗ In what form are the data collected ? (EDI, mail

in-mail out questionnaires, interviewing etc.)

About data sources:

◗ Which data sources are used?

◗ Are there any alternatives and why are they not used?

About the structure of questionnaires:

◗ Have questions been tested for clarity? Are they answerable?

◗ Are questions assessed on validity?

3 and 4. Data input and throughput

Input planning and procedures

◗ Is there a planning for the different phases of the statistical process?

◗ How is the input process organized and monitored?

◗ Have any efforts been made to optimize the input and throughput process?

◗ Are there documented procedures for non-response treatment, imputation, data editing, raising, estimation, cross-checking between data?

◗ For integration processes: is the relation between statistics and their sources documented?

◗ For data editing: are all questionnaires checked/cleaned individually and if not, what are the criteria for selection?

◗ How are sampling frame errors treated?

◗ About imputation: how are non-response gaps filled?

◗ About weighing and raising: are intermediate results calculated and how are they used?

◗ How are statistics matched with other numbers and time series?

5. Output

◗ Does the final product meet the users' needs?

◗ Are there any differences with other, related NS statistics and what has been done to minimize the differences?

◗ Are analyses about differences well documented and publicly available?

◗ Are efforts made to avoid misinterpretation of the statistics?

◗ How is the quality of the statistics presented to the users?

◗ Is a complete quality description available for the users?

◗ What is exactly known about non-sampling errors? Is this knowledge well documented?

The questions of these provisional guidelines are meant to increase quality awareness in statistical departments. In the course of an audit, a checklist for self-evaluation is given to the auditees. This is just one illustration of an important principle of auditing as we see it, namely that auditing is *not* in the first place a corrective 'policing' instrument, but rather a coaching tool to enhance the general feeling that quality is important and that auditing therefore should ultimately be perceived as a *preventive* mechanism. As such they are only a first step towards the more final quality guidelines mentioned before.

Introduction of auditing

Now we come to statistical auditing in the sense of this paper. As mentioned before, statistical auditing as it has been introduced in SN is a technique which has three purposes:

◗ to actually find out what is being done about quality management in statistical departments

◗ to generate suggestions on how to improve quality management

◗ to find out what the best quality practices are and to incorporate these into the guidelines for quali-

ty systems that will be issued by the end of 1998.

It should be stressed and this has been made clear over and over within SN, that auditing is not intended to be a form of 'policing' in order to find out where things are not going as they should go. On the contrary: statistical auditing should be (and is increasingly) perceived as a form of help and advice to achieve improvements. However, this does not mean that it is entirely innocent and harmless. If the auditors discover weaknesses and unprofessional approaches, they will certainly report these and discuss them with management. Also, in the final discussion about the audit reports, agreements are made about how to achieve specific improvements. Finally, there is a systematic follow-up to check whether the agreements are implemented.

To obtain experience with statistical auditing, two pilots were carried out in 1996. One was about the statistics on Performing Arts, the other about statistics of the Transport industry. The aim of the pilots was to better define the scope of future regular audits and to develop a set of rules on procedures and instruments. The pilot audits were done by two teams of three SN-staff each. A private consulting company, which had broad experience in auditing and quality management, was commissioned to train the auditors and to moderate the process. Auditors were selected by the Audit secretariat on the basis of their statistical and managerial qualities. The techniques applied during the audits were interviews on the one hand and analysis of documentation on the other. The findings of the audits and the recommendations made on the basis of these findings were laid down in reports.

As to the selection of auditors, the idea was that all audits would have to be done by own SN-staff. The aim was to create a 'pool' of about 25 auditors from various divisions, selected on the basis of their expertise, but to some extent also their personality. We want the auditors to come from various divisions to ensure that a variety of experiences and expertise is represented in the audit teams. The auditors do this work on a part time basis only, because we want them to remain involved in regular statistical activities as well. The disadvantage of full-time auditors would be that such people may 'lose touch' with current prac-

tices and new developments. Ideally, an audit team consists of one person who is a specialist in statistical methodology, one who is well versed in statistical organization aspects and one who has a special affinity with producing outputs. In addition, some of the qualities looked after are:

◗ good communicative skills at various levels; diplomatic skills

◗ good analytic qualities

◗ openness for change

◗ knowledge of statistical processes

◗ good editorial qualities and the ability to present results orally

It appeared to be relatively difficult to find a sufficient number of (potential) auditors, not only because of the qualities that auditors must have, but also because people who have these qualities, are usually also much in demand for other priority issues. The envisaged auditors were subsequently trained for four days, in such areas as audit philosophy, norms, rules and procedures, interview techniques, reporting and presenting audit results.

Evaluation of pilots

A first evaluation of the pilot audits showed the following main points:

◗ most auditors had liked the work

◗ they were received well in the audited sectors and cooperation of staff and management had been good

◗ the training had been enjoyed

◗ in one of the two cases, the terms of reference for the audit had not been explicit enough, which had resulted in an incomplete audit

◗ drafting a systematic audit plan for the sector to be audited was important (including questions as: who supplies/reads documentation, who interviews who etc.)

- auditing takes time; therefore it is not possible to combine it with other tasks during the audit period

- to remain distant and objective can be difficult, in particular when auditees become emotional and when the auditor is in some way familiar with the audited sector

- on the part of the auditees, more than 70% were convinced of the usefulness of audits; 90% of them had felt that the atmosphere during audit interviews had been relaxed; 71% felt that the evaluation session had been good; 90% thought that the audit report was well written and clear, but some auditees thought that the conclusions could have been harder; 65% thought that the recommendations of the audit report had been useful. It has also been noted that, in general, most auditees very much like to talk about their work and enjoy sharing experiences problems with others.

Code of conduct

As one of the results of the pilot audits, the following code of conduct for audits was agreed.

- The main purpose of statistical audits within the SN is to help identify statistical sectors what the weak and strong points of their statistical processes are and how these may be improved. In a way, audits are liking presenting a 'mirror' to the auditees.

- There will be an audit plan, as part of the management contracts between division managers and the Director-General. Each statistical process in a statistical department will be audited once every five years.

- Audits are organized and moderated by an audit secretariat, which is part of the DG staff

- Audits are carried out by teams of three auditors, selected on the basis of specific expertise. A pool of about 25 auditors will be trained and regularly employed. Their performance will regularly be monitored by the audit secretariat.

- Before an audit starts, the procedures and planning will be agreed with the department manager.

- The department manager is responsible for: the supply of proper documentation, including a list of employees and their tasks, work instructions, checklists, handbooks, existing guidelines for quality control. He/she also appoints a contact person from his sector.

- In a workshop, the audit secretariat briefs the audit-team on how the audit will be carried out. Also, the scope of the audit (including any points which deserve special attention) is formulated.

- The audit secretariat organizes an introductory meeting, in which the scope and procedures are discussed. After that an interview scheme is drafted (implying, among other things, the final selection of the people to be interviewed). The maximum number of interviews per day is three, by two auditors, because interviews are to be relaxed. Interview reports are only for auditors. However, all reports are given to auditees for correction.

- The audit team drafts first report, which is first discussed with the audit secretariat.

- One audit secretary and the lead auditor discuss the first draft with the department head and contact person.

- The audit report is subsequently discussed in a meeting with department head and auditees.

- The final audit report is then written and sent to the department manager. A copy is sent to the Director-General of SN.

- The department manager has three months to react and to draft a plan for improvements on the basis of the recommendations.

- One year after the audit has taken place a questionnaire is sent to the department manager in order to check what has been done with the recommendations.

- After every five audits, the Audit Secretariat writes a summary report about important results, which may be beneficial for other departments as well. This report is discussed by the Management Committee for Auditing and Quality Care and is

also widely circulated.

Planning of future audits

In 1997 a start was made with a regular audit program. This program is rather ambitious: two audits will take place each month (except July and August). However, so far the planning is respected[6].

It is felt that some points still require further discussion and care:

◗ The way of reporting about audit results, in particular whether more openness (so far the reports have been treated rather confidentially) may be useful in order to enhance the 'learning effects' of audits for others than the audited sector.

◗ Whether or not specific priorities will have to be set to audit certain specific aspects of statistical processes that seem to be a 'weak spot' or an urgent problem. It is now tried to expressly formulate this in each individual audit instruction.

◗ The follow-up of audits; is a plan to remedy weak points drawn up and implemented and how should the implementation be monitored?

---

[6] *One point that is not entirely in conformity with the original planning is that audits (from start to finish) take about two months, while the expectation was: one month. The reason is not so much that auditors have insufficient time (each auditor may spend 80 work hours for an audit), but that it takes more time than expected to write summaries and to formulate recommendations which are applicable for other sectors as well.*

2.2.1 Discussion C. CORRÊA GAGO- Statistical Office, Portugal

The importance of different sources of error

The paper by Mr. de Vries' and Mr. van Brakel on quality systems and statistical auditing is very rich in several ways, namely in the comprehensiveness adopted when dealing with such an evasive concept as quality, in order to identify its many meanings and facets, and in the variety of relevant topics listed as questions for auditing purposes - indeed a check-list of great pragmatic value.

All the better for the discussant, who may spare the useless effort of trying to complete the paper, as a general overlook of quality issues, in statistics. The description made of the quality audit experience going on in SN, on the other hand, is self-sufficient as an informative presentation having the additional merit of leaving us with a vivid expectation to know more about the results, as the experience proceeds and develops.

Therefore, I'll confine myself to just a few qualifications, under the form of *zooms* of selective attention to some of the many topics contained in the paper. This is because although I think all the topics addressed are relevant, I nevertheless find some of those topics much more important and more promising from the standpoint of its practical consequences than others.

A first notion, which is no less than the corner stone of the paper, deserves specific comment. The notion of *quality management*. This notion bears implications which may be much richer than meets the eye at a first glance. It implies and emphasises that quality is, or should be, in statistical production, a *management issue, a function to be managed.*

I mean: it is not an absolute optimum to be restlessly pursued, (*total quality*, as it became fashionable to say), but rather a compromise to be reasonably administrated.

Note how significant it is that Mr. de Vries, right at the beginning, when reminding us of some of the more important aspects, or *sides* as he puts it, to the

quality of official statistics, makes it in two successive batches linked by " but they should also be": official statistics *must be* relevant, timely and accurate, *but they should also be* produced in a cost effective manner and without too much of a burden for data providers.

A first notion of compromise is already here, but the paper goes well beyond that when it proposes what is indeed a very useful definition for quality: *"fit for purpose"* - the purpose of a specific user. A proposition which is illustrated with very well chosen examples.

In other words, what I think the author is suggesting is that instead of a fundamentalist search for total quality we should rather look for a level of quality which is *sufficient* - sufficient for the specific purpose, supposedly relevant.

It is while looking for that level of sufficient quality, that is, while *managing quality*, that the classical balances should be negotiated: between *timeliness* and *accuracy*, between those two and *burden* and *cost effectiveness*. In order to come out with a mix of qualities which is fit for the purpose, hence, with a statistic product of sufficient, adequate, quality.

That is that for the conceptual discussion of quality in statistics.

The paper proceeds then with the description of the comprehensive quality program adopted in 1996 by SN, laid down in the form of a *business plan* - which is very significant in itself about the approach adopted.

I'll comment briefly on the decision made by SN not to go for ISO system but instead, quote, "to adopt a more modest, pragmatic approach". I conclude, from the reading of the "provisional quality guidelines" listed in the paper, that this meant a selective rather than exhaustive requirement for formalised, written, procedures to be followed in the production process. In fact, the guidelines ask for documented proce-

dures for several critical aspects, namely the survey design, the sampling frame, the data sources used and their relation with the corresponding statistics, the planning for the different phases of the statistical process, the procedures for non-response treatment, imputation, etc., a complete quality description available, and more.

If my interpretation is correct, and I would like to know whether it is or not, SN recognises the importance of existing written procedures as a pre-requisite for quality and as a tool for quality auditing, although not making a point on having a fully documented description, a handbook, of the entire production process. I would support this "pragmatic approach", even because it coincides with our own, in NSI Portugal, which I'll be returning to for a short while later on, if I have a couple of minutes left.

But let me say right now that, in my view, having written procedures does not mean, and should not mean, to fix once and forever methodologies and rules to be followed. Of course, managing quality implies a permanent ability to introduce changes, in order to improve results. Nonetheless, without a documented set of references which are regarded, at a certain stage of evolution, as imperative norms to be followed, management control is not effective and quality evaluation may not be objective.

Now: the last question listed in the "provisional quality guidelines" deserves, in my view, another zoom. Because I think that, speaking about quality in statistics, and in spite of being the last listed, it is the mother of all questions: "What is exactly known about non-sampling errors? Is this knowledge well documented?"

Personally, I know almost nothing about the theoretical treatment which has been or is being given to this matter, and the practical use which NSI's are making of it.

I have already addressed this point in public before, which allows me the self-indulgence of quoting myself: *"Major consideration should, for example, be given to the **actual relative importance of the different sources of error** affecting the final quality of a statistical operation. Could it not be case that the sampling*

error is being increasingly by-passed by other, significantly larger, margins of error, such as errors of observation resulting from the negative evolution of the propensity to reply carefully and truthfully?

And would it not, accordingly, be much more effective from the point of view of statistical quality to invest more in a close relationship with the units of a small sample (or a small panel) rather than increase the size of samples to the limits of financial feasibility? Are we not deceiving ourselves, and deceiving users, when we announce the quality of a product based, mainly, on its sampling error?"

In fact it may be argued that, from a quality management standpoint, it is useless to reduce sampling errors, even if they are the easiest to reduce and to calculate, beyond what we may know, or reasonably suspect, to be the errors inherent to the other phases of the statistical process. This is, at least, a matter deserving investigation.

And, incidentally, it should be noted that the obvious advantages we see, cost wise and burden wise, in the utilisation of administrative sources as an alternative to surveys, those advantages should not lead us to an uncritical attitude towards the reliability of administrative sources - when they are in fact available for us, statisticians.

Are we quite sure about their quality? To what extent are they free from systematic errors and biases which may turn the cheap way into an expensive one for producers and users of official statistics?

I dare say - ignorance is daring … - that this subject, in its theoretical and practical implications, would justify specific treatment during one of our meetings. In order to compile what is known and what is being done with that knowledge.

Yes: who am I, a manager, not a professional statistician, to elaborate on these points? But it is Mr. de Vries who asks: "What is exactly known about non-sampling errors?" And Mr. de Vries is a most competent statistician and is an honourable man!

His question must be answered.

Just a few minor comments about the well structured

"general quality program" of SN, its major objectives.

A substantially reduced response burden. From our experience in Portugal we suspect that the legitimate complains from respondents are due in a larger scale to enquiries and surveys issued by institutions other than NSI or the National Statistical System. Both from the public and the private sector. We, official statisticians, must of course be attentive not to ask too much from too many, nor more than once, or once a year. But it seems doubtful whether NSI's should bear the larger effort to reduce general burden, mostly when the level of sufficient quality of our products risks to be damaged. It is very clear the usefulness of the Center for Government Surveys, created in the Netherlands, but its limits are also evident: *Government* surveys. Much better than nothing, however.

And, by the way: are we not, NSI's, also entitled to be a bit worried with the burden put upon us by so many international organisations of different descriptions who send us enquiries to answer and forms to fill? In our case, we receive almost 400 demands each year, which seems too much. In the UE, thanks to Eurostat, the most systematic and comprehensive effort for an harmonized statistical coverage is being developed. Eurostat, and now the EMI, know everything about us, or almost so. Perhaps an additional effort of rationalisation would be in order to take advantage of those intermediate institutional actors, so alleviating the burden put directly upon us. I say an additional effort, because I appreciate the effort already being done.

I must confess that, as a recent example of an avoidable duplication for EU member countries, I have in mind the ambitious SDDS exercise of the IMF, inasmuch as it goes well beyond financial and monetary matters.

About the second objective, which deals with the measurement of the effectiveness of the statistical information produced, it is a bit surprising, nowadays, the importance still given to an indicator of the sales of printed publications, which has a target for the year 2000. A brief additional comment would be welcome.

The next major objective, to have *a comprehensive quality management system* contains, as specific targets, the increase of response rates in household surveys, and a production of research papers which makes you green with envy. I would say that there is no evident correlation between response rates, a quantitative indicator, and the quality of the final product; and that the outflow of research papers is much more a result of the institutional and scientific quality reached by SN than a direct contribution to the quality of their statistical products.

As for the statistical auditing proper it appears, I repeat, an extremely promising experience whose development and results would be very useful to follow up, for many of us.

Finally, just a glimpse at what we are doing in this field, in INE, NSI Portugal. Quality became an issue of management policy, identified and organised, also since the beginning of 1996. The process is steered by a specific committee which includes as permanent members the managers in charge of coordination and integration and of informatics and information systems; production managers are called to participate as the agenda may require. Our planning and control unit, a top management staff office, has been renamed: Planning and Quality Control Office. It is, indeed, the control room of the system. It has quality correspondents, or quality focal points so to speak, in each production department, who may or may not be the departmental managers.

We distinguish very clearly two facets: the quality in the interface between INE and the public, quality of attendance if you wish, and the substantive quality of statistical products and services.

As for the first aspect, we published in October 96 a so called *Quality Chart* where very clear standards are assumed as obligations towards both statistical information suppliers and users. The fulfilment of those obligations is regularly monitored.

The current and future initiatives within our quality work program include three handbooks:

◗ A *Statistical Production Procedures Handbook*, issued in January 97, where guidelines are defined

for the sequential procedures which must be adopted at each stage of every statistical activity.

◗ A *Statistical Production Quality Control Handbook*, closely related to the previous one, where the sequential steps and methods to be applied in quality control of all statistical production activities will be described.

◗ A *Quality Handbook*, which will provide a consolidated description of INE quality policy, of the quality management system, of the organization units dealing with quality control and of the tools which support the quality work program.

In addition to those basic documents, we produce and publish on a quarterly basis a *Quality Indicators Panel*, monitoring the quality standards which should be achieved in the dissemination of statistical information (publications and other supports). And we started quality analysis, or audits, on a case by case basis, giving priority to surveys with direct questioning of the respondent.

The point is to build a quality culture. Or, to avoid being unfair, to help and support production statisticians in their strive for quality, through a clearly defined and properly instrumented quality management.

Which is, I think, very much in line with the approach adopted by SN.

2.2. Discussion: Nicholas Karavitis- Statistical Office, Greece

Audits are important but are not used for surveillance purposes.

The Greek delegate underlined the usefulness of the report because of its specific approach and the very fresh manner of regarding the auditing of statistical quality.

The Greek office would not yet have a systematic framework for quality management. Several approaches would be possible of which the right type of staff, the resources available, the cost benefit ratio would have preference.

The creation of a centre for government surveys would attempt to reduce or limit the surveys carried out outside the national institute. In Greece the statistical office had been working in this direction with co-ordinating, monitoring and auditing all statistical operations in the public sector. Plans were being developed to set up such a centre in Greece.

Relevance would be another important tool in a sta-

tistical programme. The evaluation of user satisfaction would be extremely difficult and the usual practise would be that new statistics were added to the existing ones.

Regarding audits and their code of practice, it should be stressed that audits are not policing operations in order to make them a success.

Another experience was that for the auditors to be successful they would need external as well as internal testimony and evidence. This would imply that the auditing staff should not be part of hierarchy or should even come from outside the institute itself

Concluding, the Greek delegate was convinced that all these efforts to control total quality management can lead to statistical quality, the quality of products, and to the development of a culture of quality.

Treatment of non-response in household survey

A. INTRODUCTION

One of the most serious problems a Statistical Office has to face in their surveys is the lack of information by respondents.

This may be accounted for by the absence of the survey units at the time of the interviews, by their incapacity to answer, the impossibility to find them or their refusal to collaborate.

In Spain, collaboration with the Instituto Nacional de Estadística (INE) is compulsory for all those statistics and surveys included in the quadrennial National Statistical Programme amd which satisfy the European Union requirements.

Most of the INE's household and business surveys belong to this category.

This document analyses the importance of non-response in the INE surveys, with an emphasis on household surveys; it points to the steps taken in order to reduce this drawback and to tackle the non response which subsists despite all these efforts.

B. THE SIGNIFICANCE OF NON-RESPONSE

The situation varies according to whether the surveys address enterprises or households: non-response rates are lower for the former than for the latter.

Three of the most important INE surveys may serve as examples. For the annual Industrial Survey, with a sample of 39.000 industrial establishments, the usual non-response rate is 9,97% (4,87% for refusals and 5,10% for absentees).

For the quarterly Labour Force Survey, whose sample size is 64.000 households, the usual non-response rate is 10,24% (3,59% for refusals, 6,65% for absentees). This survey is designed with a rotating sample, so that each quarter, one sixth of the sample house-

holds is replaced. For the approximately 10.600 households which are interviewed for the first time, the non-response rate is 14,66% (7,53% refusals and 7,13% absentees), whereas as for those interviewed for the second time or later, the rate is 9,24% (2,70 refusals and 6,54 absentees).

Finally, for the Household Budget Survey, non-response rate is 13,9% (10,8% for absentees and 3,17% for refusals).

The effect of non-response is more severe in enterprises surveys when units with probability one are included in the sample: the latter probably are big enterprises which sometimes represent a whole branch of activity. Consequences are less drastic where small enterprises are concerned.

In household surveys, non-response is concentrated on age group 21 to 49, where absentees are more frequent.

C. MEASURES TAKEN TO REDUCE NON-RESPONSE

Here again, a distinction is to be made between enterprises or families.

For the first, the use of administrative registers as sampling frames cut down the percentage of non-located units, even though this procedure was not designed to reduce non-response.

The purpose of the first specific step towards a lower non-response is that one and the same interviewer be in charge of every contact with the enterprise, be it for mailing the questionnnaire or claiming and editing it. Establishments are usually surveyed by mail.

Centrtalized units have been set up to claim questionnaires by telephone.

Participation of the same enterprise in several surveys

should be avoided. This is impossible to achieve in the case of big enterprises, but as their resources are larger, it seldom is a problem. Furthermore, they generally are users of the survey results.

Questionnaires have been adapted to the accounting system of the firms. For instance, those for the Industrial Survey follow the Spanish General Accounting Scheme. Currently a procedure is being built up by which it will be possible to use the enterprises' own information systems to fill in the statistical questionnaires.

Moreover, 'tailored' information is offered free of charge to the units answering the survey.

As to Household Surveys, the first step towards obtaining the families' collaboration is to send them a letter to explain the aims of the survey and to tell them that an INE interviewer is going to visit them. These calls are made by full-time personnel with a long experience in field work. In surveys requiring more than one visit, a letter is sent again before every subsequent interview.

When the family is not found, the interviewer has to call again. In the Active Population Survey (APS) a deadline of three to six weeks is established in order to increase the possibility of getting the questionnaire answered.

The use of portable computers in APS allowed to lighten the burden of respondents since part of the information is kept from previous visits.

As from 1997, telephone calls have been used for the APS (only in capital cities of the provinces) in order to alleviate the respondents' burden and to contact those families who are likely not to be at home. This is done only when the families agreed with this procedure. Throughout 1998, telephone calls will be extended to the overall sample.

An analysis has been made of the causes of refusal to collaborate. The families indicated the following:

Causes

Lack of interest

Too much trouble

Fear or distrust

Death or sickness

Not the same family in the dwelling

No causes indicated

Other causes

In the future, an effort will be made to change the factor *lack of interest*. To this end, a second letter, signed by a General Director of the INE will be sent together with an information brochure with the survey's main results. To diminish *fear or distrust*, information on the survey will be given free by telephone.

In the past, families that collaborated in the Household Budget Survey received a gift. This custom has been given up since it proved to bias the composition of the sample in favour of the poorer families, particularly when the gift in question was money.

D. TREATMENT OF NON-RESPONSE IN HOUSEHOLD SURVEYS

Despite all the efforts in fieldwork, a measure of non-response remains which has to be treated. Considering only household surveys, the INE tried the following methods:

a.- Replacing the non-response unit by another one with similar characteristics, using those 'in reserve' of the same area.

Replacement as well as selection are made easier thanks to a list of dwellings for each of the 30.000 areas in Spain. These lists may be completed with socio-demographic data from the Population Register ('Padrón Municipal') providing the identity and address of all those living in the municipalities of Spain.This is the method used for the Active Population and Household Budget Surveys; it also helps tackling first interview refusals.

b.- Attributing to a non-respondent unit the information the latter itself provided in former interviews, this being done in continuous surveys where a family remains in the sample for more than one period. This method, also applied in the APS, has proved efficient since the imputed values closely approach the actual ones.

For the APS an Evaluation Survey is conducted. Interviews are repeated for a subsample fifteen days after the first call. It facilitates the assessment of errors in coverage and data and checks the interviewers' activity. As yet, it has not been used for non-response: on the one hand, first interview non-respondents have been credited with information obtained at the second call, and on the other few refusals or absentees are interviewed during the Evaluation Survey.

Neither have weighting techniques been applied, although the INE is considering them at present. Further down, mention will be made of those analyses that availed themselves of said techniques in the Active Population Survey.

The APS is now being renovated in order to comply with the requisites of the forthcoming European Union Labour Force Survey, in January 1999. One of them is the re-weighting of NUTS II, provided that reliable data be available from outside sources.

E. RE-WEIGHTING METHODS USED IN THE LABOUR FORCE SURVEY

The main purpose of re-weighting is that the sample represents more efficiently the distribution of some particular characteristics.

Indirectly, it is also useful to correct possible flaws of the survey frame, biases in the sample selection or even inadequate weightings of the selection probabilities or the treatment of non-response.

By means of re-weighting, fresh elevation factors are obtained which improve those initially assigned to the sample units.

These renovated factors are based on reliable external sources providing information on some characteristics of the population (generally on distribution by age and gender). Thanks to re-weighting, the estimated distribution of said characteristics become equal to those stemming from other sources.

The underlying assumption in any re-weighting technique is that the auxiliary variables are highly correlated with the main variables considered in the survey.

The external source used for the LFS is the population projections by gender and age. The following methods have been implemented:

a. Post-stratification

b. Canadian or Person Matrix

c. Household Matrix

d. Calmar

Method a) is a particular case of the raking-ratio method. Methods b) and c) are based on generalized least squares.Method d) requires the minimization of a distance function for the whole sample's original and corrected weights.

Methods a), b) and c) need detailed information on the population by age and gender at the desired geographical level, that is the province in the case of Spain (NUTSIII). The Calmar method may be applied with incomplete information, i.e. with age and gender distribution up to a given geographical level and below that only with total population figures. In the experiment under consideration, account has been taken of the distribution by gender and age for the national and Autonomous Communities population totals (NUTSII).

Results as indicated in the Annexes show no great differences.

That is why the choice of a method is determined by two factors: its easy application and its requirements as to breaking down demographic projections (outside variables which are considered)

The Canadian method yields negative weights which should be corrected subsequently. The Household Matrix has yet another drawback: that of working

first with households and afterwards with persons. For both stages, projections by gender and age are needed. We therefore deemed the Calmar method with incomplete information the most appropriate one, since its external input is both the total national five-year population projections by gender and age, an absolutely reliable figure, and the Autonomous Communities population totals.

The results proved to mirror passably the provincial totals. To obtain a perfect fitting, calibration was necessary in some measure. It would be possible, however, to apply the method starting straight from the national projections by gender and age and the provincial totals, but this would make calculations much more difficult.

The results show an increase in employment and unemployent of about 4%, which is in keeping with the higher activity rate of age group 21 to 49

F. ADVANTAGES AND DRAWBACKS OF RE-WEIGHTING

Reasons for re-weighting survey data have been described above. In sum they are:

- To correct errors in the survey frame (undercoverage, outdated information, etc.) and, generally speaking, flaws at any stage of the survey design.

- To decrease errors derived from non-response, even after endeavouring to overcome this difficulty.

An inescapable condition for re-weighting is the availability of exogenous variables liable to provide a reliable population distribution which may be used as a model for the sample.

Distrust as to the reliability of these variables is the first argument against reweighting.

Another important objection is that the implementation of these methods leads to laxity in field work meant to lower non-response.

Obviously, in this context an effort should be made to improve interviews. But even with high investments of time and means, there still will be units for which it is impossible to gather information.

An additonal drawback lies in the very reliability of the re-weighted data. Periodically, the APS results are checked against those from other statistical sources, in particular from administrative registers: Social Security affiliates and job searchers at the Instituto Nacional de Empleo (National Employment Institute). As a by-product, the latter also provides information on registered unemployment.

As things are, unreweighted APS data are coherent with those of job searchers registers. Those for the unreweighted employed as shown in the APS, do not agree with the Social Security register. Re-weighting leads to inverted results: data on employment become coherent, figures on job searchers are upset.

At present, our researches on the possibility of re-weighting continue. In fact, a Working Group has been created at the Consejo Superior de Estadística (Higher Statistical Council). It is made up of the representatives for the main producers and users of labour data. One of their specific tasks is to analyse re-weighting.

2.3.1 Discussion: Frans Desmedt- Statistical Office, Belgium

Collection of the maximum of information on the respondents who do not provide the data requested.

A. NON-RESPONSE: A GENERAL OVERVIEW

Ms Guzman's contribution shows that one of the most frequently encountered problems in statistical surveys is the inability to obtain usable data for all the sections of a questionnaire from all the members of a sample. The information collected is always incomplete, irrespective of the efforts made to obtain data that are as comprehensive as possible.

A distinction should be made between two types of non-response: total non-response and partial non-response.

Total non-response occurs when an element in the sample does not reply to the survey.

Partial non-response occurs when an element participates in the survey but the data concerning certain sections of the survey are not available for some reason.

Non-response should be taken into account in each stage of a survey.

The reduction of the non-response rate upstream of the actual survey is an important part of the work. It should be included in the design of the survey, on a par with:

◗ overall costs,

◗ deadlines,

◗ required detail.

These parameters will guide the initial choice:

◗ drafting of the questionnaire (clarity, sensitive questions),

◗ method of collecting data (postal survey, telephone, existing files, etc.).

Reducing the non-response rate must also be an objective during the collection of the data, in particular through surveillance by the interviewers. Efforts to obtain the cooperation of respondents must not, however, lead to any significant loss in the quality of the data collected.

Downstream of the survey, various statistical techniques can serve to reduce the effect of non-response.

Choosing among all the methods that can be used to correct non-response necessarily calls for a reasoned analysis of its causes. Depending on the variable to be estimated and the type of non-response (partial or total), it is likely that different methods will need to be combined to achieve a usable result.

This presupposes that upstream, when the data are collected, non-responses are not simply discarded but are given careful attention.

In particular, those collecting the data should collect the maximum of information on the respondents who do not provide the data requested and the reasons for their refusal. This information is important in order, for example, to be able to correctly choose a homogeneous class in relation to the non-response.

The reasons for non-response provide key information for the treatment of the non-response. They determine whether or not the non-response was random and can therefore guide the choice of the secondary variables which will be used to define imputation or the weighting classes.

The second criterion for the choice of a method to correct the non-response is the purpose of the data collected. Depending on whether interest lies in the total of a variable or in its distribution, different imputation methods should be used.

This means that the person responsible for the methodology should have:

◗ not only a detailed knowledge of statistical methods

◗ but also detailed knowledge of the specific survey for which he or she is working.

This presupposes that the person is not involved in the survey simply when the results are compiled. He or she must be kept informed of the problems arising in the survey and be able to give useful help throughout the performance of the survey in order to prevent disasters resulting from the wrong approach to certain problems.

B. NON-RESPONSE IN CERTAIN SETS OF BUSINESS STATISTICS

Ms Guzman focused mainly on the non-response in household surveys.

To conclude I would like to make a few comments on the treatment of non-responses in two sets of business statistics.

B.1 Prodcom statistics

i) Estimate of unreturned declarations (use of external data and regression).

The estimation of unreturned declarations involves estimating the value of industrial production (Prodcom turnover).

The Prodcom turnover of a business that has not returned a declaration is estimated by regression between the monthly Prodcom turnover of the business and the monthly turnover declared to the VAT authorities.

If the turnover figures are not available, the Prodcom turnover is estimated by regression between the Prodcom turnover and the turnover of the sector.

The rest of the declaration is estimated according to the same procedure as for incomplete declarations.

ii) Estimate of incomplete declarations (imputation on the basis of ratio)

The estimate of an incomplete declaration involves the following stages:

Prodcom turnover is divided between the various products previously declared by the business in the same proportions as in the preceding months.

The value of production by product is broken down between deliveries and contract work according to the same breakdowns as in the preceding months.

The quantities are estimated according to the unit price, calculated on the basis of the business or the sector.

B.2 Survey on the structure of businesses

The survey on the structure of businesses is an annual survey of businesses. The data collected cover staff, the balance sheet and business investments during the commercial year.

i) Total non-response (reweighting on the basis of the sampling stratification).

Estimates are not systematically made for businesses which have not responded to the survey. Corrections for non-response are made by reweighting the correct declarations. Reweighting is carried out by sector of activity and size class, whilst maintaining the original stratification.

ii) Partial non-response (imputation on the basis of ratio).

Incomplete declarations are corrected (on the basis of the 1996 survey) by automatic rectification.

A declaration is rectified if the basic balance sheet data are sufficient. If not, the declaration is considered unusable and is treated as a total non-response.

Automatic rectification covers all the variables of the questionnaire, except the basic balance sheet data.

The missing data are estimated on a ratio basis and calculated using businesses within the same sector of activity and the same size class.

3.1 Bengt Swensson, Professor of Statistics University of Örebro and Uppsala University

Quality management approaches in official statistics production - some examples and reflections

A. INTRODUCTION

Probably the most well-known plan for organization for quality and productivity is the one origi-nating from one of the most influential survey statisticians ever, the late Dr. Morris H. Hansen at the US Bureau of the Census. Of course, he is known by all survey statisticians, but his fame goes far beyond this professional group, largely due to another survey sta-tistician, more known as a world-famous quality advocate, the late Dr. W. Edwards Deming. In his *Out of the Crisis* (1986) Deming sketches the plan (pp. 466-470), and he considers the principles of the plan to be appropri-ate for all types of organizations. One salient feature of the plan is: "There will be a leader of sta-tistical methodology, responsible to top management. He must be a man of unquestioned ability. He will assume leadership in statistical methodology throughout the company. He will have authority from top management to be a participant in any activity that in his judgment is worth of pur-suit. He will be a regular participant in any major meeting of the president and staff. He has the right and obligation to ask questions about any activity, and he is entitled to responsible answers. The choice of application for him to pursue must be left to his judgment, not to the judgment of others, though he will, of course, try to be helpful to anyone that asks for advice. The non-statistician can not al-ways recognize a statistical problem when he sees one." After elaborating the plan further, Deming concludes, cat-egorically: "The advantages of the plan recommend-ed here can not be questioned. It works. Any other plan that I have seen has failed to serve the best interest of the company, and has brought disap-pointment."

The term TQM (Total Quality Management), according to Walton[1], seems to originate from the US military in the 1980's as an acronym for the mil-itary's quality drive inspired by the Deming method.

It was later used by various quality consultants for their specific concepts of quality man-agement (QM). Thus there is no single meaning of the term. Deming himself did never use it - one of my col-leagues at the University of Örebro once asked, at a Deming seminar, about TQM. Deming's answer was a wry "I don't know what it is." However, TQM (with all its different mean-ings) has had a strong influence in wide circles, in particular in the US. Among basic elements usu-ally mentioned in TQM-overviews are: customer focus, teamwork, fact-based decisions, continuous process improvements, and management leadership.

This was the background, together with some con-cern for signs of deterioration in survey practice, for a few short (between half a day and one day and a half) visits by myself and Professor Carl-Erik Särndal in 1993 to the following eight statistical agencies: the Australian Bureau of Statistics (ABS), the US Bureau of Labor Statistics (BLS), the US National Agricultural Statistics Service (NASS), the New Zealand Department of Statistics (NZDS), Statistics Canada (SC), the Statistics of Income Division (SOI) of the US Internal Revenue Service (IRS), Statistics Sweden (SS), and the US Bureau of the Census (USBC). We posed a number of ques-tions to a "key QM person", e.g. a person with good knowledge of agency quality efforts[2]. Furthermore, during our visits, we met with a few Product Managers of selected statistical products. The pur-pose of our contacts was not to rank these agencies, but to try to get an idea of possibly different approaches to ensure the produc-tion and dissemi-nation of high-quality official statistics, which, inter alia, could be used in a possi-ble second edition of

---

[1] Walton, M. (1990). *Deming Management at Work.*

[2] The questions are reproduced in the Appendix. They were not meant to favor any specific quality management ap-proach; for example, no mention is made of Deming or TQM.

our book on survey sampling[3] This is a very short summary of some findings from these contacts[4], with a few reflections.

We were very kindly received at all agencies, and we met very dedicated professionals, who gener-ously shared with us their knowledge. I take this opportunity to, once again, express our sincere gratitude.

B. MAJOR RELEVANT BACK-GROUND DIFFERENCES

The eight agencies differed widely with respect to some important conditions for the production of high-quality official statistics.

◗ (a) Four agencies, ABS, NZDS, SC, and SS, were central, general purpose statistics producers. These agencies' responsibilities and functions as decreed by law differed substantially: (a1) The statistics acts regulating the activities of ABS, NZDS, and SC were comprehensive, and they ex-pressed, directly or indirectly, the respective countries' high esteem of official statistics; for ex-ample, New Zealand's Statistics Act (1975) recognizes the importance of the contribution which accurate and timely official statistics make to effective public policy, commercial decision-making, democratic processes and to an informed community. Furthermore, the Director Gener-als (the Australian Statistician, the Government Statistician, and the Chief Statistician, respec-tively) were given quite extensive power; in particular, according to the Chief Statistician, Dr. Ivan Fellegi[5], Canada's Statistics Act "is a unique statistics act - I'm not aware of any other with these particular combinations of features." It should be obvious that an agency based on this sort of statis-tical act is more likely to give the staff the feeling that they take part of a worthwhile ac-tivity, and, hence, the prospects for staff commitment should be very favorable, a commitment that is an essential prerequisite for high-quality work in an organiza-tion. (a2) The Swedish Sta-tistical Act was of a much more meager kind, in particular, it did not contain material which, by itself, might create staff pride and commitment. (b) The other four agen-cies, i.e., BLS, NASS, SOI, and USBC, were part of the US decentralized statistical system, and, hence, had more lim-ited responsibilities.

◗ "Political pressure" on the agencies to choose a spe-cific QM approach were very different. (a) In the US, as I perceived it, there seemed to be a rather strong political "encouragement" to choose a TQM approach, as evidenced by a few examples: President Bill Clinton's engagement (as a Governor) in the Arkansas quality campaign, state-ments from Clinton like "The ideas of W. Ed-wards Deming have become a powerful effective force for change in American industry. With appropriate adaptation, Total Quality Management offers the framework and the tools to be equally efficient in government", and the Government awards pro-grams (the federal Quality Im-provement Prototype award and the President's Award for Quality and Productivity Improve-ment) are all geared to TQM[6]. According to Carr and Littman, the federal government's major campaign to introduce TQM to all parts of the executive branch by 1992 had resulted in more than 50% of 2,800 federal instal-lations surveyed saying that they had quality initia-tives under way. (b) In the other countries, no such strong explicit signals seemed to have been sent out.

C. QM APPROACHES AMONG THE SELECTED AGENCIES

The answers to the questions posed to the "key QM persons" together with documents handed over to us showed that the agencies covered a broad range of approaches: from low interest for "mod-ern" quality approaches to very strong dedication, from not so strong to very strong emphasis on "traditional" qual-ity aspects.

Of the questions asked (see the Appendix) answers to questions 2-7 (block 1) give some indication of the degree of emphasis on QM in the more modern view, while answers to questions 8-10 (block 2) give some indication of the degree of emphasis on more "tradi-tional" concepts of quality. The answers can be cross-

---

[3] Särndal, C.E., Swensson, B. und Wretman, J. (1992). Model Assisted Survey Sampling.

[4] I am pleased to be able to use this occasion to present some conclusions from this project

[5] Interview in SCAN, special issue, Winter, 1993.

[6] Carr, K.D., and Littman, I.D. (1993). Excellence in Government: Total Quality Management in the 1990s.

classified according to the number of "yes" answers to the two blocks of questions. Before I present the table, I would like to emphasize that it does not rank the agencies with respect to the quality of their output - it merely gives a very rough classification of QM ap-proaches. I would also like to point out that ABS is missing in the table for reasons spelled out in the Appendix.

Number of block 2 "yes" answers ("traditional" emphasis)	Number of block 1 "yes" answers ("modern" emphasis)	
	Low	Medium-High
High	SC	SOI
Medium-Low	SS, NASS	NZDS, BLS, USBC

The table calls for some comments.

The QM approach of some of the seven agencies had a definite TQM flavor, namely, the US agen-cies, the one exception being NASS, which seemed to have considered various TQM alternatives, but "the Agency contracted for some Total Quality Management training for the Survey Quality Team which was frankly disappointing." It seemed unclear if NASS was going to continue with a TQM approach. BLS' overall approach was clearly a TQM approach, called BLSQM. However, the approach was decentralized, and participation voluntary: each organization within BLS could ac-cept or not accept to implement, which probably was a necessary requirement, since there was con-siderable internal competitiveness among BLS organizations. Examples of programs with their own QM programs and their own QM administrations were Employment and Unemployment Statistics (EUS) and Consumer Price Index (CPI). USBC were using a TQM approach. However, the key management person we inter-viewed, meant that there were clear indications of decreasing com-mitment of top management to QM, and he hoped that a new appointment at the post of Undersec-retary for Economic affairs would bring about some fresh blood in the QM process. SOI was high on both counts, partly, I believe, because it could benefit from being a division of IRS (which seemed to have a strong interest in TQM), partly, I believe, because of the director, Dr. Fritz Scheuren, a highly qualified statistician with a strong commit-ment to QM.

For the non-US agencies, the table indicates that SC did not have a formal QM Approach. In a memoran-dum of April 8, 1993, Dr. Fellegi expressed the opin-ion that "one can move successfully on all relevant fronts of the QM philosophy without necessarily embracing the paraphernalia." However, being a highly successful agency, SC had, of course, a num-ber of activities found under a "modern" QM approach. The NZDS approach was not a TQM approach, but had much of what you would expect in a "modern" quality approach. SS, finally, at our visit, was about to start a broad QM project with TQM orientation.

D. STRENGTHS AND WEAKNESSES IN THE ADOPTED APPROACH

The last question to the key management persons asked them to give their opinions on strengths and weaknesses in the chosen approach. Below are some of the answers.

BLS *Strengths*: The decentralized quality approach. The availability of a resource for QM. The quality council. *Weaknesses*: The corporate mind has not been firmly set on QM. The in-terest in QM is highly variable among pro-grams, TQM is a priority in the BLS, but it can-not be said to be a top priority.

NASS *Strengths*: NASS feels that its commitment to quality from top management on down is the key to the present success, (and that is more important to accept and follow the princi-ples of TQM than it is to formally adopt it as a management technique.) *Weaknesses*: Not formally accepting TQM in its entirety some-times lead to difficulties in explaining the NASS approach to quality with those that have formally adopted TQM.

NZDS Strengths: Making employees responsible for their output. Recognizing the customer and his/her needs. Weaknesses: Currently there is a great deal of change happening in the De-partment of Statistics. While there was little

change for 50 years, the situation is now that everything is changing. The employees now have to live with this.

SC *Strengths:* Having established closer ties to clients. A highly qualified work force. Respon-dent confidence: Statistics Canada has managed to build up trust among respondents in part through good questionnaire design. Credibility on the part of the Canadian public: Statistics Canada is viewed as a well managed bureau and enjoys a strong reputation. *Weaknesses:* A worry is the deterioration of data going into the National Accounts.

SOI *Strengths:* Involvement in/participation by customers, suppliers and Employees in the planning, managing, and improvement process. *Weaknesses:* With respect to measurement, some unanswered questions are: What to measure, when to measure, how to measure, where to measure, and who is to measure. With regard to utilization of skills, some prob-lems are: People often are not aware of skills/training they lack which prevent them from executing the quality responsibilities in a comprehensive, efficient and effective manner.

SS *Strengths:* Quality enhancement programs are actively promoted and supported by top management. Endurance in the approach to the work; if this attitude was not present, SS would fail. A clear perception of what Swedish users of statistics really want. *Weaknesses:* SS has not yet succeeded in promoting among its employees a full appreciation for and understanding of the broader view of the quality concept. Traditionally, SS employees have regarded "accuracy" as being synonymous with "quality". A necessary rethinking is, however, under way. The user perspective is now placed at the foreground. But there still remains much to be done before all employees will view themselves as providers of serv-ice to society and to specific customers, rather than as providers of "accu-rate numbers".

USBC *Strengths:* At least one thing done well: A general awareness of the customer(s) has been created, even among mathematical statisticians. *Weaknesses:* A lack of true management commitment. The thing done worst: A failure in regard to systematic measurement.

E.   A FEW REFLECTIONS

It should be noted that the "modern", more or less TQM influenced, QM approaches used by several agencies were (at the time of our visits) of a (relatively) recent date, which, of course, made it diffi-cult to sort out possible positive bottom-line effects; e.g., Hawthorne effects may well occur. (The approach of Statistics Sweden is presently under evaluation.) On the other hand, Statistics Canada's well-tried approach had (at the time of our visits) been very favorably evaluated: In *The Econo-mist's* appraisal of official statistics producers, Statistics Canada was number one (with ABS as runner-up).

It is, perhaps, somewhat astonishing that none of the agencies made reference to the Han-sen/Deming plan for organization for quality and productivity, especially since all agencies are sta-tistics producers, and since TQM is said to be inspired by the Deming method. Considering the enormous sales success of *Out of the Crisis* (my own copy from 1991 is the four-teenth printing) and considering that Deming as late as in 1993[7] continued to argue in favor of the plan (and praise Han-sen, "a leader great and good") it is not probable that top management of the agencies could be un-aware of the proposed plan. What is the explanation?

The agency that seems to come closest to the Hansen/Deming plan is Statistics Canada, with a long history of very strong statistical leadership and with many highly qualified (mathematical) statisticians. I am sorry to say that, according to my own experiences from Statistics Sweden (from two sojourns, 1968-1969 and 1977-1984), highly quali-fied statisticians have had rather modest influ-ence on major (organizational, and other) decisions at Statistics Sweden. In particular, I remember a two-hour conversation in 1984, in connection with my

---

[7] *Deming, W.E. (1993) The New Economics.*

leaving the bureau with the then-Director General Sten Johansson in which I tried to make him pay attention to the fact that during a couple of years a handful of very competent statisticians had left the bureau. My prediction was that, if nothing forceful to cope with the situation was done, Statistics Sweden would run a large risk of declining quality. As far as I remember, no such forceful action was taken. Maybe this is one of the explanations for Statistics Sweden's' fall from third to fifth place in *The Economist's* appraisals? I now pray that the new QM approach will turn out to the better.(It is alarming that out of the six agencies that answered the question "Do you consider that you have sufficient number of qualified statisticians?" three gave "no" answers.)

A final question: Is Deming, after all, right, i.e., is the Hansen/Deming plan the one that "serve(s) the best interest of the company", while any other plan will bring disappointment?

APPENDIX

Questions to "key QM person" (exclusive of a few follow-up questions) The answers (except to the last two, which are presented in the text) are given on the next page.

1. Is there a Statistics Act or similar legislation regulating your agency's activities? If YES, what mention, if any, does this act make of "Quality".

2a. Has your agency issued a Mission Statement on the general purpose of the organization?

2b. Is there a Mission Statement in regard to Quality in particular?

3. Is there a recent Strategic Plan?

4. Does your agency have a written QM Strategy (a clearly expressed commitment of upper management to QM)?

5a. Has your agency created an upper-level quality council?

5b. Has your agency appointed a Quality Management Chief Officer?

5c. Has your agency created lower-level quality councils?

6. Has your agency established Quality Goals (quality targets) for the (near) future?

7a. Has your agency adopted a Quality Definition?

7b. Has your agency issued a Quality Policy?

7c. Has your agency published a Quality Glossary?

8a. Has your agency published any document that can be characterized as a Quality Manual (a document primarily for planners and managers, in some agencies called Quality Guidelines, suggesting how to achieve quality in different stages of the production process)?

8b. Has your agency issued Guidelines on Documentation of Data Quality?

8c. Has your agency issued Release Criteria (quantitative rules for suppressing statistical data deemed to be of inferior quality)?

8d. Has your agency issued a Policy Statement on Informing Users of Data Quality and (Methodology)?

9 Has your agency started any general purpose projects (or activities) that you view as being of vital importance for achieving quality?

10a. Has your agency developed a formal procedure for recruiting personnel (university graduates) with specialized skills (e.g., mathematical statisticians, economists)?

10b. Do you consider that you have sufficient number of qualified statisticians?

10c. Does your agency have regularly scheduled course offerings (vocational training, on-the-job training)?

11a. In your opinion, what are the key success factors in your agency's approach to achieving quality output?

11b. Are there any weaknesses in your agency's approach to achieving quality output?

Answers to the questions

Question	ABS	BLS	NASS	NZDS	SC	SOI	SS	USBC
1	yes nc	yes none	yes none	yes none	yes none	yes none	yes none	yes none
2a	yes	yes	yes	yes	yes	yes	no	yes
2b	no	no	no	yes	no	nc	no	yes
3	yes	no	yes	yes	no	yes	no	yes
4	no	no	no	yes	no	yes	no	no
5a	no	yes	no	no	no	yes	yes	n.e.
5b	no	no	no	yes	no	no	no	yes
5c	no	yes	no	no	no	no	no	n.e.
6	no	no	no	yes?	no	yes	no	yes? [8]
7a	no	yes	no	yes	no	no	no	yes
7b	yes	yes	no	yes?	no	yes	no	yes
7c	no	no	no	no	no	yes	no	no
8a	yes?	no	no	no	yes	yes	no	no
8b	nc	no	no	no	yes	yes?	yes (old)	no
8c	nc	no	no	yes	yes	no	no	no
8d	nc	no	no	no	yes	yes	no	no
9	nc	nc	nc	yes	yes	yes	yes	no
10a	nc	no	no?	no	yes	yes	no	yes
10b	nc	no	yes	no	yes	yes	no	-
10c	nc	no	yes	yes	yes	yes	yes	yes

(nc = not conclusive)

The interview with the key QM person at ABS had to be interrupted, and we did not get complementary answers by mail. However I want to point out that ABS gave an impression of being a dedicated agency, with good quality efforts on the agenda. In particular, I was impressed by some of their questionnaires. At the time of our visits to the agencies, ABS had been rated runner-up (after SC) in *The Economist's* evaluation of official statistics producers.

---

[8] *According to the key QM person "employee involvement has been put on the back burner."*

3.1.1 Discussion: Claudia Cingolani, Statistical Office, Italy

Quality mandates must be supported by coherent and comprehensive strategies

By providing an interesting overview of the approaches followed by some important statistical agencies to ensure the production and dissemination of high-quality official statistics, Prof. Swensson's paper greatly contributes to the discussion on the qualitative aspects statistics. This document, in fact, gives us the opportunity to reflect on some relevant aspects emerging from the comparative analysis at the international level on the different strategies (if there are any) adopted by statistical bodies for approaching the Quality Management issue.

All the offices interviewed produce official statistics and belong to the public administration. Nevertheless, they seem to have very different approaches to the Quality Management problem and their experiences illustrate the peculiarities of each of their specific operational contexts along with the wide range of strengths and weaknesses of the approaches selected among them. Indeed, the different perceptions of the QM issue existing at the international level is due to both the lack of a common definition of the TQM concept in the statistical sector and to the organizational model within each public administration which is characterized by a dishomogeneous distribution of responsibilities and functions at the national level. This aspect seems very relevant to understanding how the institutional and legislative environment can influence Quality management strategies among countries and how basing the production of official statistics on centralized or decentralized systems can impact, more or less, on the quality management approach in the public administration.

On this point let me provide a practical example from my country where the institutional asset of official statistics is decentralized according to the Legislative Decree of 1989 on the functioning of the National Statistical System. This important legal Act, which created the basis for the decentralized production of official statistics, established that all statisti-

cal activities fall within the scope of the National Statistical Programme and that official statistics could be produced by the different bodies belonging to the system, even if under the coordination and responsibility of the National Statistical Institute.

This fragmentation of potential producers requires, on the one hand, a wider awareness of the importance of quality work in statistics as a whole and, on the other hand, that all the components of the system are able to move together in the same direction and, overall, at the same speed. The process for reaching quality standards must be supported by coherent and comprehensive strategies mainly oriented toward ensuring quality in the production and dissemination of official statistics and to improve, on the part of the public, the credibility and trust in the information produced by the system as a whole. For the fulfillment of these objectives it is necessary to improve those strategies aimed at:

one, redefining the statistical activities plan in a logical way so as to build up integrated information systems rather than concentrating actions in single statistical production cycles. The approach toward information systems should represent the key for the growth of the statistical function and, consequently, the reason for re-organizing the way in which the official statistical sector operates.

The guidelines of this organizational reform should mainly address changes aimed at:

◗ a better subdivision of functions - no longer on the basis of bureaucratic and hierarchical principles but on criteria based on horizontal functions;

◗ no longer having a static approach to procedures (which says: what should be done and how) but instead a dynamic approach (who makes what and why) focusing more on coordination aspects;

◗ and finally, a higher involvement of the staff in the

processes with less distinction between executive and controlling activities in order to improve the process of transparency, the awareness of objectives and problem solving capabilities:

two, improving the user-oriented quality of data which means that all statistical information produced should be perceived as accessible, reliable and truthful. The user-oriented quality of statistical information must be also pertinent, that is to say it must meet the information needs of users.

This general approach, however, cannot limit itself to a statement of intents, but it must meet the challenges imposed by the users' differentiation and by the increase in their needs. These two aspects are closely linked together. The quantitative and qualitative increase in the statistical information demanded does not automatically imply a proportional growth in the official statistics sector. The characteristic of official statistics as a public good is no longer to be found in its being a "natural monopoly" on production processes, but in the need to promote market efficiency and transparency.

3.2 Yves Franchet, Statistical Office of the European Communities (Eurostat)

The quality of the ESS

A. INTRODUCTION

The aim of this paper is to suggest ways in which we can improve the performance of the European Statistical System through the application of quality concepts. To achieve this we need to understand the changing environment in which we operate, recognise the features of the ESS that make it distinctive, adapt the text book quality notions to suit our own circumstances, and find ways of applying these in practice. I am particularly keen that any programme for improvement should be implemented with the full participation and support of all partners in the system. I am also anxious that we should build on the lessons learned from others and on the progress made in earlier DGINS discussions (references below). Throughout the paper I have highlighted the main messages from a Eurostat viewpoint by presenting these in Summary boxes.

A.1 Trends

Planning for the ESS over the next decade or so requires an intelligent anticipation of changes expected in many fields. I see the following ten factors as the main forces which will shape the demands made of the ESS over this period, and influence our ability to respond:

◗ The demand for European statistical information is expected to show marked future increases, driven by the needs of single currency, a growing interest in service sector statistics, developments in new statistical fields such as environmental statistics, the need of enlargement, and so on.

◗ Resources available to Eurostat and to statistical offices in Member States are expected to be level at best, possibly to fall.

◗ The growing gap between user needs and the ability of the ESS to deliver will bring an ever increasing focus on the responsiveness, internal efficiency

and effectiveness of the system as a whole.

◗ There is evidence from recent reviews and surveys, including the Mid-term report on the 1993-97 Five Year programme (the Wroe Report), that the scope for improvements in planning and delivering ESS outputs is significant.

◗ The new Five Year programme of work for ESS during 1998-2002 explicitly acknowledges the need for improvements in terms of better planning, improved customer focus, greater professionalism, improved ESS co-ordination and so on.

◗ The challenge of the single currency will bring significant new pressures in several ways. Apart from any new data requirements, the change will serve to emphasise the competitive relationship between the banking system and the NSIs and Eurostat as data providers, and may well raise the prospect of Euro-area results not necessarily built from national components. It will also give a high political profile to statistical work, not always welcome, where this impinges on decisions in the fields of convergence criteria and the growth and stability pact.

◗ Enlargement of the EU to 20 or more members will force a rethinking of ESS practices and procedures. Rules developed for six member states, already found unwieldy for 15, will prove totally unworkable for many more.

◗ The extension of powers of co-decision between the Parliament and Council envisaged under the Amsterdam Treaty will bring a shift in user focus that is difficult to predict at this early stage.

◗ Technological advance in IT and related fields will make possible major improvements in networking (e-mail, video conferencing, sharing of bulletin boards, etc), faster data transfers and processing of results, and greater scope for analysis at all stages

of the processing chain. These benefits will of course be just as available to our competitors, and failure to respond in the ESS will have a direct impact on the balance of work allocation between ESS, the banking system and the private sector.

◗ Although, in terms of promoting efficiency, Commission procedures and practices can at times seem unhelpful, there is now a climate of change and a real willingness on the part of the Commission to seek better solutions (SEM 2000, MAP 2000 etc.).

If it is to succeed, the ESS must plan ahead to anticipate the needs and opportunities arising from new circumstances in each of these areas, and avoid being caught unawares.

A.2 Standards

Before examining in detail how the aims outlined above can best be achieved, we should acknowledge the inevitable and continuous tension arising from the balancing of subsidiarity on one hand against ESS consistency and standards on the other. Whilst subsidiarity rightly requires that decisions should be pushed down to lowest sensible levels, and take full account of national differences (legal, institutional, political etc), it is important not to lose sight of the significant potential benefits of a common approach. Experience shows that these can be wide-ranging, and can include:

◗ Sharing across the ESS of improvements achieved in one MS

◗ Sharing of development costs

◗ Sharing of training costs

◗ Simplification of data transfers

◗ Assisting users e.g. the need to grasp a single scheme of seasonal adjustment rather than 15

◗ Assisting pre-entrant states who are adjusting methods and practices on joining the Community

In reaching decisions we need to take full account of all such benefits, even though some may arise in parts of the system other than our own.

A.3 Legal background

An important element in the existing ESS structure affecting quality of outputs is the legal background. The relevant Treaty Article (new Article 285 expected to be in force by summer 1998) refers to Community statistics which "shall conform to impartiality, reliability, objectivity, scientific independence, cost-effectiveness and statistical confidentiality". Reinforcing this basic Treaty requirement, the Council Regulation of 17 February 1997 on Community Statistics, often referred to as the Statistical Law, defines reliability as "the characteristic of Community statistics to reflect as faithfully as possible the reality which they are designed to represent" and refers to the need for scientific criteria in the selection of sources, methods and procedures. It also notes the need for information on coverage, methodology, procedures and sources in judging data reliability. These general requirements are specified in more precise terms in individual items of secondary legislation covering particular statistical domains.

THE EUROSTAT VIEW:

◗ Major changes will soon be demanded of all of us to cope with the combined impact of inadequate resources, single currency, enlargement, new technology and so on

◗ Now is the time to put in place mechanisms that encourage an intelligent anticipation of these events

◗ Failure to plan and respond at an early stage will result in dissatisfied users and a loss of business to our competitors.

B. THE EUROPEAN STATISTICAL SYSTEM (ESS)

B.1 Definition

A precise legal or organisational definition of the ESS is not easy to devise, but for the purposes of

improving ESS performance we can recognise the following elements:

◗ Eurostat and the National Statistical Institutes (NSIs)

◗ National government departments engaged in statistical work (agriculture, employment, etc.)

◗ Departments of banks and financial institutions engaged in work contributing to national or international statistics

◗ Regional bodies contributing to national or international statistics

◗ Bodies akin to those listed above, but operating in EFTA or in pre-entrant states

◗ Organisations and agencies with varying degrees of autonomy from central government engaged in national or regional statistical work (mapping agencies, customs agencies etc.)

◗ International bodies contributing to European statistics (UN, OECD, ILO, etc.)

◗ DGs of the Commission engaged in European statistical work

In summary, the ESS covers all elements of the system which collects, processes, analyses and disseminates Community statistics.

B.2 Mission and roles

In the absence of any agreed Mission statement for the ESS, I have taken the Eurostat Mission statement as a starting point and propose the following:

"to provide a high quality statistical information service to the European Union i.e. to national governments and parliaments, to international institutions, particularly those of the EU; to the business community; and to other users of statistics at regional, national and European levels"

In terms of objectives and roles this implies contributions in the following areas:

◗ Agreeing and implementing appropriate standards

and methods - this extends to cooperation with international organisations and seeking necessary new European legislation.

◗ Planning, monitoring and evaluating the ESS work programme with particular attention to coordination of the overall network and the meeting of quality targets

◗ Providing the European institutions and Member States with the information needed to implement, monitor and evaluate Community policies

◗ Disseminating our statistics to the European public at large including business enterprises

B.3 Outputs

The general notion of product - result of a production process - applies to the ESS without major difficulty. If considered as a whole, the ESS produces mainly **public statistics**, i.e. quantitative assessments of phenomena to which some explanations and comments are added and **analyses**, i.e. interpretation of the quantitative assessments using general accepted methods (demographic analysis, economic analysis, etc.). The extent to which Statistical Offices engage in analytical work varies a great deal between the ESS partners. Both statistics and accompanying analysis are widely disseminated, free or for a price far below the cost of production.

Answers to ad-hoc questions, which are the services delivered by the units of the dissemination department in the Statistical Offices of the ESS when answering questions raised by users (for example datashops) are in our analysis limited to the service of translating users' needs into statistical terms. The subsequent service - transmitting statistics and / or analyses for the request is attached to the corresponding statistic. For example, the transmission of existing statistics or the production of new tables on unemployment from existing data is considered to be part of the product "unemployment statistics". Therefore, there is not necessarily one-to-one correspondence between products and units of the Statistical Offices.

Large scale services: this covers specific production

activities that may be requested by a particular user from the public administration (ministry for example) and that are not necessarily widely disseminated. In addition, this generally supposes that this kind of user provides the ESS with a specific budget. This may involve the ESS as a whole, as very often the so-called partnership with Commission's DGs or other agencies, and relate to service deliveries like complete organisation of a specific survey, production of a database, methodological support, statistical or even support to economic analysis, production of specific publications not needed directly by the other users of the ESS (example: 1995 Internal Market survey and database). This may as well involve only one Statistical Office[1] of the ESS.

In addition to these final products, the ESS produces some products mainly for its own intermediate consumption in the production process of statistics and economic analyses:

a) *norms relating to concepts*: they refer to the definition of statistical units, classifications and characteristics and are generally enforced by EU and national legal acts. Examples are the definition of statistical units for enterprise statistics, NACE Rev.1, and definition of characteristics for Extrastat and Intrastat (definition of the "statistical value" of goods). Systems of accounts like ESA can be considered as a system of norms of the three types above to which are added norms for production (see below).

b) *norms for production:* these are generally purely recommendations (subsidiarity rules) and have generally have no legal status. A typical example is the methodological manual on business short term indicators.

c) *registers*: they are essential for probability survey sampling and co-ordination of surveys.

d) *training and exchange of know-how*: services delivered by the TES program or other training departments of the Statistical Offices, technical public meetings organised by the Statistical Offices, R&D activities in statistics (including the production of statistical packages). Example: organisation of TES courses, meetings in Eurostat on seasonal

adjustment, SUPCOM projects, DOSIS program, Methodological Conference of the ONS, CAL-MAR or CLAN package.

B.4   Customers

The traditional wording used by statisticians is neither clients nor customers, but users (see for instance "Statistics Canada policy for informing users"). Clients may be used if the financial dimension of the provision of service is important and if the delivered service is highly specific. However, most statistical products are delivered not on a tailored base, but via general media (standard publications or databases), and do not give exclusivity to any user, which is in fact against the ethics of public statisticians. Furthermore, the product is not consumed by the user and generally does not lose any value regardless of whether it is used immediately or later by other persons. Ignoring the copyright aspects, statistics can be used, stored in databases and reused (even re-sold) to other users. What is sold is generally not the information, but the right of access to information.

The users of the ESS can be classified into several categories:

◗ **institutions:** at the European level, at the national, regional or local level

◗ **media:** they are the main link to the general public and play a key role in the general image of the Statistical Offices.

◗ **social actors:** employers associations, trade unions

◗ **researchers, students**

◗ **businesses:** either for their own market research activities (for very big enterprises) or because their deliver some consultancy services in the information sector.

◗ **internal to the ESS:** Many derived statistics depend on results produced by other units (for the calculation of ratios or more complex indicators such as aggregated accounts). These users are not important in themselves, but play an important

---

[1] *we use here the wording "Statistical Office" to mean either one of the National Statistical Institutes, National Statistical Agencies, Statistical Departments of Ministries or Eurostat.*

role as nodes of the network connecting final users to each of the products.

B.5 Markets

Table 1 below gives an overview of a schematic description of the "market" for public statistics. It could be further broken down by domain (agriculture, services, social statistics, etc.) and by type of statistics (structural or short term). The setting of priorities raises in particular the issue of the diversity of stake holders in the ESS, in particular of the diversity of budget providers but also the diversity of objectives: a European (respectively national, regional) Office has more interest in European (respectively national, regional) aggregated statistics than in other types of statistics.

Fixing priorities becomes then a very sensitive issue. We have introduced in the table suggestion for Eurostat's own priorities. Aggregating priorities has always been a very difficult issue: the official procedure in the EU is mutual consensus and in case of failure, use of voting rules with qualified majority[2].

Also important is the distinction between direct users, who use statistics and analyses of the ESS by buying the publications directly from the ESS dissemination departments, and indirect users who use ESS statistics or analyses reproduced in other reports, in general databanks, or in statistics or economic analyses derived from primary products from the ESS but re-disseminated by other organisations (e.g. OECD or private companies). In the second case, it is of course much more difficult to know these users and to investigate their satisfaction levels.

We should also distinguish between current users and future users, between users that are important because they require an important share of the resources to be satisfied, and others because they can influence budget allocations.

[2] in the case of statistics, as soon as the Amsterdam Treaty is ratified. In the short term, there are still some domains with simple majority ruling (art. 213 of the Treaty).

Table 1: Existence and strategic importance of the users

Institutional Users:	COM	EMI	EU Parliament	Other EU Institutions[3]	Ministries	Parliament	Other national Institutions	Regions	Local authorities
Products	Europe	Europe	Europe	Europe	National	National	National	National	National
Statistics									
Statistics									
Answer to ad-hoc questions									
Large-scale services									

[3] Committee of Regions, Court of Accounts, Court of Justice, EIB...

Legend	Negligible	Weak	Medium	Important	Very important

Table 2: Existence and strategic importance of the users

Users: Social Actors	International Media	National Media	Regional Media	FEBI, FEBS	Other actors[4]	Information Centres	National Federations of businesses	Political parties
Level of competence	Europe	National	National	Europe	Europe	National	National	National
Products						Important	Medium	Medium
Statistics	Very important	Very important	Very important	Medium		Important	Medium	Medium
Analyses	Very important	Very important	Very important	Medium		Important	Medium	Medium
Answer to ad-hoc questions	Important	Important	Important	Very important		Very important	Very important	Medium
Large-scale services								

[4] Trade Unions, ...

Legend	Negligible	Weak	Medium	Important	Very important

Table 3: Existence and strategic importance of the users

Users: Others	Businesses exc. Banks, Insurance, Consult.	Consult. Co.	Banks, Insurance Co	Re-disseminators	Research, Education	General Public	Internal Users[5]
Products							
Statistics	Medium	Medium	Medium	Important	Weak	Medium	Medium
Analyses	Medium	Medium	Medium	Important	Weak	Medium	Weak
Answer to ad-hoc questions	Weak	Weak	Weak	Weak	Weak	Weak	
Large-scale services					Important		Important

[5] since many statistical products depend from each other, internal users should not be forgotten since final users are not aware of the importance of these intermediary products.

Legend	Negligible	Weak	Medium	Important	Very important

B.6 Features

In devising ways of ensuring that the ESS meets its goals, we must acknowledge special features of the ESS which can affect the applicability of management models or solutions, particularly where these have been derived primarily from private sector experience. In particular:

◗ The ESS lacks the unified command structure, decision making machinery and budgetary process normally associated with a single organisation.

◗ Its relationships with its main suppliers and main users are affected by the absence of pricing, charging or normal commercial markets for many of its data inputs and data products.

◗ Although useful initial progress has been made in clarifying concepts of data quality, there are many areas of ESS output where those concepts are not yet fully applied, and the level of understanding of these matters amongst users is often low.

◗ In many areas, the absence of external independent tests to judge the quality of data products (as might be applied for example to a car or to a PC) means that careful analysis of the total process chain producing the data becomes of particular importance.

◗ Given the complexity of roles played by Member States, the SPC, the Council and DGs of the Commission in terms of data suppliers, data users, competitors, influence on Eurostat budgets etc., any simple notion of independent suppliers and customers is clearly untenable.

I conclude from these features that the ESS is best viewed as a complex "network venture" of the type often referred to in management literature in the context of "virtual organisations" or "networked enterprises". The main message for all partners in the ESS network must be our inter-dependence. Our customers will have limited interest in whether data weaknesses or delays arise in one part of the system or another. The image of all component parts benefits if the final product is appreciated, and suffers if it is not. We must make sure that the very diversity of our combined European experience acts to our benefit, through a constant search for best practice and an insistence that this is spread widely without delay.

> ## THE EUROSTAT VIEW:
>
> ◗ We need more debate and agreement on what the ESS is and how its Mission should be formulated and interpreted
>
> ◗ We need a clearer understanding of the ESS's products and users, and a shared view on their relative importance
>
> ◗ Whilst the ESS structures do not reflect those of a normal commercial business, they are not unique and could be better understood through the growing management literature on "networked enterprises".

C. QUALITY CONCEPTS

C.1 EFQM model

In considering quality concepts appropriate for application across the ESS, we must recognise several factors:

◗ The notion of quality (total quality, TQM or whatever) is very broad, touching on all aspects of an organisation's activity.

◗ It is therefore particularly important to clarify terminology and concepts if debate is to be meaningful. "Going for quality" may mean better data precision to one participant, better project planning to another, fewer delays in results to a third, and so on.

◗ As with most management subjects, the field is littered with buzzwords of the month and gurus selling their latest books. We need an agreed distillation of what is relevant, helpful and high priority for the ESS before serious work can go ahead.

◗ Even where quality concepts are agreed, there may well be differing views on how these can best be actioned in each element of the partnership. Existing applications, for example, differ in whether the appointment of a Quality Manager is necessary, whether a management board member should have special responsibility for quality, the degree of reliance on consultant support, the value of self-assessment, the relevance of particular models, and so on.

Despite these challenges, it should be possible to agree quality concepts that will assist in improving ESS work, to

agree practical applications that will put them into effect, and to agree priorities for action. Taking as a starting point the approach of the European Foundation for Quality Management (EFQM), we can recognise eight dominant themes that pervade all quality management literature:

◗ The focus on customers. In the ESS context I regard this as vital, despite (perhaps because of) complications mentioned above regarding absence of charging and complexity of user relationships

◗ Supplier partnerships, in particular the need to build "trust and appropriate integration" with supplying agencies.

◗ The development and involvement of our staff, in particular the importance of communication and staff development.

◗ The systematic management of processes and facts, in particular for ESS better documentation of processes and better understanding of linkages and flows between MS and Eurostat.

◗ Continuous improvement and innovation, in particular ensuring a culture for the ESS that encourages original thought and improvement through benchmarking.

◗ Leaderships and consistency of purpose, in particular for ESS better decision making through the Statistical Policy Committee and Working Groups.

◗ Public responsibility

◗ Results orientation, in particular for the ESS tackling the substantial problems in the statistical field of measuring output and productivity.

In conjunction with these eight dominant themes, EFQM proposes a model for use as a framework for assessing organisational performance, either for a formal EFQM award or for self-assessment purposes. I have adopted this model (outlined in Annex A) as a starting point for the application of quality concepts to the ESS in the remainder of this note.

C.2 ESS adaptation

The EFQM model is designed for application to single organisations rather than to networked enterprises. In the light of experience with the model in Eurostat over the past year or so I suggest the following adaptations to tailor the model more to ESS circumstances:

◗ The box labelled People management (in Annex A)

should be interpreted instead as Network management ie handling the relationships between different elements making up the overall ESS.

◗ Similarly, the box labelled People (ie staff) satisfaction should be interpreted as Partner satisfaction ie whether the component elements are content with decision making processes, information flows, allocation of tasks and so on.

◗ Lastly, we need acknowledge that in the context of statistical work the box for Impact on society (covering impact other than through the output of statistics) is of only marginal importance and can for practical purposes be ignored.

THE EUROSTAT VIEW:

◗ There is an urgent need for a common vocabulary and framework for considering ways in which the ESS could be improved

◗ The quality approach in general, and EFQM model in particular provide a useful starting point, but these require adaptation to ESS circumstances

D. ESS STRENGHTS AND WEAKNESSES

Having confirmed the extent and nature of the ESS and outlined the quality concepts to be applied, it should be possible to attempt an initial assessment of ESS strengths and weaknesses. Suggestions are set out in Annex B as a basis for further debate.

THE EUROSTAT VIEW:

◗ We need a better grasp of the current strengths and weaknesses of the ESS, in order to plan effective improvements

◗ Its main strengths include the depth of knowledge and experience shared by its staff, the close liaison maintained through SPC, Working Groups etc and the legal framework available through European legislation

◗ Its main weaknesses include resource restrictions, weak formulation of overall strategy and ineffective evaluation of progress

E. QUALITY OF STATISTICAL OUTPUT

E.1 Quality of a product: definition

Quality is defined in the ISO 8402 - 1986 as:

"the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs".

We have already outlined above what are our statistical products and who are our users.

In general three approaches are possible to assess the quality level:

(i) conduct surveys about the satisfaction of customers

(ii) assess internally the quality, using criteria that reflect users needs

(ii) practice benchmarking, i.e. compare the production structure with other structure which have a good reputation as far as quality is concerned.

Approach i) is effective if it takes into account aspects such as: the capacity of users to assess quality, the relations between producers and users that could bias the collected opinions (many users do not pay a realistic price for statistics, and there might be conflicts of interest between some users and the statistical offices), the need to have a long term policy and to have a supply policy.

Approach ii) requires objectivity in conducting the self assessment, something is not necessarily easy to achieve. It depends as well on the criteria chosen for the self-assessment; these should reflect the determinants of on users' satisfaction.

Approach iii) is efficient if the entities with which comparisons are made have sufficient similarities with the activities of the European Statistical System.

Obviously, a combined use of the three methods above is recommended in order to reach a sufficient level of confidence in the assessment.

Since many statistical products are inter-dependent (surveys depend on registers, short term statistics need information drawn from structural statistics for the weighting, national accounts need a wide variety of all basic statistics), the quality of one product cannot be totally isolated from the quality of others. Users' satisfaction has to take into account indirect use of the corresponding statistics; quality measurement of a given statistic has to integrate the quality assessment of primary statistical products involved in the process.

E.2 Quality image

The quality of complex products is difficult to assess. This requires competence, time and access to information on the quality of the component products. Some users have then a vague and general idea about quality, sometimes not clearly related to objective information. This is quite common for goods and services in general, and is partly the domain of investigation of advertisers.

Many users may have no clear idea of the concept of quality for a statistical product, since the concept is very difficult to understand. The concept of accuracy is especially difficult to grasp, since by nature there are errors in all statistical products. These come in part from the sampling procedure, but also from the data collection phase where any single response could be erroneous, voluntarily or involuntarily. Again, the classical dichotomy between true/false is not adequate, and one rational tool to tackle this complexity is the probabilistic model. Adopting definitions as outlined in section 5.3.1 below, it is clear that similar problems occur with relevance (how can one single concept apply to all units of a population?), coherence (how can a user make his opinion when he faces incoherent figures emanating from different surveys and, how can he make his own opinion about comparability of statistics in the European Union?).

Therefore, what prevails in general is rather the vague quality image of the Statistical Office that produces the corresponding statistics. This is the important concept in practice. It means that any obvious error communicated to users who are not culturally close to the Statistical Community will probably damage the confidence in other statistics produced by the same Statistical Office (this applies as well to any source of information in general), unless the communication on quality is sufficiently well managed. So information on the quality of statistical products for users is a crucial element to secure the quality image.

E.3 Quality of statistics

E.3.1   A common definition of quality of statistics

Many organisations are involved in the definition of quality in statistics or in data in general. The ISO is devel-

oping new norms on quality of statistics particularly in the domain of Quality Control (Wilrich, 1997), MIT has launched the Total Data Quality Management program in 1992 (MIT, 1997), a new Data Quality Journal and a Newsletter (Data Quality News, 1997) delivered their first issues end of 1997. Statistical Offices must establish and maintain contacts with these think-tanks, particularly regarding data on socio-economic phenomena.

In our view, the quality of statistics can be defined with reference to several criteria:

◗ relevance of statistical concepts
◗ accuracy of estimates
◗ timeliness and punctuality
◗ accessibility and clarity of the information
◗ comparability of statistics
◗ coherence
◗ completeness.

Although not measures of quality, the resources available for the production of statistics and the form-filling burden placed on respondents act as constraints on quality. When assessing the ability of a Statistical Office to comply with quality guidelines, it is necessary to take these into account.

What do these components mean precisely for us?

Relevance

A statistic is relevant if it meets users' needs. The identification of users and their expectations is therefore necessary.

Accuracy of estimates

Accuracy is defined as the closeness between the estimated value and the (unknown) true population value. Assessing the accuracy of an estimate involves analysing the total error associated with the estimate.

Timeliness and punctuality:

Most users want up-to-date figures which are published frequently and on time at pre-established dates.

Accessibility and clarity of the information

Statistical data have most value when they are easily accessible by users, and are available in the forms users desire and are adequately documented. Assistance in using and interpreting the statistics should also be forthcoming from the providers.

Comparability of statistics

Statistics for a given characteristic have the greatest usefulness when they enable reliable comparisons of values taken by the characteristic across space and over time.

Coherence

When originating from a single source, statistics are coherent in that elementary concepts can be combined reliably in more complex ways. When originating from different sources, and in particular from statistical surveys of different frequencies, statistics are coherent insofar as they are based on common definitions, classifications and methodological standards. The messages that statistics convey to users will then be coherent, or at least will not contradict each other.

Completeness

Domains for which statistics are available should reflect the needs and priorities expressed by the users of the European Statistical System.

Clearly, completeness can only be assessed at the general level of the statistical system. Coherence and Comparability apply to groups of statistics. Only relevance, accuracy, timeliness and punctuality, accessibility and clarity of the information apply to one particular statistic.

There is a trade-off between the different components of quality, especially between timeliness and accuracy, accuracy and comparability through space[6], relevance and comparability over time, relevance and accuracy, coherence for large domains and relevance for sub-domains.

The above breakdown of quality into components is not unique. Other organisations have proposed other breakdowns for the joint use of statistics:

1) *United Nations (1983)*. Nine points overall, including:

◗ 'Comparability over time'

◗ 'Comparability with other statistics'

2) *Statistics Canada (1992).* Nine points overall, including:

◗ 'Comparability over time'

---

[6] *Because comparability through space may impose common concepts that are more difficult to measure than national - but non-comparable - concepts.*

- 'Comparability with data from other sources'

3) *Statistics Netherlands (1995).* Section II.2. on 'The need for coherence' lists seven aspects:

- additivity of statistics referring to different areas of economic activity, e.g. fixed capital formation in trade and manufacturing industry;

- comparability of statistics on different subjects, e.g. employment data from labour surveys and turnover statistics from production statistics;

- consistency of the outcomes of short-term statistics with those of annual statistics;

- consistency of the outcomes of regional statistics with those of national statistics;

- continuity of the outcomes for period t with those for period t+1;

- comparability of national statistics with international statistics;

- conformity of concepts and outcomes of business statistics with the System of National Accounts (SNA) and the European System of Accounts (ESA).

4) *Statistics Sweden (1994)*: four components are singled out: Content, Time, Reliability, Accessibility. Content includes comparability with other statistics and Time includes comparability over time. A very user-oriented approach with great importance attached to a quality declaration.

The following analysis of quality has been chosen by Eurostat. Since Eurostat is the Statistical Office of the European Commission, it has to attach much attention to comparability. Since all the technical work is co-ordinated by thematic committees or working groups, coherence has to receive specific attention because the co-ordination of the work achieved by different working groups is a challenge. Completeness emphasises the effort of providing statistics in an increasing number of domains, following the extension of competence of the European Union.

E.3.2 A common vocabulary for tackling quality:

In order to be in a position to measure the quality of statistics, we have to define without any ambiguity an EU glossary defining precisely all technical terms that

are necessary for the definition of quality measurement rules. For instance, terms like non-response, under-coverage, misclassification, variance, etc. have to be defined.

E.3.3 A common way to report on quality of statistics

In order to measure the quality of the statistics produced by the ESS, it is necessary to have at our disposal an objective and standard evaluation of all statistics produced by the ESS.

Each of the quality components has to be decomposed into observable and quantifiable items. When this quality report is available for all statistics, it will be possible to aggregate the information item by item, or statistic by statistic, by attaching to each survey and to each item a given weight.

This quality evaluation should be done by the Statistical Offices in close co-operation, and would be, in addition, aggregated at the ESS level. This requires of course an agreement on common guidelines for quality measurement, a structure for training, and a structured exchange of information.

The contents list for the proposed quality report, covering 62 items in all, is summarised in Annex C.

In practice, reporting on completeness makes sense only at the top level of each Statistical Office, i.e. the SPC in the case of the ESS.

E.3.4 The quality of other products

The second type of **final product** of the ESS is **analyses**. It could be assessed either by the satisfaction of users or by benchmarking. An objective self assessment of quality seems difficult to conduct.

As for ad-hoc questions, it is relatively easy to build an assessment system. Satisfaction surveys are easy to conduct provided there is a register of questions received, and benchmarking is quite possible. Objective assessment can be achieved by specific questions in order to check the efficiency of the staff and of the organisation.

For large-scale services, since the Statistical Office is in the traditional frame of a supplier/client relationship, quality evaluation is relatively easy.

The quality of **intermediate products** should also be assessed, since they are keys to the quality of the final

products, and since the ESS is not a single entity but a network of entities, they deserve much attention. Eurostat will put forward guidelines for example for self-evaluation of registers and norms in the coming months.

E.3.5 Improving the quality

Providing more relevant information

In a society with a high rate of change, statisticians have to adapt at a rapid pace. The necessity to harmonise existing statistics, budget restrictions and complaints about the burden on respondents represent major challenges to Statistical Offices. It should however be clear that before the EU statistical system is harmonised, it will have to adapt. The challenge is probably to design revision procedures for the regulations that are less time consuming, and that allow an efficient communication with partners (delegates of respondents like FEBI and FEBS).

Improving the accuracy of key statistics

This aspect of quality is often less visible and has a varying degree of importance for users. It is however crucial for public statistics for many reasons:

- public decisions have to be made on data that is as exact as possible

- public information is more likely to be checked by users than statistics that are not widely disseminated

- statistics produced by Statistical Offices are more likely to be checked for coherence by users, due to the extent of the public Statistical System and the number of different surveys including questions on very similar concepts, which raises indirectly the issue of accuracy.

Even if some users seem not to care much for accuracy when asked, especially when they would have to pay the price for improvements, none of them is ready to admit that they use statistics that are not reliable. For instance, many budgetary cuts in statistical institutes have been accompanied by quality initiatives in order to attempt to preserve their quality image. The ESS cannot afford the risk of publishing inaccurate statistics. In particular, all statistics used to monitor European policy, which are exposed to the political debate, have to be extremely reliable. The only way to keep rational debates at the European level is to base judgements on facts, and accurate statistics are the only tool to present objective facts in the political debate. Therefore, the ESS has to keep and

develop where necessary high standards of accuracy for its statistics and to make this accuracy visible.

Public Statistical Offices are the only organisations which can guarantee high levels of quality for socio-economic statistics. They are the only ones in a position to manage registers effectively, to limit non-response, to make coherent statistics by cross-checking data, to design efficient editing and imputation methods. Developing communication on these aspects justifies the use of public funding and reinforces considerably the quality image of the Statistical Offices. In an era of increased exchange of information, where more and more databases re-disseminate information without caring for the content, Statistical Offices have to indicate clearly this difference.

Progressing in all these directions deserve a special effort and an organised project: the ESS must embed this in a **general quality assurance project.**

Improving timeliness of statistics:

The ESS has a clear problem with timeliness: priorities and capacities of Member States are different; this leads to ESS statistics that are in general too late, especially for European aggregates. Considering the economic power of the EU, and especially after the implementation of the EMU, this has to be dramatically improved. More estimations will be done at the EU aggregate level. The procedures of data transfer between Statistical Offices have to be improved, and more estimation procedures have to be implemented at the national level for selected statistics of major importance.

Providing information that is more accessible and clearer

The ESS could improve its image in the EU by putting forward more analyses, adding value to the available information. We could still improve our communication structure, dealing better with language barriers and established circles of information.

Improving intra-EU and extra-EU comparability

Intra-EU comparability is very slow to implement, and substantial differences may stay in the long run if we do not act. When it is not possible to harmonise statistics totally, either for budgetary reasons or because it would increase the burden on respondents, some estimates of divergence should be elaborated using modelling and pilot studies. Extra-EU comparability has to be taken

more fully on board by the ESS. Eurostat or other Statistical Offices should do the necessary work and transmit the results to other Statistical Offices.

Improving coherence:

ESS Statistics arise from such complicated procedures that it is extremely difficult to achieve coherence when developing new projects. Clear improvement has to be achieved in different phases:

◗ improve communication on possible incoherence (be pro-active and not reactive)

◗ set-up systems for ex-post reconciliation of divergent important statistics

◗ in the long term, revise the agreed legislation in order to incorporate better coherence.

THE EUROSTAT VIEW:

◗ Over recent years useful progress has been made in establishing a framework within which we can judge the quality of national and international statistical products.

◗ We now need to gain wide acceptance of these proposals and to see them applied across all the ESS outputs

---

F.   Quality of the services

Improved data quality, as defined above, whilst clearly of paramount importance to ESS success, is far from the whole story. In the following section, making use of the EFQM model structure, I suggest other aspects of ESS performance which must be tackled if ESS is to meet its customers' needs.

F.1 Leadership

In the absence of a unified command and decision making structure associated with a single organisation, improved leadership for the ESS must be sought through:

◗ Better working of the SPC

➤ more effective use of "partnership" arrangements

➤ more emphasis on strategic issues

➤ fewer "tours de table" where there are no new

issues of substance to uncover

➤ better preparatory work in WGs

◗ Better working of Working Groups and Task Forces

➤ hold meetings only when they are really needed

➤ more attention paid to agreed strategy e.g. Five Year Programme

➤ more thorough preparation of material for SPC

◗ Clearer, agreed, guidance from ESS senior staff on quality issues

F.2 Network management

Staff commitment and motivation on ESS issues should be pursued through:

◗ Greater use of common training programmes

◗ More information on ESS developments e.g. through the Sigma publication.

◗ More open target-setting at national, unit and personal levels to encourage transparency on who is expected to do what.

◗ Staff attitude surveys that are co-ordinated across the ESS as a whole.

◗ Encouragement of networking e.g. common access to Cybernews and national equivalents

F.3 Policy and strategy

Attention to strategy orientation of ESS at all levels should be encouraged through:

◗ More reference to Five Year and annual programmes as day to day work tools

◗ Increased role for cost-benefit analysis by:

➤ routine availability of comprehensive ESS cost data (Eurostat, MS and data suppliers)

➤ agreement on how benefit assessments should be tackled

F.4 Resources (including staff)

Tackling ESS resource management across the entire ESS system is very much in its infancy. Improvements should

be sought through:

◗ Greater interchange of staff at all management levels

◗ More seminars and training courses offered on subjects of common ESS interest

◗ Sharing of experience on how analysis and documentation of statistical processes should be tackled

◗ Full ESS documentation of all key ESS processes

◗ Joint application of benchmarking activity to identify "best in class"

◗ Particular attention to statistical methods that have wide application across ESS work (methods relating to registers, selection, grossing for non-response, handling non-response bias, identification of outliers, disclosure testing and cell combination of results, presentation of results, electronic distribution, construction of index numbers, determination of trends and so on)

◗ Particular attention to data handling at the interface of MSs and Eurostat activities to eliminate unnecessary changes of medium, data formats etc.

F.5 Partner satisfaction

Little in concrete terms is known of the ways in which components of the ESS perceive each other, or how they perceive the barriers to wider ESS integration. Progress should be made on the following fronts:

◗ Surveys of satisfaction levels of ESS component organisations, concentrating on the SPC leadership role, the contribution of Working Groups and Task Forces, and the need for effective sharing of information.

◗ Evaluations of major ESS projects, past and current, to shed light on why some are more successful than others, and how these assessments are viewed differently in different MSs.

F.6 Customer satisfaction

According to recent critical reviews and surveys, a major area for improvement, to be achieved through:

◗ A shared view on who our "customers" are, and their relative importance to ESS success

◗ Joint surveys of customer satisfaction

◗ A renewed effort to make the CDIS and CEIES com-

mittee structures more effective

◗ More attention to common software for electronic dissemination in order that users have to learn one rather than many schemes

F.7 Business results

For reasons noted above, most NSIs and Eurostat face major difficulties in presenting "business results" in terms of measured value or volume of outputs, or overall financial results. Improvements should be sought through:

◗ More effort in producing comprehensive ESS accounts (revenue and costs by each work area)

◗ Shared effort on how output and productivity of statistical systems can be assessed.

THE EUROSTAT VIEW:

◗ In the eyes of many users the level of service supporting our figures is just as important as the data quality itself

◗ We therefore need information on how this service level is perceived that is regular and comprehensive

◗ There is scope for major improvements in the way we plan, authorise, monitor and evaluate ESS activities

G. EUROSTAT'S QUALITY INITIATIVE

In late 1996, Eurostat launched a Quality initiative, working with consultancy support. Inevitably, much of the early months was taken up with preliminary discussions and pilot testing of a range of approaches in order to judge what best suited Eurostat's needs. The next major task was preparation of a Corporate Plan outlining Eurostat's mission, confirming our strengths and weaknesses, and setting objectives and actions for the future. The resulting programme of work, to be implemented over the next two years, is outlined in Annex D. A brief summary of action in each main development area follows.

G.1 Statistical outputs

G.1.1   Statistics

As outlined above, Eurostat has agreed a definition of

quality of statistics, and launched a series of internal quality reports (see Annex C). It has appointed a Quality Manager to co-ordinate several programs that deal with quality of statistics

◗ In parallel, it has incorporated in the proposals for new pieces of legislation the mention of quality reports. This proposal has been accepted without difficulty by the Member States

◗ A Working Group on Quality of Business Statistics has been created, to make a technical proposal - to the preparatory SPC Committee on structural business statistics - on a definition of quality of business statistics, a glossary on quality of business statistics, a standard for quality reports both in structural and in short term statistics. There is an agreement on the proposal as long term objective, but the proposed implementation schedule has to be adapted to Member States' resources. Up to now, the internal decisions of Eurostat on quality of statistics are coherent with what has been discussed with Member States: Eurostat benefited in this domain from the efficient support of Statistics Sweden and the UK Office for National Statistics.

◗ A pilot project for the implementation of such a quality report was launched at end 1997 with Commission funding (SUP.COM 1997 budget) and is to be run by ONS, Statistics Sweden, and the Universities of Southampton and of Bath (UK). It will benefit from the support of Professor Skinner and Professor Draper.

◗ A model for simulations on quality optimisation, started in 1995, and run by Statistics Netherlands, is due to finish during the first half of 1998. Funded by the project SUP.COM 1995, this model sets a promising basis for a Total Survey Design tool.

◗ A first presentation of quality reports has been made to the Member States' delegates of the Working Groups "Statistics of Tourism" , "Statistics of Salaries and Labour Cost" and "Environment Statistics".

◗ The directors of Social Statistics of the ESS met in a traditional think tank session at Mondorf, in early 1998, and discussed the topic of quality in social statistics (Eurostat, 1998)

◗ A first TES course on "Quality of Statistics" is being prepared by Eurostat and Statistics Sweden and will be held in November 1998.

◗ An attempt to organise purely technical groups, on a voluntary basis, has been made in Seasonal Adjustment. It seems to work well and to prove that scale economics are possible within the ESS.

G.1.2 Other products:

Eurostat internal working group is tackling the problem of definition and measurement of the quality of statistical norms.

G.2 Service quality

G.2.1 Rolling reviews

A thorough review of each area of ongoing work is to be carried out, scheduled in order that each area is covered during the period of the 1998-2002 Five Year programme. Topics to be addressed are summarised in Annex E.

G.2.2 Customer satisfaction

This element covers four related aspects of customer satisfaction levels:

◗ Surveys of users in DGs of the Commission, to be conducted on a pilot basis in autumn 1998 and repeated in 2000 (to inform the 1998-2002 mid-term review) and in 2003 (to inform the final review).

◗ Surveys of Member States as data users, on the same timetable.

◗ Surveys of Member States as ESS partners, covering a wider variety of co-ordination issues (leadership, information flows, meetings etc.) on the same timetable.

◗ Conduct of individual meetings, as reflected in assessment sheets completed by those attending.

G.2.3 Cost-benefit

This element, to be overseen by the Cost-benefit Task Force, will be targeted on progress in three areas:

◗ Comprehensive ESS-wide cost data

◗ An improved understanding of how benefits should be identified and valued

◗ An insistence that cost-benefit methods and subsequent evaluation should be applied rigorously to all ESS new work

G.2.4 Project management

At present, Eurostat has no agreed guidelines on how major projects should be initiated, documented, costed, authorised, monitored or evaluated. The main output of this element will be guidelines on how this should be tackled.

G.2.5 Management aspects

The aim of this element will be to put in place:

◗ A management training programme covering new management styles sought throughout Eurostat to meet the requirements of the Quality initiative

◗ Training events covering technical aspects of the initiative: quality concepts, cost-benefit methods, project control etc.

◗ Regular surveys of staff attitudes.

G.2.6 Business plans

Unit business plans, known as PDUs (Plan de développement d'Unités) are being introduced across all Eurostat Units, covering the following aspects of business planning for each work area:

◗ Identification of customers and their needs

◗ Definition of operational objectives and targets

◗ Documentation of key processes

◗ Construction of appropriate work indicators

G.2.7 Indicators

Eurostat has identified indicators that will shed light on the extent to which each objective of its Corporate Plan is being met (Annex F). This element of the initiative involves the definition, setting up, data collection and reporting of these indicators.

THE EUROSTAT VIEW:

◗ The Eurostat Qualistat initiative offers a means of achieving improved performance across all aspects of Eurostat activity

◗ Eurostat is an important but relatively small element in the ESS as a whole, and the real benefits of a quality approach can only be secured through an extension of the initiative to the whole ESS

H. ESS ACTION PROGRAMME

The Qualistat initiative outlined above inevitably focuses on Eurostat rather than ESS-wide activity. I am keen that, wherever feasible, the initiative should be broadened to take an ESS perspective, and to encourage participation across all ESS component organisations. The following proposals show how this might be done.

H.1 Improving the quality of statistics

The ESS should:

◗ attach progressively a **quality report to all statistics** and intermediate products: this should be implemented progressively and be achieved around 2005 for statistics, starting by all statistics that have been harmonised. As for intermediate products, this should be tackled in order to prepare better the negotiations of international intermediate products at the world level.

◗ **progressively improve the quality** of statistics along an agreed plan in particular as far as timeliness, clarity, comparability and coherence are concerned. The systematic review planned for the monitoring of the next five years program should be accompanied by a proposal for specific quality improvements for all statistics, choosing for each statistic the components that deserve a specific effort.

◗ better **adapt to users' needs**: the surveys of user satisfaction, already conducted by some Member States, should be better co-ordinated with those envisaged at the European level, to be launched by Eurostat in 1998-9. This information is essential to secure the relevance of the different statistics and the completeness of the ESS.

◗ **develop more analyses** of the produced statistics, in a co-ordinated way: this is essential for the valorisation of the statistics that are available, but requires much

more timely statistics.

Many of these objectives require of course either more resources or a better organisation of work.

H.2 Improving the quality of the service

◗ The **Rolling review** exercises (Annex E) should be launched with an explicit remit to consider ESS-wide activity in agreed areas: staff members, costs, form filling burdens, and progress on major projects.

◗ The **cost-benefit work** should be treated as high priority in all component organisations, with a clear target of comprehensive, coherent, cost information to be assembled without delay. This should cover the official costs of conducting statistical work, and the respondent costs (form filling) placed on businesses.

◗ We must co-operate more in the agreement and application of **project management** procedures covering all major ESS projects. Considerable resource in the Statistical Offices, and particularly in the NSIs is devoted to work on data collection, editing, imputation, estimation of aggregates that is done in parallel between the 15 Member States. It is high time for **scale economies**, and for common development of programs and projects. Does it make sense to test independently in each NSI new packages, new computer organisations? Subsidiarity does not mean inefficiency: in this domain too, alternatives should be developed more rapidly: LEGS could be put in place for the methodological issues that are decentralised by Eurostat, but also for common centres of interest on NSIs. What Eurostat has tested with SUP.COM and DOSIS projects, i.e. the co-operation of institutes of different EU nationalities in the methodological and research domain could be extended to technical projects, on a voluntary basis by Statistical Offices sharing a common interest in a project and ready to fund this joint venture.

◗ We must also co-operate more in extending existing arrangements for joint training and methodological work. The aim must be to:

◗ improve the transfer of know-how between Member States. Clearly, in every domain one or a few statistical offices have more experience and more efficient methods and practice. Bringing the discussions in the ESS at a more technical level can achieve substantial improvements. Eurostat will assist in identifying these centres of excellence and in organising the exchanges.

It is however necessary to guarantee that these exchanges are of interest for all participants and that there is a strong will to learn and to put into practice new methods.

◗ **improve the training of Statistical Offices' staff** (statisticians, economists, data base managers and sociologists): clearly, training is an issue, especially for an activity that has very specific activities that are not shared by the market sector. An exchange of training practices, of training courses, using in particular the new possibilities (CDs, WWW servers of training courses) would be extremely efficient. Furthermore, it would encourage the development of language skills and develop further a common European statistical culture.

◗ support better research in statistics in the EU in order to reduce the delay behind North America: we have to state objectively that Europe has to catch up as well in this research domain. The ESPRIT research funds attributed to statistics give to the ESS a possibility to have some influence on the support of specific research in statistics. It would stimulate more efficient teaching, the development of better methods and tools, and improve the image of the ESS vis-à-vis the scientific community.

THE EUROSTAT VIEW:

◗ Our first general requirement is for a substantial debate involving all ESS contributors on what ESS quality improvements we can realistically expect to achieve in the medium term. I hope that Stockholm DGINS will provide a good launch for this.

◗ The second general requirement is for some formal apparatus to take the initiative forward. This could build on existing structures or be quite new - a small group answering initially to the SPC or to the Partnership group perhaps.

◗ Finally, we need make sure that progress is reported promptly to all who need to know, and that this is fully evaluated as a guide to future quality planning.

The challenge may be large and daunting, but conversion of the ESS into a high performance quality organisation will make the effort worthwhile.

References:

Working together as partners in European statistics - Proceedings of the 80th DGINS conference London May 1994

Strengthening of the European statistical system - Proceedings of the 81st DGINS conference Meersburg May 1995

EFQM Award Information brochure - European Foundation for Quality Management 1997

EFQM Self Assessment Guidelines for the Public Sector - European Foundation for Quality Management 1997

Benchmarking Next Steps Executive Agencies Dr Martin Samuels UK Cabinet Office 1997

Data Quality News (1997),
http://web.mit.edu/tdqm/www/intro.html

Eurostat (1996), 'Proposal for a quality report on structural business statistics', Document n° D3/Quality/96/04 final.

Eurostat (1998), Statistical harmonisation and quality, document for the Mondorf seminar.

MIT (1997), TDQM Research program,

http://web.mit.edu/tdqm/www/intro.html

Statistics Canada (1992). Policy on Informing Users of Data Quality and Methodology. Policy manual, Statistics Canada, April 7, 1992.

Statistics Netherlands (1995) Reference Manual on Design and Implementa-tion of Business Surveys. Koeijers, E. and Willeboordse, A. (editors). First draft, on request of Eurostat.

Statistics Sweden (1994) Quality definition and recommendations for quality declarations of official statistics (authors: Chris Denell, Eva Elvers and Bengt Rosén).

United Nations (1983) Guidelines for quality presentations that are prepared for users of statistics. Statistical Commission and Economic Commission of Europe, Conference of European Statisticians, Meeting on Statistical Methodology, 21-24 November 1983.

Wilrich P.T.(1997), The application of the international standard 5725 on accuracy of measurement methods and results, Proceedings of the ISI 97 Conference, tome LVII, book 2, 335-338.

ANNEX A

EFQM model

Note: percentages indicate relative contribution to overall performance

Leadership 10%

People management 9%

Policy and strategy 8%

Resources 9%

Processes 14%

People satisfaction 9%

Customer satisfaction 20%

Impact on society 6%

Business results 15%

Enablers 50%

Results 50%
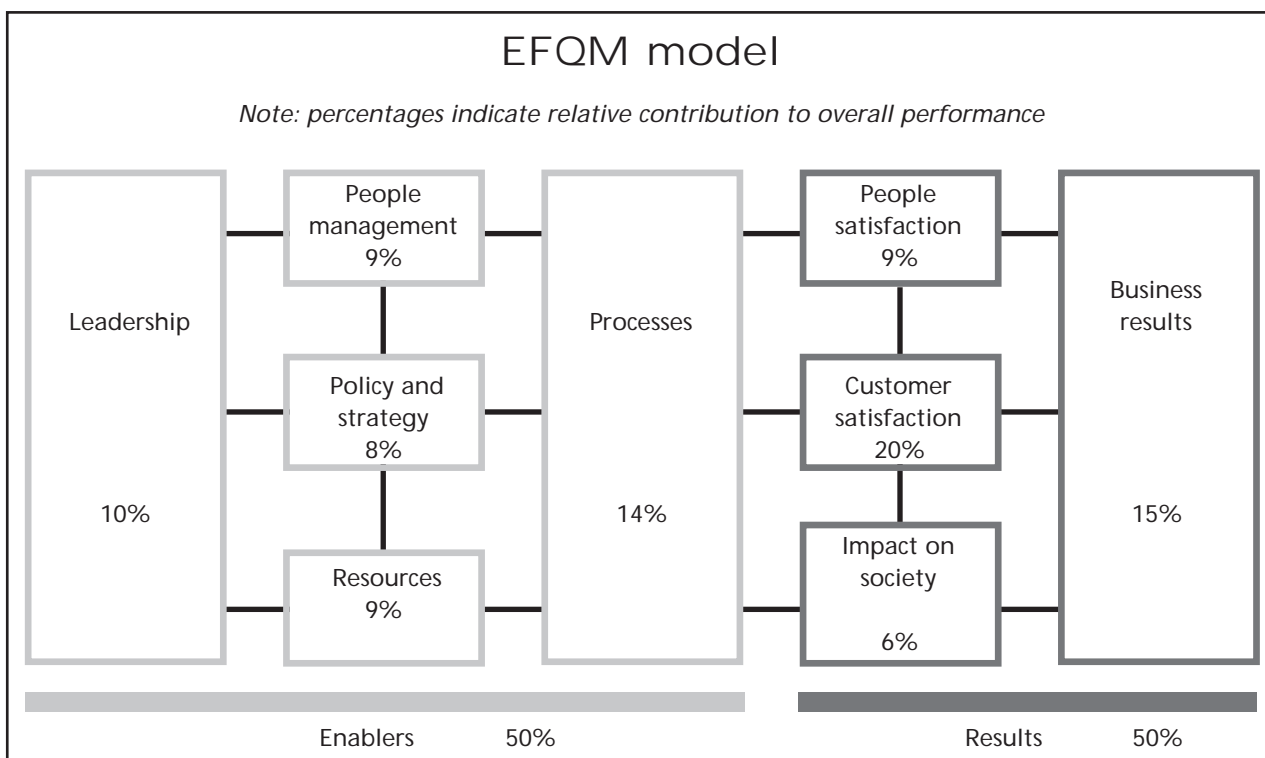
ANNEX B

ESS PERFORMANCE -
MAIN STRENGTHS AND WEAKNESSES

	Strengths	**Weaknesses**
1. Leadership	◗ Clear legal framework and comitology rules are in place ◗ SPC and DGINS meet regularly ; Partnership Group assists	◗ SPC and DGINS strategic roles should be strengthened ◗ No agreed approach to ESS strategy ◗ No agreed approach to quality issues
2. Policy and strategy	◗ 5 year work programmes are in place ◗ Annual work programmes are in place ◗ SPC and DGINS provides opportunities for sharing views on strategic issues	◗ Lack of ESS identity ◗ No clear shared vision of ESS future between partners ◗ Cost-benefit concepts are unclear ◗ Evaluation is patchy ◗ Direct DG funding (Partenariat) can distort priorities ◗ Little detailed knowledge in Eurostat of statistical work carried out in DGs ◗ NSIs are concerned that ESS is over-influenced by the Commission and that in some areas Commission needs are overstated
3. Network management	◗ Opportunities for frequent contact through meetings (SPC, DGINS, CDIS, CEIES etc) ◗ LEGs scheme	◗ Working groups and Task Forces could be reduced in number or used to better advantage ◗ Little is known of overall form filling burdens ◗ More input from DGs on data requirements and priorities would be welcome
4. Resources (staff, financial, information etc)	◗ END scheme ensures secondment of staff from MS to Eurostat (currently around 55 A grades) ◗ The TES scheme provides joint training opportunities ◗ Close liaison between MS and Eurostat through committees, seminars etc. ◗ Wide ranging experience and expertise across ESS staff	◗ Personal work programming and review arrangements are variable across ESS components ◗ Staff development provision is variable ◗ No regular staff attitude surveys ◗ A systematic approach to allocation of financial resources requires improved budgetary information from all partners ◗ Strict annuality in Commission budget process hinders resource allocation ◗ Total intellectual resource (e.g. on systems and methods) could be shared better ◗ Resources across the ESS are generally inadequate to meet all user needs

	Strengths	Weaknesses
5. Processes	▶ Advanced IT systems ▶ Reliable catalogues of source data exists for some domains e.g. national account	▶ Inadequate monitoring data for main processes and lack of clarity over success criteria ▶ Too little is known of overall data quality of ESS products ▶ No regular reviews of ongoing work ▶ Process documentation is partial ▶ Weaknesses at MS-Eurostat interface through lack of standards ▶ Scope for better project management ▶ Scope for more sharing of best practices
6. Partner satisfaction	▶ Image survey conducted 1997	▶ Absence of regular indicators on opinions of ESS partners
7. Customer satisfaction	▶ CDIS and CEIES meet regularly	▶ No co-ordinated survey of customer satisfaction levels ▶ Scope for closer contact with main users ▶ Doubts over effectiveness of CDIS
8. Business result	▶ None	▶ No comprehensive ESS cost or sales data ▶ No indicators or main performance fields eg timeliness ▶ Doubts on quality of some data inputs from MS

ANNEX C

CONTENTS OF QUALITY REPORT

1. RELEVANCE

1.1 Users description

1.2a Users' needs origin

1.2b Evaluation of Users' needs satisfaction

1.3 Relevance for users

2. ACCURACY

2.1. EVALUATION OF ERRORS

2.1.1. Sampling errors

     a) Use of a probability sample?
     b) Availability of detailed methodological note on the sample design
     e) Coefficient of variation of the totals
     f) Coefficient of variation of the corresponding growth rates
     g) Indicators of reliability for non probability sample?

2.1.2. Non sampling errors

2.1.2.1. Frame errors

     a) methodological note
     b) Summary on under-coverage/ over-coverage/ classification errors

2.1.2.2. Non response errors

     a) rates of total non response
     b) breakdown of non-responses into sub-categories (like refusals, not contacted etc.)
     c) patterns of non response
     d) Estimated effects on the reliability
     e) method(s) used for adjustments
     f) partial non-response rates
     g) references to methodological notes

2.1.2.3. Measurement errors

     a) errors due to the reporting unit
     b) errors due to the medium (questionnaire, interview)
     c) errors due to the interviewer
     d) reference to some methodological note

2.1.2.4. Processing errors

     a) data capture error rates
     b) codification error rates
     c) errors during the edit phase
     d) other errors
     e) existence of a detailed methodological note

2.1.2.5. Model assumptions

a) small areas estimation

b) seasonal adjustment

c) robust estimation

d) calibration

e) other

f) methodological note

2.2. OVERALL ACCURACY

a) qualitative judgement

b) methodological note

3. TIMELINESS AND PUNCTUALITY

a) total delay between reference period and publication

b) delay due to data collection

c) delays due to other production phases, broken down by phase (data capture, data editing, imputation, tabulation, validation, writing of official comments, preparation of dissemination support).

d) punctuality of the production.

4. ACCESSIBILITY AND CLARITY:

4.1. Accessibility

4.1.1. Forms of dissemination

a) Reference database of the Statistical Office(WWW)

b) Short publication (2 to 4 pages, like Statistics in focus)

c) Paper dissemination (normal collections)

d) Horizontal publications (Statistical yearbook, Overview of short term statistics...)

e) Others (CD-ROM, special publication, diskettes provided on request...).

4.1.2. Documentation

a) State other sources of information on the statistics and the production process.

b) existence of a methodological document available to users.

4.1.3. Clarity of the publication

a) existence of comments

b) analysis of the results

c) use of pedagogical tools (graphs, diagrams.....).

4.2. Information services

a) further help to understand and analyse the statistics available

b) provision of supplementary tables on paper or diskettes.

c) bibliography on advanced use of the information by economists

5. COMPARABILITY OF STATISTICS

5.1. Comparability over time

      a) brief description of the changes
      b) quantitative estimates of the changes.

5.2. Comparability over space

      a) references to descriptions of the difference of the - statistical concepts (units, characteristics, classifications)
      b) references to descriptions of the difference in the estimation process (sampling and non-sampling aspects)
      c) reference to existing quantitative assessment as well as to summaries of the differences.

      d) extra-EU comparability.

6. COHERENCE WITH OTHER STATISTICS

      a) between short term and structural statistics
      b) between business statistics and social statistics
      c) national accounts and survey statistics
      d) horizontal statistics and thematic/sectoral statistics.
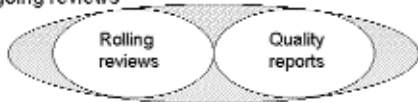

Note: in total 62 items

Note that reporting on completeness makes sense only at the top level of each Statistical Office, i.e. the SPC in the case of the ESS.
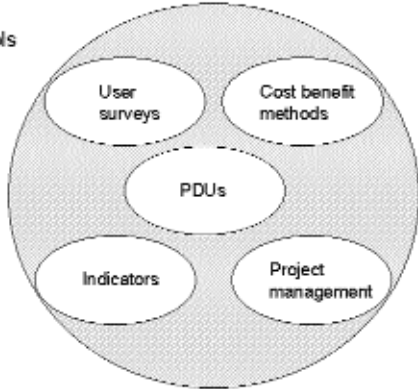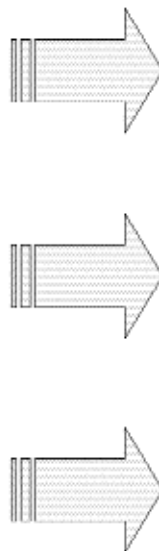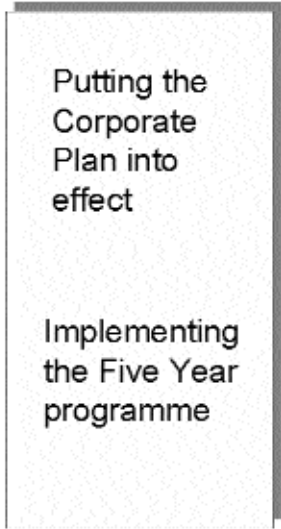
ANNEX D

ANNEX E

CONTENTS FOR ROLLING REVIEW

1. Background

◗ Organisation and staffing
◗ Requirement in 5 Year Programme and in recent annual programmes, and progress achieved
◗ Recent relevant indicators e.g. timeliness (MS delays and Eurostat delays)
◗ Recent reports on data quality
◗ Dissemination issues
◗ Relevant projects: objectives and progress against timetable and against budgets
◗ Conformity with Eurostat standards eg are computing arrangements in line with Eurostat strategies; is dissemination via the Reference data base?

2. Users

◗ Identification of key users
◗ Visits to key users
◗ Identification of uses to which figures are put
◗ Identification of benefits arising from use of data i.e. impact of improved decisions on criteria for programme evaluation
◗ User preferences for more data, different data, less data, more timely data, etc

3. Costs

◗ Definition of cost data that is available for recent past and projected future (coverage, treatment of internal services, computing, allocation of overheads etc.)
◗ Eurostat costs
◗ MS costs
◗ Responder form filling costs

4. Value for money

◗ Efficiency savings over recent years and projected
◗ Scope for savings through:
  - elimination (e.g. use of other sources)
  - better use of statistical methods (e.g. improved sampling)
  - modification (e.g. simpler data sets)
  - less detail (e.g. fewer breakdowns)
  - less frequent (e.g. quarterly to annual)
  - better use of computing facilities (e.g. selection of software)

5. Action

◗ Proposals for change
◗ Action programme and timetable
◗ Evaluation criteria
◗ Proposed evaluation data to check on progress

EXAMPLES OF POSSIBLE EUROSTAT INDICATORS

Output

Usage of data (accesses to reference data base)
Volume of press attributions
Data shops income

User satisfaction

Satisfaction levels measured in user surveys
Evaluation of meetings

Staff

Staff satisfaction survey results
Staff absence rates
Volume of training activity

Product quality

Contents of quality reports
Timeliness of data outputs (distinguishing delays in MS and delays in Eurostat)

Other

EFQM self assessment results
Reference data base coverage of data outputs
Reference data base timeliness

3.2.1 Discussion: Johan Allen: Statistical Office, Germany

Quality management : subsidiarity must take precedence over seemingly simple standard solutions.

1. Quality is a concern shared by Eurostat and all National Statistical Institutes. The quality of statistical results is crucial to the credibility of European statistics and hence to their image as supplier of the information on which EU policy decisions of far-reaching importance will be based. Despite the many levels at which the term "quality" can be understood, this statement has always been true of official statistics and is hardly a recent invention. Yet the fact that we are now taking an intensive look at quality is no accident. There are reasons for our concern about quality, and not only where Community statistics are concerned.

2. Freshness and quality deficits arise from the difficulties faced by Member States even when attempting to meet the existing legal requirements with stagnating or dwindling resources. The efficiency gains achieved through the use of modern data-processing and information technologies should not be overestimated. The freshness of the data disseminated ultimately depends on the resources and capacity of those producing the data. However comprehensive the technology, it is no substitute for a constant and increasingly stringent setting of priorities within European statistics in an attempt to meet the ever-increasing requirements of the European Commission. The data requirements laid down in Community legislation must take far greater account than in the past of the Member States' ability to implement them. The Multiannual Statistical Programme 1998-2002 offers an initial opportunity for this: it must be implemented consistently and culminate in a review of existing legislation.

3. There are objections in both principle and practice to the transferability of Eurostat's quality project to the whole of the ESS in the form of a "networked enterprise". Our joint efforts to achieve high-quality statistical results must not turn into *de facto* supervision of the NSIs by Eurostat. This applies even if the ESS is having to compete to satisfy Europe's demand for information, as Mr Franchet suggests in his paper.

4. The analogy of the "networked enterprise" seems to me too technocratic and simplistic. Because of the complexity and the independence of the partners involved, the transfer of management concepts to the whole of European statistics can succeed only in certain selected aspects and if responsibility rests with the Member States ("bottom up"). The seemingly straightforward standardisation of survey techniques or processing procedures must be questioned in the light of subsidiarity considerations and cost-benefit aspects and, in most statistical domains, is either methodologically impossible or pointless.

5. The subsidiarity principle is one of the cornerstones of our European Statistical System and gives the NSIs a central role. Neither "quality improvement" nor the technical simplification of procedures and methods in ESS practice constitutes adequate grounds for jeopardising it.

6. On behalf of German official statistics I readily accept the invitation to work with Eurostat on the quality of European statistics. We might, for instance, consider intensifying the exchanges of experience which have already begun between Eurostat and some NSIs (such as the Office for National Statistics) not only by documenting the quality concepts applied in official statistics, but also through specific benchmarking projects aimed at transferring and adopting best practices in certain areas. Mr Franchet's paper contains some essential contributions to the clarification of concepts and gives valuable insights into Eurostat's experience.

7. The choice of quality management concepts for a Statistical Institute should hinge on the experience of Eurostat and the other NSIs. The Federal

Statistical Office is making quality management efforts at various levels. In collaboration with the Statistical Offices of the Länder we are continually examining and improving the freshness of Federal data as part of deadline monitoring; an early warning system gives the alert at a stage when rapid intervention can still help. An ideal model for the Federal Statistical Office has been built up in a multi-stage process which will involve all staff. The Office's personnel policy stresses individual responsibility and teamwork as well as fostering a wide range of training opportunities.

8. The quality reports suggested by Eurostat are a means to an end and must not become an end in themselves. Their scale and degree of detail must be examined for cost-effectiveness in each and every case and they should if necessary be reduced to a manageable size, which may differ from one statistical domain to the next, the main consideration always being the purpose of the reports and the possible consequences of different levels of quality. Here, too, subsidiarity must take precedence over seemingly simple standard solutions.

9. The idea of stronger user-orientation deserves our firm support. The existing procedure of automatically handing over to the ESS the task of fulfilling requests from users who are often remote from statistical theory and practice is extremely unsatisfactory, however. The ESS must regard its service role increasingly as one of consultancy and be more actively aware of its methodological competence than used to be the case. This means that the information demand of Commission and other user groups arising from policy objectives must first be translated into statistical concepts. If the demand is to be met efficiently and quickly, the existing data supply must form the basis for the statistical consultancy task of the ESS. The priority here must be given to ex-post harmonisation of the data available in the Member States. The National Statistical Institutes must be included from the start in any discussions with users, including those at European level.