

The International Conference on Quality in Official Statistics, Q2001

An international conference on quality issues in official statistics, Q2001, was held in Stockholm on May 14-15, 2001. The conference was jointly organised by Statistics Sweden and Eurostat. During the conference the results of the Leadership Group (LEG) on Quality were presented and discussed. The LEG was led by Statistics Sweden and its members came from eight National Statistical Institutes and Eurostat, and its mandate was to investigate how the European Statistical System and its components can be improved through Quality Management and other efforts.

A conference invitation was extended to the survey community at large to assure that also other work of relevance to the LEG would be presented at the conference. The conference attracted some 350 participants from 37 countries and consisted of 39 sessions where invited and contributed papers were presented. It was evident from the presentations that Quality Management is attracting more and more interest in statistical agencies around the world and that there is a lot to be learned from experiences others have made. This CD contains almost all of the papers presented at the conference.

Lars Lyberg, Ph.D.
Chair, LEG and Q2001
Statistics Sweden



Stockholm, Sweden

Quality Conference, May 14–15 2001

Abstracts

Session 2, The Auditorium

10:20 LEG Chapter on Quality Framework

by Marta Haworth, ONS, U.K.

The purpose of this chapter is to set out the framework for the European Statistical System (ESS) and to lay a foundation for the following Chapters of the Leadership Group's Report which deals with specific aspects of quality management the ESS in more detail. The Chapter discusses concepts and definitions relevant to quality and quality management. It discusses total quality management, process and data quality management and also quality measurement and reporting. It also discusses how model frameworks such as the European Foundation for Quality Management Model (EFQM) which can be used in statistical organisations and what benefits this approach brings about.

The Chapter presents the Leadership Group's vision of the framework for managing quality in the ESS in future and discusses key components of it: user need assessment, focus on data quality, the attributes of data quality list developed for the ESS, addressing user needs and priorities, assessment of quality standards reached and dissemination of results and metadata and user support. The vision is set against summary information on the frameworks which currently operate within the ESS and which were surveyed by the Leadership Group in the course of its work.

Continuous improvement in statistics and services, measurement of dataquality standards against the criteria of relevance, accuracy, timeliness and punctuality in disseminating results, accessibility and clarity of results, comparability, coherence and completeness are also discussed.

10:40 LEG Chapter on Survey of Quality Practices in National Statistical Institutes

by Lilli Jappec, Eurostat and Statistics Sweden

As part of the work in the Leadership Group on Quality, interviews were carried out during 2000 with 16 National Statistical Institutes (NSI) to



Stockholm, Sweden

find out about quality practices used. This paper summarises the main findings from this survey. The paper will cover issues such as management models, quality reports, leadership and staff, customer/user orientation, strength and weaknesses of the European Statistical System (ESS), and methods used to ensure data quality in statistics. It will also discuss some of the problems that the ESS is faced with in its ambition to achieve high quality.

All NSIs have some kind of quality effort going on. Most of these efforts are conducted on an ad-hoc basis. Only a few NSIs have adopted a systematic approach. There is, however, a growing awareness that quality is something that we have to worry about and actively work on to achieve. It will not come automatically. Today eight NSIs within the ESS have adopted a management model for their quality work, but most of them have just started. Despite cultural differences within the ESS there are many areas in which the whole system would gain from more cooperation. The paper will discuss some of those areas.

Even in areas where NSIs and the ESS have a history of quality thinking such as data quality, there are many unsolved issues to address. In some areas there is already work going on, e.g., quality reports on the ESS level are underway. In other areas however, such as methodological handbooks or recommended practices, there is a great demand from the NSIs for further development. The results from the survey give some useful insights in which areas to start this work.

11:00 Statistical Institutes in the Modern Public Administration – What Might They Look Like?

by Jelke Bethlehem, Statistics Netherlands and Pentti Pietilä, Statistics Finland

NSIs are facing several major threats and opportunities in the future. The new and rapidly changing demand for information from customers and the ongoing large modernisation process in the Public Sector will give a pressure and possibilities. Real e-goverment is just to come as a demand and recent development in information communication technology offers new kind of tools to work. The shortage of labour and the shortage of new necessary competences can also cause severe limitations. To meet these things succesfully also NSIs have to reconsider the way they work. The introduction of a new mindset and the use of



Stockholm, Sweden

new management tools are very necessary tools to help organisations to transform themselves to more future oriented, responsive and customer focused. Key issues to success will be a new kind of leadership and strategic planning system, which are based on competence management and new leadership skills. Making organisations to learning institutions will be necessary to keep them flexible not forgetting empowerment and teamwork in releasing the full potential of staff. Also knowing much more of the needs of the customers and recognising real customer intimacy are critical success factors. NSIs could be great places to work, because challenging, international and modern IT-focused jobs are very popular among applicants. The process thinking instead of traditional stovepiping could be more used. Identifying, analysing and managing core processes and organising work around them would probably be the most demanding but the most useful operation in improving performance excellence. The wise use of modern information technology around core processes could change current systems in a drastic way. Anyway finally just the intercorrelation between different improvement areas will give the best fruits and improvements in performance results giving also better value for customers.

The paper presents some trendynamics to face. Also the paper presents some kind of a theoretical picture of a top quality NSI. The paper is not a scientific one. It is more an artistic drawing of the huge possibilities organisations have at their disposal.

Co-operation around these things and using common frameworks like Quality Awards and Common Assessment Frameworks might very strongly improve the performance excellence not used only technically but understanding the real meaning of the new mindset. Becoming a top quality organisation needs to be fulfilled huge efforts. Maybe 10-15 years goal oriented and permanent efforts like European quality winner 2000 in Public Sector, Inland Revenue Office, has shown. Many quality award winners, however, have said that the process has been the best investment ever done!



Stockholm, Sweden

Session 3 Room 456

10:20 LEG chapter on Dissemination of Information

by Werner Grünewald, Eurostat

Though not explicitly mentioned in the terms of references of the Leadership Expert Group (LEG) on quality, a crucial element of any kind of quality improvement in the ESS is a good management and dissemination of information. (Dissemination of) Information as such is not part of the definition of quality though there are close links. The paper deals with core issues of dissemination of information in the European Statistical System (ESS) in general.

Dissemination of information has a variety of dimensions with different importance for the ESS. The LEG decided to restrict its activities to core dimensions for the ESS, i.e. the level of information flows, the type of information and the organisation of information flows.

The results will be presented in the paper. They include an overview over the different dimensions and present partly new ideas on information as well as on their organisation. Another element is a list of current good practices from the EU Member States represented in the LEG that are recommended for further discussion and potential use in the Member States.

10:40 On Circulation of Information – Necessary but not Sufficient

by Donal Murphy, Central Statistical Office, Ireland¹

Eurostat's Leadership Expert Group (LEG) on Quality has identified six levels of information flows within the European Statistical System (ESS). The circulation of information on best quality practices and developments at each level is essential for the continued improvement of the quality. This is basically determined at the level of the national statistical organisations (NSOs) that produce the statistics. The infor-

¹ Former head of the Irish Central Statistics Office (CSO), 1991-2000



Stockholm, Sweden

mation flows involving them are, therefore, particularly important. These are considered individually. The circulation of information on quality practices (and other relevant statistical matters) is only a first step. Because of the volume now becoming available electronically and the danger of overload it is not sufficient to only circulate information or give access to its sources. An effective information management system is needed to identify what is relevant and to make it accessible to employees. What really matters is the enhancement of employee knowledge. A process of knowledge management, currently a buzzword in the business world, would facilitate this. Although the approach has been challenged statistical agencies would benefit from adopting a knowledge management strategy that focuses on providing a sharing learning work environment to extend the availability and use of existing knowledge and the creation of new knowledge at all employee levels.

11:00 Quality Web Site at the U.S. Census Bureau

by Deborah Stempowski, U.S. Bureau of the Census

As a part of the Quality program at the U.S. Census Bureau, we needed a vehicle to better communicate and share knowledge about quality throughout the organization. We saw a need to establish an intranet site that would provide users with the Census Bureau's quality standards, guidelines, and best practices while also providing a repository for documents that others could use in developing their quality components.

Over the last year, we have begun to build and expand this website. This presentation will provide a demonstration of the U.S. Census Bureau's website and provide some background about how we developed the site using Oracle WebDB. While it is a work in progress, we would welcome any suggestions and recommendations for improvements as well as discussions of how other conference participants are communicating their quality programs via the web or other means.



Stockholm, Sweden

Session 4 Room 357

10:20 Questionnaire Development, Adaptation and Assessment for ESS1

by Janet Harkness, ZUMA Mannheim, Germany

The paper presents procedures to be used in developing translations and assessing equivalencies for the projected European Social Survey, First Round (ESS1). The proposal submitted to the EU for the ESS sets out a seven-stage plan for developing and assessing the questionnaires needed to field the first round of the ESS in the proposed fifteen countries. These include statistical rounds of assessment. The paper presented here describes an organizational framework and the key text-based procedures required to produce quality products (good survey item translations) as well as text-based procedures to assess the quality of translations. The discussion is set within the context of a) the constraints of a tight methodological budget b) the short time available, c) the implications of the considerable cultural and linguistic diversity involved, d) complementation of statistical and text-based insights. Procedures considered include probe interviews, think-alouds, back translation, de-centring and committee discussions. Reference is also made to research on quality monitoring of adaptations and translations carried out for the International Social Survey Programme.

10:40 Predicting Response Quality with a Simulation Model

by S. Draisma, Department of Social Research Methodology
Vrije Universiteit Amsterdam

The quality of answers given to survey questions remains a hot issue. The simulation model RESPONSE (Dijkstra, Draisma & van der Zouwen, in: *Journal of Mathematical Sociology*, vol 20 (2-3), 1995) was developed to investigate, simulate and predict response errors in answers given to factual binary survey questions. The model estimates proportions correct, incorrect and non-substantial answers given to survey questions, and may thus be used as an evaluation tool for the quality of question wordings. In this paper, the usefulness of the model for the practice of



Stockholm, Sweden

survey research is examined. The degree to which RESPONSE generates better estimates of valid answers is the subject of the paper.

The conceptual model of RESPONSE consists of three question characteristics or input parameters. These are: the difference in *social desirability* of the response alternatives, the *certainty* a respondent has about the correct answer and the *salience* (or importance) of the question.

The simulation program offers several options dependent on the type of input: parameter values may be estimated, or the degree of response error, or the distribution of empirical answers, or the proportion 'valid' answers.

A procedure is developed to facilitate the use of RESPONSE in survey practice. Obtained answer distributions of binary survey questions, together with externally estimated parameter values, are used as input for the model. The simulated output, i.e., generated validity tables for binary questions, renders an estimation of the true score distribution of the sample of respondents that answered the question. This output should be a better estimate of the true score distribution over the answer alternatives than the obtained answers (empirical answer distribution). To obtain values on the question parameters, sessions of jury evaluations are organized, in which experts in the domain of survey research provide judgments on the three question characteristics. A regression procedure was used to calibrate jury evaluations of the pertinent question characteristics. These regressed jury evaluations, together with empirical answer distributions, are used as input for RESPONSE. The model estimates the true score distribution. This estimate can be compared with the obtained answers: RESPONSE forms a valid quality control tool if the estimated true score approaches the 'real' true score distribution more closely than the obtained answers.

For the application of the procedure, a database is required with 'key-questions' for which certain validating information is available. Regular survey questions may be compared with these key-questions to get an estimate of the size of response error. Information about the true score distribution of key-questions is necessary and was obtained in different ways. Firstly, questions were obtained in 'record check' studies, in which given answers can be compared to 'true scores'. Secondly, information is derived from the distribution of a variable in a population, covered by



Stockholm, Sweden

official statistics. Finally, additional information about the respondent sample, as well as special questioning techniques are used. The database can be used by survey researchers to compare their questions with the validated key-questions and have RESPONSE predict response error for a question with specific values on the question parameters.

11:00 Assessing the Quality of Questionnaires: A Comparison of Three Methods for the 'ex ante' Evaluation of Survey Questions

by Johannes van der Zouwen*, Willem E. Saris**, Stasja Draisma*
and William van der Veld**

(*Vrije Universiteit Amsterdam **University of Amsterdam)

The proper wording and formatting of questions included in a questionnaire for a survey, is crucial for obtaining data with sufficient quality. It is therefore necessary to assess and improve the adequacy of these questions.

Most of the existing evaluation procedures assess the adequacy of the questions after being posed in a pilot study, or even in the main study. Examples of these 'ex post' evaluations are debriefings of interviewers, analyses of the interactions between respondent and interviewer, and statistical analyses of the obtained responses searching for response effects and response bias. These ex post procedures may present good estimates of the adequacy of the questions, but this information often comes too late, that is, after the survey, or the pilot, has been held.

However, there also exists methods for the 'ex ante' evaluation of questionnaires, permitting the designer of the questionnaire to assess beforehand, that is before the questions are actually posed, the adequacy of the proposed questions, and to redesign the question if necessary. In this paper three methods for the 'ex ante' evaluation of drafted questions are described and their assessments compared.

The first method is the Survey Quality Predictor (SQP) developed by Saris and Van der Veld. A computer assisted expert system analyses characteristics of a question, like length of the question, number of response alternatives, response modes, and computes on the basis of



Stockholm, Sweden

these characteristics estimates for the reliability and validity of the responses to be obtained by this question. The information stored in the expert system comes from meta-analyses of MTMM-studies of the effects of question characteristics on reliability and validity.

The second method is a computer simulation model, RESPONSE, that computes on the basis of a few characteristics of the question, the relationship between the obtained response distribution and the true score distribution, that is, the degree to which the answers given are informative about the true scores of the respondents.

The third method is the Task Difficulty Score (TDS) consisting of 12 items, that relate to the degree of ambiguity and cognitive complexity of the task the respondent has to fulfil in responding to a particular question.

These three methods will each be applied to a common set of regular survey questions. These questions differ with respect to their topic, the degree to which response categories differ in social desirability, the format of the question, the number of response alternatives, the modes of responding, the use of show cards and memory aids, etc.

The outcomes of the assessments of these questions with each procedure will be compared and differences between these three evaluations will be explained. The three ex ante assessments will also be confronted with outcomes of ex post evaluations of the same, or equivalent, questions, using various indicators of data quality like response validity, reliability of the responses, completeness of the response, absence of response effects, and comparability of the responses given by different respondents.

The paper concludes with suggestions for rules for applying and combining different assessment procedures.



Stockholm, Sweden

11:20 **Validation of the Dutch Version of the Illegal Aliens Scale by Using the Three Step Test Interview**

by Kees van der Veer & Tony Hak, Vrije Universiteit Amsterdam,
The Netherlands

The 20-item Illegal Aliens Scale, which was developed by Ommundson and Larsen at Oregon State University, has recently been translated into Norwegian, Danish and Dutch. In this paper we discuss the results of a validation study of the Dutch version of this scale. The main method of investigation in this study was the Three-step test-interview which consists of the following steps, in that order:

- a. 'Think aloud' during the task, i.e. when completing the questionnaire
- b. 'Cognitive' interview after the task, i.e. after completion of the questionnaire
- c. 'In-depth' interview about the concepts measured by the questionnaire.

A range of results of this validation study will be discussed:

- problems regarding terminology – e.g. what is an 'illegal alien'?
 - problems regarding the formulation of items – e.g. double negation
 - problems regarding response categories – e.g. the meaning of 'uncertain'
 - response effects of different 'response strategies' used by respondents
- It will be discussed also whether and, if so, how these results could be used to revise the scale.

Session 5 Room 353

10:20 **Estimating Variance from a Completely Enumerated Stratum**

by Raoul Depoutot and Pascal Rivière, INSEE, France

Estimating variance is a complicated problem in business statistics, even if there is no sampling error. It is the case when data come from administrative sources, for example.



Q
2001

Stockholm, Sweden

The method of estimation of the variance obviously depends on imputation methodologies.

If we use mean imputes or hot deck imputes : the stratum has to be treated as a sampled stratum with the sample consisting of the non-imputed units. In that case the estimation of the variance is approximately correct and easy to compute. The idea is to act as if the subsample of respondents was the true sample.

In the regulation on structural business surveys, Eurostat proposes to use this sampling variance formula to approximate the variance, whatever the imputation methodology. It generally gives a raw proxy of the variance, except in the completely enumerated stratum.

In that part of the population, business statisticians do not use mean imputes nor hot-deck. Historical imputes (imputing n value by $n-1$ value of the same unit) are far better, even if there is a bias due to the evolution from one period to another. In that case, in theory, there is no variance as the imputation is purely deterministic. Such an approach is not satisfactory: intuitively, there is a variance.

To estimate a variance that is more conform to intuition, the idea is to assume a model for the variable of interest x , and to calculate the variance of **the difference between the estimated total and the true total**. This model-based approach is described in the Deville - Särndal methodology. The model can be an additive model or a multiplicative model.

In the paper, we analyse two kinds of imputes :

- historical imputes : $n-1$ value, for the same enterprise and the same variable
- historical ratio imputes : $n-1$ value, but this time multiplied by the growth rate of an auxiliary variable, for the same enterprise

In both cases, we computed the variance on the main sectors of economic activity in France, in a stratum defined as « the units with more than 50 employees».

We show that the «first approximation» (acting as if the respondents were a sample) strongly overestimates the variance. Why ? because in business statistics, the relative dispersion of quantitative variables in the



Q
2001

Stockholm, Sweden

take-all stratum (defined as above) can be extremely high: between 100 % and 1000 %.

The paper compares the variance estimates obtained with different kinds of models, and shows that it is necessary to define a model that explains the target variable if we want to have a good estimation of the variance.

Bibliography

DEVILLE, J.C., SÄRNDAL, C.E. (1994), Variance estimation for the regression imputed HT estimator, *Journal of official statistics*, 10, 381-394

10:40 An Access Panel for Official Statistics – How the Set of Methodological Tools May Be Further Developed

by Dr. Sabine Bechtold, Federal Statistical Office of Germany

An access panel could be an adequate means to enhance the efficiency of household surveys of official statistics that are conducted without compulsory response. The panel would serve as a basis for sampling. The access panel database would include the addresses and basic variables of households selected on a representative basis who are prepared to participate in official statistical surveys with voluntary response. An access panel of that kind would constitute itself a sample of the population which could in turn be used to draw samples for household surveys without compulsory response. The decisive characteristics of an access panel would be the following:

- It would exclusively cover households who are potentially prepared to provide information.
- regularly updated basic information would be available on those households.

As a result, an adequately structured access panel would have the following major advantages:

- The overall period required for a survey could be reduced since, based on an access panel, the recruitment of a given number of households prepared to provide information would take by far less time. This is in particular true for samples which are designed to cover only a specific



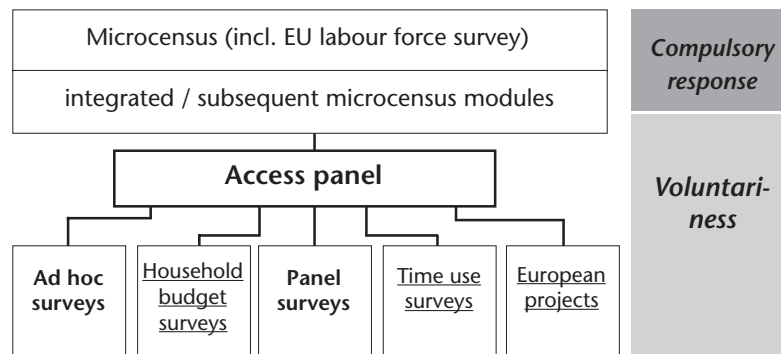
Q 2001

Stockholm, Sweden

group of the population (e.g. low-income earners). The basic variables stored would be available to directly contact households belonging to that group without a preceding survey being required to identify those households.

- The cost of surveys could be reduced since higher degrees of exhaustiveness could be reckoned with and, consequently, less households would have to be contacted.
- As stratification would be possible based on the variables stored, samples of a particularly high quality could be drawn by random sampling. That qualitative advantage would even be strengthened by the chance to make a comprehensive, statistical correction of the volunteer bias.

The preconditions for an access panel of official statistics to constitute an adequate basis for high-quality samples are that (1) such panel will be based on a random sample and (2) detailed information will be available on both the households participating and those not participating. An access panel constructed from the microcensus would fulfil the above preconditions. Here, households which have just completed their four-year maximum participation in the microcensus would be asked whether they would be prepared to take part in future surveys of official statistics. Besides, the households concerned would have to agree that, for the purpose specified, their addresses and variables previously inquired in the microcensus would be stored in the access panel database. The households recruited in this ways would constitute the access panel





Q
2001

Stockholm, Sweden

As part of an overall system, an access panel could serve as a basis for all sample surveys of households and individuals which do not require compulsory response and for which a sample size considerably smaller than that of the microcensus would be considered sufficient. The range of surveys could hence cover ad hoc surveys for special purposes of the Federation and the Länder, household budget and time use surveys, household panel surveys for national and European purposes and other European statistical projects. By defining basic variables, which would be part of any of the surveys, the results of the individual surveys could be matched in a consistent way. Ad hoc surveys regarding specific issues could then rely on a known data structure as a result of which their informative value would be by far higher than that of unrelated surveys.

11:00 **Building a Simulation Model to Compare Types of Sampling for Price Collection in Terms of Accuracy and Costs**

by Rosa Bernardini Papalia, Dipartimento di Scienze Statistiche, Università di Perugia, Italy and Cristina Cecconi, Istituto Nazionale di Statistica, Servizio Statistiche dei Prezzi

In this work, we propose and study a simulation model able to compare, in term of accuracy and costs, sampling designs in constructing consumer price indices.

Specifically, the paper deals with quality measurement in price collection. It is aimed at viewing the production of a consumer price index as a process composed by different interconnected sub-processes.

Each of these sub-processes gives its contribution to the global measure of quality according to the quality dimensions, the criteria and the indicators chosen to represent it collectively. Adopting this *process orientation*, for each sub-process a specific identification of quality dimensions is required.

In this context we focus our attention on the sampling design choice. Differences in sampling practices are analyzed by considering two



Stockholm, Sweden

sampling dimensions: (i) the item dimension, and (ii) the varieties dimension.

According to this perspective, it follows that the collection of data plays an important key role in the general framework of the computation of a consumer price index. From this point of view, differences in sampling designs can be regarded as possible sources of non comparability among the quality characteristics of the produced indexes.

In order to compare different types of items sampling methods, in terms of quality measurement of a process, a simulation model is proposed. It is tested that a quality estimation choice could imply a clear preference in the items sampling design choice in computing a consumer price index. We are also interested in analyzing the optimum sample size in term of the criteria used to determine the number of price quotations for different items.

Using scanner data, different sampling designs are considered: (i) simple random sampling, (ii) stratified sampling, (iii) systematic sampling and (iv) cut off sampling. We therefore analyze two quality dimensions: accuracy and costs.

Accuracy essentially means evaluating data ability to represent the phenomenon. In this context it is measured computing different standard errors (SE) of consumer price indices in different sampling designs with respect to: (i) fixed items sample size, and (ii) fixed varieties sample size. The idea is to choose the sampling design that gives the smallest SE of price indices.

For the second quality dimension, we refer to costs for price collection. We are interested in determining sample sizes in different sampling designs, with fixed standard errors values of consumer price indices, and in choosing the sampling design that produces the smallest sample size. With reference to the sampling dimensions considered, the sampling design with smallest number of items and varieties is chosen, for a given result in term of accuracy.

According to a process view, Systems Thinking and the *ithink* software provide a framework to compare different sampling design for price



Stockholm, Sweden

collection by simulating the alternative strategies, oriented to cost reduction and to accuracy improvement, here considered. The advantage concerns the possibility to represent continuous-improvement changes of sampling design methods and to evaluate the effectiveness of alternative criteria such as variance of the price indices and cost of the price collection.

Key Words: Sampling Designs; Consumer Price Indices; Systems Thinking; Scanner Data.

Session 6 Room 361

**10:20 Total Quality Management in Statistics
Finland – Experiences of Information Service
Unit**

by Reija Helenius and Pirjo Liewendahl, Statistics Finland

This paper describes the quality work experiences of the information service unit in Statistics Finland.

The quality work is a long and demanding task. It demands continuous improvement. In Statistics Finland the quality work began with Weststat consults in 1996. After that several quality tools have been experimented with and the experiment still continues. Understanding and realizing fully the concepts and the methods of the quality management system have demanded time and practice. When the quality work commenced in information service unit using and applying the quality methods were quite mechanical. However, now more emphasis is placed on the customer orientation.

Customer orientation has two important components. At first you have to be willing to determine customer's needs. Then you have to train your staff to adapt to these needs and apply them to the product development. In practise it means that you have to adapt your processes to the customers processes.

The quality work has to be systematical. In information service unit we have a quality work programme to ensure that each of our subunits have:



Stockholm, Sweden

1. their own business idea (why we do this service)
2. a strategy (where we are going)
3. a system of measuring (how we have managed in our strategy)
4. an education plan for the staff (to maintain the continuous improvement)
5. their own quality project (to improve everyday work).

All of these points are naturally based on the goals of Statistics Finland but to bring these goals to the grass roots level is not always so easy. This is a challenge for the management because strategic aims have to be concrete and forwardlooking at the same time. The educational plan is based on the research which we did in the summer of 2000. We pinpointed our main strengths and recognized areas in which we have to give special training to the staff.

110 people are employed in the information service unit and we have several subunits (for instance library of statistics, statistical databases, press office, customer training and regional services). Our units have enjoyed different levels of success in the application the quality work principle. For this reason motivation and inside marketing are important. The follow-up system which tells quickly and clearly how we have managed in attaining our goals is also significant.

In addition to the "official" quality programme we would like to develop the quality of everyday working life. Many details are involved how the customer experiences the quality. In the same way by developing these small elements, for instance working methods, it can be demonstrated to the staff that the quality work is usefull for everyone. For the everyday quality work we have written some common best practise lists for the meetings, time management, customer service standards and the working environment.

10:40 Some Quality Issues in Statistics Norway

by Hans Viggo Sæbø, Jan Byfuglien and Randi Johannessen,
Statistics Norway

Statistics Norway has like other National Statistical Institutes worked with quality for many years, but we are now about to start a systematic approach, inspired by the TQM principles and the work on this in



Stockholm, Sweden

several NSIs such as Statistics Sweden. But since our systematic quality work has just started, most of the paper is based on earlier experiences.

The paper considers some issues that have been given special attention in Statistics Norway during the past few years, or areas where we, to some extent, do things differently compared to most other NSIs. This applies particularly our user orientation, which is characterised by a large volume of free general dissemination of data on Internet for the general public directly or via media. For this purpose we have actively recruited journalists to work in Statistics Norway.

We have worked on the documentation of product quality in general over the past few years, including, amongst others, a system of systematic description of statistics on our web-service. A similar system exists in other Nordic countries.

Statistics Norway is one of few Statistical Institutes that has a relatively large research department. Hence, we will focus on the importance of internal research activity in order to maintain and improve several quality aspects in specific subject areas as well as our overall statistical products.

Management of quality issues should be closely linked to the overall system for management and reporting. This is especially important since there is an increasing need to focus on indicators for measuring overall performance both for internal reporting and follow up, as well as for external reporting. Some experiences in this direction will be reported.

The paper sums up the status so far and plans for our systematic quality work. Even if Statistics Norway has a relatively high standing and a lot of quality improvement work has been going on, there is a need for continuous improvements in user satisfaction, product quality and work efficiency, also due to cuts in governmental grants and increased market funding and competition. The organisation is ready for a more systematic quality approach, in which existing and planned quality activities will be included and adapted to a new framework.

User needs will be the point of departure for our systematic quality approach. Traditionally, we have concentrated quality work on product



Stockholm, Sweden

quality, often on accuracy. During the last years we have covered other dimensions of product quality (e.g. relevance, timeliness and punctuality, comparability and coherence, accessibility). Good product quality is necessary for good user satisfaction, but to obtain product quality at a reasonable cost process quality is important.

So far we have not worked much on process quality which will be an important aspect of the systematic approach. The area of CPI is an exception. In this field we have carried out a quality project which aims at a formal system for continuously measuring, monitoring and evaluating customers needs to improve the quality of the CPI. The improvement process encompasses quality of the CPI as a product as well as the underlying production process. The work covers development of quality measures like estimators for different biases, sampling variance, number of price quotations, number of missing prices, production time, punctuality etc. The quality work on the CPI is presented in the end of the paper.

11:00 **Quality Management in Statistics Denmark – A Pragmatic Approach**

by Lars Thygesen, Statistics Denmark

Statistics Denmark's new corporate Strategy 2005 gives high priority to quality in statistical products, and consequently to quality management. However we have decided not to embark on TQM which seems too ambitious for our situation.

Instead we define a limited number of *quality projects (QP)*, each focusing on one statistical area. The QP will endeavour to look into all aspects of quality of the statistics. The Danish quality concept is similar to Eurostat's definition of quality in statistics, based on five dimensions of quality of the products. A full set of quality declarations for all products is available on the Internet. Product quality is the result of process quality. In the QPs, all processes are examined and compared to standards for best practice, if established.

While the five dimensions are important, we recognize that quality ultimately depends of the fitness to use, i.e. the relationship between the product and users' needs. The paper will discuss how to engage users in the process.



Stockholm, Sweden

Quality is the responsibility of all staff, each being responsible within her own sphere. This has implications for the organisation of quality management. It may be a problem that quality projects tend to be imposed from the top of the organisation. How can we commit each staff member to quality work, how can we give incentives, how can we ensure that subject matter statisticians and methodologists feel they own the projects? It is believed that constant management attention to the task is one essential prerequisite.

The QPs result in a report to the management, containing a catalogue of the problems encountered, as well as suggestions of projects to remedy those problems. It is expected that crosscutting conclusions will emerge, leading to new standards for the production of statistics.

Session 7 Room 359

10:20 Counting and Estimation: Methodology for Improving the Quality of Censuses

by Stephen E. Fienberg, Department of Statistics and Center for Automated Learning and Discovery, Carnegie Mellon University, USA

In most countries census data are recognized as being fallible, i.e., subject to various errors of undercounting and erroneous enumeration. Part of the difficulty faced by statistical agencies is that the public tends to think of the census as an exercise in counting. When the census process is more properly viewed as one of estimation then statistical issues come to the fore, including bias-variance tradeoffs. This presentation will focus on the methodology used in the 1990 and 2000 U.S. decennial censuses for attempting to correct for census error through adjustment. In 1990, 1 in 10 residents of the U.S. were not properly counted. By the time of the conference we will know whether the 2000 census had a similar level of error and the extent to which adjustment appears to correct for the error. The presentation will include some of the highlights of the Census Bureau's evaluations of its Accuracy and Coverage Evaluation Survey and the resulting dual-systems adjusted data scheduled to be released at the block level in March, 2001. In particular it will include discussions of how one can examine empirically departures from the assumptions underlying the dual systems method. The presentation will also discuss some of the quality considerations surrounding the U.S. decision to



Stockholm, Sweden

adopt (or not adopt) the adjusted data, implications or censuses in other countries, and some related estimation problems growing out of the use of multiple lists.

10:40 The Use of Demographic Benchmarks to Ensure Census Data Quality

by Kirsten K. West and J. Gregory Robinson, Population Division, U.S. Census Bureau

The census of the United States enumerates its population where it legally resides on Census Day. The goal is to have an accurate and complete count. This goal is not easily achieved. Evaluation studies have indicated that both over- and undercounts have been associated with the censuses reaching many decades back. In this paper, we discuss how demographic benchmark analysis can be used to monitor coverage and data quality. Automation of data collection and data processing and high powered computers allow for this shift in focus to “real time” evaluations. If the analysis results flow back to management in a timely manner, remedial actions can be implemented to improve coverage and enhance data quality.

The use of demographic benchmarks to evaluate census data is not new to the U. S. Census Bureau. Demographic Analysis is a well known evaluation tool with a history that dates back to the 1950's. It has accompanied every census since then. Historically, the demographic approach measures census-to-census changes in coverage of the population. In recent years, plans have evolved for expansion of the program. The vision is to produce coverage estimates on a timely basis and to extend the scope of demographic coverage indicators below the national level. The vision also includes the use of demographic benchmarks, such as housing unit estimates, population estimates, sex and age ratios as tools to assess coverage early in the census process.

Experiences from the Census 2000 are used to illustrate the approach and its utility. Timely assessments of the address list using housing unit benchmarks developed independently of the census proved particularly useful to monitor the outcomes of census field operations. The better the address list, the better the chance of a complete coverage of housing units in the census and hence, the better the coverage of the population. In addition to the housing unit benchmarks, the population benchmarks



Q
2001

Stockholm, Sweden

were employed. The population benchmarks serve to ensure that data processing steps are applied appropriately and as intended preserving the quality of the data and its internal consistency. When the census totals become available, the population benchmarks also provide an assessment of the completeness of coverage and changes in coverage patterns from previous censuses.

Demographic benchmarks have been developed for high levels of geography such as the nation, the states and the county levels. As we move to lower levels of geography the benchmarks are either non-existent or of unknown reliability. Thus, to pinpoint areas in need of data review, it is necessary to develop benchmarks for lower levels of geography such as the tract level. We advocate that research be devoted to the development of demographic benchmarks for all levels of geograph

11:00 **Quality of Immunization Histories Reported in the National Immunization Survey**

by Meena Khare (National Center for Health Statistics),
Michael P. Battaglia (Abt Associates Inc),
Shannon Stokley (National Immunization Program),
Robert A. Wright (National Center for Health Statistics),
Vicki J. Huggins (Abt Associates Inc)

Keywords: RDD, telephone survey, vaccination coverage estimates, best value

The National Immunization Survey (NIS), sponsored by the Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services, measures vaccination coverage among children aged 19-35 months in the United States, in the 50 states, and in 28 large urban areas and monitors the progress towards the vaccinespecific coverage goals of Healthy People 2000 and 2010. NIS, a large on-going telephone survey, consists of a list-assisted randomdigit-dialing (RDD) telephone survey of the parents/guardians of the eligible children, followed by a mail survey of the children's immunization providers to obtain immunization histories from medical records. Currently, provider reported histories are used as the 'gold standard' to produce coverage estimates in the NIS.

In the 1999 NIS, household interviews were completed for 34,442 age-eligible children, and consent to contact the child's immunization



Stockholm, Sweden

provider(s) was obtained from the respondent for 28,936 (84%) children. A total of 22,521 (65%) children had 'usable' histories reported by the immunization provider(s). Approximately 83% of the 'usable' histories were obtained from single providers and the other 17% from two or more providers. Two types of dates are collected as part of the immunization histories from the household and the provider(s), which affect the vaccination coverage estimates: date of birth (DOB) and vaccination dates. For a small proportion (less than one percent) of children with 'usable' provider histories, discrepancies were observed in vaccination dates between the household- and the provider-reported dates and/or between dates reported by two or more providers for a single child. This paper provides a summary of the ongoing quality control procedures used to ensure the validity and accuracy of the immunization histories collected in the NIS and compares the quality of the immunization histories reported by the household respondents and the immunization providers. This paper also discusses a potential new method to determine the 'best' vaccination value by supplementing the provider information with the household-reported 'shot card' histories.

11:20 **The Statistical View**

by Isabel Bozzino, Director of the Institute of Statistics of Andalusia (IEA), Spain

We have just started a new century, as well as a new millennium, and what we have agreed on calling the post-modern paradigm is now close to be contemplated as an useless artificiality.

The penetration capacity that the post-modern paradigm has had in the history of the Sciences and of the Cultures has been, nevertheless, not only considerable but also beneficial. What post-modern theory and practice has taught us is that the prestigious "truth" is only – as Nietzsche established and liked to say – a rhetorical effect, more than institutional, for we always use language and, obviously, act within a certain frame of political-discursive conditions and assumptions. This way of thinking, called "weak" by some, against the previous and "strong" notion of truth as something undoubted, has made its way through with a lot of changes concerning geopolitical and socio-economic frameworks, living arrangements, family, population dynamics, and also ideologies.



Stockholm, Sweden

The role of Statistics, and especially of the Official statistics, in this context of changes is, at least, peculiar. Because if, on the one hand, she should give testimony of the changes that take place, on the other hand she obviates at least three things:

- 1) Their official character gives relevance to those social aspects that the official view considers outstanding while, possibly, she leaves others in dimness. For example, it is quite ironic that in order to be part of the activity rates from which the employment statistics are made it is necessary to have looked for work. This means that there is a wide rate of submerged activity composed by the real people that are not registered as active in the employment offices.
- 2) In the second place, there are certain realities for which, by virtue of their novelty, measure and analysis instruments have not been adapted yet.
- 3) And, finally, the own Statistics are far from being the paradigm of the scientific objectivity, since the answers, mainly in some fields in which the subjectivity is more present, are not so much objective as ideological. The statistics of the happiness, according to which it is considered that more than 70% of the people are happy, are, from this point of view, the significant ones.

Surprisingly, our “post-modern” condition has made us doubt the truth in such a way that today we can benefit from a viewpoint that makes us suspect but does not confuse the official consent with the truth of the things.

Therefore, and from that viewpoint of suspicion, so post-modern as it was said before, the question is how to make statistics of quality. I believe that the answer is that it is necessary to clean the statistical optics. This report will try to expose this idea to the light of some examples.

Session 8 The Auditorium

13:50 LEG chapter on Data Quality

by Håkan Linden, Eurostat, European Commission, L-2920 Luxembourg

Keywords. Quality concepts, quality reports, quality indicators, legislation's on quality, process quality



Stockholm, Sweden

In March 1999 the SPC decided to set up a Leadership Expert Group (LEG) on Quality with the main objective to improve quality within the European Statistical System. Many different aspects of quality have been tackled. This paper presents the key elements of the LEG writings on data quality.

The quality concepts used in statistical organisations have changed during the last decade, and the most dominant approach today is based on the ISO8402 norm from 1986 which states that quality is “the totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs”. Even if different statistical organisations use a different breakdown of quality into components, there is globally a very good convergence around the main themes. The Eurostat’s definition of quality is based on seven main components (relevance of statistical concepts, accuracy of estimates, timeliness and punctuality in disseminating results, accessibility and clarity of the information, comparability, coherence, and completeness) and one of the recommendations of the LEG is that this definition should be used for reporting product quality.

However, these dimensions are not easily measured. Some can be direct measures of error of sources (for example response bias) and indicators of process quality (for example response rate), but often one is forced to use proxy measures or qualitative assessments. This is due to lack of available methodology and software, and time and cost constraints.

The increased needs of survey comparability, comparability between National Statistical Institutes, and for better descriptions of the European Statistical System, stress the importance of standard sets of measures and indicators for reporting product quality. Work on such standardised quality reports is underway in several countries. Examples are the development of business survey reports for French official statistics, and Sweden has a rule stating that every survey should be accompanied by a quality declaration, and some surveys or survey systems in the U.S. have produced so called quality profiles. Eurostat has a Standard Quality Report, which has been discussed and agreed in the “Working Group on the Assessment of Quality in Statistics”.

Furthermore, there are different ways of implementing quality reports, and this depends most of the time on the purpose of the reports. For



Stockholm, Sweden

example, some NSI's use the reports internally as quality assessment tools, and others use short reports with main indicators for informing users about the quality of the results. However, it is becoming more and more common that Commission Regulations are concerned with the criteria for the evaluation of quality of specific subject domains. Two examples are the regulations on Structural Business Statistics and Labour Costs Statistics.

Thus, the starting point is to measure and report, but the quality of the output can only be guaranteed if there is good quality in all the underlying processes. By improving the process quality the product quality will follow. This presentation will also give a description of the underlying processes which have an impact on the product quality.

14:10 A Systematic Approach to Quality Measurements and Presentations

by Eva Elvers and Lennart Nordberg, Statistics Sweden

Many national statistical agencies have adopted a user-oriented quality concept. This means that the users should make quality assessments of the statistics they (intend to) use. To make this possible the producer has to supply quality declarations containing useful and user-friendly information. Even the presentation in itself can be difficult to make, since there are so many different kinds of users with very different needs. Obviously, the producer needs facts for the presentation and has to measure all quality components.

Measurement of quality is, of course, vital also for use within the agency. This is the only proper way by which the production processes can be evaluated and improved, using resources in a cost-effective way. It can be difficult and expensive to estimate some quality components such as accuracy, and in particular non-sampling errors.

The main issue addressed in this paper is how to meet all these different requirements in a systematic way. The suggested approach includes work on different "levels": At one extreme, a set of simple quality indicators is developed and measured recurrently (sampling variances, non-response rates etc). At the other extreme, deeper and more demanding studies are made on a non-recurrent basis of some of the



Stockholm, Sweden

more difficult quality components, such as e.g. measurement errors. The two extremes should be combined in a balanced way, with the indicators as potential warning signals and with regard to survey-specific characteristics.

This quality measurement work should result in a system of documents with different levels of detail, from short non-technical descriptions via more detailed declarations to technical or research reports. The Internet is very suitable to form links between these different sets of information.

14:30 **Managing Data Quality: The Accuracy Dimension**

by Gordon Brackstone, Statistics Canada

Under a definition of data quality that is based on *fitness for use*, the management of accuracy in a statistical agency becomes a necessary but not sufficient condition for the production of data of acceptable quality. Accurate data may not be useful if they are late, inaccessible, uninterpretable, incoherent, or on the wrong topic. The degree of effort and attention devoted to managing accuracy must be balanced against the importance of managing all the other dimensions of data quality that are equally necessary for data to be useful.

With that proviso, the paper will examine the essential elements of managing accuracy in a statistical agency. Some of these elements operate at the corporate level, and therefore have a potential impact across all programs in the agency, while others apply at the level of each individual program. Upper limits of achievable accuracy are, to a large extent, set when broad resource levels are initially allocated to programs, or subsequently revised. The process by which such resource allocations are made is therefore a crucial component of the management of accuracy. Within the budgetary limits set for individual statistical programs, the management of accuracy may be considered under the three broad headings of program design, implementation and monitoring, and accuracy assessment. The paper will describe approaches for managing accuracy in each of these three key activities, with illustrations drawn from Statistics Canada's experience. The rapid technological change that agencies are currently experiencing has several ramifications for accuracy management which will also be emphasised where they arise.



Stockholm, Sweden

Session 9 Room 456

13:50 **LEG chapter on Assessment Tools**

by Michel Blanc, INSEE, France and Max Booleman, Statistics Netherlands

The purpose of quality management is continuous improvement, formulated in clear objectives to achieving sustainable excellence in all aspects of performance. For making improvement programs it is important to know the weak points within an organisation. Measurements are necessary to get this knowledge.

Such measurements can be made through management models, like the one proposed for instance by the European Foundation for Quality Management (EFQM). This model distinguishes nine criteria. User satisfaction, Staff satisfaction, impact on society and business results (*results*) are achieved through leadership, policy and strategy, people management, resources and processes (*enablers*), leading ultimately to excellence.

This model, which can be used particularly for purposes of self-assessment, is applicable to all kinds of organisations. Some National Statistical Institutes have already adopted it. And a special framework, inspired by EFQM model and called CAF (Common Assessment Framework), has been developed for the public administration of the European Union countries.

Concerning the four 'result criteria', information about the effectiveness of the organisation can be got by measuring, that is constructing a set of indicators. And it is necessary also to straighten out the whole internal organisation, by asking questions about the five 'enabler criteria'. The LEG has specially considered two experiences, one about staff satisfaction, starting from the case of ONS (UK), the other about statistical audits, with the example of Statistics Netherlands.

As business results and satisfaction of users depend to a great extent on the motivation and the professionalism of the employees, it is important to measure the appreciation by the staff. A staff survey enables to obtain information about staff perceptions and attitudes to issues concerning their jobs, their line managers, their opportunities and prospects, the



Stockholm, Sweden

organisation as a whole, internal communication and their training and development. Of course you must be able to design action plans based on the results of the survey.

A statistical audit aims at evaluating the quality of the statistical production processes and the statistical results, and to propose ways of improvement. A number of questions must be addressed when carrying out such an audit: who are the auditors, how are they trained, which method do they use (for instance can they use some quality guidelines as a basis?), who is audited, which kind of report is produced, for whom is the report intended, etc.? At Statistics Netherlands there are audit teams and auditing is mainly an instrument for assisting statistics producing sectors with the evaluation of strengths and weaknesses and an aid when formulating proposals for improvements.

Apart from these examples, the presentation will attempt to review different assessment tools experienced in the European Statistical System, such as benchmarking, user satisfaction surveys or product quality indicators.

14:10 The Role of Peer Reviews in the Quality Management in Official Statistics in Switzerland

by Carlo Malaguerra, Swiss Federal Statistical Office, Neuchâtel

The paper describes the recent initiative of the Swiss Federal Statistical Office (SFSO) to have the Swiss statistical system assessed by two peers from another country. There was no acute crisis that made such an uncommon step unavoidable. But the SFSO felt that an independent expert view on whether the current structural features of the Swiss system are in line with or rather an impediment to an effective system of official statistics based on the UN fundamental principles, and whether this system is fit for meeting present and future needs, could form a solid foundation for some important decisions and initiatives to be taken soon.

The review was carried out between November 1999 and May 2000, and published in June 2000.

The review can be seen as one part of the whole work realising Total Quality Management (TQM), including the processes for continual



Stockholm, Sweden

improvement. This is an important contribution to the satisfaction of an organization's customer, personnel and all involved parties. The review qualifies the SFSO as professionally sound organisation, and its staff as competent. At the same time, there is scope for improvement with respect to many aspects of the statistical system and to some aspects of its main actor, the SFSO. The paper specifies recommendations how these shortcomings could be overcome, the main recommendation being that a national statistical agency with a much clearer status of independence, and in charge of official statistics at both federal and regional levels, be established.

The paper argues for peer reviews to become a regular feature for assessing the systemic elements of a statistical system, as a necessary complement to assessing the quality of individual statistics or of a range of statistics in a given subject area.

14:30 **Steps Towards a Framework for Assessing Data Quality**

by Carol S. Carson, Statistics Department, International Monetary Fund, USA

The paper describes the data quality assessment framework that is emerging in the IMF Statistics Department and aims to stimulate further discussion of the framework, the modalities of its application, and adjunct materials. The framework described synthesizes elements covering quality of the institution—the superstructure for producing and disseminating statistics—and quality of the individual statistical product. The framework comprises a generic assessment framework and specific assessment frameworks for the main aggregates used for macroeconomic analysis. The generic framework, which brings together the internationally accepted core principles/standards/or practices for official statistics, serves as the umbrella under which the dataset-specific quality assessment frameworks are developed. So far the work on the dataset-specific frameworks cover national accounts, balance of payments, analytical accounts of the central bank, government finance statistics, and produce prices. The paper will focus on the latest developments—e.g., new dataset-specific frameworks, experiences in field tests, consultations with nonstatisticians, and adjunct materials such as guidance notes.



Stockholm, Sweden

Session 10 Room 357

13:50 Quality Program at the U.S. Census Bureau

by Cheryl Landman (CASRO), Gia Donnalley (DIR) & Cynthia Clark(DIR), U.S Bureau of the Census

Quality has always been an important component of the U.S. Census Bureau's surveys and censuses. The Census Bureau's reputation largely depends upon the quality of its products and their value to users of its census and survey-based data. Many organizations are measured against our quality. A tactical component of the Census Bureau Strategic Plan is to "Establish a comprehensive program that builds excellence from the design through implementation and dissemination phases with built-in features for product and service assessments/evaluations." In support of the Bureau's strategic plan, the Methodology and Standards Directorate plans and directs statistical and methodological research and development for survey and census programs. Quality is a critical component to the success of our surveys, programs, products and services and is an integral part of the Directorate's mission.

Numerous areas at the U.S. Census Bureau have quality assurance and quality control programs and initiatives but there was a lack of sharing and lessons learned from these initiatives. Over the last year, the U.S. Census Bureau, as part of business process improvements, has implemented a Quality Program to coordinate and facilitate a corporate approach for quality.

The Quality Program functions as an umbrella project that relates the different efforts throughout the Bureau. This approach aids in identifying overlapping efforts, sharing information and filling in any gaps that exist. The Quality Program includes building excellence through innovative techniques and technologies, evaluations and improvements in our business processes, statistical research and usability studies as well as identifying measures and benchmarking standards for bureau-wide usage.

The goals of the Quality Program components are to design processes, establish standards, tools, checklists, etc. This is not planned to be an enforcement mechanism. Program areas monitor their own processes



Stockholm, Sweden

and programs to ensure compliance. The Methodology and Standards Directorate partners with program areas to meet their specific needs that are not covered by the generalized tools and processes. The Methodology and Standards Directorate provides training, develops tools, conducts research and sponsors seminars to educate employees about quality. The directorate developed a web site that will serve as the central source for quality information.

This presentation will give an overview of the U.S. Census Bureau's Quality Program and highlight several quality initiatives that are currently in progress or that have been completed. It will address how we are improving our Quality Program through the identification of current and best practices, sharing knowledge, and training components.

This presentation will also discuss the future plans as the U.S. Census Bureau continues to develop as a 'center of excellence'.

14:10 Quality Manual for a Statistical Agency: The Case of Finland

by Leena Hietaniemi, Statistics Finland

The current state of quality work at Statistics Finland will be described in the paper. Over the next few years the emphasis will shift to implementing into our daily work the methods of continuous quality improvement. The *Guidelines for Quality Improvement in Official Statistics* (Statistics Finland 2000), a manual recently produced for the quality improvement of statistics production processes, products and services, will be used extensively for this purpose. The goals, contents and production process of the manual will be described. We will also discuss the implementation of the quality goals included in the manual, and also the plans for the development of the manual. The manual will be distributed in paper format and it will be available in the intranet of the agency. The manual will be subject to periodic updating. A total of 40 experts of the agency have contributed to the preparation of the manual. The work has been co-ordinated by the Quality Guidance Team of the agency, a group including representation from different statistical and support units and top management of the agency, and also the so-called Quality Guides.



Stockholm, Sweden

14:30 **Quality Management in Statistics and Performance Indicators**

by Maria João Zilhão and Margarida Madaleno, National Statistical Institute, Portugal

Key words: Quality in statistics; Performance indicators; Quality management; Statistical production and dissemination.

This paper will focus the importance of performance indicators in Quality Management in Statistics.

Performance indicators are important in every organisation, as they provide relevant and updated information to support decision making by top and middle management

In a TQM and a process-oriented approach, performance indicators are used to evaluate and monitor efforts and results aiming at a better performance and efficiency. In statistical systems these indicators should not only be oriented by the different attributes which define Quality in statistical production and dissemination, but also include some information prepared for users, suppliers, employees and other.

An integrated set of performance indicators is an essential part of the management information system, but its implementation is a difficult task although crucial to generate improvements in the organisation. It is also expected to face some constraints and resistances. Some examples of performance indicators, which are used in INE-Portugal, will be presented, such as "Quality Indicators Panel", indicators used to measure and monitor the performance against our objectives and plans, and also indicators used for human resources and financial management.

14:50 **Introducing Total Quality Management (TQM) in Statistics Austria**

by Ewald Kutzenberger, Werner Holzer and Norbert Rainer, Statistics Austria

Like many other Statistical Institutes also Statistics Austria has started a



Stockholm, Sweden

comprehensive project to introduce TQM as central part of the new management and enterprise policy. The new Austrian Federal Statistics Law which entered into force on 1 January 2000, has set a new and modern framework under which official statistics is to be elaborated and disseminated in Austria now. Furthermore, this new legal framework has already set some of the important features of a TQM concept as a standard for Statistics Austria.

The TQM-project developed in Statistics Austria is based on the EFQM model and tries to profit from experience in Eurostat and in other Statistical Institutes as far as possible. Generally, the switch to a comprehensive quality management is viewed as a permanent process in which all staff members are to be involved. The concept centres around the usual five main areas: customer satisfaction, product quality, staff orientation, reduction of respondent burdens and efficiency.

While the implementation of the projects has started only recently, it is to be emphasised that it is our basic philosophy that all five areas of the TQM concept have to be dealt with simultaneously and none of the areas must be neglected at all. Only in that way a permanent process of changes and overall quality management can be launched and it is guaranteed that the targets set are well balanced. It is the aim of this contribution to present the TQM concept of Statistics Austria in more detail and to inform about the ongoing and planned activities and projects.

Session 11 Room 353

13:50 Estimating Sampling Errors for Movements in Business Surveys

by Susan Full and Daniel Lewis, Office for National Statistics, U.K.

Key words: sampling error; movements; jackknife.

Measuring and reporting on the quality of statistics is an important but complex task. One of the key measures of quality is accuracy, which has two key components: sampling and non-sampling errors. Sampling error is a direct measure of quality rather than just a quality indicator.



Stockholm, Sweden

Sampling errors of levels of key statistics are routinely calculated and published for ONS business surveys. However, it is intended to extend this to more statistics but to achieve this a number of issues have been identified that need to be addressed. These issues include estimating sampling errors for movements and the estimation of sampling errors for complex statistics that are often based on several survey sources. There are also the issues of the effects of non-response, outliers, misclassifications and births and deaths on sampling error estimates.

For many surveys and variables, it is the change between periods, rather than the level, which is of most interest. Estimating the accuracy of the estimated change between periods would be relatively straightforward if the target population and survey sample remained the same from one period to the next. However for ONS business surveys this is generally not the case. There are changes in the survey sample due to sample rotation and changes in both the population and survey sample due to births and deaths. There will be changes due to reclassifications both in terms of size and industry although the business register is managed so that the majority of these reclassifications are only introduced once during the year.

This paper reports on the progress made within ONS to develop suitable methods and tools to estimate the sampling errors of movements. Two approaches have been taken: the first using jackknife variance estimation and the second using Taylor linearisation via Statistics Sweden's CLAN software. The paper will discuss the data requirements for each method, assumptions made and the advantages and disadvantages of the two approaches. The methods implemented will be suitable for number-raised and ratio estimation as these are commonly used within business surveys. Further work will be necessary if these methods are to be implemented for other estimation methods. Some of the challenges of the project have been to develop methods and tools that can be applied across a wide range of surveys in order to minimise costs and to ensure consistency of sampling error measurement. This paper also explores some of the issues that have been addressed to ensure that the methodology implemented is understandable to both the survey statisticians and users of the statistics. Another issue is keeping the balance between implementing a statistically correct method and a method which makes some assumptions but can be implemented easily so that sampling errors of movements can be produced on a regular basis, alongside survey estimates, rather than as an annual process.



Stockholm, Sweden

14:10 Variance Estimation for the Scientific Use File of the German Microcensus

by Ulrich Rendtel, J.W. Goethe Universitaet, Frankfurt/M, Germany and Bernhard Schimpl-Neimanns, Zentrum fuer Umfragen, Methoden und Analysen (ZUMA), Mannheim, Germany

An aspect of quality of official statistical data that is increasingly considered important is their utility for social science research. In Germany, microdata from the microcensus are released to the scientific community in the form of so-called factually anonymous Scientific Use Files. However, usage of such files for the purpose of statistical analyses is not without problems. The paper's focus is on the problems of variance estimation, arising both from the specific sampling design of the German microcensus and from the procedure by which Scientific Use Files are generated from the original data. The problems are discussed in detail, and a practical solution is suggested, by which improved variance estimation can be attained.

For the first time the Scientific Use File of the microcensus 1996 provides anonymized sample information. However, not all relevant information is at hand. After a short presentation of the sample design of the microcensus and the selection procedure of the Scientific Use File, a solution is presented that uses the available design information in an efficient way. Following Särndal et al. (1992) the variance of population parameter estimates can be estimated for totals, means, and ratios. The variance estimates based on the Scientific Use File are compared with results from the Federal Statistical Office. Furthermore, we develop the regression estimator (group mean model) for the poststratification of the microcensus estimates. The results show that variance reduction by poststratification is not relevant for the Scientific Use File. However, large differences between the poststratified data and the original data pose the question which of the two population estimates is biased. We also investigate the behaviour of the linear regression of the design effect, which is frequently used as a makeshift for variance estimation. Generally, the variance estimation, by means of design effects, gives reasonably proper results. But in detail, we find both considerable over- and underestimation. Given the relevant design information, it is no longer necessary to rely on such coarse approximations.



Stockholm, Sweden

Therefore, we propose that variables for stratum, clustering, and poststratification should be released with the data if there are no confidentiality concerns. This would considerably improve variance estimations based on Scientific Use Files.

Key words: Microcensus, Variance Estimation, Factually Anonymous Data, Scientific Use File, Stratified Multistage Survey Data, Post-Stratification.

14:30 Has the Use of the Master Sample Contributed to Data Quality and Cost Efficiency? A Comparison Between the October Household Surveys of 1997 and 1999.

by Seble Worku, Statistics South Africa, Pretoria

Between 1994 and 1998, independent samples of households have been drawn for the annual October household surveys carried out by Statistics South Africa. For the first time in 1999, a new master sample design based on the 1996 census was used. In this paper, a comparison between the sample designs used prior to 1999 and the sample design of the master sample will be made. In particular, emphasis shall be placed on the contribution of the master sample to data quality and cost efficiency. Also comparison between the relative precision of the estimates of population characteristics obtained from the October Household Surveys of 1997 and 1999 will be made.

Session 12 Room 361

13:50 Enhancement of Statistics Provided by the Bank of Japan

by Kouji Kuwahara, Research and Statistics Department, Bank of Japan

The Bank of Japan collects and compiles various statistics for the purpose of researching financial and economic conditions. Among these statistics, the Research and Statistics Department (hereafter, the RSD) compiles and releases monetary statistics (Money Stock, Flow of Funds Accounts, etc.),



Stockholm, Sweden

price statistics (Wholesale Price Indexes, Corporate Service Price Indexes, etc.), and corporate statistics (*Short-term Economic Survey of Enterprises in Japan <Tankan>*). Based on the understandings that “statistics are public goods,” we are making every effort to provide “reliable statistics” as the statistics compilation section of the central bank.

The statistics compilation section of the RSD has released several explanations about the revisions and enhancements made to the statistics. In this statement, we present the basic stance of the RSD in terms of collecting, compiling, and releasing the statistics and summarize the revisions and enhancements of the statistics (refer to appendix 1 for further details).

14:10 Quality in Short-term Statistics: The Euroindicator Database of Eurostat

by Klaus Reeh, Eurostat and Francesca Nuges, ZSDA

The literature about quality in official statistics focuses mainly on survey statistics. This paper, however, raises a fairly new challenging issue: the measurement of data quality in time series statistics.

This paper describes in an introductory chapter the main features of a recently created Eurostat database: a collection of monthly and quarterly business cycle indicators for the Euro-zone, the European Union and the EU Member States.

A second chapter deals with the main public and private clients of this new database: the ECB and the ESCB, the business cycle analysis community with their researchers, advisers and commentators. Thereafter their needs for infra-annual statistics are derived from their respective missions and interests. Moreover, the likely emergence of new needs in this field is equally discussed.

A third chapter sketches the procedure currently in place for evaluating the quality (adequacy) of the database in view of the needs identified before. This evaluation procedure, built on a series of quality indicators, is activated every month. These indicators focus on aspects identified as important for the clients: relevance, timeliness, completeness, accuracy, length, availability, and/or clarity. The adequacy of these indicators is discussed and possible ways for improving them are explored.



Stockholm, Sweden

Moreover, Eurostat has recently launched its Qualistat initiative. This requires regular reporting on aspects of quality to all those involved and concerned. The reporting scheme for the different addressees on the basis of the evaluation procedure for the Euroindicator database is also outlined.

A final chapter will discuss the concept of quality for time series under various angles: quality of national series and EU aggregates; quality and flash estimates, quality and revisions, quality and data frequency, quality and different uses of the data, quality and data transformation,

14:30 **Quality Assurance by Benford's Law**

by Arne Bråten, Central Bank of Norway

In our work at the Statistics Department in the Central Bank of Norway we seek to maintain and improve the quality of our products. To ensure high quality statistics we are dependent on using efficient methods to guide us in the production process. It is important to have precise automated checks for plausible errors, in order to direct our attention and resources to observations with the highest probability of containing errors. As a natural part of our quality assurance we are investigating new statistical methods for this identification.

By using raw-data from our Balance of Payments system we will investigate the feasibility and benefits from introducing quality checks based on Benford's Law.

Benford's Law describes the probability of the first digit of a number to have a specific value. It applies to a very wide range of sources of numbers, e.g. the length of rivers, population counts, tax returns, atomic weights, and amounts on invoices and other accounting entries. The probability is given by the function;

$Pr(d) = \log(1 + (1/d))$, where $d=1, 2, \dots, 9$



Stockholm, Sweden

The following probabilities is derived from the formula:

Value of first digit	Probability
1	0,30103
2	0,17609
3	0,12494
4	0,09691
5	0,07918
6	0,06695
7	0,05799
8	0,05115
9	0,04576

It is counter-intuitively that there is a probability of 0,6 of the first digit being 1, 2 or 3 in such a wide variety of numbers. One would expect the probabilities to be more uniformly distributed. This fact has been utilized in several applications where the target has been to identify sets of transactions and entries with a high probability of being faulty or containing errors.

The first paper on this theory is by Newcomb, "Note on the Frequency of Use of the Different Digits in Natural Numbers", *The American Journal of Mathematics* (1881) 4, pp. 39-40. The same thoughts were described by Benford in 1938, "The Law of Anomalous Numbers" *Proceedings American Philosophical Society* 78, pp. 551-72.

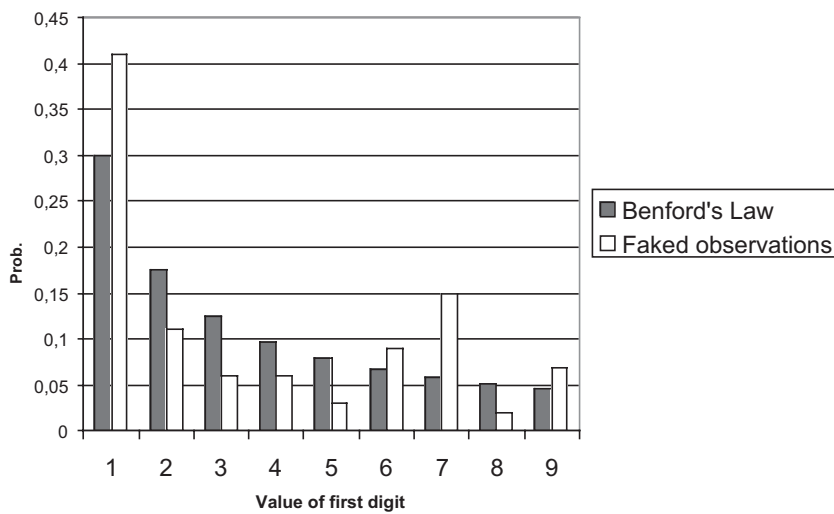
Our idea is to match the expected probabilities derived from Benford's Law with data reported to us by banks and other reporting units. We will then check how well the data conform to the distribution given by Benford's Law. A significant deviation will initiate further inspections (manually) of the involved transactions with an aim to improve the correctness.

Different levels of aggregation or sub-populations for the applicability of Benford's Law are proposed: across several reporting units at one point of time and across several points of time for one reporting unit.



Stockholm, Sweden

An example which would initiate further investigations, is given in the graph below:



A person in our staff produced a set of 100 numbers off her head, these data are used in the graph above. The difference between what we observe and what we would expect from Benford's Law is quite large. The high frequencies of digits 1 and 7 sticks out, suggesting something is suspicious. By applying standard statistical tests we can decide if the deviation from Benford's Law is significant. If so, we can conclude the string of numbers does not originate from a true source. In a real application further investigations would be conducted, which could result in either corrected numbers (there was identified an error) or an improved understanding of the material (there was not identified an error).

If the proposed quality check described above yields beneficial results in identifying errors correctly and not triggering false alarms in excess, we will consider it likely to recommend the use of Benford's Law in the process of compiling Balance of Payments statistics.



Stockholm, Sweden

Session 13 Room 359

13:50 Benchmarking System Development at INSEE

by Christophe Alviset, INSEE, France

System development at INSEE has undergone a series of transformations over the last 10 years in three key areas: organization, methods and technology. Organization has changed in 1990 and 1996 to ensure better project success, methods have been set up in 1990, evolving over the 10 year period and undergoing a review in 1999 to contribute to better choices, better quality and hopefully shorter projects. Technology has moved from mainframe to client-server to web-based and we have a variety of architectures to support.

There are currently 80-100 projects underway, 200-230 applications in production and 250 full-time equivalent busy with projects and maintenance. At the corporate level two questions do not receive a proper answer: are our projects taking too long? Are we reasonably on time and on budget in our projects?

These complex questions can be tentatively answered using a benchmark of our system development activities against that of similar organizations. We are in the process of choosing a service company to do this for us.

Measuring system development is not easy in that there are few tangible products and repetitive activities. However some progress has been made over the last 20 years and typical measures include the number of function points (a standardized measure of output of development and size of projects and applications) and the number of defects in addition to cost and duration. This in turn allows to have measures of productivity and efficiency. To be applied however, the very concepts of projects and applications must be standardized between organizations, including the delicate notions of when a project starts and when it ends.

Detailed analysis by project phase and by type of activity is however the real added-value of the benchmark as it pinpoints the areas for improvement. The benchmark will have three objectives:



Stockholm, Sweden

- to evaluate how we spend our time and money over one year: is our system development adequate and properly organized? Is the ratio of statistical means to system development means acceptable? How are we doing in outsourcing? How much maintenance and new developments do we do compared to similar organizations? What is the size of our portfolio of applications?
- to study a number of projects over their complete duration: size of the projects, duration and cost broken down by project phase, do we exceed time and cost budgets by a typical amount?
- to study the quality of applications: size, number of defects, response time, user satisfaction

From all of this will be derived recommendations on improvement. We hope in this fashion to gain a better understanding of our portfolio of applications and projects and have a new point of view on what areas need improvement.

The benchmark will unfold in the first quarter of 2001 and the early results will be presented at the conference.

14:10 Integrated Review System – Supporting Process Quality in U.S. Census 2000 Estimation Processing

by Randal ZuWallack, Susan L. Atha, and Philip M. Gbur, U.S. Census Bureau, Washington, DC 20233-7600

Key Words: Documentation; Official Statistics

The premier official statistics released by the United States Bureau of the Census are arguably the population counts from the decennial census. In an effort to support process quality in the U.S. Census 2000 estimation process, the Integrated Review System (IReS) reviews, compares, and documents intermediate results in real time. The IReS provides key information for analyzing the reasonableness of the results, thus ensuring proper implementation of the estimation process as well as providing a means for reviewing specific components of the process. The key information is at the macro level in the form of selected rates and distributions, which can be tracked through the various stages of the process. The tracking will be done separately on a geographic path and on a demographic characteristics path. In addition to the tracking methodology, another function of the review is to create summary files. These files will provide a history of the estimation process to allow



Stockholm, Sweden

readily available data for investigation of questions that arise during subsequent operations while producing documentation as a by-product of the process. Thus, the IReS will simultaneously support quality in the estimation process and generate documentation in support of the release of population counts from the U.S. Census 2000. This paper provides a brief overview of the Census 2000 estimation process, the design and implementation of the IReS, and an initial evaluation of the system's effectiveness.

14:30 INPS Statistical Data Bank – The Process of Normalisation for the Data in the Firms' Data Bank

by Antonietta Mundo and Natalia Orrù, National Institute for Social Security (INPS), Italy

INPS' objectives and strategies in the field of statistics are aimed at an ever wider use of the administrative managerial files, not only for cognitive purposes of the institutional kind, but also for informational needs of more general interest. The creation of wide-ranging statistical data banks is favoured; consultation of the same via internet is free and accessible to everyone.

Through its creation of Observatories on the world of work, our Institute has set itself the objective of offering an ongoing and permanent set of detailed and continually updated statistics, processed from entire universes of economic subjects; it will be able to supply a complete map of entrepreneurial Italy, the labour market, the employment structure, the migratory flow and the payment system, so as to offer a dynamic contribution to the task of evaluating the productive transformations in progress.

INPS Statistical Data Banks collect detailed data on the following phenomena: businesses, employees, domestic workers, self-employed workers, INPS pensions and INPS administrative management.

It is intended to examine the quality control procedure for the data utilised in one of the aforesaid Data Banks: the one dealing with entrepreneurial businesses, to be exact.



Stockholm, Sweden

The control technique utilised is based on the following axiom of economic theory: “under ideal conditions, employment in a productive unit shows a marked inertia over time“. The economic and operational reality of the businesses has led to the utilisation also of the criterion of “consistency over time“ of the number of employees in a firm: i.e., the constancy of the temporal covariance of the employment data. In particular, the four main steps in the process of normalising the data in question are as follows:

- 1) to calculate the employment percentage variations from one month to the next, identifying the variations that exceed the conjectural structure of temporal covariance;
- 2) to check whether it is possible to make the series of observations from a few successive months more coherent, not taking into account the high variation;
- 3) to consider data responsible for anomalous variations unacceptable;
- 4) to deal with unacceptable data considering them by linear interpolation or drag in the case of an end of series.

The previous stages of the normalisation process are developed by creating special indicators to record all the cases where it is necessary to improve the quality of the datum and by employing suitable statistical techniques to improve the quality of the datum deemed unacceptable.

The main indicators employed in the process are as follows:

- percentage rate of variation;
- breadth of interval on which the variation is calculated;
- size of firm;
- employment variation rates.

Other normalisation processes are also employed for payments.

14:50 The Optimal Use of Data Sets and the Quality of Indicators in a Process Oriented Statistical Office

by Max Booleman, Statistics Netherlands

Keywords: repeated weighting, quality index, numerical consistency



Stockholm, Sweden

Introduction

In Statistics Netherlands (SN) the statistical processes have undergone radical changes. Within the departments the divisions are responsible for a certain stage of the statistical process for a wide range of statistics. These stages start with the collection of data and conclude with the dissemination of the results. This clustering of similar activities, which used to take place within various departments, means Statistics Netherlands a more efficient way of working. Further efficiency gains are envisaged by obtaining access to a variety of administrative registrations containing relevant data for statistics. This has the added value of reducing the administrative burden of respondents. Statistics Netherlands has developed a design based strategy aiming full numerical consistency of its statistical estimates, based on a matched data file containing all available register and survey data about the target population. The key element of this estimation procedure is repeated recalibration of survey data sets, i.e. a new set of raising weights is derived each time additional estimates are being produced. This procedure will allow taking account of any related statistics obtained in earlier rounds of estimation. This paper presents an overview of the activities in the Department of Statistical Methods of Statistics Netherlands to develop product quality indicators based on such an estimation strategy.

Session 14 Room 461

13:50 **The Shortening of the Production Process in a Swedish Enterprise-based Survey¹**

by Heather Bergdahl, Statistics Sweden

Constant demands are put on statistical organisations to speed up the production process of economic and labour market statistics. The organisations would most probably be able meet these demands provided with sufficient extra resources. With given resources, however, the quality of the survey or the staff would suffer. The question arises: can the production process be sped up maintaining the quality level of the survey with given resources?

¹ This abstract refers to the report "The Shortening of the Production Process in the Short Term Employment Survey", SCB 2000, by Heather Bergdahl, Joakim Haak, Jan Hörngren and Sören Klippfjell.



Q 2001

Stockholm, Sweden

A project was set up in the fall of 1999 at Statistics Sweden in the department for Labour and Education Statistics to investigate if the production process of an enterprise-based survey could be sped up without risking the quality level of the survey. Two survey managers, one statistician and three production assistants participated in the project group. The focused survey is called "Short-term employment statistics" which collects various data from establishments on the number of employees, absenteeism as well as personnel turnover on a quarterly basis. At the commencement of the project, figures were published approximately 7 weeks after the referenced quarter. At this time, there were 12 production assistants working with the survey along with a survey manager.

The project recommends a reorganisation of the survey team into production groups, an idea that was developed together with the production team. With this type of organisation, follow-up is facilitated and a greater readiness/flexibility is created for unpredictable events. Also people with different strengths and weaknesses can work more efficiently together towards a common goal.

A second recommendation is based on an experiment that was conducted in April, 2000 with a Reminder postcard. The experiment was integrated with the ordinary reminding process. The results show that the implementation of a Reminder postcard brings about a steadier inflow of forms and consequently a more even workload for the production assistants. A larger portion of the material comes in sooner and closer to the reference day which is considered an advantage. Interestingly, the inflow of forms was sped up for the identical population even for the following quarter when no Reminder postcards were sent out. A conceivable conclusion is that positive effects are noted even for collection periods in the future. The project recommends therefore an implementation of Reminder postcards for the collection periods of the 1st and 3rd quarters.

There is no doubt that the production process will be shortened if these recommendations are followed.

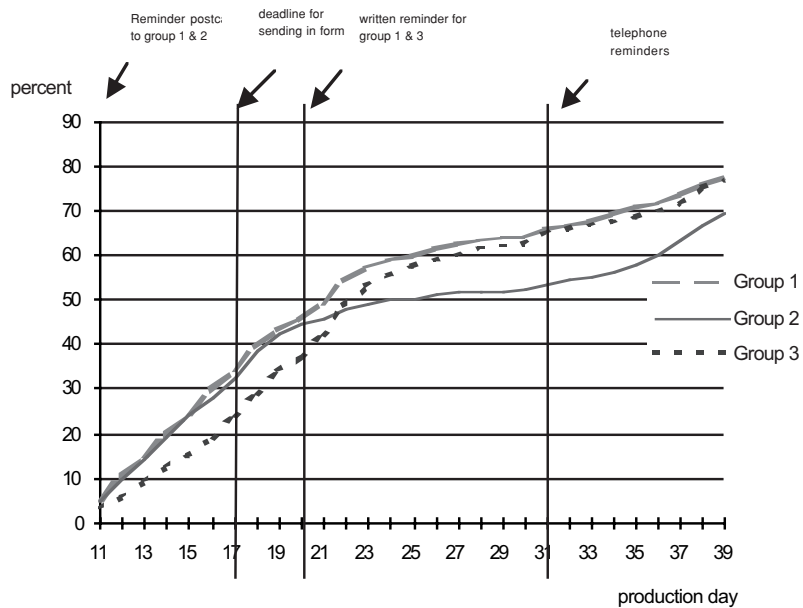


Q 2001

Stockholm, Sweden

The experiment

Diagram 1. Accumulated proportion of registered forms in percent of the total number of establishments per experiment group (time period 12 april 2000 – 26 april 2000)



A Thank-you card was sent out on production day 10 to a randomly chosen group of respondents who had yet not sent in their form with respect to the collection month of March. This amounted to 3031 establishments who were subdivided randomly into 3 different groups.

- Group 1:* received a Reminder postcard and written reminder with an extra form (1067 establishments – 35%)
- Group 2:* received only a Reminder postcard (897 establishments- 30%)
- Group 3:* received only a written reminder with an extra form as is the custom (1067 establishments –35%)



Stockholm, Sweden

14:10 Improving the Quality of the Index Numbers Production in the Basque Country

by Cristina Prado, Lourdes Llorens, and Marina Ayestarán,
EUSTAT – the Basque Statistics Institute, Spain

This presentation aims to explain the renovation process that EUSTAT – the Basque Statistics Institute is carrying out to improve the economic and social trends statistical production. This renovation focuses on the index number generation, such as the Industrial production index, the Industrial price index and the Domestic trade index. The process mainly consists of the definition of a completely integrated and automated system including all the stages involved in these statistical operations. Such stages include not only the information sending to the enterprises of the Panel but also the validation, imputation and calculation of the index numbers.

This paper gives a system description summary, underlining the multiple advantages provided by a completely integrated system. We also try to describe in detail the renovation and incorporation of some processes closely related to the statistical production quality improvement, such as data validation process, no-answer imputation, and process automation allowing the enterprise panel representation control. The latter feature will be subject to a more detailed presentation, due to the importance and the effort involved in the basis renovation. Therefore the “Index Basis Change” is built. The integration of both the index production and the basis renovation in one system provides a more efficient management of human and material resources.

The improvement of the different processes shown in this paper is a logical consequence of the statistical production experience, on one hand, and on the other is due to the analysis implementation over the last years. The mentioned procedures have been incorporated to the system thanks to a general renovation of this.

EUSTAT has been benefited by the incorporation of fully integrated and automated systems and by the improvement made in the procedures. Those benefits mean an increase in the quality results, an improvement in the publication date and an increment in the technicians' productivity, allowing a more efficient distribution of human resources.



Q 2001

Stockholm, Sweden

Finally we will try to evaluate other procedures, about to be implemented, such as the data collection by Internet and by computerised telephonical assistance.

14:30 **Improved National Accounts**

by Karl-Gustav Hansson, Statistics Sweden

In late years Statistics Sweden have been heavily criticized for its economic statistics and national accounts by almost every user, especially the Central Bank, the Institute of Economic Research, and economic analyzers. In short, one say that our published figures are too late, misleading and wrong. Actually, this kind of critique has been around since the 1940th, which especially become apparent when you look at the international literature. For accurate economic analyzing, modeling, forecasting and the statistics have to be fast, accurate, entirely consistent and meaningful. That is not the present situation. There are number of reasons for that. An important one is that we have not yet properly used available possibilities to gather and compile basic data, neither have become aware of the new insights how to produce entirely consistent and fast figures.

Today we do not have to make a choice between fastness and quality. We can choose both. Today we do not have to wait for proper and correct price and volume indices, especially the lacking one in service production. Right away we can start an implementation procedure for a proper compilation of such. The needed knowledge and experience for that are available. For example it has been shown that it is possible to come up with more or less entirely consistent figures of high accuracy in a week after the end of the month. The main implication of these insights is that we rather instantly can gather and compile economic statistic simultaneously in current and fixed prices. Thus, it is quite possible to consolidate entirely consistent monthly national accounts from the real side within a month of high accuracy, especially for the present branch structure. Which mean by one importing thing, that we totally can avoid the non-accuracy between our present quarterly, regional and yearly national accounts.



Stockholm, Sweden

Session 15 The Auditorium

16:05 LEG chapter on Quality and Customers: A result of a Complex Dialogue Between User and Producers of Statistics

by Walter Radermacher, Thomas Körner, StBA, Wiesbaden
Michel Blanc, Alain Desrosières, INSEE, Paris

When it comes to quality, there are few other statements we could as instantly agree upon as upon the demand that official statistics has to be “user oriented”: “It is for users to decide. They are the people who determine quality criteria“ heralds a recent publication on quality in official statistics. Statisticians, it is argued, are “no longer ‘number freaks’ in a world of their own“ but become “managers of statistics“ who stay “in constant touch with those who make decisions“.

Staying in touch with the users could, however, be a more difficult assignment as it seems at first sight. Improving user orientation is a complex issue - and this is especially true for official statistics. In few other areas the categories of customers and users are as manifold. Seldom is the interrelationship between users and producers as complex as it is in official statistics. In this contribution, we would like to make this complexity a bit clearer and to show some ways of dealing with user orientation.

Accordingly, the paper deals with a characterisation of the inter-relationship between users and customers (the user-producer-dialogue). Some fundamental concepts are introduced which help to explain what quality management and quality improvement mean in this context. Quality, it is argued, can only be attained in an optimisation process in which an optimum mixture of quality items has to be found. Finally, some instruments of optimisation are named. For the case of statistical councils - important instruments in this respect - the current situation in the European National Statistical Institutes (NSIs) is briefly outlined.



Stockholm, Sweden

16:25 What Do Users Want to Know about Quality, What Should They Know, and What Can We Tell Them ?

by Susan Linacre, Office for National Statistics, U.K.

Quality in statistics, as elsewhere, is frequently defined as fitness for use. However this implies a need for the user of statistics to be able to assess the fitness of a particular statistic for each proposed use. This paper focuses on the issue of assessing fitness for use of statistics, whose role this assessment is, and how it might be achieved.

It looks at a segmentation of users into different groups in terms of their wants and needs for information on quality. It builds on the authors experience in Australia and the UK on how statistics are used, and the expectations and needs of users for statistics and for information on the quality of those statistics. The paper then sets out a number of the quality measures used by statistical agencies, and approaches taken to describing quality. It assesses these against the perceived needs of different groups of users, identifying issues and suggesting some ways forward.

16:45 Customer-driven Quality, Revisited

by Rich Allen, National Agricultural Statistics Service, USA

In a 1994 paper, prepared for the American Statistical Association Winter Meetings, the National Agricultural Statistics Service used the term customer-driven quality to describe its ongoing communications with data users and other customers. That paper summarized some 40 examples of improving quality through timing, content, or format changes which had been made in the previous three years in response to input from customers—without new legislation or funding. This paper discusses the continuing evolution of customer contact mechanisms as the world has largely shifted to electronic communication and cites lessons learned after installing new features. It will provide many new examples of customer-driven quality and highlight two important case studies. One, involving a time of day change for the release of many statistical reports, will document Agency efforts to work with and placate two divergent customer viewpoints. The second will



Stockholm, Sweden

describe procedures to provide immediate release news services
improved access to reports while maintaining absolute security.

Session 16 Room 456

16:05 LEG chapter on CBM and Minimum Standards

by Mats Bergdahl and Lilli Japiec, Statistics Sweden
Margarida Madaleno, INE, Portugal, Marina Signore, ISTAT, Italy,
and Ioannis Tzougas, National Service of Greece

During the course of the LEG several tools for standardising and improving the quality of statistical surveys have been identified and discussed, and their respective attributes evaluated. Among the tools are Current Best Methods, Minimum Standards, Quality Guidelines and Recommended Practices. These tools aim to improve the quality of statistical processes and products both through standardisation, by removing unjustified variation, and by presenting good proven approaches to apply in a given situation. The LEG has found that they all have their respective merits and drawbacks according to the following attributes:

- How it will inspire to excellence
- How it will reduce unjustified variation
- How it will work on a European level
- How easy it will be to construct
- How easy it is to understand
- How communicable it is to customers
- How uncontroversial it is

Some of the tools have been found more useful for specific surveys, others for individual organisations and still others for the European Statistical System. The presentation will cover the definition of the tools, the description of their attributes and the recommendations of the LEG for future work.



Stockholm, Sweden

16:25 **Good Practices in Official Statistics** **An element of statistical quality**

by Willem de Vries, Deputy Director, United Nations Statistics Division¹

The Fundamental Principles of Official Statistics, adopted by the United Nations Statistical Commission in 1994, give general guidelines for 'good statistical behavior'. Their focus is on such issues as independence, relevance, objectivity, impartiality, professional standards, accessibility, accountability, confidentiality of individual data, compliance with international standards, etc. To make the Principles more transparent and concrete, the United Nations Statistics Division has recently been collecting examples of statistical practices across the world. Its Good Practices web site (a searchable database that is open to the public since March 2001) now contains hundreds of text documents about a wide range of statistical policies, from dozens of countries, and it is still growing. This paper explores the implementation of good practices to the extent that the UN has received information about them. A tentative and not really surprising conclusion is that those statistical offices that have the reputation of producing good, high-quality statistics have indeed over the years developed a set of good statistical policies and that, in addition, most of them advocate these policies publicly. There are, on the other hand, many countries where the implementation of good practices seems to be at a rudimentary stage and where few practices, if any, have been made explicit.

¹ This paper represents the personal views of the author and not necessarily the policies of the United Nations.

16:45 **Revising Statistical Standards: An Exercise in Quality Improvement**

by Marilyn M. McMillen, Chief Statistician, US National Center for Education Statistics

The Statistical Standards Program (SSP) of the U.S. National Center for Education Statistics (NCES) provides methodological and statistical support to the Center, as well as to federal and nonfederal organizations that engage in statistical work in support of the mission of NCES. This



Stockholm, Sweden

Program develops standards for procedures to ensure the quality of statistical surveys, analyses and products, and consults and advises on the implementation of standards for all Center projects.

Although NCES has been guided by written statistical standards for some 12 years, they were last revised in 1992. Given the number of advances in the field of survey management, data collection, processing, and analysis in the intervening years, SSP is currently leading an agency-wide project to produce an updated set of statistical standards. This project includes 15 working groups and 54 staff members (half the agency).

The topics covered range from survey management, file formats, data documentation, graphic displays and product dissemination to data confidentiality, disclosure risk control, nonresponse, imputations, multiple comparisons, and variance estimation. The process we are following includes three levels of internal review, one round of external review, and finally an independent review by an expert panel convened by the U.S. National Institute of Statistical Sciences.

To date, the efforts of the working groups have uncovered a number of difficult issues in areas such as disclosure risk control, the measurement and use of response rates, the analysis of nonresponse bias, and the imputation of data from longitudinal studies. This report will outline the steps involved in this standards revision process and discuss the issues and proposed resolutions for some of the data quality issues uncovered by the revision process.

Session 17 Room 357

16:05 The Color of Culture or Is Asking for One's Race a Culture Measure?

by Peter Ph. Mohler, ZUMA Mannheim, Germany

A. Culture and social context

Attitudes, opinions, behavior, and other characteristics of individuals, groups or societies cannot be taken as absolute values. They are always embedded in a larger context ('environment' of a 'system' in Luhmann's terminology): individual characteristics are embedded into those of



Stockholm, Sweden

groups and the society at large, groups into the cluster of groups around as well as the society and the society itself into clusters of other societies or the 'world system'. Such contexts are measured in standard surveys with Back Ground Variables (BV - demographics, face sheet data). Among the concepts and constructs measured by such BV are 'minority and majority' and 'dominant (power wise)' group very important. For instance, GENDER can be seen as a construct indicating 'membership to a dominant/non-dominant group in a traditional society'. RELIGIOUS DENOMINATION can be seen also along the lines of majority-minority. The question rises whether 'culture' might be also such a societal context or, whether culture is the top level context not directly measurable (one should mention here, that Parsons segregated culture from society, while Luhmann does not).

B. Culture in survey research

'Where do you want to go to abroad' answer by a young American who answered 'to the West Coast, there is another culture', or, as one researcher said in a scientific discussion 'some African American are related to a specific syndrom' – with the response by a young American researcher, whether she got a kind of a disease, because of her skin color. These are but two of the evidences, showing how near all the misses about 'abroad' and 'other' culture are in our everyday as well as survey speak. Moreover, while people are asked about their 'race', 'ethnicity' or 'citizenship' – survey researchers quickly blow up such measures into a fuzzy notion of 'different cultures'.

C. The color of culture

The quick jump from race (i.e. color and 'looking different', in naive official terms), ethnicity (i.e. a undefined mixture of color, origin, language etc.) and citizenship to culture obscures the multi-dimensionality of the concept. What 'Afro American', 'Latino American', 'Asian American' and 'Other Americans' actually measure in surveys are minority-majority characteristics, inclusion and exclusion and so on. The usage of such measures for 'culture' filters out the cultural variations within groups (as seen already by F. Kluckhohn in the early fifties of the last century) as well as the cultural equivalence across the color-demarcation.

The paper will investigate up to date usage of measures of race, ethnicity etc. as proxies of culture in the survey literature and criticise them in the light of general social and cultural theories.



Stockholm, Sweden

16:25 Modelling of Comparability and Coherence of International Statistics

by Raoul Depoutot, INSEE, France

The concern for the quality of statistics has grown rapidly during the last decade, and is now quite widely spread among national and international statistical offices. Spatial and particularly International Comparability is one of the components of the quality concept, recognised both by national and international organisations (Champsaur (1997), UN (1983)). Unlike the other components (accuracy, relevance, accessibility, timeliness, coherence, comparability through time, completeness) - Eurostat (1996a) - for which much theoretical and/or applied work has been conducted, the situation for international comparability is less favourable. The existing work is classified in the paper into three categories:

- a uniform approach used by some international organisations to conduct surveys (Kish (1994)), which consists in designing exactly the same survey in several countries
- a “meta-data” approach, to document with many details all discrepancies between national statistical production processes, statistic by statistic
- a “subsidiarity” approach, which consists in adopting common definitions of the statistics to be produced, either with few or no constraint on the production process or the level of quality achieved or with the obligation to report on quality (Eurostat 1996b).

We show then that the concern for quality requirements in international comparability leads to develop a system based rather on micro-data than on aggregate statistics, supporting the view expressed by Laaksonen (1995).

Building on this, we develop a theoretical model approach of international comparability in order to define *comparability adjusted series* and *lack of comparability*, focusing mainly on differences in concepts. Methods for considering both sampling errors and estimation error in the econometric model are then referred to. We propose as well an approach



Stockholm, Sweden

to robustness and sensitivity analysis for international comparisons. Elaborating on former work (Depoutot, 1998), it is shown lastly that the proposed approach for comparability applies as well to *coherence* and that quality measurement is essential for *quantitative coherence adjustment*.

In particular, the impact of divergent concepts is isolated from the impact of differences in the measurement methods. For coherence, non-linearity of the coherence constraint (like for instance coherence for seasonal adjusted and non-adjusted series) is shown to be a specific third factor.

It is shown as well that due to large dimension of the problem for international comparisons, it is cost-effective to: 1) use EU references as common references for international comparisons 2) use extensively model-based estimations for the adjustment for non-comparability/ non-coherence.

16:45 **The European Social Survey: Substance and Methods**

by Roger Jowell (National Centre for Social Research, London, UK)
Peter Mohler (ZUMA, Mannheim, Germany)
Ineke Stoop (Social and Cultural Planning Office, The Hague, Netherlands)

Five years ago the European Science Foundation (ESF) started the preparation of the European Social Survey (ESS). Whereas Eurostat harmonizes and collects behavioural and factual data across the EU's member states the focus of the ESS is the systematic study of European citizens' attitudes, attributes, and behaviour relating to a core set of economically, socially and politically relevant domains. After an extensive period of preparation, methodological and substantive studies and consultation the funding for the first wave of the ESS seems secure. The survey costs of the ESS will be provided by each of the 15 participating countries whereas the EU will fund the central organizing costs (design, co-ordination, methodological research, quality control, archiving, etc.). The preparation of the ESS has now reached a final stage, headed for launching in 2002. The resulting data will be available at no charge to the scientific community for substantive and methodological research.



Stockholm, Sweden

The questionnaire will comprise a core set of well-tested items on ethnic and national identities, political trust, party affiliation and voting behaviour, religious affiliation, value orientations, media consumption, experience of unemployment, economic hardship and crime and a number of socio-demographic items. It will also contain two rotating topic-specific modules for which researchers will be invited to submit proposals. In addition contextual data at a national and European level will be collected.

A large and innovative venture such as the ESS requires extensive methodological and quality control. The implementation and effect of the questionnaire, sampling design and of non-responses will be routinely monitored and assessed. This will aid interpretation of the data and ensure that high standards and comparability are maintained. In addition to this routine assessment, methodological research will be carried out designed to extend knowledge (and ultimately to improve best practice) concerning the effects of nonresponse and the role of the interview in this process.

Cross-national surveys are beset by country-specific differences in methodological and procedural habits - such as preferred modes of interviewing, particular approaches to sampling, considerable variations in response rates and how to measure them, different methods of training interviewers, specialised socio-demographic classifications, the absence of certain 'standard' conceptualisations and so on. It goes without saying that international comparisons based on these sorts of data are highly problematical. One of the primary aims of the ESS is to meet that problem head-on with a view to attaining the sorts of standards in European research that are routinely attained in the best national surveys.

The paper will focus on methodological standards and research that will be incorporated in the ESS to ensure high levels of data quality and cross-national comparability.



Stockholm, Sweden

17:05 **CHINTEX Project Description**

by Roland Günther, Federal Statistical Office of Germany

Research Framework

CHINTEX is a shared-cost research project supported by the European Commission's Fifth Framework Programme for Research and Technological Development. It applies to the second key action "Information Society Technologies" (IST-Programme) and there to the cross programme action "New Indicators and Statistical Methods" (CPA.4).

Objectives

The harmonisation of European surveys run by national institutes is a task of growing importance for the European statistical system. The European Community Household Panel (ECHP) switched in three countries from input harmonisation with strict comparability of questionnaires and statistical routines to ex-post harmonisation with reduced comparability of questionnaires and statistical routines. This switch of method was caused by the stop of three national sub-samples and their replacement by ongoing household panels run by other institutes. It is the overall objective of CHINTEX by means of this unique data situation to clarify if it is necessary to have centralised, standardised survey instruments to achieve harmonisation and comparability or if this objective can also be achieved by ex-post harmonisation, by which independent national sources are satisfactorily converted to common concepts, definitions, survey questions etc. Furthermore, the project investigates important hypotheses about the data quality of panel surveys (non-response, reporting errors and panel effects) which are of general interest for survey statisticians.

Description of Work

The project starts with an assessment of the attained level of harmonisation in the three national conversion projects with respect to variable conversion and application of unique imputation rules and weighting schemes. In a second step a taxonomy of conversion problems for non-trivial cases shall be established to produce results of general interest. The project will demonstrate exemplarily how statistical tools may be used. Another branch of harmonisation is the use of common rules for imputation and for the construction of weights. The variability of population estimates according to these two sources is investigated.



Stockholm, Sweden

Furthermore we will look upon the impact of field related factors that cannot be harmonised ex-post. Their potential impact on non-response, quality of income data and estimation of statistical models will be examined to check whether long running panels have lower non-response and a better quality of income data. On the other hand panel attrition is regarded to introduce a bias for the estimation of statistical models. CHINTEX will investigate this hypothesis by comparing estimates from the old panels with those from new ones and by using register information in the case of the Finnish ECHP sample. This register information will also be used for a quality assessment of the ECHP imputation rules with respect to income. In addition, a cross-validation approach is used to exploit the information of the parallel panels. If the evaluation indicates a poor performance of the ECHP imputation rules, alternative suggestions shall be made. In this respect CHINTEX will concentrate on the longitudinal aspects of imputation. The project will organise three workshops open to academic and official statistics to disseminate and discuss its results.

Expected Results

The project establishes new methodological framework for the conversion of data that can be used for the ECHP but can be generalised for other surveys. It presents empirical evidence on the effect of switching from input to ex-post harmonisation and analyses the effectiveness of the ex-post approach in producing comparable information. Insight on the impact of panel duration on the quality of the data is given and important hypothesis about possible panel effects are tested. The project presents an evaluation of the ECHP imputation rules on income and suggests alternatives whenever necessary.

Research Consortium

The project is carried out in co-operation of national institutes of official statistics and research institutes with high expertise in the field of panel surveys.

- Statistisches Bundesamt, Institute for Research and Development in Federal Statistics, Wiesbaden, Germany (co-ordinator of the project)
- Johann Wolfgang Goethe Universität, Institute for Statistics and Mathematics, Frankfurt (Main), Germany
- DIW (German Institute for Economic Research), German Socio Economic Panel Study, Berlin, Germany
- University of Essex, Institute for Social and Economic Research, Colchester, United Kingdom



Q 2001

Stockholm, Sweden

- CEPS/INSTEAD, Household Division, Differdange, Luxembourg
- Statistics Finland, Social Statistics, Helsinki, Finland in co-operation with Åbo Akademi University, Department of Economics and Statistics, Turku, Finland

Data Base Used

CHINTEX uses a data base which is unique in multiple aspects. For the three years 1994 to 1996, parallel panel surveys on private households had been carried out in Germany, the United Kingdom, and Luxembourg. In each country, one panel was part of the input harmonised European Community Household Panel (ECHP) used by official statistics, and the other panel was a sample for scientific use. After the stop of the three ECHP sub-samples, the national scientific panels had to be converted to become part of Eurostat's harmonised ECHP data base. In Finland, ECHP data are available for the waves 1996 until 1999. The 1996 wave was merged at the individual level with register information on income. CHINTEX is going to extend the register merge for the entire period.

Work-plan

The overall duration of the project is from January 2000 to December 2002.

The project is subdivided into seven research and three supporting work-packages. The seven research work-packages are:

- "Assessment of the Attained Level of Harmonisation" (WP 1),
- "Non-trivial Conversions and their Methodology" (WP 2),
- "Imputation and Weighting Homogenisation" (WP 3),
- "Non-Response" (WP 4),
- «Quality of Income Data» (WP 5),
- "Panel Effects" (WP 6) and
- "Imputation (WP 7).

The supporting work-packages are concerned with:

- "Assessment and Evaluation" (WP 8),
- "Dissemination and Exploitation" (WP 9) and
- "Project Management" (WP 10).



Stockholm, Sweden

Session 18 Room 353

16:05 Data Quality as Function of Time, Survey Technique and Structural Characteristics of Enterprises

by C.Abbate and D.Filipponi, ISTAT, Italy

In 1999 a survey on almost 400.000 Italian enterprises was carried on using a complex survey technique with the aim of reducing the non response rate. Thus, over a period of seven months, using five different techniques it was achieved a response rate of 75%. All the data collection states have been supported by a monitoring activity through which it was possible to learn the response delay and the number of soliciting for the 310.000 respondent enterprises. Strong attention was placed in all phases of quality data control, registering for each enterprise the number of errors and missing values. This information allowed to obtain a measure of data quality for each record. A model for the data quality was build using the response delay (recorded in days), the soliciting technique and other structural characteristics of the enterprise (such as economical activity, number of employees and geographical location) as explanatory variables.

16:25 Algorithms for Prioritising Respondents in Phone Surveys

by Bjørn Steen Larsen, Peter Linde and Birger Madsen, Statistics Denmark

This paper deals with algorithms for prioritising respondents in telephone surveys. Every organisation carrying out telephone surveys has to decide when to phone a specific respondent. In many cases this is based on random assignment, simply because no prior information (except on geography) is available.

In Statistics Denmark much information is at hand about each respondent prior to the interviewing phase. We thus have information on age, sex, education, business sector (if respondent has a job), income, marital status (cohabitation), household size, type of family (are there children



Q 2001

Stockholm, Sweden

living at home) and of course geography (urbanisation).

All this information can be used to get a prior estimate of probability of nonresponse in specific time periods for each respondent. Information from the historical data on the interviewing process is crucial in this context.

Having estimates of probabilities of nonresponse at specific time periods, a linear programming problem can be formulated. This problem maximises the total number of interviews obtained subject to resource requirements.

16:45 Attrition Bias in the Bank of Italy's Survey of Household Income and Wealth

by Anna Giraldo, Enrico Rettore, and Ugo Trivellato, Dipartimento di Scienze Statistiche, Università di Padova

The Bank of Italy (BI) runs the Survey of Household Income and Wealth (SHIW) since 1965. SHIW provides information on income, saving, consumption expenditure and real wealth of Italian households, as well as on household composition and labour force participation. Due to its characteristics and to the BI liberal policy in the release of the micro-data, most research on income and wealth distribution in Italy is carried out using SHIW data.

The survey has been conducted on a yearly basis up to 1987 and on a biennial basis since then. To allow for dynamic analyses, in 1989 BI introduced a longitudinal component into the survey. While in the 1989 and 1991 waves sample units were included in the panel conditional on their willingness to further participation as assessed at the previous survey round, in 1993 the design of the panel sample went through a major revision: at each point in time a sub-sample is randomly drawn from the pool of households who just entered the survey and added to the panel sample. By now, the total sample size is about 8.000 households while the panel sample size at each point in time is roughly one half the total sample size.

The purpose of this paper is to investigate the degree of representativeness of the SHIW panel sample. We focus on the four waves panel 1989-1995. On releasing the data, BI provides the user with some broad *caveats*



Stockholm, Sweden

about the quality of the overall sample, in a cross-sectional perspective. On the other hand, very little is known about the panel sample representativeness over time. Here we are mostly concerned with the effect of the attrition-driven non response on the quality of the sample. By reducing the sample sizes and by potentially biasing the sample, attrition is one of the major problems plaguing panel surveys. In the four-wave SHIW panel 1989-1995 the sample size is as large as 2,187 in the first wave and reduces to 827 in the 1995 wave, with an overall response rate of about 38%.

We deal with three specific issues. Firstly, we aim at detecting whether the attrition plaguing the panel was concentrated in some specific wave or was evenly spread over the four waves. We estimate sequential logits for the probability to be still in the sample at time t , $t=1991, 1993, 1995$, conditional on being in the sample at time $t-1$ and on some observable characteristics of the sample units. By means of this analysis we also detect which variables drive the attrition process.

Secondly, as a by product of the sequential logit analysis, a weighting adjustment to compensate for the bias induced by attrition is derived, the weights being based on the logit regression coefficients. We compare the estimated means of some variables (income and consumption) as resulting from our weighting procedure to those resulting from using the weights suggested by BI.

Thirdly, we test whether the 1993 new design improves the quality of the panel sample, by reducing its bias.

Session 19 Room 361

16:05 Reporting Quality in National Statistics

by Pam Davies, Office for National Statistics, U.K.

Keywords: reporting quality; measuring quality.

Measuring and reporting on the quality of statistics is a complex undertaking. There are several dimensions to the concept of quality, many potential impacts on the quality of the statistics, and typically no



Stockholm, Sweden

comprehensive measures of quality. There are also a number of quality conflicts – different users will regard different components of quality as more or less important. Just as users' needs for statistics varies, so their needs for quality information will vary. Some users will be interested in the technical details of the methods underpinning a statistic whereas others will need much less detail. The challenge to official statisticians is to balance the needs of a wide variety of users, to set up systems that produce good quality statistics, to measure the quality of those statistics, and to disseminate information on quality to users.

A major advance in quality reporting was achieved in the recent project *Model Quality Report in Business Statistics*. The project developed a detailed description of the methods for assessing the quality of surveys, with particular application in the context of business surveys, and then applied these methods in some example surveys to evaluate their quality. The work was specified and initiated by Eurostat and was undertaken by a consortium of the UK Office for National Statistics, Statistics Sweden, the University of Southampton and the University of Bath. This project focussed very heavily on Eurostat as the customer. Some of the conclusions of this work may need to be adapted for statistics that are not based on business surveys, and also to take account of the needs of a wide range of different users.

Reporting quality has been addressed in a number of other ways within the UK. The publication of the *GSS Statistical Quality Checklist* provided a checklist of information to be provided about official statistics. The SQC recommends that a range of information is provided along with the published statistics. This information enables users to judge whether statistics are fit for purpose. For many national statistics a wide range of information on quality is made available to users. This information is made available via StatBase – the UK's on-line database that holds a large selection of UK official statistics. Although the information is wide ranging, it is not as comprehensive or detailed as that recommended in the *Model Quality Report in Business Surveys* project.

The UK plans to build on these and other developments and develop a strategy for quality reporting that takes account of a wide range of users, as well as the needs of other stakeholders in the survey process. This paper will review the information on quality that is currently produced and made available for national statistics; discuss the developments that



Stockholm, Sweden

are needed to implement a consistent approach to quality reporting across national statistics; and record progress towards developing an approach for quality reporting.

16:25 Correlates of Quality on an Establishment Survey

by David Cantor, W. Sherman Edwards and Kerry Levin, Westat, Inc., USA

While there is an increasing awareness that these data are subject to many of the same errors as household surveys, there is also reason to believe that unique sources of error contribute to the quality of the information. Edwards and Cantor (1991) provide a framework within which to understand the response process, partially based on the information processing model that is popular for household surveys. Similarly, Tomaskovic-Dewey, Leiter and Thompson (1995) provide a framework that proposes unique processes within organizations that lead to decisions to participate on establishment surveys. There has been very little research, however, that has examined empirical measures of quality on an establishment survey and how these measures relate to theoretical models such as these. The purpose of this paper is to draw data from several establishment surveys to assess quality along two dimensions – nonresponse and measurement error. Analysis will assess quality, and examine correlates of the quality measures. Implications of these results will be discussed within theories of organizational theory nonresponse (e.g., Tomaskovic-Dewey, Leiter and Thompson, 1995) and measurement error on establishment surveys (Edwards and Cantor, 1991).

One analysis will describe the patterns of nonresponse, as well as the potential for nonresponse error. Establishment surveys are marked by relatively numerous contacts with the respondent, while household surveys typically result in relatively quick refusals. The analysis for the paper will characterize this pattern for an establishment survey conducted by telephone and assess whether these patterns have implications for nonresponse error. The analysis will relate the level of effort used to make contact to key survey variables of interest. Characteristics of the establishment will be correlated to the difficulty of making contact to



Stockholm, Sweden

better understand nonresponse patterns and how it relates to the decision to respond to a survey.

The second analysis will examine correlates of measurement error. A number of characteristics of respondents and organizations have been hypothesized to affect the quality of information collected on an establishment survey (Edwards and Cantor, 1991): 1) the structure and complexity of the organization, 2) the motivation to participate on the survey, 3) the available knowledge to provide the information and 4) the retrieval process the respondent may use to provide the answers. The analysis will examine whether quality is related to: 1) the respondent's relationship to the record system (respondent function and level), 2) the level of effort to complete the interview (motivation), 3) the size of the establishment and 4) the type of information used to respond to the survey (e.g., use of records, consulting with others within the organization). These analyses will be conducted across two different types of questions. One concerns whether the establishment is covered under a particular law. This requires that the respondent knows about the law and applies it correctly to the location's situation. Quality will be judged using the size of the establishment, which can be used to determine whether the respondent is correct. The second type of question requires the respondent to retrieve information from records. Quality on this item will be judged by whether or not an answer is provided.

16:45 An interactive System to Improve Quality in Official Statistics

by Martino L., De Santis A., and Salvi S., Italian National Institute of Statistics

The issue of improving quality of statistical information has recently been addressed by many national statistical institutions. According to this need, Eurostat has issued some guidelines that summarize the basic principles to be followed when carrying out a survey¹. Such guidelines advise to use the most appropriate sampling strategy, data collection technique and estimator according to time and budget constraints and to the targets of information. All these ingredients help preventing non-representativeness, resource shortages and low precision before the data are collected. After data collection, a set of procedures to check microdata and to provide quality indicators are needed. These include correcting



Stockholm, Sweden

for inconsistencies, detecting outliers and measurement errors and adjusting for missing data.

The Italian National Institute of Statistics, following Eurostat directives and the users' request for information of better quality, started revising methodologies of surveys in various areas. The agricultural sector has been deeply involved in this process. An integrated system of agricultural statistics, that include conjunctural and structural surveys, along with surveys on related sectors such as food-industry and environment protection, has recently been designed based on a set of existing statistics². The main aim of this activity was improving efficiency of the estimates and therefore the quality of the final information. To this scope the sampling survey on farms structural aspects - providing information over type of farming, use of land, type of management and labor force among others - has been completely redesigned³ so that a set of other surveys at the present time represent sub-samples of it. In addition a computerized system able to support checks and control procedures has been developed with the aim of providing a computerized tool that is user-friendly, flexible and time-saving.

The use of generalized and integrated systems is strongly suggested when managing data coming out from a survey. They allow a reduction of the resources devoted to check procedures using macro-editing interactive and automatic techniques. In addition they allow documenting the entire process and providing quality indicators. Much work has been done in this area especially by national statistics institutions (e.g GEIS developed by Statistics Canada⁴).

Within the process - carried out by the Italian National Institute of Statistics - of renewing the agricultural statistics sector, the computerized system AGAIN (acronimous of the Italian "Analisi e Gestione Automatica delle Indagini") has been set up using SAS/AF. The software architecture has been designed to support the following main purposes: some basic utilities, estimation procedures, data check and control, quality indicators. The last two modules are devoted to get a better quality level by minimum data changes and to report about it. The main advantages in using this software rely on its portability (since it only requires to have SAS license) and its adaptability. Even though the system has been tested on data collected within the milk and dairies survey, it will be quite easily adapted to more complex ones such as the farms' structure survey.



Stockholm, Sweden

¹ Eurostat (2000): "Assessment of quality in statistics". Item 4 of the agenda "Standard quality report". Luxembourg 4-5 Aprile 2000

² Schirinzi G. (1998) Lineamenti per una ristrutturazione del sistema di statistiche agricole in Italia, Documento Istat.

³ Benedetti (2000): L'impostazione di un sistema integrato di rilevazioni sulle aziende agricole: soluzioni adottate problemi aperti. Seminario dal titolo "La revisione metodologica delle indagini nel settore primario". Roma: Istat, 24/5/2000.

⁴ GEIS development team (1991): Generalized Edit and Imputation System Operations Users' Guide. Statistics Canada, Research and Generalized Systems Subdivision Technical Report.

Session 20 Room 359

16:05 Using Total Quality Management to Improve Spanish Industrial Statistics

by Pedro Revilla, National Statistical Institute of Spain

The TQM approach used to produce industrial statistics in Spain is presented in this paper. The starting point is asking customers about our statistics products. Then, specific programs and actions are implemented trying to improve customer satisfaction. The following projects have been launched as a direct consequence of expressed customers needs: improving timeliness; improving a service of tailor-made information; implementing a system of integrated microdata that allow to mix data from different surveys; and offering individual data tailored to the reporting enterprises as a "payment" for answering questionnaires.

16:25 Quality Control of ICR Data Capture for the 2001 Census of Agriculture

by Walter Mudryk, Statistics Canada

The data capture of the 2001 Census of Agriculture will be performed using relatively new technology called Intelligent Character Recognition (ICR). This approach to data capture will combine automated data capture (which uses character, mark and image recognition) with manual data capture of unrecognizable fields by operators who will 'key from image' using a 'heads-up' data capture technique. Since the potential for



Stockholm, Sweden

unacceptable levels of error exists at various stages of this process, a Statistical Quality Control methodology has been built into each critical step of this operation.

There are essentially three main stages to the overall Data Capture operation: document preparation, scanning & validation, and the actual capture of the questionnaire data by the ICR system (which includes automated and manual capture from scanned images). Since all three stages depend on one another and each process may contribute to significant errors down the line, a quality control procedure is being built into each stage. Documents that are not properly prepared for scanning could have a negative effect on the quality of the images being produced by the scanning operation. Poor quality images will be more difficult to capture. Also, errors will occur with the systems inability to interpret the data correctly (i.e. substitution error) as well as with operators who are not properly trained in the heads-up capture component of the operation. The QC methodology for each component ensures that the associated errors are being controlled at their source.

This paper will describe this new approach to capturing the Census of Agriculture data as well as the quality control and assurance methods and procedures that have been built into this operation to ensure a high degree of accuracy in the captured data.

Key Words: Statistical Quality Control, Intelligent Character Recognition (ICR), Data Capture.

(Format of Paper)

1. Introduction

- Brief Description of Census of Agriculture with Survey Objectives
- Objectives of ICR Data Capture Process
- Process Flow Description
- Detailed Flow Chart

2. Quality Control of ICR Processes

- Specific Area's of QC Application – Flow Chart
- QC Methods for Document Preparation
- QC Methods for Scanning
- QC Methods for Machine Recognition
- QC Methods for *Heads-up* Capture



Stockholm, Sweden

3. Conclusion

- Preliminary Test Results
- Measurements & Evaluations
- Future Directions

16:45 Planning the Quality of the Automatic Coding Process for the Next Italian General Population Census

by S. Macchia, S. Mastroluca, and A. Reale, ISTAT – Italy

For next Italian General Population Census, which will take place next year on October the 21st, Istat is evaluating the possibility to adopt some innovations in order to: *i*) improve the efficiency of informatic procedures, *ii*) reduce costs and time, *iii*) provide a higher quality of data. Main process innovations concern a new data capture system, consisting of optical character recognition (OCR) and intelligent character recognition (ICR), and automated coding system for textual responses.

The paper concerns in particular the automatic coding process and its impact in terms of quality. The software selected for this purpose is ACTR (Automated Coding by Text Recognition), developed by Statistics Canada. It has already been tested inside Istat in many occasions and also for the two Census pilot surveys (October 1998 and April 2000). In the First Pilot Survey both manual and automatic coding have been applied in order to compare the two processes; only two variables have been considered for this test: 'Occupation' (type of work done in a job) and 'Industry' (branch of economic activity). Methodological studies have been performed to assess the improvement of quality, both in terms of *efficiency* (percentage of responses automatically coded) and *accuracy* (percentage of responses correctly coded manually or automatically) and indicators of time and costs operation have been produced: the results will be described in the paper.

The test was encouraging, so, for the second pilot Survey, automatic coding has been extended to other three variables: Italian municipalities, citizenship and educational qualifications. The quality levels achieved have been measured: as it will be shown, as far as Occupation and Industry are concerned, they are coherent with the first test; for the other variables they are similar to those obtained by other National Statistical Offices.



Stockholm, Sweden

The paper will also describe the process defined, in view of next Census, aimed at measuring the quality of automated coding, taking into account the new hypothesized data capture system. The Census development programme settles that the data capture process and also the coding activity will be probably in outsourcing; Istat will provide the dictionaries for the automated coding to the external company. It has been planned that for the alphabetical strings we will not measure the quality of OCR/ICR recognition process (as it will be done for tick boxes, numeric fields and pre-print codes), but directly the *efficiency* and *accuracy* of automated coding processed by the provider. In particular a set of indicators have been defined and it will be verified, on a sample basis, that results obtained by the company are coherent with those obtained by running ACTR on the same sample.

Finally some words will be spent on another aim which has been defined in order to maximise data quality and to reduce manual coding activity for cases not solved by the automated coding system: it has been planned to develop a system of automatic imputation which considers the set of codes that ACTR proposes when it is no able to assign a unique code.

Session 21 Room 461

16:05 On the Quality of Opinion Surveys in Dutch Market Research

by Harry van den Berg, Vrije Universiteit Amsterdam, The Netherlands

Market research constitutes an important branch of applied social research. Up till now, very little is known about the quality of market research. In order to improve this quality different aspects of the practice of market research have to be scrutinized such as the procedures used in developing questionnaires, the quality of questionnaires, the procedures used in processing answers of respondents and the quality of inferences from these answers.

This paper is focussed on a specific type of question wording very often used in opinion surveys in the field of market research: the assertion. In this question form the respondent is asked to tell whether (s)he agrees or disagrees with an evaluative assertion about a specific topic.



Stockholm, Sweden

In the paper the main results of a methodological study concerning the use of assertions in opinion surveys will be presented. The study is focussed on the following research questions:

- 1) Which question forms are used in opinion surveys in the field of market research and how important are statements as a question form?
- 2) To what extent do statements and the composition of questionnaires with statements fulfil methodological criteria concerning the question wording of statements and the embedding of statements in a questionnaire?
- 3) To what extent are inferences from answers of respondents to statements - as reported by market researchers - justified according to methodological criteria?
- 4) Which procedures do market researchers use in designing questionnaires, selecting specific question forms such as statements, and constructing statements?
- 5) In order to be able to answer these questions, twenty market research organisations were selected to cooperate. They were asked to send a set of questionnaires and related research reports that could be viewed as representative examples of the kind of market research conducted by them. This request resulted in about fifty questionnaires that are analysed in depth. Besides interviews were held with senior market researchers who are responsible for designing questionnaires.

The results of the in-depth analysis of questionnaires, research reports and interviews with questionnaire designers are illuminating in several respects.

Firstly, the main methodological problems in the practice of designing questionnaires with statements could be identified. A remarkable result is that some of these problems are underestimated or even neglected in handbooks of survey methodology.

Secondly, in reporting the outcome of opinion surveys based on questionnaires with statements, there is a general tendency to present different types of over-interpretations, i.e. conclusions that are insufficiently warranted by the data.

Thirdly, the routines and procedures used by market researchers in designing questionnaires do not contain sufficient checks and possibilities for checking and testing questionnaires in order to prevent methodological weaknesses that constitute a severe validity-risk.



Stockholm, Sweden

A more general conclusion is that from it is important and fruitful to study the practice of applied social research in order to overcome the traditional cleavage between developments in the field of methodology and developments in the field of applied social research.

16:25 Interviewers and Survey Quality: Error Analysis and Prevention

M.G. Muratore, S.Vitaletti

The goal of this paper is to focus on the relationship between interviewers and quality of a survey, its positive and negative implications, the possibility to measure them and to strengthen or to prevent them. In 1997- 98 the Italian Statistical Institute (Istat) carried out the first victimisation survey: "Citizen's Safety Survey". Data collection was centralised, because of the use of C.A.T.I. technique (50.000 telephone interviews were collected during 4 months). This was very useful for the continuous training strategy (Muratore, Quattrociochi 1998), for the daily monitoring system (based on quality indicators) (Quattrociochi, Sabbadini 1998, Muratore 2000) and, above all, for the deep relationship created between interviewers and researchers (Muratore, Sabbadini 1999). This framework, based on the motivation of interviewers, produced good results: households have been available to participate to the survey and the refusal rate was going down as long as the survey were being conducted.

Therefore, it's essential to point out the interviewers' role, to analyse their behaviour and their influence on data collection phase.

Different means are used to do this:

1. an evaluation survey gives us an important and substantial feedback about interviewers' opinions and feelings on different aspects: the content and the instruments used for the training with reference to efficacy and effectiveness to achieve established aims, the difficulties met with the households and the interviewees, the relationships established with Istat staff, the feelings experienced. Above all these last aspects will be very useful to better the training schedule (the second Safety's Survey is planned for 2001-2002).



Stockholm, Sweden

2. the textual analysis of open-ended questions, i.e. the interviewers' note and difficulties met in each interview, offers very important suggestions about the content and the interview's process. About the content is essential to check possible ex-post errors declared by interviewers and to understand better the performance of the questionnaire. About the process it's useful to be aware of the survey settings: households needs and lifestyle, which hours are better to interview them, which causes are indicated for the interruption or for the appointment.

3. quality indicators on each interviewer, and their evolution during the survey. The indicators are very different; response rate, refusal rate and indicators of non contact, too, vary very much among interviewers. A deeper analysis of these aspects is necessary to be aware of possible problems and to solve them. For instance, it's possible to automate some kind of telephonic outcomes or to know how an improvement in training could be effective on reducing the refusals.

4. analysis of the information produced by the preceding survey. This is implemented through the new theory of the symbolic objects, using the technique of symbolic marking. This technique permits, on the base of the first survey of "Citizen's Safety", to analyse and to plan soft check and hard check to limit interviewers' discretionality and to check the truthfulness of the interviewees answers.

16:45 Quality Management of Closed Questions in a Survey Among the Elderly

by J.H. Smit, Department of Sociology and Social Gerontology, Vrije Universiteit Amsterdam, The Netherlands

Since 1990, the concept of Total Quality Management is slowly finding its way in survey research. The view that survey research can be described in terms of a production process enables us to divide the process in different phases, where each phase should meet certain quality norms in order to produce high quality data at the end of the process. When we roughly divide the process in: (1) design (sampling and questionnaire construction), (2) fieldwork (data collection) and (3) data analysis and report, it is clear that quality management in phase 2 is rather difficult when the fieldwork consists of face to face interviews at the respondent's residence.



Stockholm, Sweden

An important expedient for quality management during such a fieldwork is taping the interviews with the purpose to evaluate the interaction between interviewer and respondent. By evaluating the tapes, the fieldwork staff may obtain an impression of the overall performance of interviewers and of the questions that cause problems for (certain) interviewers and respondents. An obstacle however, for using tapes as a quality control measure for closed questions is the fact that questions of which response alternatives are presented on a show card are common practice in survey research.

A necessary condition for evaluation of the quality of interviews with tapes is that some interaction between interviewer and respondent occurs. By using show cards such interaction is suppressed. A show card may reduce interviewer-related mistakes by making the task for the interviewer less complex, compared with reading out the response alternatives. However, it also reduces the opportunities for interviewers to get hints from the respondent that a question is not well understood, and the chosen response alternative might not be a valid indication of the real quantity one is trying to measure. Furthermore it also leads to less opportunities for the fieldwork staff to evaluate interviewer behavior and ambiguity of question content. Therefore using show cards with closed questions is certainly a trade off in terms of quality management.

Within the framework of the Longitudinal Aging Study Amsterdam 2000 elderly respondents (58-88 years of age) were interviewed and permission was asked to tape each interview. Also a split ballot experiment was carried out in order to explore the differences between the two alternative ways of presenting response alternatives (1000 respondents with response alternatives on show cards versus 1000 with response alternatives read out).

Results of the request to tape the interviews showed that more than 95% of the respondents agreed to this kind of quality management tool. Refusals were not associated with respondent or interviewer characteristics. Although less possibilities for quality management exist for questions with show cards, preliminary analyses showed that data quality obtained with show cards was not less compared with questions in which the response alternatives were read out by the interviewer. In fact there was some evidence that especially socially undesirable response alternatives were chosen more often when show cards were used.



Stockholm, Sweden

Tuesday, May 15
Session 22 The Auditorium

9:05 LEG Chapter on Documentation

by Marina Signore, ISTAT, Italy

The presentation will describe the work done by the LEG on documentation as well as the recommendations for the Member States concerning documentation.

Two main purposes of documentation have been identified, namely:

- To ensure and improve quality
- To better use and understand data

In fact, documentation (for example check-lists or manuals such as “Current Best Methods”) is a starting point for ensuring the quality of processes since it is a tool for standardising the way of producing and for eliminating unnecessary variations in the outputs. At the same time, documentation is a tool for improving the quality of current activities since it makes the producers aware of the strengths and the weaknesses in the processes and in the products. The other major purpose of documentation is to help the users in retrieving the required information and in making comparisons and statistical analyses. To this purpose, users should be provided with metadata as well as with quality measures.

In general, documentation concerns all the activities carried out by National Statistical Institutes (NSIs) including the production of statistical information and other processes which support such an activity. Another aspect that should be taken into account is that different types of users of documentation (which are both producers and users of statistical information) have different requirements. The needs of each users’ typology differ both with regard to the content and to the degree of detail to which a given subject is treated.

The presentation will illustrate the different types of processes which are suitable for documentation and what type of information should be provided for each type of process as well as to satisfy different users’ needs.



Stockholm, Sweden

It should be noted that most NSIs have to face the problem of lack of time for documenting as well as the associated costs. Therefore, there is a need to find means which facilitate the documentation activity such as information systems. Information systems might be quite costly to implement since they require a design and a test phase and proper expertise. On the other hand, they present many advantages such as the possibility to achieve a high degree of standardisation and comparability, the opportunity to reuse information already produced and finally to support people in the documenting activity. Such tools will therefore make the documentation activity cheaper in the long run. References to information systems in place in the LEG countries will also be provided.

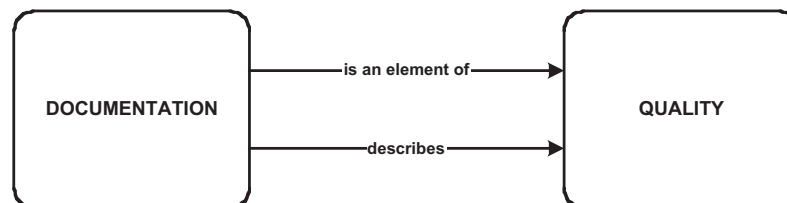
The presentation will also include a review of the approaches followed by the LEG countries with regard to documentation as well as the strategies which have been defined for the next coming years. The common issues as well as the actual diversities will be highlighted.

9:25 Documentation and Quality in Official Statistics

by Bo Sundgren, Statistics Sweden

Documentation describes something, e.g. a physical site, a process, or a set of data. In connection with official statistics, documentation describes a statistical information system, a statistical process, or a set of statistical data (macrodata or microdata).

There is a double relationship between documentation and quality. On the one hand, documentation is an element of quality. If we want to discuss the quality of something, be it statistics or something else, we need some kind of description of that "something" to start from. On the other hand, a documentation could describe the quality of something, e.g. the quality of some official statistics, or some statistical data.





Stockholm, Sweden

Both aggregated statistics (so-called macrodata) and observation data (so-called microdata) are data. Thus documentations of macrodata and microdata are descriptions of data, or “data on data”, so-called metadata. Descriptions of processes may also be metadata, since the description of a process may often indirectly describe data emanating from the process as well. For example, a sampling procedure determines the precision of estimates that are based upon data from the sampling procedure.

There are four major purposes of documentation in connection with official statistics:

- Communication of meaning and quality of aggregated statistics. What is the meaning of these statistics? Which are the uncertainties of these figures? Are these statistics from country A comparable with these statistics from country B, and if not, how can they be made more comparable? Which methods of statistical analysis can be applied to these statistics, and how should the results be interpreted?
- Communication of meaning and quality of statistical microdata. Can these microdata, originally collected for purpose A, be reused for purpose B? Can these microdata be combined with those microdata? Is it possible to perform a longitudinal study on available yearly generations of these microdata?
- Communication of specific knowledge about a specific statistical process or system of processes, e.g. for training purposes in connection with new recruitments, or in connection with auditing activities. How is the labour force survey conducted at Statistics Sweden? How do you calculate uncertainty measures in this survey?
- Communication of general knowledge about a general statistical process or a general statistical information system. Which are the current best methods and practices for (i) frame construction; (ii) data editing; (iii) confidentiality control; (iv) coding; (v) non-response compensation? Which is the current best method for designing the information systems of a statistical office in accordance with the data warehouse concept?

The communication in connection with documentations is usually indirect, since the sender and the receiver are typically separated both in time and space. This requires a documentation to be formulated in a relatively well defined language, which is well understood by both the sender and the receiver, despite their separation in time and space.



Stockholm, Sweden

This paper will discuss conceptual issues like those exemplified above, as well as more practical aspects of how to organise the documentation work of a statistical organisation.

9:45 Designing and Implementing a Documentation System in INSEE

by Dominique Crosnier and Philippe Domergue, INSEE, France

Historical background

Origin

During the eighties, setting up of a policy of increased mobility of INSEE'S executives. Quality benefit expected from this mobility, through accumulation of various experiences.

Consequently, necessity to proceed to a rationalization of transfer methods concerning the know-how, for two reasons :

- avoid its loss, which could entail a loss in quality ;
- reduce the time for the executives to reach an operational level in operation and avoid to always reinvent everything.

Implementation

In order to reach these objectives, a research is performed by the Institute. It draws its inspiration from the methods used in industry to document production processes.

It leads to the elaboration of a documentation method of statistical operations and to the development of a computer tool, running under MVS, which allows its implementation.

Nothing is imposed to the producers of statistics : the method and the tool are used on a voluntary basis.

Result

Disappointing in comparison to what was hoped for and wished. Recourse to this tool to organize and transmit Know-how is not very frequent. It does not succeed in becoming imperative among statistics producers.



Stockholm, Sweden

Which is more, when it is used, it is without having recourse to all the organization opportunities it provides with. The outcome is not very different from an unorganized storage using paper. The know-how transfer is carried on in a more or less traditional manner.

Analysis

Analysis phase to determine the reasons of this relative failure.

Conclusions are the following :

The theoretical model is suitable. Problems are elsewhere. One can identify three of them :

- high cost of elaboration for producers ;
- little incitement to organize documentation ;
- little user-friendliness of the tool.

As a conclusion, one favours the technological aspect as a blocking factor. So, theoretical model and voluntary basis principle are kept, while the tool is rebuilt in order to :

- decrease the cost for producers who remain the main actors of the elaboration of documentation ;
- offer a flexibility enabling to handle various cases ;
- offer new functionalities which provide services susceptible to be a greater incentive to producers (particularly as information dissemination is concerned) ;

New tool (end of the nineties).

Characteristics

The tool is developed in a micro-computer environment. It is integrated to the work station. It is linked with other office-automation tools, the elements of which can be stored in the documentation-oriented data base.

It offers functions for implementation and publishing of documentation:

- automated generation of documents (files or paper) ;
- automated generation of HTML pages.



Stockholm, Sweden

Result

Method and tool are more widely used than in the past. The number of documentation bases increases regularly and the fields of statistical production covered diversify.

The concern is double :

- know-how organization and transfer being easier, there is a quality benefit in the production process ;
- ease of providing the users with the information elements necessary to the understanding of produced data provides a quality benefit through improvement of dissemination.

Conclusion

The voluntary bases principle being maintained, the elaboration of documentation being centralized, the use of the system does not provide for a thorough standardization of statistical production.

Nevertheless, the development of its use will contribute in improving statistical production coordination (consistency in the use of concepts and classifications for instance), which increases the quality of production and dissemination processes of statistical data.

Session 23 Room 456

9:05 LEG chapter on Strengths and Weaknesses of ESS

by Werner Grünewald, Eurostat

In its proposal for the Leadership Group (LEG) on quality, Statistics Sweden mentioned a wide range of topics to be covered in the work of the group, amongst others the "Identification of weak points in the current ESS". There is not whatsoever doubt that the ESS still suffers from a series of weaknesses. The day-to-day experience of the LEG members as well as the group discussions however showed that the ESS is not only characterised by weaknesses but also gained certain strengths. The group therefore decided to widen the scope of the topic and to establish an up-to-date inventory of the strengths and weaknesses of the ESS. The results of the considerations are presented in the paper.



Stockholm, Sweden

The LEG considers the inventory as input to future systematic improvement actions. As not all strengths and weaknesses are considered to be of the same importance for the ESS as a whole some kind of prioritisation is needed. The paper will present those strengths and weaknesses for which improvement actions should be launched first. Also some ideas about the administration of the work will be given.

9:25 Quality of the European Statistical System: The Statistical Dividend of European Integration

by Paolo Garonna and Tiina Luige, UN/ECE Statistical Division

We would like to develop the following points in the paper:

1. International dialogue has been since the beginning of official statistics and continues to be a very powerful tool for enhancing statistical quality. How European integration has contributed to statistical quality.
2. Sources of strengths and weaknesses of European statistics: the dynamics of innovation and statistical development, the remaining gaps.
3. Within the EU statistical system: need to integrate the different components of quality as seen by users, producers, and in relation to products, production processes, data collection, dissemination and analysis.
4. The environmental factors for quality enhancement: institutions, education, interaction with the media, scientific authority, etc..

The relationship between the EU statistical system and the global system.
The challenge of Pan-European statistical integration: looking beyond the EU enlargement.



Stockholm, Sweden

Session 24 Room 357

9:05 Comparing the Effect of Different Imputation Methods for Units with Large Amounts of Item Non-response: A Case Study

by A. Cirianni, M. Di Zio, O. Luzi, A.C. Seeber, ISTAT, Italy

In the case of large scale surveys conducted by National Statistical Institutes, the integration of item non-responses represents a critical problem because of its impact on the quality of final results. This problem is particularly important in sample business surveys, generally affected by large amounts of total and/or partial missing data. A number of different methods have been proposed in literature for compensating for missing information, depending on the characteristics of both the surveyed phenomena and the target population.

A specific problem is represented by units affected by large amounts of missing information: these units could be either considered as total non responses, or their missing data could be integrated by using different techniques. In the first case, these units are handled at the estimation level by applying some re-weighting techniques for compensating for the non-respondents. In the latter case, different kinds of imputation techniques can be adopted at the data processing stage for integrating missing data using the observed partial information.

The effect on final results of different approaches of handling this kind of respondents is a crucial element to be taken into account when choosing among them. In this context, in this paper a comparative evaluation of different strategies of dealing with units affected by large amounts of item non-response is presented in the case of the 1997 Italian Labour Cost survey (LCS). The LCS is a periodic sample survey that collects information on employment, worked hours, wages and salaries and labour cost on about 12.000 enterprises with more than 10 employees. In the 1997 data, for about 800 respondent businesses the problem of dealing with poor direct or indirect questionnaire information on target variables arose. The strategy adopted for handling these units consisted of integrating their missing data by applying an *ad hoc* deterministic imputation method based on a hierarchical use of economic indicator



Q
2001

Stockholm, Sweden

estimators. The imputations were carried out among pre-defined imputation classes determined on the basis of businesses economical activity, number of employees and localisation on the Italian territory.

In the paper three methods for compensating for the missing information are compared: the currently used deterministic *ad hoc* method, a nearest neighbour donor imputation method, and a re-weighting approach where units are considered total non-responses.

The evaluation of methods is carried out by measuring their effects on the estimates and distributions of the target survey variables, both at the survey sample level and for the subset of partially respondent units under study. In particular, the variables used for the evaluation were Total number of employees, Total of worked hours, Total of wages and salaries and Total labour cost.

9:25 How to Improve the Quality of an Economic Variable in a Household Survey: A Simulation Study Through Multiple Imputation

by C. Quintano, R. Castellano, and A. Regoli, Istituto di Statistica e Matematica Facoltà di Economia, Istituto Universitario Navale, NAPOLI, Italy

According to the standard definition of the European Statistical System, the accuracy is one dimension of the concept of quality in statistics. To assess the accuracy means to study the total error associated with the estimate.

It is well known that one of the factors affecting the non sample error component is the presence of non responses. Traditionally, in the literature a distinction is found between weighting and imputation methods to compensate for non responses. The latter are most used to deal with item non responses and many different imputation techniques have been developed over the years.

The imputation methods are well considered from data analysts as they can produce complete data sets on which they can apply standard techniques. But, sometimes, they forget that the “basic objective” should be to get better estimates for the parameters of interest.



Stockholm, Sweden

Of course, the difficulties are linked to nature of the variables and, in a survey, the experience shows that the financial and economic variables are likely to be missing or underestimated.

For this reason, we focus on the data from the Bank of Italy's Survey of Household Income and Wealth which provides detailed information about household income and financial activities. The 1998 data are collected on 7147 households and in the released data set there are no missing values. This condition permits a quality analysis of the total household income by means of a simulation study. To this purpose, missing values have been artificially generated in the income variable and the effects of a multiple imputation procedure are compared with respect to: i) the original complete dataset; ii) the dataset net of cases with missing income values; iii) the dataset with missing income values filled in by a single imputation method.

Conditional on the distributional assumptions, the missing mechanism, the missing rate and the imputation method, the results are then evaluated on the basis of: i) the bias of parameter estimates (mean, median, first and third quartiles); ii) the bias of variance estimates; iii) the confidence interval width.

Several combinations of types of distribution and missing mechanisms are tested.

Moreover, the multiple imputation methods produce a decomposition of the variance in a sample component and an extra sample component, the latter representing the variability due to the nonresponse mechanism and to the imputation procedure.

In this framework, the incidence of extra sample variability on the total variability can be considered a further evaluation element.

9:45 Multiple Imputation In Government Agencies

by Donald B. Rubin, Harvard University, Boston, USA

Multiple imputation for nonresponse is, I believe, the only general purpose statistical technique that can validly handle missing data



Stockholm, Sweden

problems of the type that government agencies face. The reason why is that it separates the issues of creating a multiply-imputed data set, which is generally very demanding, from the issues of analyzing the resultant multiply-imputed data set, which are simple. Thus, an agency can be responsible for the difficult part, and the user only has to handle the easy part, which involves the repeated application of standard complete-data software. Current applications in the US involve NHANES and other data sets at NCHS, FARS at DOT, NMES of AHCPR, SCF at the Federal Reserve, CES at BLS, and various data sets at the Census Bureau. Although still requiring fairly specialized software to create the imputations, many options now exist, including commercial software (SOLAS) with plans for SAS and perhaps others to enter the arena shortly, and freeware, including Joseph Schafer's routines. The focus of this presentation will be a highly complex application: the multiple imputation of NMES. This data set has very detailed medical expenditure information, by type of service and disease category, and thus has seen extensive use in the US tobacco litigation.

Session 25 Room 353

9:05 The French Statistical Council (CNIS) Origin, Assignments and Role in Quality

by Michel Blanc and Alain Desrosières, INSEE, France

The services of official statistics are in contact with social actors in at least two different ways, either as «respondents», who complain of the burden imposed to them by various administrative forms or questionnaires, or as «users», who want more and more statistical information. Thence a certain contradiction or paradox (that psychologists call «double bind»): «Don't ask us too many questions, but give us more answers».

To go beyond this contradiction the idea has come out in France, in the 1970s, to gather two modes of consulting social actors, separated before in two committees created respectively in 1951 and in 1964, into the same organisation. It appeared at that time that there were many debates, in the process of economic and social planning, about economic information, leading to the idea that official statistics had to meet the needs of many different social actors, not only of the government. Thus a Statistics National Committee (CNS) was created in 1972, which became the



Stockholm, Sweden

National Council for Statistical Information (CNIS) in 1984. The CNIS organisation is inspired from committees for planning modernization, and consists of representatives of administration, trade unions, employers' organisations, etc. This institutional model for consulting and discussion guarantees a good representativeness and a good legitimacy.

CNIS has an executive committee («Bureau») composed of representatives of these organisations. Thirteen thematic committees or task forces (called «Formations») review, discuss and adopt statistical projects which will be carried out in the national statistical system («Système statistique public»). So CNIS, which is an advisory body, plays a very important role to assess the *relevance* of statistical operations. It is to be noted that this occurs in an upstream phase of the process, and not only in ex-post phases of surveys, for instance through customers satisfaction surveys.

CNIS gives advice on an annual program of official statistics, but also on a five year statistical program. At this occasion a large consultation is conducted, at national but also at local level. During the last consultation there was much emphasis on the need for rapid and geographically detailed data (*timeliness and punctuality*).

Another function of CNIS is to discuss which surveys are mandatory. This has an influence on the response rate, and then the *accuracy* of the results. Two committees belonging to CNIS examine respectively refusals to reply to mandatory surveys and applications for access to certain business statistical files (an aspect of *accessibility*).

In 1994, due to the increase both in the demand for relevant and cost efficient statistics and in the reluctance of enterprises (specially small enterprises) to fulfil statistical questionnaires, a new committee was created within the CNIS : the Quality Label Committee. In order to get CNIS approval, a survey project must satisfy two sets of criteria, which are examined in two steps:

- 1) «*relevance*» («opportunité» in French) : the survey meets an actual need, and there is no information already available ;
- 2) «conformity» to criteria of *accuracy, accessibility, clarity, comparability, confidentiality*, etc., and minimization of the response burden.



Stockholm, Sweden

CNIS deals with many questions related to the quality of statistical information, both in a perspective of data quality and of process quality. In particular the *relevance criterion*, which generally gets poor consideration in spite of the fact that it appears first, receives a particular attention. Thus the Council can be seen as a place where producers of statistics and representatives of social partners, considered both as users and respondents, agree on what *good official statistics* are and produce *standards of good practices*.

Of course the system is not free of weaknesses, among which one can mention a lack of spontaneity and of overview in the debates, but it has a real power of proposition (for instance a survey on the homeless, which will be launched soon, has been entirely initiated from CNIS) and it contributes to give credit to official statistics.

The question of the way social standards of good practices (enabling official statistics to be perceived as legitimate by social actors) are defined is probably insufficiently raised in the EU, in particular because the implementation of TQM models draws more attention towards other kinds of standards or management practices.

9:25 Some Data Quality Issues in Statistical Publications in Poland

by Jan Kordos, Central Statistical Office/ University of Ecology and Management, Poland

Since 1989 the Central Statistical Office of Poland has started to harmonisation of official statistics with the European Union standards. Basic nomenclatures and classifications have been adopted, the programme of surveys was considerably changed, new surveys were launched, new methods were applied and the methodology adjusted to the requirements of market economy.

Some complete statistical reporting of previous system stopped and new sample surveys started, such as labour force survey, agricultural surveys, business surveys, enterprise surveys, and several ad hoc sample surveys connected with living conditions, health care, time use, etc. Household budget survey has been adjusted to new requirements.



Stockholm, Sweden

However, data quality issues of these surveys are important problems to be solved. Some attempts in this direction has been started.

It is generally accepted that users of survey data need information about a survey's quality to properly assess survey results. It is very well known that there are many dimensions to survey quality and the measurement and presentation of this information is no easy task.

The pressure from users on quality as expressed in users' satisfaction surveys focuses mainly on timelines and availability, which does not encourage specific investigation in accuracy. There is as well strong sense among public statisticians that it is difficult to make quality understandable to non-specialists, and that it is of the professional responsibilities of statisticians to decide on necessary quality levels. The studies presented in this paper illustrate the range of the Central Statistical Office of Poland practices in reporting information on error sources in surveys.

The sources of error that affect survey data quality – sampling error, coverage error, nonresponse error, measurement error, and processing error and their measurement are described in a number of texts books published in Poland which provide a summary of methods used to measure error sources.

The author conducted three studies to help characterise current practices for reporting sources of error in publications. The first study reviewed publications and data releases of 10 pages or less („short-format reports) issued by five division of the Central Statistical Office: (i) enterprise, (ii) labour, (iii) living conditions, (iv) agriculture and environment, (v) production and services. The second study reviewed selected „analytic publications“ from the same divisions. These are printed reports resulting from a primary summarisation of on-time survey or an on-going series of surveys. The third study reviewed articles and publications of methodological nature devoted to some aspects of data quality, e.g. presentation and interpretation of sampling errors, coverage errors, non-response errors, measurement errors. In final remarks the author stresses the need for improvement of current practice in data quality presentation and concludes that similar situation may be observed in other countries in transition.



Stockholm, Sweden

9:45 Improving the Quality of Products and Services to Census Bureau Customers

by Joanne Dickinson and Gloria Gutierrez, U.S. Census Bureau

The Census Bureau has undertaken several initiatives to improve the quality of its products and services to customers. A recent review of Census Bureau quality policies and related quality assurance measures resulted in the establishment of a new office to review external communications. This paper discusses the functions of this office and initial plans. It also reviews several ongoing activities designed to inform employees about customers, their needs and their importance to the organization. Events and communications throughout the year include an annual Customer Service Week that includes customer service awards, games, presentations, training, and an exhibit contest that allows offices to illustrate their customer orientation. Information about best-selling products and leading customer groups, new product releases, and how these products are used are communicated regularly to employees. Internal surveys are used to measure changes in employee awareness of customers and agency attitudes and perceptions.

10:05 Partnership – the Keyword in Quality?

by Martin Collins & Wendy Sykes, London, U.K.

In this paper, we would seek to build on previous contributions to the discussion of quality. While we have new work to report, we intend to maintain the style of those previous contributions – speculative thinking rather than hard facts; food for thought rather than prescriptions for behaviour.

At the 1995 International Conference on Survey Measurement and Process Quality, we contributed a paper on the definition of quality from the user's point of view: "Survey quality: the user's view". We argued that basic statistical qualities, and systems for Quality Management, needed to be augmented by strong customer relationship skills and management. Here we were borrowing language and concepts from 'Marketing'.

The paper was developed for publication in the Journal of Official



Stockholm, Sweden

Statistics in 1999: "Extending the definition of survey quality". The development concentrated on establishing a context – especially based in the monograph and proceedings of the conference.

Since then, we have completed a project for ONS Social Survey Division: "Defining quality in survey supply". This involved semi-structured interviews with eight major survey customers. These interviews confirmed the fundamental importance of technical values in survey research, and of effective quality management systems. But such factors define only the 'acceptable machine'. Among the public sector users of SSD services, SSD is the model. Others, including large market research suppliers, have learned to follow the SSD standards, and can often compete on price. Differentiation – other than in terms of price – is not easy.

Clues to possible sources of differentiation arising from these interviews can best be expressed as questions about a supplier, or potential supplier:

Do they really understand our objectives?
Do they know the policy environment?
Can they offer something more than just data?
Are they researchers, rather than merely a data factory?
Can they communicate – at all levels?

All such questions can be summarised in one concept – 'partnership'. When a supplier takes on a project, they take on also the customer's problems and objectives. Customers know that problems will arise. They may not understand fully the technicalities, so solutions remain in the hands of the supplier. Still, they would rather know about problems, and contribute to solutions, than be kept in the dark.

Research suppliers (corporate or individual) can be resistant to the concept of partnership, preferring to concentrate on data collection and avoid involvement in issues of data use. But some customers may also prefer clear separation of functions. In either case, the wish for separation may on occasion be well-founded. But, more generally, we believe that it tends to arise from a naïve 'positivist' view of research, and a belief that objectivity is somehow threatened by lowering the barriers.

Partnership – a close and continuing relationship defined by more than



Stockholm, Sweden

the formal terms of contract – is the focus of our paper. We will examine both its key features and the conditions under which it is most (and least) likely to be established. Meanwhile, we plan to extend our work through interviews with suppliers about the contribution of customers.

Session 26 Room 361

9:05 Measuring Quality in the U.S. Census 2000 Dual System Estimator Using a Total Error Model

by Rita J. Petroni, US Bureau of the Census

The U.S. Bureau of the Census uses dual system estimates (DSEs) to measure census coverage error and may use them to determine U.S. congressional districts. The DSE is based on data from the original census enumeration and a quality check survey (i.e. the Accuracy and Coverage Evaluation (A.C.E.) survey) conducted after the original enumeration. This paper discusses the total error model which is being implemented to measure the quality of the U.S. Census 2000 DSE.

The DSE is subject to various components of nonsampling error, in addition to sampling error. The total error model attempts to measure these components of error. In particular the model includes model bias (i.e. correlation bias), contamination bias, ratio estimator bias, measurement error (i.e. matching error, errors in assigning residency status, fabrication error, processing error, data collection error, and error from inconsistent poststratification), and random error.

Measures of the component errors are based on a wide variety of techniques. For example, correlation bias estimation uses demographic analysis and bayesian modeling. Contamination bias estimation is based on comparisons of census data collected in A.C.E. blocks and in non A.C.E. blocks. Ratio estimator bias is estimated using a jackknife or bootstrap procedure. Some components of measurement error are measured from an evaluation followup (EFU) survey which recollects residency data in a sample of A.C.E. blocks. Modeling techniques are applied to improve the accuracy of random measurement error. Sampling variance is estimated using a jackknife technique. Missing data error is estimated using sensitivity analysis.



Stockholm, Sweden

A total error model was also implemented after the 1990 census. The model for Census 2000 incorporates modifications to reflect DSE methodological differences and modifications to improve the 1990 model. For example, the 2000 total error model includes error from inconsistent poststratification as an additional source of measurement error and contamination error. The number of clusters in the 2000 EFU survey is over twice the number in the 1990 EFU survey. We use improved methodology to obtain correlation bias estimates.

This paper describes the 2000 total error model and its components and the estimation of each.

9:25 Improving the Quality of Data Obtained from Administrative Sources: Modelling Errors with Information Obtained Through Survey Data

by Piero Demetrio Falorsi, Alessandro Pallara, Raffaella Succi, ISTAT, Italy and Aldo Russo, Università di Roma Tre, Italy

Keywords: *Administrative Data, Non sampling errors, Validation Survey, Misclassification Errors.*

Recently an increasing attention has been devoted to data obtained from administrative sources as a means for improving quality of information, especially for business, farms and institutions.

There are some important advantages deriving from the use of the administrative data:

- the statistical information may be available at relatively little additional cost from several different sources;
- the timeliness is often better than can be achieved in complex surveys;
- information is obtained without additional loading on respondents;
- because of detailed coverage of administrative data they can be useful for obtaining estimates for small domains of statistical units.

Nevertheless the use of administrative data involves a number of problems, among others: elementary units in the administrative databases often do not correspond to the definition of the statistical unit suitable for the analysis; the population in the administrative data base



Stockholm, Sweden

does not match exactly the target population because of coverage and measurement errors; when producing estimates for domains there may be problems deriving from misclassification of administrative units. In this paper we address the problem of producing estimates of parameters for a population of units reported in administrative registers which are affected from misclassification errors. These errors may introduce bias in the estimates for:

- the *entire population*, when the classification variables defines the inclusion in the target population;
- *subpopulations of units* (domains), defined from different values of the cross-classification variables.

In order to improve the quality of the estimates, we propose a methodology based on the comparison between administrative data and data obtained from a validation survey. This comparison is used for different aims (Kuha and Skinner, 1997; Ekholm and Palmgren, 1987):

- (a) evaluating the quality of the classification variables in the administrative register;
estimating the parameters of statistical models defining the relationships between the errors of the classification variables and other variables in the administrative source;
- (b) estimating the probabilities of misclassification errors of different subpopulations of units of the administrative register;
- (c) using the results of (b) and (c) for: improving the quality of the estimates and measuring their uncertainty.

We present results of an application of the proposed methodology to data of the Italian Business Register using the information obtained from the *Short-Form Survey of the Intermediate Census on the Economic Activities*.

Main References

- Kuha J. and Skinner C., (1997), "Categorical Data Analysis and Misclassification", in *Survey Measurement and Process Quality*, pp. 633-670, John Wiley, New York.
- Ekholm A. and Palmgren J., (1987), "Correction for Misclassification using Doubly Sampled Data", *Journal of Official Statistics*, 3, pp. 419-429.



Stockholm, Sweden

9:45 **On the Use of Bayesian Approach to Estimate Response Errors in National Statistical Institutes**

by Brancato G, Fortini M, and Pichiorri T, Italian National Statistical Institute - ISTAT

One of the major sources of error in a self-administered survey is represented by the response error. As well known, this error may cause an increase in the variability of the results and may introduce a bias.

Usually, the methods used to estimate the response error are based on data gathered through a reinterview. The estimation of the response variance requires an independent reinterview conducted under the same conditions of the original survey, while the estimation of the bias requires a more accurate process allowing for an unbiased observation. With such procedures the evaluation of the response error of a survey turns out to be rather costly.

Bayesian methods to estimate the measurement error have been introduced in literature, although to our knowledge never applied in the field of the estimation of response error within a national statistical institute. However, it seems that such an approach could be explored in the official statistics framework, because it does not necessarily require the reinterview, as long as some information on the errors is available.

This paper reports an application of bayesian analysis to estimate response errors on data from the pilot survey of 2001 Census of Household and Population, using information on the errors derived from the previous Census. In addition, a reinterview with reconciliation process has been carried out on the pilot survey of 2001 census, allowing to estimate response errors with non bayesian models. Therefore, a discussion on results from the two approaches will be presented. The opportunity to further investigate the bayesian method will also be considered.

Essential References

Bailar BA. (1968). Recent research in reinterview procedures. *Journal of the American Statistical Association*, 63, 41-63.



Stockholm, Sweden

Biemer PP, Forsman G. (1992). On the quality of reinterview data with application to the Current Population Survey. *Journal of the American Statistical Association*, 87, 915-923.

Gaba A, Winkler RL. (1992) Implications of errors in survey data: a bayesian model. *Management Science*, 38, 913-925.

Sinclair MD, Gastwirth J. (1996). On Procedures for Evaluating the Effectiveness of Reinterview Survey Methods: Application to Labor Force Data. *Journal of the American Statistical Association*, 91, 961-969.

Session 27 Room 359

9:05 U.S. Census 2000 Evaluation Program Quality Assurance Process

by David L. Hubble, U.S. Bureau of the Census

As with previous censuses conducted by the U.S. Census Bureau, there is an extensive effort made to evaluate the quality of Census 2000. Over 140 evaluations will be conducted. The evaluations will cover many topics, including address list development, promotion and partnerships, data collection, data quality, field operations, coverage improvement, processing, and coverage measurement.

And while the evaluations are meant to help the U.S. Census Bureau measure the quality of Census 2000 and help plan for the 2010 Census, the question of how do we ensure high quality evaluations exists. The methods employed and data sources required are nearly as diverse as the evaluations themselves. With scores of employees involved in developing and implementing these evaluations, how do we ensure that the highest standards are being applied consistently?

To address these issues and as part of the U.S. Census Bureau's Quality Initiatives Program, we developed the Census 2000 Evaluation Program Quality Assurance Process. Our objective is to make quality integral to everything we do. In general, the quality assurance process provides overall guidelines and requirements for conducting the evaluations. Specifically, the process is to build quality assurance procedures into all stages of the evaluation, from project development and implementation through review and issuance of the final report.



Stockholm, Sweden

For the most part, the quality assurance process for the Census 2000 Evaluation Program is the formalizing of many Abest practices[®] that many staff may already be employing. The quality assurance package is comprised of the following 5 components: Study Plan with Methodology, Specifications, Computer System Design and Review, Development of Software and Procedures, and Analysis and Report Writing. Within each component, step by step guidelines are provided with detailed instructions for all those with a role in the process. The roles range from evaluation manager to Census Bureau executive staff. Best practices for each role are embedded in the detailed instructions, or included as separate attachments. Also, a checklist was developed to assist evaluation managers in controlling and monitoring the quality assurance process. Finally, training sessions were conducted with evaluation managers and supervisors to review each component of the quality assurance package and address questions.

This presentation will give an overview of the Census 2000 Evaluation Program Quality Assurance process and materials. Early implementation observations will also be discussed.

9:25 Data Quality Control System for the 2001 Italian Population Census

by Angela Chieppa and Fernanda Panizon, ISTAT, Italy

Quality is one of the major issue for the next Italian Population Census which will be carried out by ISTAT in 2001: the way of collecting Census data is still based on the traditional method of using Census questionnaires, which are home delivered through about 100 thousand designated enumerators. But many innovations will hopefully improve managing the field work, data processing and dissemination, that is to say the accuracy and timeliness of the Census data.

The main innovations will be:

- Monitoring the field work and collecting the provisional data. A strong effort is made by ISTAT in preparing and organising Census operations, which will be centrally monitored, in each different phase, especially using computer nets and web facilities. The hard work of revision and summarisation of data, done by the Statistical Offices of Municipalities, will be simplified and helped



Stockholm, Sweden

through automation for checking the calculations and coherence of provisional data to be sent to the Central Census Office.

- Data capturing and coding using new technologies. ISTAT is taking into consideration the adoption of innovative technologies for data capturing and coding. These innovations - if definitively adopted - will have important consequences on data collection methodology.

- Implementation of a comprehensive informative system for data processing, projected in order to satisfy the informative needs of the data quality control system (which is the major issue of our paper). Regarding data processing a large relational database is being prepared to assure integrity and consistency of the data procedures. For the validation of the data, edit and imputation methodologies will be largely automated, using both stochastic and deterministic approaches to assure the most coherent final data.

We first consider data quality control system from the point of view of the data: both checking and monitoring data in their modification during the processing.

Different kind of data controls are to be performed:

- quantitative controls on coverage of collected data, especially by means of auxiliary forms. Auxiliary forms, filled by Statistical Offices of the Municipality, contain aggregate data on population and houses and have to be compared with the results from questionnaires to validate them.

- qualitative controls, such as the comparison with other external information and explorative analysis for the macro-level, and the identification of incompatibility for the micro-level. These controls are mainly aimed to separate systematic errors from random ones and to choose the appropriate action for the editing and imputation phases (deterministic or stochastic approaches). Among external information we consider, with a particular attention, data coming from a recent current survey on public register of population: those data are going to be used to check the coherence of Census questionnaire in terms of structure of population by sex, year of birth and marital status.

From another point of view, we are interested in monitoring the execution of automated procedures, through the analysis of specific



Stockholm, Sweden

indicators (as number of imputations, percentage of errors by variables or by observations).

Each phase of data processing is going to be controlled and it will produce specific “register of quality” for the storing of indicators, such as the kind of checks, the results of controls, the corrective actions and the final outcomes. The evaluation of data quality will be obtained analysing the whole set of those registers, and building on them both analytic and synthetic indicators.

Data quality control system is now being tested with the data of a pilot survey (April 2000): some results will be presented in the paper.

9:45 Quality Assurance Challenges in the United States' Census 2000

by David Whitford and Jennifer Reichert, U.S. Census Bureau

The United States' Census 2000 operations were more innovative and complicated than those of previous censuses. We made use of many advances in automation, and technical progress introduced the need to hire contractors to undertake specialized duties. The increased automation and contracting presented many new challenges to ensure quality for Census 2000. This paper will, by presenting these challenges and our solutions, provide an representative overview of the entire quality assurance (QA) program for Census 2000.

Timing of the U.S. census is very tight. The Census Bureau is required to produce state population totals nine months after census day and, using its coverage measurement survey, adjusted counts within one year. So, to meet those restrictions, all operations must be implemented and completed quickly, and QA must be effective and have prompt turnaround. One QA challenge was to get timely information to supervisors, do prompt checks of “suspect” work, and provide reports to managers at headquarters quickly.

For interviewing operations, we used two QA tools: a random reinterview to check each enumerator's work at the outset and an administrative reinterview program to continuously target enumerators needing attention.



Q 2001

Stockholm, Sweden

Both reinterview programs used automated selection to ensure accuracy and efficiency. For the random reinterview, the system selected random cases from every enumerator's work for reinterview. For the administrative reinterview, the system compiled statistics such as vacancy rates, delete rates, and one-person household rates for each enumerator. The automated selection system then used statistical process control methodology to flag enumerators who were "out of control" for those statistics.

In the Census 2000 coverage measurement survey, the random and administrative reinterview programs were further automated by making use of a Computer Assisted Personal Interviewing (CAPI) environment. Completed interviews were downloaded to headquarters each evening. The next day, supervisors received QA summary data on their laptop computers.

After interviewing, we matched the coverage measurement interviews with the census interviews. In Census 2000, this operation was paperless and, therefore, very quick. We specified a very complex, completely automated QA system from which we could derive the operation's outgoing quality.

In Census 2000, for the first time, respondents were given the opportunity to seek help or respond to the census by telephone. This operation was massive – contractors needed to handle 11 million calls spread out amongst 13 contractors over approximately 13 weeks. The quality challenge was to monitor the results of the contractors' QA programs by gathering information from 22 sites and using that information effectively.

Finally, all census data were captured using imaging, Optical Mark Recognition (OMR), Optical Character Recognition (OCR), and keying. The Census Bureau contracted out these immense and highly technical tasks. The Census Bureau needed to develop a QA program that would ensure that the end product of the automated data capture system was acceptable. We developed a system that monitored a sample of the images and data from the contractor's output to determine if the captured data were accurate.



Stockholm, Sweden

10:05 **Data and Process Quality in the Italian Consumer Expenditure Survey**

by P. Calia and C. Filippucci, Dipartimento di Scienze Statistiche, Università di Bologna, Italy

Italian National Statistical Institute (Istat) is developing data quality control strategies. In this context the present paper deals with the production process of consumer expenditure data with particular emphasis on the field operations.

The starting point was the framework proposed by Morganstein and Marker (1997) for a continuous quality improvement plan even if one of the final goal of our work is to analyse his feasibility, in other words, our aim is finding out arising problems when such framework is applied to a specific and quite complex survey. Is to be stressed that data collection is not carried out directly by Istat but by municipalities involved in the sample.

The problem is faced starting with an accurate description of all phases and operations and their sequence. For each phase is identified and analysed procedures, standards and available documentation, the subjects involved, conformance of operations to the procedures and problems in practical implementation. Moreover we tried to evaluate the above aspects in terms of some quality dimensions chosen among those more frequently used in the literature (Brackstone, 1999) and appropriate to the field operations.

The second step is the selection of the key variables of the process to keep under control and suggesting some strategies for measurement.

Then measurement capability of the process is investigated.

On the basis of the results obtained a tentative proposal to establish the basis for a continuous monitoring of the process is presented.

References:

- Brackstone G. (1999), Managing data quality in a statistical agency, *Survey Methodology*, vol. 25, pp 139-149.
- Morganstein D., Marker D. A. (1997), Continuous quality improvement in statistical agencies, in Lyberg et al. (eds), *Survey measurement and process quality*, New York, Wiley, & sons Inc



Stockholm, Sweden

Session 28 The Auditorium

11:05 LEG chapter on The interrelationship of different Quality Management Mainframes

by Mats Bergdahl, Max Booleman, Richard van Brakel and Werner Grünewald

The objective of this paper is to inform about different basic Quality Management approaches used in (statistical and non-statistical) organisations, aiming at the improvement of their performance. Different frameworks will be presented, like Quality Control, Total Quality Management (TQM), the EFQM model, the ISO 9000 family, the Business Process Redesign (BPR) approach, the Balanced Scorecard and Six Sigma. The presentation will include a description of their strengths and weaknesses and shows to which extent the different concepts overlap.

11:25 New Tools for a Strategic Quality Improvement – Statistics Finland's Experience

by Heli Jeskanen-Sundström, Statistics Finland

In order to make real and long-term quality improvements in the organisation, the strategic meaning of quality has to be recognised. The future challenges cannot be met by using the old strategy tools. The Total Quality Management (TQM) approach has become more and more common in the National Statistical Institutes. TQM put emphasis – not only on statistical quality of products but more widely on management system as a whole, on processes and on people working in these institutions. The essential part of the TQM approach is to make self-assessment of the whole organisation and its development regularly. Statistics Finland started to use the Quality Award frame as an TQM assessment tool in 1998.

At the same time, we started to develop the new strategic management system based on the Balanced Scorecard (BSC) –method originally launched by Kaplan and Norton. Quality Awards and BSC have been born in different places, but actually they work very well together. According to the BSC, the success of the organisation depends - not only on its economic and financial resources and potential - but on its ability



Stockholm, Sweden

to combine economic goals to the strategic goals related to customer satisfaction, internal processes as well as competence and renewal of its staff. BSC is an excellent management tool – applicable also on unit level - for communicating and implementing the strategic goals and necessary operations in the whole organisation.

The paper will discuss the ongoing process of integrating both the TQM and the new strategic management system into a single entity. It shows how the self-assessment can support the strategic planning and how the TQM based quality work has become an essential element of the overall annual and longer-term planning process of Statistics Finland.

11:45 Achieving an ISO Quality Standard for Survey Research

by Bill Blyth, TaylorNelson Sofres, U.K.

This paper describes the progress that has been achieved towards the agreement of the specification for an International Quality Standard for survey research

The concept of quality in survey research contains a number of dimensions. Collins and Sykes* identify distinct dimensions of design, resource, process management, service standards and relationship in their paper. Whilst some of these dimensions are difficult to define or are specific to a particular survey others are more general and definable. In particular this applies to the dimensions of process management and service standards.

This paper describes the development, over the last twenty five years within the commercial survey research area, of formal standards covering process and service standard quality.

From early standards in the UK and Netherlands specific to personal interviewing, it will explore how the effects of more formal management, quality assurance, increased internationalisation and the impact of legislation have led to a wide acceptance of the need for a formal international standard.

It will discuss which aspects of the survey process are suitable for this



Stockholm, Sweden

approach and will describe the difficulties in drafting and agreeing minimum standards.

It will draw on the European Quality Standard that has been agreed by EFAMRO (the European Federation of Market Research Organisations) to give examples of the coverage of such standards.

It will explore how the need to have external assessment of a consistent quality can be achieved both between accreditation agencies and between countries.

It will describe the steps that are now underway to move to a formal ISO Quality Standard that will have global coverage. It is believed that this would be the first service industry specific quality standard and thus will act as a template for others.

It will discuss the implication of such a standard for the ways in which international surveys will be carried out.

It will discuss the relevance of such a standard for National Statistics Agencies or other suppliers of Official Statistics.

*Collins and Sykes, Extending the Definition of Survey Quality; JOS 15, 1, 1999

Session 29 Room 456

11:05 Measuring Customer Satisfaction

by Claes Cassel, Statistics Sweden

Customer satisfaction has been in focus for some years. Earlier this was mostly a concern for the private sector where satisfied and loyal customers obviously play a crucial role in generating profit for the companies. Measuring customer satisfaction has thus become a major concern for customer oriented companies. It should be measured in such a way that it makes it possible to suggest measure for increasing the customer satisfaction and loyalty. This is often achieved by using an approach involving structural equation models with latent variables. The



Stockholm, Sweden

model is a visualisation of the underlying process of how customer satisfaction and loyalty is formed.

In recent years the services offered by official authorities and institutions such as for example health care institutions, elderly care institutions and child care institutions have become the object of studies using a similar technique. In this case the focus of interest has shifted from customers to users. Measuring customer satisfaction here means evaluating the users perception of the quality of the services offered. The paper gives a presentation of the technique used and some studies conducted by Statistics Sweden. The Swedish Customer Satisfaction Index is described as well as some specific institutional study.

11:25 Measuring Customer Satisfaction with U.S. Census Bureau Products

by John Kavaliunas and Gloria Gutierrez, U.S. Census Bureau

Since 1993 the Census Bureau has conducted over 40 surveys to measure customer satisfaction with various products or to evaluate various customer-related programs. Our purpose in doing so is to continually upgrade and improve the quality of our products and services, determine how customers are using them, and to develop ongoing measures of customer satisfaction. In addition to traditional surveys through the mail, we have also used the Internet, focus groups, and beta testing with groups of users.

In 1999 and 2000, the Census Bureau participated, along with 30 other government agencies, in the American Customer Satisfaction Index, to develop an overall measure of customer satisfaction with our products and services that would be comparable to similar indices in the private sector.

This paper discusses our experiences and lessons learned in

- o identifying customers
- o asking the right questions
- o taking customer surveys
- o gathering market intelligence
- o monitoring trends, user preferences, and overall customer satisfaction

translating findings into actionable quality improvements.



Stockholm, Sweden

11:45 **Did You Enjoy Our Service? Measuring Dissatisfaction in Customer Satisfaction Surveys**

by Peter Ph. Mohler, ZUMA Mannheim, Germany

A. Importance of customer satisfaction surveys

It is common knowledge among practitioners and theorists of quality management that it is much cheaper to keep customers than to get new ones and, more important, that nothing is in the long run as damaging as mouth-to-mouth propaganda of dissatisfied customers. Hence, knowing about your customers' demands, critique and satisfaction is a crucial element of modern management. The core issue of customer satisfaction surveys is, actually, not 'satisfaction' but 'dissatisfaction'. Managers are interested in weaknesses of their own business in order to counter problems before they result in customers exit. Today, the notion of customers is no more confined to business and industry. Government, non-profit organizations, state-agencies, universities etc. they all want to know about their voters, sponsors, clients, students etc. Customer satisfaction surveys is one of the most important means to acquire information about customers. They may already outnumber other sorts of surveys.

B. Measurement problems

'Did you enjoy your meal?' - how often does one respond 'no, not at all'? General laws of communication (politeness, face saving, small talk) do not allow to use such a simple and straightforward question as a reliable measure of satisfaction or dissatisfaction. It is well known that in almost all cases Lickert-type rating scales are skewed in customer satisfaction surveys to the positive side. Another problem may be, that the client is asked about a product or service which she/he is not using recently or had no problems in the last weeks. Hence, problems might be forgotten. Thus proper timing of the interview is at least as an important issue as asking the right questions.

C. Benchmarking

The one-eyed is the king among the blind. In other words, all results for a specific agency or company are relative to the overall performance, image and expectations in a specific service sector. Satisfaction scores with politicians might be lower than with show stars, to give an extreme example. Benchmarking, the comparison of one agency's scores with the top leaders in a certain sector requires good and robust indices. Such



Stockholm, Sweden

indices are routinely provided by professionals in the customer satisfaction survey business. However, in quite a number of cases, the index construction is opaque, even a top company secret. Without knowledge about the ingredients, mechanisms and methodological considerations such benchmark indices are suspicious, indeed.

The paper concentrates on technical and methodological issues in measuring and benchmarking customer dissatisfaction by means of surveys.

Session 30 Room 357

11:05 Measuring and Reporting Data Quality in U.S. Federal Government Data Collection Programs

By Lee Giesbrecht et al, Bureau of Transportation Statistics, USA

In 1996 the United States' Federal Committee on Statistical Methodology (FCSM) established a subcommittee to review the measurement and reporting of data quality in federal data collection programs. The issues contained within this broad mandate are complex. Measuring the quality of survey data takes on different meanings depending on the constituency. Different data users have different goals and, consequently, different ideas of what constitutes "quality." Similarly, the reporting of quality can be implemented quite differently depending on the type of data product being produced.

The FCSM subcommittee, whose membership represents the experiences of twelve U.S. statistical agencies, approached this topic by asking: 1) What measurement methods do federal data collection programs use to assess sources of error?; 2) To what extent do federal data collection programs report information on sources of error to the user community?; 3) How does reporting about error sources vary across different types of publications and dissemination media?; and 4) What information on sources of error should federal data collection programs provide and how should they provide it? The subcommittee completed three studies that focused on reporting sources of error in each of three types of data products; short-format reports, analytic reports, and the Internet. These studies were presented at a conference on U.S. government statistics in November 1999.



Stockholm, Sweden

In its present work, the subcommittee has completed a working paper that discusses data quality in terms of the measurement and reporting of various error sources that affect data quality: sampling error, non-response error, coverage error, measurement error, and processing error. The subcommittee's report discusses the measurement of the error, providing a brief discussion of the measurement techniques and methods used; then the nature and extent of current practices for reporting information about the error source; and, finally, its recommendations for measuring and reporting about survey error. This paper will summarize the results of the studies on current reporting practices and discuss recommendations to improve reporting of information on the various sources of error.

MEMBERS OF THE SUBCOMMITTEE ON
MEASURING AND REPORTING THE QUALITY OF SURVEY DATA:

Daniel Kasprzyk, Chair, National Center for Education Statistics (Education)

Dale Atkinson, National Agricultural Statistics Services (Agriculture)

Judy Conn, Center for Disease Control and Prevention (Health and Human Services)

Charles Darby, Agency for Healthcare Research and Quality

Howard Hogan, Bureau of the Census (Commerce)

Nancy Kirkendall, Energy Information Administration (Energy)

Lee Giesbrecht, Bureau of Transportation Statistics (Transportation)

Marilyn McMillen, National Center for Education Statistics (Education)

Renee Miller, Energy Information Administration (Energy)

Chris Moriarity, National Center for Health Statistics (Health and Human Services)

Dennis Schwanz, Bureau of the Census (Commerce)

W. Karl Sieber, National Institute for Occupational Safety and Health (Health and Human Services)

Antoinette Ware-Martin, Energy Information Administration (Energy)

John Wolken, Federal Reserve Board

Graham Kalton, Senior Advisor, Joint Program in Survey Methodology and WESTAT



Stockholm, Sweden

11:25 **Quality Assessment of Statistics in Eurostat**

by Harald Sonnberger and Håkan Linden, Eurostat, European Commission, L-2920 Luxembourg

Keywords. Data quality, quality reports, legislation on quality, European R&D framework programmes, current best practices.

Since 1994 Eurostat has developed its own approach to the quality measurement. It has in particular benefited from the experience of many organisations and from the explicit support of several national statistical offices of the European Union. Now, Eurostat has a structure for the management of quality in statistics and a single reference framework for the definition of quality in statistics and for the reporting on quality. However, many activities have to be undertaken for the measurement and evaluation of the quality of statistics.

This paper reviews some of the activities undertaken and key challenges in the European Statistical System (ESS) on the assessment of the quality of statistics.

In 1998 Eurostat modified the terms of reference of a working group on the quality evaluation of structural business statistics, which consists of methodologists of the National Statistical Institutes and which had been established in 1995. The role of the group is to harmonise the definition of quality in statistics, to standardise quality reports, to deal with methodological problems for measuring the quality of statistics, and to co-ordinate the work on quality and all activities related to quality within Eurostat and the National Statistical Institutes.

In parallel, many activities have been going on in the ESS. Some of the milestones that had or will have particular influence for the work on the assessment of quality in statistics are:

- The establishment of Task Forces for assessment of quality and quality policies.
- The Commission Regulations concerning the criteria for the evaluation of quality of specific subject domains.
- The results of the work of the Eurostat's Leadership Group (LEG) on Quality, which develops recommendations on how systematic quality improvements should be launched and implemented in the ESS.



Stockholm, Sweden

- The Eurostat quality initiative, Qualitat, which includes a module on Eurostat internal quality reports that, should be provided for all domains.
- The Statistical Programming Committee (SPC) recommendation that the quality of estimation methods should be certified and co-ordinated by the Working Group on Quality Assessment.
- The Framework Programmes. The research projects into statistics that are funded under EU's research programmes.

Since the theoretical framework for the assessment of quality in statistics have been agreed, the priority of the work of Eurostat, with the Member States, is to establish standard methods for assessing the quality. This implies communication and co-ordination between the actors of the European statistical system and harmonised solutions to the methodological problems encountered. The planned network of methodologists from the NSIs and Academia will play an important role for development of manuals of current best practices.

11:45 **Statistical Quality Assessment in Korea National Statistical Office**

by Dong-Myeong Lee and Aelee Shon, Korea National Statistical Office

The purpose of this paper is to describe the progress of KNSO's (Korea National Statistical Office) work for assessing the quality of statistical data and for institutionalizing the procedure of assessment. Recently, the quality of data is a crucial issue for many statistical institutions in the national and international level. KNSO also has been recognized the importance of statistical quality and made an effort to improve accuracy and confidence of data. However, attention to data quality is highly emphasized since 1997 with the financial difficulty in Korea. The socioeconomic situation under the financial crisis needed the reliable data related to employment and other major economic sectors. Therefore consideration to data quality is more gathering strength as a central matter for the management of a KNSO than ever before.

Even a number of statistical agencies have offered presentations containing the definition of statistical quality, but they have yet to reach consensus on the conceptualization of the term. In this context, this paper



Stockholm, Sweden

deals with the definition of quality firstly. Based on general reviewing the definition, five dimensions of quality have been identified: relevance, accuracy, timeliness, accessibility, and comparability cost efficiency. For each dimension of quality consists of several quality elements. And each element is composed of selected indicators.

This paper presents the dimensions and indicators that have been developed and applied tentatively in Korea. It also introduces the procedure of institutionalization and the statistical quality assessment system in Korea. KNSO established a new quality assessment system in 1999 and conducted several pilot assessments. From the results, it describes the Customer Satisfaction Survey method and procedure that has been done by KNSO. In the final section, it introduces the lessons learnt from the quality assessment practices in Korea.

Session 31 Room 353

11:05 Data Validation Through Measurements on Conditional Distributions of a Logically Related Group of Variables.

by George Petrakos and Gregory Farmakis, Liaison Systems SA & Panteion University of Athens

In the domain of official statistics, information is delivered to the final information consumer at the end of a several-stage life cycle starting at data collection from various possibly non-homogeneous sources (i.e. questionnaires of various forms, administrative sources etc.). No matter what the information quality assurance processes are at the downstream stages (such as aggregation, compilation of metadata or publication and delivery), it is the quality of these collected primary data, which inherently defines the quality of the information to be delivered.

Data validation, i.e. the process of detecting erroneous, or probably erroneous, values within statistical data sets, is normally done through the application of validation rules, such as ranges or deterministic correlations among variables. The presented research work on the contrary, is based on the concept of probabilistic validation treatment, based on the conditional probability distribution of multi-dimensional



Stockholm, Sweden

random variables. After the probability space is defined, the conditional probabilities are calculated from historical data and/or prior information, and an estimation of the underlying probability structure is derived. Thus, a generic and consistent validation treatment is presented, which can be applied regardless of the semantics of the statistical variables, eliminating the overhead of identifying and formally defining variable-specific validation rules. These concepts, the suitable statistical methodology as well as the software architecture to implement it are presented in this paper.

Moreover, the application of the presented methodology has to be supported by intelligent data mining tools, capable of (i) calculating the probability distributions of the multi-dimensional variables based on “clean” data sets (ii) processing the data sets under inspection against the derived distributions. Obviously, due to the significant number of dimensions as well as the important volume of statistical data, this process can be demanding in terms of both computing and data management, therefore requiring a specifically optimised software architecture. The software architecture presented here concerns: (i) the design of a novel data warehouse architecture, which is “aware” of the existence of multi-dimensional variables, inherently underlying the usual record based structure of the statistical data sets; (ii) the design of the required flexible data mining tools capable to determine the probability structures of the variables; (iii) the design of a validation engine capable of applying the validation procedure.

The methodological and software concepts presented here are being implemented and validated within the INSPECTOR project, funded by the EC's 5th Framework Programme.

11:25 **Quality of Very Large Databases**

by William E. Winkler, Bureau of the Census

Keywords: record linkage, editing, data mining

Analyses and data mining of large computer files are affected by the quality of the information in the files. For large population registers and for files that are created by merging two or more files, duplicate entries must be identified. Duplicate identification can depend on record linkage software that can deal with name, address, and date-of-birth data



Stockholm, Sweden

containing many typographical errors. Quantitative and qualitative data must be edited to assure that mutually contradictory or missing items are changed automatically and quickly. This talk describes computational methods that are suitable for groups of files where individual files contain between 1 and 500 million records.

11:45 **The Quality Control of the Statistical Process and the SODAS Software – A Case Study**

by Carlos Marcelo, INE, Portugal

Keywords: SODAS, symbolic objects, multivariate data analysis and quality control.

SODAS (Symbolic Official Data Analysis System) software package has been developed in the context of the *European Esprit Program*, from 1997 to 1999.

SODAS is a set of statistical methods for analysing and visualising complex, aggregated, relational and higher-level data. It extends classical methods of exploratory data analysis to the case of the so-called *symbolic variables*, which may have values such as *subsets* of categories, *intervals* or *frequency distributions*. In the classical approach, only a *single number*, *category* or *level* is allowed.

The aim of the SODAS software is to facilitate the use of numerical symbolic data analysis techniques in statistical offices and companies, and consequently demonstrate that these techniques meet several user needs: analysis of data with complex structure; better explanations of statistical results; concepts and metadata representation, manipulation and analysis.

The software environment includes generic tools for storing, querying and updating *symbolic objects* and a collection of data analysis methods. Symbolic objects form the basic data structure for representing complex data, metadata or knowledge to be analysed. Data analysis methods, dedicated to the symbolic objects, include: univariate descriptive methods, clustering, decision-tree construction, discrimination and factorial analysis.

This paper presents an application of SODAS to the quality control of the statistical process.



Stockholm, Sweden

12:05 **Action Workflow Analysis: A Methodology for Statistics Production Process**

by Stefania Bergamasco, Marina Gandolfo, Susanna Terracina, and Antonio Toma, ISTAT, Italy

Keywords: statistical production process, process modeling, workflow analysis, continuous improvement, business process reengineering

In the perspective of process improvement projects process modeling is a fundamental concept. Modeling work processes and associated data and information needs are used in two ways: first to define, understand and communicate the way functional processes currently work (the AS-IS condition); and then, following improvement analysis, the way these same processes could or should work (the TO-BE condition).

The main purpose of this paper is to discuss the application of Action Workflow Analysis (AWA) in order to design the statistics production processes. We have applied this kind of process analysis to a statistical production process chosen as case study. The case study has been focalized on the "labour force survey" with a particular attention on "the new interviewers network".

Session 32 Room 361

11:05 **Quality Performance Measurement in the Computation of Consumer Price Indices**

by Rosa Bernardini Papalia and Pierluigi Daddi, Dipartimento di Scienze Statistiche, Università di Perugia, Italy

In this paper we propose a systematic approach to measuring quality performance connected to the computation of consumer price indices (CPI). The idea is to describe a general strategy able to evaluate quality of CPI by translating principles of "organizational performance" into main macro-phases of the consumer price indices production process (CPIPP). From this point of view the objective is to optimize each macro-phase process in order to improve CPI production process quality. The result is improved processes as measured by improved performance.

According to this prospective we propose process performance



Stockholm, Sweden

indicators associated with Strategic, Tactical and Operational “Zones of Measurement“ described in Dransfield, Fisher and Vogel (1999). In this context, ‘Strategic’ relates to how the outside world judges the success of CPI, while ‘Tactical’ relates to how the goals of the measurement process are to be implemented. Finally, ‘Operational’ relates to the implementation of CPI.

We start by identifying the main CPIP macro-phases: 1) definition of the nature of a consumer price index; 2) identification users of a consumer price index; 3) definition of scope of the index; 4) definition of calculation of elementary aggregate indices in term of kinds of goods or services, types of outlets and regions, formula; 5) definition of weights; 6) definition of sampling methods of: geographical localities, outlets, items and varieties in each outlet; 7) definition of standard methods for collecting and processing price data; 8) definition of CPI diffusion strategy. In order to analyze the impact of each phase in the CPIP we determine a minimal set of Key- Performance Indicators (KPI’s) or key process output variables that collectively predict the quality performance of the CPIP. The KPI’s are grouped according to the Strategic, Tactical and Operational “Zones of Measurement“. At the first level they give an external success measure while from the Tactical and Operational point of view they represent an evaluation with respect to the objectives of the organization (National Statistical Institute) and with respect to the implementation of the CPIP, respectively. We then proceed by identifying adequate In-Process Variables (IPV’s) or Process Input Variables at strategic, tactical and operational levels that allow us to acquire an efficient knowledge of the Key-Performance Indicators previously defined. It follows that there are different IPV’s for each macro-phase of the CPIP.

The final aim of the paper is to develop predictive models for the main success measures based on the KPI’s. A predictive model represents a particular dimension-oriented quality approach as accuracy improvement or costs reduction. Each model, enable to simulate of some changes in the CPIP verifying the quality improvement. For the analysis of the models it is necessary to build scenarios for each strategy, to define in-process variables for each scenario and to study the behavior of the model under the different scenarios. The final evaluation of success measures relative to the compared scenarios is given in the term of Key Performance Indicators of interest.

The advantage of this strategy is related to the possibility of identifying



Stockholm, Sweden

the pros and cons of each phase and also key-problems inherent to particular quality dimensions. To describe and to improve the quality performance in the computation of consumer price indices, a set of measures is required, grouped in categories depending on their purpose and the idea is to obtain a global measure of success for the CPIPP.

Key Words: Performance measurement; Quality improvement; Key performance indicators; Consumer Price Indices; Official Statistics

11:25 Factors Affecting Quality of Statistics on Environmental Expenditures by Companies in Belgium

by Bruno Kestemont, Statbel, Belgium

A new generation of statistics is needed for sustainable development policy, as it is the integration of various disciplines like economy, sociology and environment. Due to the emergence of new «inter-disciplinary» concepts in statistics the quality of answers has become more important than the statistical quality of the surveys needed. The paper analyses the effect of the different enquiry phases on the quality of statistics for a limited number of pilot economic-environmental variables. 3000 Belgian companies were interviewed on these variables within the framework of the Structural Business Surveys. The status and underlying methods of these surveys varied according to the used parameters and the company types: mandatory or facultative, simplified or detailed, census or sample, variables with high or low occurrence, paper questionnaire with or without a follow-up by telephone.

The statistical error is high for environmental investments, mainly due to the sampling method and the low occurrence rate. Calibration methods could improve the quality of the results for this kind of variables. The answer («measurement») error is high for new concepts like «current environmental expenditures» or «investments in integrated technologies». Telephone calls to firms not having declared any environmental expenditure resulted in an important rate of corrections of the original paper answers.

In order to get an optimal quality of results for a given budget, statisticians must always balance the largest sources of errors: measurement



Stockholm, Sweden

errors and statistical errors. On the one hand, heavy questionnaires accompanied with telephone calls, provide better quality answers but also prohibitive statistical errors after extrapolation. On the other hand, simplified questionnaires, sent to a large sample provide low statistical errors but also a lower quality of answers. Even for apparently popular concepts like «environment» or «waste», a period of «vulgarisation» of their meaning in statistics is necessary to reach a satisfying quality.

11:45 Quality Management of the Processes Relating to the Production of Consumer Price Indices (CPIs): The UK Experience

by David Fenwick & Graham Tippen, Office for National Statistics, U.K.

The National Statistician has a public duty to ensure proper quality assurance of all UK National Statistics including the monthly Retail Prices Index. The latter is the main domestic measure of consumer inflation used both for the management of the economy and for the indexation of state benefits and index linked gilts. Once published it is never revised. The basic data used to construct the RPI is also used in the construction of the Harmonised Index of Consumer Prices.

The quality framework for the RPI is defined as “a process of continuous improvement, systematically evaluated against customer requirements” and achievement of this goal is assessed through an evaluation of accuracy, timeliness, efficiency and relevance measured against pre-set targets.

In this context ISO 9000 has provided a key element for a formal quality management system for the monthly production process. It has therefore become an integral part of an ongoing programme designed to maintain and improve the quality of the United Kingdom’s Retail Prices Index. Other elements include benchmarking against other National Statistical Institutes and early in 2001 the use of the EFQM Excellence Model.

This paper reviews the meaning of quality in the context of statistical outputs, the different approaches available to quality management, the potential these have to improve business performance and how the latter can be measured.



Stockholm, Sweden

It then goes on to discuss the use of ISO 9000 as a quality management tool in a statistical production area and examines its value as a driver and controller of quality. The paper also describes some of the benefits that have accrued more generally from using a formal Quality Management System and conversely addresses some of the disadvantages. Business performance is considered in the context of the ultimate quality of the consumer price index.

Within this framework the paper also reviews the role of benchmarking a CPI against CPIs produced by other National Statistical Institutes. It then looks at the benefits that can be achieved from application of the EFQM Excellence Model and how the latter can be used in conjunction with and to assist in adhering to ISO 9000 standards. The question is then raised whether there is scope for using these techniques for statistics such as the European Harmonised Index of Consumer Prices where comparability between countries is particularly important.

Session 33 Room 359

11:05 Developing New Quality Indicators in Social Surveys

by Lucy Haselden and Amanda White, Office for National Statistics, UK

Mean Square Error is often seen as the best overall measure of the quality of survey data but it is extremely limited in its uses because it relies heavily on a 'true' population figure being available to compare the data with and can only be produced once the survey is finished. Response rates are often quoted as a more practical alternative: they give a good indicator of how reliable the data is likely to be and can be produced in a timely way. However, response rates only examine one aspect of survey quality, albeit an important one and are often recorded differently across different surveys and organisations.

In Social Survey Division of the UK Office for National Statistics, we have been looking for alternative ways to report on quality. In order to achieve this we have mapped out all the processes involved throughout the lifetime of a survey, from sampling and questionnaire design through to producing a finished report. Having done this, we have been able to



Stockholm, Sweden

find quality indicators for many of these processes. By using this approach, we hope to be able to appraise any changes to our processes as well as informing our customers of the quality of the work we carry out. The paper will look at how we have approached the search for quality indicators and describe some of the indicators that we propose.

11:25 Review of Data Quality Indicators in Health Surveys

by Johanna Laiho and Kari Djerf, Statistics Finland

The purpose of this paper is to review and compare data quality indicators used in different major health surveys that concentrate on general population at national level. The outcome of this review is then related to a broader discussion of the development of survey quality indicators taking into account the specific needs of health surveys.

The development of meaningful, comprehensive and informative data quality indicators to meet the needs of the data collection organisations and the data users is a major task requiring continuous development and review. In recent years, much effort has been made to start developing, tailoring and implementing guidelines for data quality in major data collection organisations. However, huge variation persists for example in terms of how data quality is understood, how the data quality indicators are derived, and in the level of integration. In addition to this, there are unsolved questions between balancing the survey specific needs and applying general recommendation when aiming for comparability of data quality measures over time, across surveys and survey organisations and for cross country comparisons.

In major health surveys, compound survey designs, complex data collection structures and use of mixed data collection modes impose even further demand for inclusive and adequate data quality indicators. Even though; health surveys tend to be cross sectional surveys, the data collection consists often of multiple stages, which may cause further data quality problems and forms an essential part to be monitored. Commonly, in health surveys data is collected in the first stage by professionally trained interviewers followed by one or more visits by health professionals like nurses or doctors. In addition to the interview



Q
2001

Stockholm, Sweden

data, components like self-completion questionnaires, anthropometric measurements, medical samples and tests are obtained. These add to further stages to the data collection. More over, the interviewer, nurse and doctor effects may contribute to the error of the survey estimates.

We review how all these aspects are taken into account in different health surveys and how differently the multiple aspects of data quality are emphasised.

11:45 Quality Measurements of the Canadian Discharge Abstract Database

by Ann Brown and Julie Richards, Canadian Institute for Health Information

The Discharge Abstract Database is one of the key data holdings held by the Canadian Institute for Health Information (CIHI). The Institute is a national, not-for-profit organization responsible for providing health information and for developing tools to strengthen and improve Canada's health information system. In 1994, CIHI was mandated to develop and maintain the country's comprehensive health information system serving to shape policies and decisions by the public and by leaders in the health care sector.

The Discharge Abstract Database contains acute care discharge data from most Canadian hospitals. The data generated are essential for determining, for example, the number and types of procedures, the relative consumption of services, and the length of hospital stays.

CIHI is conducting a data quality study of selected clinical administrative data from the Discharge Abstract Database (DAD). This study is evaluating and measuring the accuracy of the DAD by returning to the original sources of information across Canada and comparing this information with what exists in the CIHI database. This study, the first of its kind at a national level, will occur annually, over a three-year period. The study is aimed at measuring discrepancies, identifying sources of error, and providing users with reliable information about its data quality.



Stockholm, Sweden

The study features a two-stage sample design, with a representative sample of facilities and abstracts. The reliability of the estimates of the coding differences will be calculated for selected health and perinatal indicators.

This paper describes the methodology of the DAD data quality study and the quality measurements available from the first year of the study.

12:05 Design and Estimation Innovations in the Medical Expenditure Panel Survey for the Measurement of Health Care Expenditures and Quality

by Steven B. Cohen, Center for Cost and Financing Studies, AHRQ, USA



Efforts are underway in the Department of Health and Human Services towards the development of a national health care quality reporting system. The purpose of the reporting system is to provide an annual profile of the nation's quality of care and to help measure improvements over time. Quality is often defined as meeting customers' expectations. Consequently, the quality reporting system will need to include a comprehensive set of indicators that characterize several dimensions of patient satisfaction and consumer satisfaction with providers, health plans and access to care. This paper will focus on statistical and methodological design innovations in the Medical Expenditure Panel Survey (MEPS) to ensure the underlying objectives, subject to cost constraints, are realized. More specifically, a detailed discussion of the criteria that were considered in the identification of medical conditions that would be given special emphasis for healthcare quality measurement in the MEPS is provided. Attention will be given to sample design strategies, estimation issues and data collection strategies to improve the quality of resultant survey data, particularly in terms of precision of survey estimates and sample yields. Finally, the methodologies used to identify a set of additional patient satisfaction and health care quality measures for inclusion in the survey are also discussed.



The Medical Expenditure Panel Survey was designed to produce national and regional annual estimates of the health care utilization,





Q
2001

Stockholm, Sweden

expenditures, sources of payment and insurance coverage of the U.S. civilian non-institutionalized population. The MEPS includes a survey of medical providers, to supplement the data provided by household respondents. The design of the MEPS survey permits both person based and family level estimates. The scope and depth of this data collection effort reflects the needs of government agencies, legislative bodies, and health professionals for the comprehensive national estimates needed in the formulation and analysis of national health policies. The survey is sponsored by the Agency for Healthcare Research and Quality (AHRQ) with co-sponsorship from the National Center for Health Statistics (NCHS).

The MEPS collects data on the specific health services that Americans use, how frequently they use them, the cost of these services and how they are paid, as well as data on the cost, scope, and breadth of private health insurance held by and available to the U.S. population. MEPS is unparalleled for the degree of detail in its data, and its ability to link health service medical expenditures and health insurance data to the demographic, employment, economic, health status, utilization of health services, and other characteristics of survey respondents. Moreover, the MEPS provides a foundation for estimating the impact of changes in sources of payment and insurance coverage on various economic groups or special populations of interest, such as the poor, the elderly, veterans, the uninsured, and racial and ethnic minorities. This paper will also focus on the statistical and methodological design innovations to the survey to support the measurement of health care expenditures at permit accurate national estimates.

Session 34 The Auditorium

14:35 LEG chapter on Implementation of Quality Management

by Margarida Madaleno, INE, Portugal

Implementation of Quality Management Systems in NSI's is a strategic decision to be taken by the top management. It is clearly a top-down process, but the involvement of the staff is also a necessary pre-condition for success.



Stockholm, Sweden

It is a long-term commitment and implies a change on NSI's focus: from a quality control of the product to customer-orientation based on a new culture for the organisation and a set of core values such as continuous improvement, benchmarking, teamwork, innovation, and others.

A sequence of characteristic steps in the implementation of Quality Management can be outlined though facing some obstacles that might emerge.

The LEG discussed the main guidelines for setting up a Quality Management System and acknowledged that most NSI's have ongoing quality improvement projects but work is not always systematic. However some NSI's have already made an option for a particular Quality Management model and self-assessment tools in order to achieve higher stages in quality and performance.

The presentation will cover the main aspects that should be taken into account regarding the Implementation of Quality Management Systems and the recommendations of the LEG for future work.

14:55 Implementing Continuous Quality Improvement (CQI) in a Statistical Organization

by David A. Marker and David R. Morganstein, Westat, USA

Many organizations are actively involved in improving quality. While improvement in methods and technology are common, improved work practices and quality management are less frequent. This presentation will describe key steps in successful implementation of CQI in these often ignored two areas. Examples will be given from statistical agencies in the United States and throughout Europe. Topics will include the roles of managers and internal facilitators, quality improvement teams, a customer focus, employee reward structures, current best methods, and other methods found to be helpful.



Stockholm, Sweden

15:15 Improving Process Quality – From Control to Motivation

by Walter Radermacher and Thomas Körner, Federal Statistical Office, Germany

The improvement of statistical products in most cases requires changes in the processes with which those products are produced. In this sense, improving quality almost in every instance has to make sure that processes are in good shape and continuously adapted to the requirements of the users, the agency, and the staff. But process improvements are important in yet another respect: Most National Statistical Institutions have to take more assignments while the number of the staff is decreasing. This requirements necessitates process optimisation and rationalisation.

The paper argues that process optimisation often not was successful because there was too much control and examination and too less room for innovation and creativity. Improving the process quality of NSIs does not only involve documentation and standardisation. Also a climate of change and openness has to be created. This important function sometimes has been neglected by traditional tools for process optimisation.

Using the example of a new scheme for analysing and improving processes developed by the Federal Statistical Office Germany, we show one possible way of process evaluation. In this scheme (started in October 2000), “objective“ and centralised measurement of processes and the number of staff required has been replaced by a procedure in which the people themselves - during an EFQM self assessment of the department - set priorities for the department and their units. Subsequently, selected processes are analysed and improvement plans are worked out in a decentralised way. For every step a standardised method of analysis with templates is provided in order to minimise the additional work. Key factors of success and possible risks will, finally, briefly be outlined.



Stockholm, Sweden

Session 35 Room 456

14:35 Implementing Quality Indicators in a Business Register

by Pascal Rivière, INSEE, France

This paper seeks to establish a general framework for the implementation of quality indicators in a business register. It proposes simple principles for the specific construction and dissemination of the register's quality indicators, from different points of view.

How is a quality indicator for a register constructed ?

It can be characterised by 5 basic components :

- **time** : as the register is not a fixed object in time, a register quality indicator is initially characterised by a reference date or a reference period.

- **scope** : an indicator is applicable to a given set of units. This scope is thus characterised by a unit type (enterprise or establishment) and generally forms a subset of the overall set of register units, except in the particular case of undercoverage indicators, for which the scope is necessarily larger.

- **sub-populations** : the quality indicators must be applied to one or more divisions of the scope, considered relevant to the analysis, that constitute sub-populations of interest. It is still useful to create global quality indicators, throughout the scope, but this is never sufficient because of the extreme heterogeneity of the world of businesses.

- **variable** : a quality indicator must be applied to a given variable of the register. For example, if the variable is the status (active / non-active), the indicator will be a rate of falsely active units.

- **criterion** : to construct a quality indicator, it is necessary to have an appraisal mode, unit by unit, for the quality of the variable. But there is no exhaustive reference for determining, unit by unit, whether or not the



Stockholm, Sweden

data is valid. In effect, if there were such a reference, it would go into the register! The means of appraisal are : register's external information sources (administrative sources, surveys, etc.), specific examination (register surveys), internal or temporal consistency, use of metadata (age of the value).

The previous framework refers to the quality of the contents of the register. There are other aspects of the quality of a business register, for which indicators can be defined :

- the quality of the framework : it should not be forgotten that a register, prior to being «filled» with units, is also a structure, a framework, that of its database: it is a sort of «empty shell» into which the units are inserted.

- the quality of the register as a tool : a register is accessed practically, by using a certain amount of software. The reliability of this software, its performance level and its user-friendliness are extremely important in judging the overall quality of the register, as the user perceives it.

- the quality of the register as a service : the register is not only a database. Register clerks provide a service, that forms part of the quality of the final «product» as a whole. Are the people calling the register clerks served properly on the telephone? Is the provided information clear ? Is it supplied within the correct deadlines?, etc.

Once the quality indicators have been defined, they have to be disseminated. Potentially, however, there is a huge number of possible indicators, numerous tables, etc. Finally, which strategy should be adopted?

Let us adopt two ideas:

- the disseminated set of indicators should be short and readable
- it is necessary to standardise the indicators so that they are comparable

The key lies in the readability of what is published. The quality indicators are not results that need to be published as such but tools for further works and tools for exchanges with users, information for decision-making, with the precise aim of improving quality.



Stockholm, Sweden

14:55 Measuring the Quality of a Business Register

by Mark Pont, Office for National Statistics, U.K.

Key words: business register, quality improvement, quality measurement

It is important to measure the quality of a business register to enable users to have an appreciation of quality – in the case of a business register it may be difficult for users to make their own assessment of quality as the gold standard is not easily defined. Moreover, the business register is perhaps the most important tool in the NSI, underpinning the production of national accounts and a wide range of other economic statistics. Quality measurement is also important for those responsible for managing the business register. In conjunction with users, required quality levels can be determined. Then, the available resources can be allocated in the most cost-effective way to improve quality in those areas where improvement is most needed.

Several recent developments in ONS have aimed to address the issues of quality measurement and improvement of its register, the Inter-Departmental Business Register (IDBR). A strategy group including representatives of the IDBR and customer areas has been set up to oversee quality measurement and development initiatives. This built on work undertaken in a Eurostat-funded Register Quality Survey during 1998 which established some basic quality measures for the IDBR.

More recently, an audit of the IDBR has recently been undertaken as part of the National Statistics Quality Review programme. The review carried out an extensive user consultation exercise to determine the aspects of quality most important to customers, evaluated whether the IDBR met those needs, and set out to improve the level of information available on quality. We recommended a range of quality indicators that could be used to describe the quality of the IDBR. These indicators were developed from our consultation with users, and the work of the review to evaluate the quality of a range of the main updating processes (including reviewing the quality of input data sources). The initial range of quality measures proposed covers the areas of completeness of coverage; accuracy and coherence of auxiliary information contained on the IDBR; currency of information; availability of contact details (name,



Stockholm, Sweden

address etc); quality of inputs and processes; quality of the IDBR's data analysis and sample selection services; and, customer satisfaction.

This paper reports on these recent quality initiatives, concentrating on the development of the quality indicators based on the review of the IDBR. We will describe the purpose of the measures proposed, the use to which they will be put, and how they will be measured.

Session 36 Room 357

14:35 Metadata (A Statistician and Survey Methodology Viewpoint) and Quality

by C. Easley Hoy, U.S. Bureau of the Census

Keywords: Metadata, organization of survey topics, quality reports

Reports of survey quality and a survey error profile primarily depend upon the existence, accessibility, and avoidability of survey metadata. Metadata can be organized a number of different ways depending upon the author, user, and its planned usage. Organization administrators, information technologists, regional offices, survey planners and coordinators, statisticians, survey methodologists, subject matter analysts, and programmers can all contribute and use various aspects of metadata for different purposes. Such objectives include documentation, research, analysis, and evaluation of the efficiency and quality of survey data and survey operations.

This paper offers a statistician and survey methodologist viewpoint of organizing metadata for authors of metadata and users of metadata to produce survey quality reports. This viewpoint of metadata considers content, survey operations and functional responsibility in organizing the survey planning, specifications, collection instrument, procedures, processing, estimation, and sources of error information. Also it provides some standard definitions of terms to clarify the metadata. This organization of metadata provides logical places where authors can register and store metadata/prior research on survey topics. It provides reasonable places where users/researchers can retrieve metadata or prior research for the same respective topics. It provides sensible places where



Stockholm, Sweden

new research results can be stored and found. Therefore, this organization of metadata enables researchers and analysts to evaluate survey quality more easily and prepare summary quality reports that also become part of the metadata.

14:55 **Quality Profiles and Data Users**

by Pat Doyle and Cynthia Clark, U.S. Census Bureau

Quality profiles have evolved over the years into documents survey methodologists find quite useful. They are produced periodically, usually at considerable cost and effort, and focus on a broad array of issues surrounding the quality of surveys and samples. They capture many sources of nonsampling as well as sampling error and their production has often led to significant improvements in the surveys or underlying samples. As a result, they constitute a significant contribution to the state of the art in survey methodology.

On the other hand, the quality profiles are often targeted to survey methodologists rather than to secondary data users and are produced infrequently thus quickly become out of date for ongoing data collection systems. In general, they focus heavily on errors and leave a negative spin on even the best surveys in the business. The style of writing is technical reflecting the emphasis on the survey methodologist as the audience. Often, they present results for a given survey in a vacuum rather than in context of other surveys from which similar estimates can be obtained.

As data producers, we have an obligation to make sure our data are well-used which requires that their strengths and limitations be well-understood by data users. Thus we need to extend the audience and scope for our quality profiles beyond the survey methodologists to all data users.

Data users could benefit from documents that are more frequent, more relevant, and less technical in presentation. They would also benefit from understanding the quality of estimates from a given survey in the context of similar estimates from all sources (including administrative data). Users do not expect any estimate to be error free but do need to know if a given estimate from a specific source would result in a bad



Stockholm, Sweden

decision. Users recognize that the best estimate for a given decision may not come from the survey with the smallest sampling error, if the questions in that survey are far less precise than a more-targeted survey with larger sampling error.

Recognizing the need to inform all users of data quality rather than just methodologists, the demographic programs within the U.S. Census Bureau recently instituted a new reporting requirement for surveys. This requirement calls for routine production of quality profiles, integrated into the survey process, and written with all users in mind. Launched in fall 2000, the initiative calls for routine production of a quality report as part of the survey process by 2004. The plan specifies the need for a summary of all facets of quality improvement research such as questionnaire testing, quality control procedures, and field efforts to improve response rates. It requires the presentation of traditional survey and administrative measures relative to other surveys of similar type. Finally, it recommends a style of presentation appropriate for the general user community.

15:15 **Quality Profiles in U. S. Statistical Agencies**

by Daniel Kasprzyk, National Center for Education Statistics, USA
and Graham Kalton, Westat, Inc., USA

A quality profile for a survey program is a review and synthesis of available information on the sources and magnitude of errors in a survey. Even though during the last twenty-five years, the emphasis has changed from „errors in data“ to „quality of data,“ the concept has not changed radically since its first major implementation by Brooks and Bailar (1978) on the U.S. Current Population Survey. This paper will review the objectives and uses of a quality profile from two points of view - the data producer and the data user, and compare quality profiles available from U.S. survey programs. While the development of a quality profile seems obvious and straightforward, a number of difficulties typically occur. The difficulties include insufficient documentation and information on most error sources. The paper will discuss these issues in the context of recent quality profiles on the U. S. Census Bureau's Survey of Income and Program Participation (SIPP) and the U. S. National Center for Education Statistics' Schools and Staffing Survey (SASS). Finally, an assessment of the usefulness of quality profiles will be given, and some thoughts on



Stockholm, Sweden

extending quality profiles to include broader concepts of quality (timeliness and relevance, for example) will be offered.

15:35 How to Secure Quality of Statistics from the Internet

by Alf Fyhrlund, Statistics Sweden

The increasing volume of statistics on the Internet is a function of the transition of society from focus on industrial production and services into information issues. Production and distribution of goods and services demand for more information among companies and organisations. There is a need of statistical information that can contribute to business intelligence and knowledge management. In the information society it is necessary to sort out different dimensions of quality of data from a user's point of view.

The **source of information** is of course a very important quality dimension. It is essential to distinguish between official and non-official statistics on the Internet as well as statistics directly from the producer and from intermediates. In this context it is also important to identify if international statistics is harmonised with reference to concepts and definitions or only presented as a collection of national series. It is evident that such information is part of the meta data; otherwise an important quality dimension is missing.

Meta, macro and micro data related to statistics are components in a **new functionality for publishing** in different data formats on the Internet. Data warehousing with multidimensional databases (e.g. "cubes") are nowadays in a growing extent used by big companies, but are since long time a traditional way of organising information among national statistical institutes (NSI's). OLAP (on-line-analytical processing), data-marts and data mining are examples of methods, which without systematic meta data handling are marketed by software consultants to private firms.

Statistics databases of the NSI's or international organisations are often based on some kind of meta data structure, which is also used for searching the information. The **organisation of meta data** is a very



Stockholm, Sweden

strategic aspect of the quality of statistics since the interpretation of the information will judge if statistics really are relevant for a specific purpose. There are different kinds of meta data like definition of concepts and classification categories. Web-techniques and electronic publishing permit hyper-links between micro/micro and meta data.

Another aspect of securing the quality from a user perspective is that efficient *search functions of statistics* are facilitated by relevant meta data linked to macro and micro data on the Internet. Meta data are most efficiently searched in catalogued structures. But there are also possibilities of website-specific free text searching of meta data in order to find the relevant statistics.

At last *the interactivity between the producer and the user* of statistics is facilitated on the Internet. The user can get more direct information on specific queries if communicating directly with the producer. This interaction is made more efficient when meta data on tables can contain e-mail address to the responsible production statistician. Indirectly this interactivity can also be performed via statistical networks, call-centres, statistics consultants and even through education or marketing of statistics. In all these contexts the Internet environment is an effective way to stimulate interactivity between user and producer, which might even improve production quality in the long run.

Session 37 Room 353

14:35 Assessing Customers' Satisfaction in Multi-mode Dissemination of Statistical Information

by Claudio Quintano, Giovanni Barbieri, Gerardo Giacummo, Silvia Milozzi

The new technologies used by Istat (Statistics Italy) provide for a new way of data collection which increases quality and reduces the burden on respondents. These new technologies are used also for the dissemination of information and the improvement of its fruition with respect to customers' needs.

In this effort, Istat is part of a larger context which includes also other areas of the Public Administration, where growing attention is paid to



Stockholm, Sweden

users-customers as the focus of the provision of services characterised by:

- a wide range of objects: from the classical consultation and acquisition of statistical data, to the fruition of preliminary data, to the access to microunits files;
- by a large number of information transfer modes that complement the traditional one on paper, to be read or bought, such as floppy disks, CD-ROMs, DVDs, Web-sites, etc.

In this way, the data producing institution abandons the traditional role of producer of public services with its own supplying methods, to which the citizen-user (today called user-customer) had to adapt, given the public nature of the products offered.

Today one of the qualifying features of the service provides for the frequent control of customer satisfaction, as is usually done by private enterprises, in order to adjust statistical information to customers' requirements, in terms of relevance for and compatibility with the needs of research and to the new methods used to obtain information, stimulating at the same time the adoption of new technologies by the public. These are the premises that led Istat and the Istituto Universitario Navale di Napoli¹ to conceive this project for ascertaining customer satisfaction on products and services offered by Istat.

The project aims at monitoring the customers' degree of satisfaction for the products presently offered and their propensity to new methods of dissemination of statistical information, which in turn will help to identify the potential public not yet reached by existing methods.

In the past few years, in fact, the most important NSIs have been using strategies based on "electronic first, paper second", increasing the access to on-line information, especially for current economic indicators. Istat too accelerated the release of on-line data in its new web-site, built as a portal for accessing statistics, organised by sector, and representing an area for data dissemination and contact with users.

¹ An experimental study of contribution of the IT to the integration of economic statistical sources and the improvement of data quality and information dissemination: "Qualità delle statistiche economiche e nuove strategie di formazione dei dati " financed by Ministero dell'Università e della Ricerca e della Tecnologia for 1999 and directed by Professor C. Filippucci (Università Bologna)



Stockholm, Sweden

This project aims at evaluating if and how this radical change is accepted by users. To this purpose, users were classified in three different groups, according to their requirements and expectations:

- subscribers of periodical publications;
- customers of Istat's Rome data shop;
- web-site users.

Three different surveys were directed to the three groups:

- 1) telephone calls to a sample of subscribers, to ascertain their use of paper products, their degree of satisfaction and any complaints related to on-line products;
- 2) interviews to a sample of data-shop customers (assumed as weaker in their ability to access other forms of data dissemination) to understand their difficulties in finding data, to evaluate the adequacy of products and services currently provided with their actual needs and to assess the problems they might encounter in moving from assisted service to a self-service approach;
- 3) a survey of web surfers (assumed as computer literate customers) in order to understand their needs, degree of satisfaction and expectations.

14:55 The Contribution of Cognitive Aspects of Survey Methodology to Social Services User Surveys in England

by Olwen Rowlands, Office for National Statistics, U.K.
and Hazel Qureshi, University of York, U.K.

The issue of user satisfaction with public services is becoming increasingly recognised internationally. At the same time the use of user feedback to evaluate services is now part of the official approach to assessing performance of public sector social care services

Following a White Paper entitled 'Modernising Social Services' there is a UK government requirement that from April 2000 onward all local authority social services directorates carry out local user satisfaction surveys. The eventual aim of the government department responsible for setting and monitoring standards within social services is to have nationally comparative data on the experience of all social services clients groups in England.



Stockholm, Sweden

The paper will report on the second stage of a two stage project to design a 'model' questionnaire and methodology to be used by local authority social services departments to measure the experiences and attitudes of a range of clients and, where appropriate, their carers.

A key feature of the methodology of the second stage will be the use of expert review and cognitive testing to design:

- questions about services and attitudes to them that users can understand and answer;
- questionnaires which service users can complete accurately; and to
- explore methods of enhancing response to local authority social services user surveys.

The paper will also consider some of the salient issues that need to be taken into account if nationally comparable measures of service user experiences are obtained from local surveys administered by bodies with variable research skills and differing organisational structures.

15:15 New Statistical Measures for New Economic Conditions – Development and Use of Pan-European Quality Indices and Customer Satisfaction Measures

by Jan A Eklöf, Stockholm School of Economics, Sweden

In this paper we will focus on the development and use of Quality indices and Customer Satisfaction Measures for studying how these may be used in obtaining additional information for:

- Price development
- Productivity analysis
- Opening of traditionally monopolistic markets

Focus will be on drawing inference and further develop the methodology started under the Pan-European initiative named European Customer Satisfaction Index (ECSI), and how this is now further developed into an integrated European Performance Satisfaction Index system (EPSI).

Special focus is to be given to questions and challenges in using and comparing results from systems like ECSI where objective and user



Stockholm, Sweden

experienced sets of information are combined. This means that differences in perceived satisfaction and loyalty for studied industries, and how these may be interpreted in terms of:

- Differences in expectations;
- Company to company variation;
- National cultural aspects; and
- Varying quality in service delivery.

As a special aspect, the combination of this kind of information on the “intangibles“ with more traditional economic statistics into regular information gathering and dissemination systems will be considered. The presentation will be finalized with a discussion on possibilities and limitations in the chosen approach and how it may be improved for offering further usefulness in cross-country analysis as well as for industry and company specific benchmarking (for example in using balanced scorecard reporting systems and using the EFQM Excellence model in strategic development.

Session 38 Room 361

14:35 Evaluation of Survey Data Quality Using Matched Census-survey Records

by Amanda White and Stephanie Freeth, Office for National Statistics, UK

The decennial population censuses in Britain provide a rare opportunity to evaluate data quality on major government household surveys. Following the last three censuses in Britain, survey nonresponse on major government household surveys has been investigated by linking addresses sampled for surveys taking place around the time of the Census to individual Census records for the same addresses. Thus all the Census variables are available to compare the characteristics of nonrespondents and respondents, estimate nonresponse bias and evaluate methods of nonresponse adjustment.

A more ambitious study is planned in connection with the 2001 Census. Knowledge of factors contributing to survey nonresponse has advanced considerably in the past ten years suggesting enhancements to the basic study design, while more powerful and sophisticated statistical



Stockholm, Sweden

modelling techniques are now available for the analysis. The 2001 study builds both on our own earlier work and on the work of Groves and Couper (1998), whose book draws extensively on the results of a census match study carried out following the 1990 US decennial census.

The study will involve 8-10 household surveys which vary significantly in design and response rates. Features of the study include:

- More extensive use of auxiliary area data than in 1991
- Information about interviewers' call details and response outcomes
- Interviewers recording data about the selected unit, interaction with respondents etc. when making contact
- Survey of interviewers providing data on interviewer characteristics and attitudes to be linked to the dataset
- Multilevel modelling to separate interviewer, household and area level effects

The paper will describe the proposed approach to investigating the relative contribution of different factors affecting nonresponse and the implications for both reducing and adjusting for nonresponse. Thus the study will provide a major quality measure for British social survey data.

14:55 Response Rates as a Measure of Survey Quality

by Peter Lynn, National Centre for Social Research, UK

Johanna Laiho, Statistics Finland

Jean Martin and Roeland Beerten, Office for National Statistics, UK

Response rates are an important indicator of survey quality. They are perhaps more likely to be reported than any other survey process quality measure. They are often compared across surveys, across survey organisations and over time. Yet, these comparisons are often distorted by differences in the way that survey outcomes are defined and recorded, differences in the way that response rates are calculated and differences in the extent and nature of published contextual information. Often, these differences are simply unknown to persons not involved with the surveys in question.



Stockholm, Sweden

Furthermore, response rates are used as indicators of a number of different aspects of survey quality - for example, interviewer performance, field performance, survey performance, and population coverage quality. For each purpose, there are alternative ways to define and calculate the rate.

To provide meaningful survey quality indicators, response rates should be defined and calculated in standard ways, with appropriate documentation of rules and definitions on particular surveys. Some progress has been made in this area in the USA, where the American Association of Public Opinion Research (AAPOR) has drawn up standard definitions applicable to RDD telephone surveys and face-to-face surveys where one person is selected per household. In Europe there are currently no standards in this area. Moreover, for various reasons the US standards cannot be adopted directly.

This paper describes a project to develop standards for defining and reporting survey outcomes and response rates in the UK. We will describe alternative definitions of response rates and will express views on their appropriateness for different purposes. We will illustrate the empirical impact of adopting alternative definitions, using outcome data from a number of surveys, including field tests of our proposed outcome schema.

Key dimensions of response rate definitions will be discussed, including:

- Weighted or unweighted rates
- Treatment of partial responses
- Treatment of cases of uncertain eligibility
- Treatment of multiple levels of response
- Separation of contact from co-operation

Furthermore, the meanings of response rates depend not only on the method of calculation of the rate but also on the definitions of the underlying outcome categories. Some important issues in defining outcome categories and implementing those definitions will also be discussed.



Q

2001

Stockholm, Sweden

15:15 **Correlates of Nonresponse**

by Rina Campoprese, Sante Orsini, ISTAT, Italy

Since 1993, about 24.000 Italian households have been selected every year by Istat to be part of the survey on Living conditions. Data on non response by cause are therefore available to tray and answer the following questions:

What are the levels and trends of non response rates?

Is there any relationship between non response and errors in sampling frame?

Is there any relationship between non response and size of municipalities?

What are the differences among geographical areas (regions)?

Do households of different size show different propensity to non response?





Q
2001

Stockholm, Sweden

