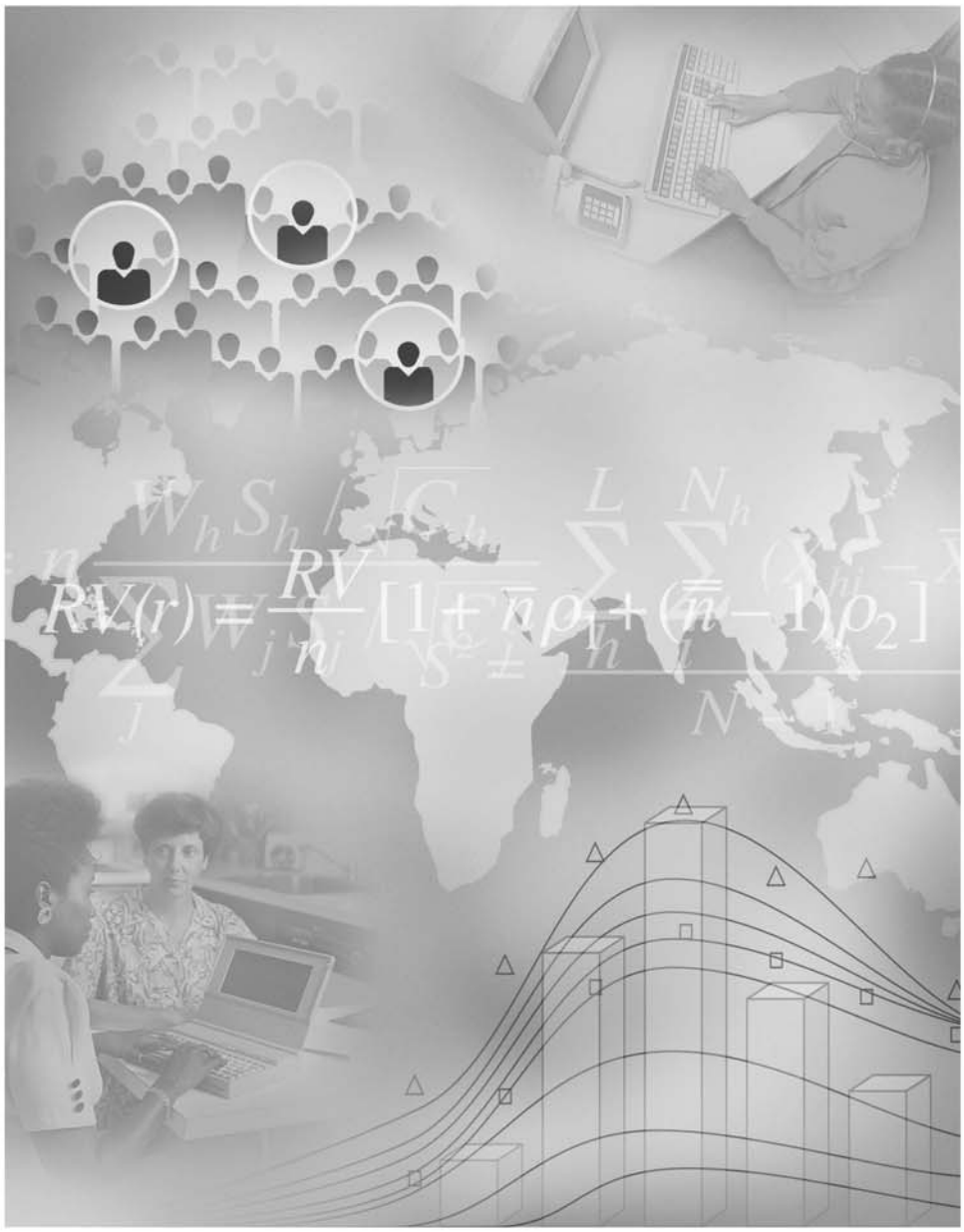
the Survey Statistician

The Newsletter of the International Association of Survey Statisticians

No. 54 **July 2006**



International Statistical Institute

Institut International de Statistique

the Survey Statistician
The Newsletter of the International Association of Survey Statisticians
No. 54 July 2006

Editor
Steven G. Heeringa

Section Editors
John Kovar — Country Reports
James Lepkowski — Software Review
Anders Christianson — Ask the Experts

Production Editor
Gail Arnold

Circulation
Claude Olivier
Anne-Marie Vespa-Leyder

The Survey Statistician is published twice a year in English and French by the International Association of Survey Statisticians and distributed to all its members. Information for membership in the Association or change of address for current members should be addressed to:

Secrétariat de l'AISE/IASS
c/o INSEE-CEFIL
Att. Mme Claude Olivier
3, rue de la Cité
33500 Libourne - FRANCE
E-mail: claude.olivier@insee.fr

Comments on the contents or suggestions for articles in *The Survey Statistician* should be sent via e-mail to sheering@isr.umich.edu or mailed to:

Steven Heeringa
Institute for Social Research
426 Thompson St., POB 1248
Ann Arbor, MI 48106-1248 - USA

In This Issue

No. 54, July 2006

Letter from the President

I would like to highlight four ways in which IASS members can support the work of the Association and encourage you to make suggestions on any or all of these topics.

Firstly, as mentioned in the last Newsletter, our Nominations Committee, under the leadership of Beverley Carlson (carlsonb@un.org), is seeking potential candidates to run for the positions of President-Elect, Vice-President (two positions), Scientific Secretary, and Council members (6 positions) in the 2007 elections. If you would like to suggest suitable candidates to lead our organization in the future please contact Beverley or any other member of the Committee.

Secondly, I am pleased to announce that Council member Leyla Mohadjer has agreed to be the IASS Program Chair for the 2009 ISI session in Durban. We need to formulate our program proposals for that session by June 2007 in time for discussion by the program Committee during the 2007 ISI session in Lisbon. It is time to start thinking about topics that should be on our program in 2009. If you have suggestions please pass them on to Leyla (leylamohadjer@westat.com).

Thirdly, we are continuing to seek entries for the 2007 Cochran-Hansen award for young statisticians from developing countries. Since many prospective candidates for this prize may not yet be members of IASS, we depend on members to identify and encourage young statisticians to apply. Please consider whether there are any qualifying young statisticians among your colleagues or students who should be encouraged to submit an entry. Council member, Sarah Nusser (nusser@iastate.edu) can be contacted for further information.

Finally, we are still short of Country representatives for some countries. We are publishing the list again in this issue. If your country is not mentioned in this list and you would like to help IASS, please contact Vice-President, John Kovar (john.kovar@statcan.ca) to volunteer your services or to find out what would be involved.

Many of you will have heard already of the death of Joseph Waksberg in January. Joe was a pioneer among survey statisticians and a long-time IASS member. In addition to lending his name to a widely-used technique in telephone surveying, Joe enjoyed two long and distinguished careers, first at the U.S. Bureau of the Census and then at Westat. Our sympathies go out to his family, friends and colleagues.

Preparations for the 2007 ISI meetings in Lisbon, and associated activities, are continuing on schedule. As usual we are planning to organize short courses around the 2007 meetings. Our main challenge is to ensure that students from developing countries are able to attend these courses and we are seeking sources of financing for this purpose. If members have any suggestions in this regard, please make them known to our Scientific Secretary (lilli.japec@scb.se). In connection with the 2007 meetings, ISI has approved a satellite conference on Small Area Estimation that IASS supported. The conference will take place in Pisa, Italy on September 3-5, 2007,

Finally, I was pleased to meet many IASS members while attending the recent Q2006 (European Conference on Quality in Survey Statistics) in Cardiff, Wales. IASS organized a panel discussion on the Role of Survey Statisticians in Managing Data Quality as part of the Conference.

As always, I welcome any comments or suggestions from you on IASS activities.

Gordon Brackstone
President, IASS

Report from the IASS Scientific Secretary

In this issue of The Survey Statistician you will find information about ongoing IASS-activities such as the Lisbon short courses, a report from the International Conference on Telephone Survey Methodology II, as well as a report from the IASS Lisbon 2007 Programme Committee and information about future conferences sponsored by IASS. You can also read about a new journal of relevance to our field, Survey Research Methods http://esra.sqp.nl/esra/journal/surveymethodology. This is one of three new journals that have evolved during recent months, the other two being The Journal of Privacy and Confidentiality, www.jpc.stat.cmu.edu and The International Journal of Public Information Systems, www.ijpis.net .

Future meetings of interest to IASS members include:

- The International Conference on Methodology of Longitudinal Surveys (MOLS 2006), Essex, UK, July 12-14, 2006. Website: http://www.iser.essex.ac.uk/ulsc/mols2006/

- The International Conference on Establishment Surveys (ICES III), Montreal, Canada, June 18-21, 2007. Website: http://www.amstat.org/meetings/ices/2007

- The Second Baltic-Nordic Conference on Survey Sampling, Kuusamo, Finland, June 2-7, 2007. Website: http://www.mathstat.helsinki.fi/msm/banocoss/

- Small Area Estimation SAE2007, satellite meeting to the Lisbon ISI Conference - Pisa, Italy, September 3-5, 2007. Contact: Monica Pratesi (m.pratesi@ec.unipi)

Please do not hesitate to contact me if you have suggestions or questions regarding any of the IASS activities.

Report from the International Conference on Telephone Survey Methodology II

In January around 350 people gathered in Florida for the second International Conference on Telephone Survey Methodology. The conference, chaired by Clyde Tucker, U.S. Bureau of Labor Statistics, covered many interesting topics such as, organization of call centers, sampling and weighting, interviewer effects and new technology. In his plenary speech, Clyde Tucker addressed the rapid changes in both telecommunications and the social and political environment and how these changes have affected and might affect the way we do telephone surveys in the future. Mode effects were discussed in a number of presentations and results from studies comparing telephone surveys with, for example, web and telephone surveys were presented. One of the future challenges for many western countries is how to design mixed-mode surveys in a way that mode effects are minimized. Another interesting topic that was addressed was oral translations in multi-lingual surveys. It was illustrated how quality can be affected in a negative way when oral translations are used.

There were also four short courses offered as part of the conference. The topics were survey quality, sampling, questionnaire design, and multilevel analysis. The invited papers presented at the conference will be published in a Wiley monograph early 2007.

Lilli Japec
IASS Scientific Secretary

IASS Short Courses at the 56th ISI Session, Lisbon 2007

The International Association of Survey Statisticians has developed a program of short courses to be offered just prior to or after the 56th Session of the International Statistical Institute (ISI), which will be held in Lisbon, Portugal, August, 2007. Funding assistance for supporting statisticians from developing and transition countries is being sought.

The following courses will be offered.

Course A: Workshop on Survey Sampling
Presented by: Colm O'Muircheartaigh and Steven Heeringa.

The workshop will focus on practical aspects of sampling for household surveys. It will start

from basic principles and build up to complex stratified multi-stage sample designs. It will cover the main sampling techniques and also such issues as sampling frames, weighting, and imputation. It will end with an introduction to variance estimation with complex sample designs.

Course B: Variance Estimation in Complex Surveys
Presented by: Wayne Fuller, Kirk Wolter, F. Jay Breidt, and Anthony An

The purpose of this course is to provide training in variance estimation in complex surveys for survey statisticians, especially those from developing countries. The course will cover methods of estimating variances for statistics such as means, proportions, ratios, and regression coefficients. Variance estimation for imputed samples will be discussed. Both linearization and replication methods will be presented. The use of computer software for computing variances of statistics from complex sample designs will be demonstrated and instruction will be given in practical applications. About one-half of the course will be devoted to computer implementation.

Course C: Workshop on Editing and Imputation of Survey Data
Presented by: John G. Kovar and Eric Rancourt

Surveys and censuses conducted by national statistical agencies, research institutes and other survey organizations suffer from various degrees of nonresponse even under ideal conditions. In order to try to alleviate the problems caused by nonresponse, editing and imputation methods are usually applied. Since the process of editing and imputation is time and resource intensive, care must be exercised in controlling the efficiency as well as the effectiveness of the methods. The aim of this short course is to provide the students with an introductory level description and discussion of the methods of prevention, detection and treatment of nonresponse. Evaluation of such methods and their impact on the survey outputs will be highlighted. Existing edit and imputation software will be compared. Numerous examples will be provided to illustrate the material presented.

Course D: Introduction to Survey Quality
Presented by: Paul Biemer and Lars Lyberg

The course is designed for a broad audience that includes experienced survey researchers and practitioners who would benefit from a better understanding of the survey quality as well as others with little prior experience in survey methods. It provides an introduction to the concepts, principles and terminology used in the study of survey quality. Accuracy is but one dimension, albeit a critical one, of a multi-dimensional framework that embodies survey quality. The course examines the goals of survey design with particular emphasis on minimizing total mean squared error of key survey estimates subject to costs constraints as well as constraints on some quality dimensions. The sources of survey error are discussed within this schema, focusing on four major sources: coverage error, nonresponse, data processing error and measurement error. Methods that are most often used in practice for evaluating the effects of the survey error will also be covered.

Course E: Statistical Disclosure Control
Presented by: Anco Hundepool, Eric Schulte Nordholt and Peter-Paul de Wolf

The purpose of this course is to provide the participants with an understanding of the methodological aspects of Statistical Disclosure Control, to train them in solving problems on this topic and to demonstrate the ARGUS software. The meaning and impact of Statistical Disclosure Control can only be appreciated in the light of practical problems and policy related issues. Therefore, some attention is also paid to such topics without putting heavy emphasis on them. Topics covered include theory and methods on microdata, exercises on microdata, demonstration of Mu-ARGUS, theory and methods on tabular data, exercises on tabular data, demonstration of Tau-ARGUS, legal issues, on-site facilities and remote access.

Course F: Design and Analysis of Repeated Surveys
Presented by: David Steel and Craig McLaren

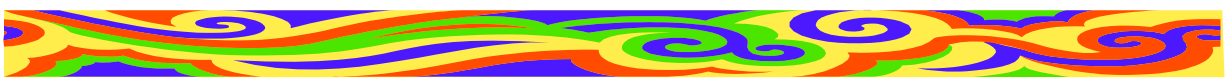
This course will consider the interaction between the design of a repeated survey and the methods used for estimation and analysis. The choice of rotation pattern will be considered in terms of the impact on the estimation of levels and changes. Composite and other forms of estimators will be reviewed and the interaction between design and estimation explored. Estimation of seasonally adjusted and trend estimates from repeated surveys will also be considered.

For further information, please contact:
Lilli Japec
IASS Scientific Secretary
isi2007_IASS_shortcourse@scb.se

Report on IASS Lisbon 2007 Programme Committee

The IASS Programme Committee, led by David Steel, has put together a number of sections for the ISI meeting in Lisbon 2007. Below you can find the section topics and names of organizers.

Title	Primary	Joint With	Name	email
The impact of new information technologies: on survey research design; on a totally new information production model	IAOS	IASS	Jonathan Palmer	jonathan.palmer@abs.gov.au
Teaching of Survey Statistics	IASE	IASS	Steve Heeringa	sheering@isr.umich.edu
Information integration: statistical theory when combining and using multiple data sets in concert	IASS	IAOS	Dean Judson	Dean.H.Judson@census.gov
Designing and Updating Longitudinal Samples	IASS		Vijay Verma	verma@unisi.it
Statistical disclosure control of survey microdata	IASS		Mark Elliot	Mark.J.Elliot@manchester.ac.uk
Using multiple modes to collecting data in surveys	IASS	IAOS	Peter Lynn	plynn@essex.ac.uk
Confidentialising tables and data with geographically fine breakdown	IASS		Larry Cox	LCOX@CDC.GOV
Prioritising non-response follow-up to minimise mean square error	IASS		Cynthia Clark	cynthia.clark@ons.gov.uk
What Censuses and administrative sources can tell us about Non Sampling Errors?	IASS		Ray Chambers	ray@uow.edu.au
Measuring and reporting quality of small area estimates	IASS		Daniella Cocchi	daniela.cocchi@unibo.it
Randomisation-assisted model-based survey sampling	IASS		Phil Kott	pkott@nass.usda.gov
New methods of sampling	IASS		Yves Tille	yves.tille@unine.ch
Opinion Polls: Do they do more Harm than Good?	IASS		Murrary Goot	murray.goot@mq.edu.au
How ISI can encourage donor and international organizations to strengthen their own statistical capacities	IASS	IAOS	Ibrahin Yansaneh	yansaneh@un.org



Australia

Paul Sutcliffe

The ABS conducts a regular program of **National Health Surveys** (NHS) run every three years. This is complemented by a six yearly National Aboriginal and Torres Strait Islander Health Survey (NATSIHS). The first NATSIHS collected information from over 10,000 Indigenous Australians over 10 months in 2004-5. Results will be available from the ABS website from mid April, and will enable comparison to the NHS run over roughly the same period, as well as smaller samples of Indigenous people that were included in the 2001 and 1995 NHSs.

The survey collected information on social and demographic characteristics, health status, risk factors, service use, immunization and women's health. The ABS used a screening approach to identify Indigenous people in non-remote areas of Australia, and enumerated within a sample of remote communities to enable comparison between non-remote and remote areas. As well, a module of questions on Social and Emotional Wellbeing (SEWB) was included for the first time, drawing on selected questions from the Kessler 10 and SF-36 instruments. The ABS will be evaluating the quality of the SEWB results and their suitability for comparisons with results for the non-Indigenous population, where available.

A Survey of Mental Health and Wellbeing is planned for mid 2007 and will be conducted on behalf of the Australian Government Department of Health and Ageing. It will enable comparison with international data as well as the last survey on this topic, which was conducted in 1997. This survey will be developed using the Composite International Diagnostic Instrument (CIDI), which has been used in a number of mental health surveys in other countries.

For more information about the Heath Surveys Program, please contact Catriona Bate at: catriona.bate@abs.gov.au

The ABS conducted the **Survey of Foreign Currency Exposure** in respect of 31 March 2005, with the assistance of the Reserve Bank of Australia, to provide additional information to that available in the International Investment position, on the mitigating impact of hedging activities on foreign currency exposures.

The survey sought to cover all enterprises with significant foreign currency denominated balance sheet items and/or significant foreign currency denominated receipts and payments from trade.

The information collected included:

- foreign equity assets;
- foreign currency denominated debt assets and liabilities;
- foreign currency denominated receipts and payments from trade expected after 31 March 2005;
- the principal value of outstanding derivative contracts with a foreign currency component; and
- policies on hedging foreign currency exposure.

Results from the survey were released in Foreign Currency Exposure, Australia (cat.no. 5308.0) on 22 November 2005. For more information about the Survey of Foreign Currency please contact Wendy Raedt at wendy.raedt@abs.gov.au

The Australian Bureau of Statistics is collaborating with a range of Australian government agencies to develop a **National Data Network** (NDN). The NDN is seen as a key initiative in Australia's National Statistical Service by making statistics produced across government agencies easier to find and access, in addition to driving improvements in data quality and documentation.

The NDN will create a distributed library of data holdings relevant to policy analysis and research. Data owners will publish their statistical data to a node on the network, where they will be able to use the NDN to set controls over who can access this data. The nodes are connected to a central hub that contains a library of metadata with entries for each dataset. Researchers and analysts can search across the library and discover data of interest according to controls set by the data owner. Over time, the network will also see the development of a range of statistical services. The NDN is currently in development stage to test the concept and refine the working model. The network is expected to go into full production early next year.

For more information about the National Data Network please contact Mark Lound at mark.lound@abs.gov.au

Brazil

Giuseppe de Abreu Antonaci and
Pedro Luis do Nascimento Silva

The **International Seminar on Editing and Imputation**, which took place at IBGE, the Brazilian Institute of Geography and Statistics, in Rio de Janeiro in November 2005, aimed at promoting an opportunity for discussion and exchange of ideas in the area of editing and imputation. It focused on the dissemination of statistical methods and computing systems for editing and imputation of data derived from surveys, censuses and administrative sources in the social and economic areas. It had 225 participants, mostly from Brazil, but also from Argentina, Australia, Canada, Chile, England, Mexico, Mozambique, Spain and the USA. There were 27 papers presented in a posters session, four papers given as oral presentations, a short course on *Editing and Imputation*, and 10 invited conferences. The invited speakers were Adão Hentges, Beatriz Mendes, Djalma Pessoa and Lúcia Barroso from Brazil, Marcel Bureau and Jean-François Beaumont from Canada, Philip Kokic from Australia, Ray Chambers from England and Víctor Guerreiro from Mexico. The material used in the presentations can be found at www.ibge.gov.br/sici/english/index.htm.

Most official statistics agencies around the world face the challenge of producing up to date information for detailed geographies, with good precision and at reduced costs. The development of **Alternative Methods for Demographic Censuses** is one of the main attempts to meet this challenge. Two seminars to discuss this subject have already been organized by IBGE, "Instituto Brasileiro de Geografia e Estatística", and the INEGI, "Instituto Nacional de Estadística Geografía e Informática" from Mexico, with the support of INSEE, the U.S. Census Bureau and the UNSD. From 29 to 31 of May 2006 the **3rd Seminar on Alternative Methods for Demographic Censuses** will take place at IBGE in Rio de Janeiro. For more information please visit www.ibge.gov.br/seminarios_metodos_alternativos_censo.

IBGE is responsible for periodically organizing the **National Conference of Statistics** and the **National Conference of Geography and Cartography** (CONFEST and CONFEGE, in Portuguese). This year the 5th CONFEST and the 4th CONFEGE will be held as part of the **2nd National Meeting of Producers and Users of Social, Economic and Territorial Information** that will take place at IBGE, in Rio de Janeiro, from 21 to 25 of August 2006. This event will give the opportunity to IBGE and other producers of statistical and geographical information in Brazil to interact with theirs users and get feedback about information produced as well as about new demands, thus improving the cooperation between users and producers to enhance information production, dissemination and usage. For more information concerning the event please visit: http://www.ibge.gov.br/confest_e_confege/.

IBGE has initiated a project to study the introduction of an **Integrated System for**

Household Surveys. The two main features of this project are the development of a Master Sample and the integration of the Monthly Labor Force Survey and the Annual Household Survey. Issues related to questionnaire and sample design (including rotation pattern) are now under discussion. A first version of the project is due to appear at the end of July 2006. For information on this project, contact IBGE via its Directorate for Surveys.

Canada

John Kovar

Following the recent Census of population in Canada, May 16, 2006, Statistics Canada will be conducting a **program of post-censal surveys** as in previous censuses. Five surveys are expected: the Participation and Activity Limitations Survey (PALS), the Survey on the Vitality of Official Language Minorities (SVOLM), the Maternity Experience Survey (MES), the Aboriginal Children Survey (ACS) and the Aboriginal Peoples Survey (APS). The survey frame for these surveys is built directly from the Response Database of the 2006 Census, and the samples are drawn according to answers provided to a set of screening questions and to other demographic characteristics. Among other advantages, this approach has proved to be a cost-effective means of collecting information for relatively rare and scattered populations and of producing good quality estimates for small domains. It is also possible to effectively coordinate sample selection to minimize response burden. Except for Aboriginal communities where personal interviews are carried out, the data collection mode for these surveys is telephone interviews. Computer assisted telephone interviewing is used for PALS, MES and SVOLM. Data collection is scheduled to take place during the fall of 2006 for all the surveys. For more information, you can contact Jean-Pierre Morin, Chief methodologist, Social Survey Methods Division, Statistics Canada, at morinjp@statcan.ca .

The **Monthly Wholesale and Retail Trade Survey** (MWRTS) is conducted in order to provide monthly sales and inventory estimates at various Canadian province and industry levels in the Retail and Wholesale sectors. The survey underwent a full-scale redesign for the first time since 1988. Monthly estimates based on the first phase of the redesign were released in June 2004 for April 2004 reference month. The new MWRTS is based on the North American Industry Classification System (NAICS). The coverage of businesses with no employees has also been significantly enhanced. Response burden and collection costs have been reduced by taking advantage of administrative data, in particular the Goods and Services Tax (GST) program data. The smallest businesses are not surveyed directly but an estimate for their total contribution is derived from the administrative data source. An improved sample design is used for the rest to determine which units are subjected to sampling. Overall, the total number surveyed units each month has been reduced by 30%. The second phase of the redesign aimed at making further reductions in response burden and survey costs. In this second phase, MWRTS followed the recent examples of the Monthly Restaurants, Caterers and Taverns Survey and the Monthly Survey of Manufacturing. It consisted in replacing the process of collecting data directly from the respondents with data modeled from GST returns. This was applied to a subsample of the sample, and only for simple units (units that represent only one establishment). This resulted in an additional reduction of 15% in the number of units that are surveyed directly compared to the old MWRTS number. Monthly estimates based on the second phase of the redesign were released for the first time in December 2005 for the October 2005 reference month. For more information, you can contact François Brisebois, Chief methodologist, Business Survey Methods Division, Statistics Canada, at brisfra@statcan.ca .

Survey Methodology is a refereed journal published twice a year by Statistics Canada. It contains articles dealing with various aspects of statistical development relevant to a statistical agency. The year 2005 saw the launching of *Survey Methodology* on-line. Each new issue is now available in electronic format on-line at no cost. Currently, only a few of the past issues are up but the plan is to

have all the issues from 1993 and on eventually available. Older issues in scanned format are also available in pdf format upon request. For more information, to subscribe or to submit a paper, please contact John Kovar, the Editor in Chief of *Survey Methodology*, Kovar@statcan.ca or at smj@statcan.ca. More details are also available on Statistics Canada's internet site at www.statcan.ca.

Hungary

László Mihályffy

Until 2004, the **trends in tourism in Hungary** could only be analyzed by observing the supply side – having information from accommodation establishments and travel organisers. In addition, information on the number of Hungarian travellers abroad and that of international visitors coming to Hungary from the border guards were also available. Those methods of data collection were not suitable to obtain information on a number of items (demographic, social characteristics of tourists, motivation, and expenditure) that can be collected only by interviews. As a result, in June 2004 a regular and comprehensive tourism statistical system was introduced. It consists of three surveys:

1. Travel habits of Hungarians
2. Travel habits of Hungarians travelling abroad
3. Expenditure of foreigners visiting Hungary.

The implementation of the surveys was sponsored also by the EU Commission. The results are published in several publications and sent to several organisations.

1. Travel habits of Hungarians

The sample of this survey contains 25 thousand households in each quarter. The basic unit of the sample is the dwelling, the address of which is selected from the list of the 2001 population census in a random way. It is a split-panel survey where a given household should report on every trip made during the four or eight quarters. About one half of the sample is used for four quarters, while the rest for eight quarters.

The questionnaire consists of three separate parts. In the first part the characteristics of the household and its members are recorded. These data are asked only once for every household. Later on we follow only the changes in the number of the household members.

In the second part of the questionnaire we collect information on the same-day trips.

In the third part, a separate page should be filled in for each trip made by at least one member of the household. It contains data about the month of trip, length of stay, means of travel, type of accommodation establishment, purpose of trip, activities made on the trip, organisation of the trip and expenditure.

2. *Travel habits of Hungarians travelling abroad*

The introduction of this survey was necessary because the survey "*Travel habits of Hungarians*" had not supplied sufficient information on the trips abroad (outbound tourism). As such, this survey can be considered as a supplement of the former.

In this survey two thousand Hungarians are interviewed monthly at the border. The questionnaire contains data on the month of trip, length of stay, means, type of accommodation establishment, purpose and organisation of the trip and expenditure.

3. *Tourism expenditure of international visitors in Hungary*

Non-resident visitors are interviewed when crossing the Hungarian border on their departure and report on the expenditure and consumption structure. Each month about 10 thousand questionnaires are completed.

There are questions in the questionnaire on the residence, age, and sex of the visitor, purpose of the trip, length of stay, type of accommodation establishment, visited regions, organisation of the trip and expenditure. For more information, contact Csaba Kovács at csaba.kovacs@office.ksh.hu

Italy

Claudio Quintano

We describe below some of the **main innovations in products and processes of the Italian National Statistical Institute**.

In the field of social statistical production for the period 2004-2005 the new survey on Interventions and Social Services of Municipalities represents the basis for monitoring public-social expenditures and overall welfare services supplied in Italy. This survey is carried out together with the Ministry of Economy and Finance and is web-based. A further innovation is to be found in the continuous Labour Force Survey that substitutes the previous quarterly survey. The new LFS is carried out through a mixed CAPI (Computer Assisted Personal Interviewing) and CATI (Computer Assisted Telephone Interviewing) technique supported by a network of 300 interviewers directly managed and coordinated by the Italian NSI. In addition, the new Eu-Silc (European Union Survey on Income and Living Conditions) sample survey is being issued for the first time and some results are expected within the end of the current year. This survey contains a panel element and therefore allows a valuable cross-insight on social exclusion; it enforces an EU regulation and aims to supply useful comparable data both to carry out the traditional analysis on income distribution and poverty (in terms of income) and to integrate this analysis with indicators on quality of life.

As far as economic statistics are concerned, attention should be focused on the first issue of the Large Enterprises Local Units Survey (IULGI) that was started in October 2004 and will be concluded in 2005. IULGI's fundamental and strategic aim is the updating of the local units lists belonging to each large enterprise and of the related stratification variables of enterprises and their local units. This survey is partly web-based; the data collection by the regional offices of the Italian NSI is particularly accurate at territorial level. This allows for obtaining high response rates. The ASIA-Local Units archive is implemented through the integration of the IULGI survey with administrative data obtained from small- and medium-size enterprises; in this way data similar to those obtained through the General Census on Industry and Services are available every year. During 2004 the fourth EU survey on ICT use and electronic trade in enterprises with at least 10 persons employed from the manufacturing and services sectors was carried out. Moreover, two pilot surveys were implemented: the first on the use of ICT technologies made by banks and insurance companies; the second on knowledge management with special attention on innovation and scientific research. Both surveys meet the need for collecting data harmonised at international level.

As to integrating data from different sources, the study of statistical matching techniques for the reconstruction of missing information has been carried on and simulation studies on iterative reproportioning techniques have been carried out.

With regard to data dissemination, it should be stressed that the Institute new web site (www.istat.it) was implemented and is currently on line. The site is particularly user-oriented: information areas can be accessed more directly, some relevant data can be viewed on the home page and the search-engine has become a powerful and efficient orientation tool. The new site was issued last June 2005 and provides the user with a huge amount of data for free. More specifically, the online catalogue allows the user to find all Institute publications by viewing a summary informative card and by exporting, if available, an electronic version of the requested volume. Through the Istat site it is possible to access various Data Warehouses both related to a specific area (population, culture, health, justice, economy) and cross-sectional (whereby all main statistics are represented by subject); the available Data Warehouses disseminate user friendly data and metadata and allow the creation and export of tailor-made tables.

For more information, please contact Nadia Mignolli at mignolli@istat.it

A **new Italian quarterly survey on job vacancies and worked hours** was launched. The European Council Regulation on Short Term Statistics (N. 1165/98 as amended by N. 1158/2005) requires the Member States to

produce worked hours indicators for NACE Rev. 1.1 sections C-F (Annexes A and B), while the European Council Regulation on a Labour Cost Index (N. 450/2003) requires hourly labour cost indices for NACE Rev. 1.1 sections C-K. Furthermore, a European Council Regulation on job vacancies is currently under discussion. A new quarterly business survey was launched in Italy in the second half of 2003, aiming at producing quarterly indicators of job vacancies and worked hours for firms with at least ten employees in NACE Rev. 1.1 sections C-K.

The survey sample includes 8,000 firms, selected on the basis of a stratified simple random sampling design. Economic activity, size in terms of number of employees and geographical localisation has been considered as stratification variables. All firms with at least 500 employees are included in the sample. Data are collected partly via CATI interviews (80-85% of all respondents) and partly via self-completion of an online questionnaire (posted on the Istat's web site). Reminders are sent via fax and e-mail to firms that should respond online. Respondents are, on average, around 75% of all contacted enterprises. The data editing and grossing up to the population total procedures are still being developed. However, a provisional version of these procedures exists that allows to transmit provisional and confidential data on job vacancies and occupied posts to Eurostat, for the calculation of European job vacancy rates. Data dissemination should begin during 2006, at least for what concerns job vacancies.

For further information on the survey methodology or data products, please contact Marina Sorrentino at mrsorren@istat.it

Another new quarterly survey was launched in Italy, this one on **Repair of motor vehicles and motorcycles**. The European Council Regulation on Short Term Statistics (N. 1165/98 of 19th May 1998, amended by the Council Regulation N. 1158/2005 - Annex D) requires the Member States to produce turnover indicators for many market services activities, including repair of motor vehicles (NACE 50.2). The new quarterly survey concerning repair was launched in Italy at the beginning of 2004, aiming at producing quarterly turnover indexes with base year 2000. The survey has been designed on the basis of a balanced sampling approach, where the balancing variable was the yearly turnover drawn from the ISTAT business register ASIA. A preliminary stratification has been carried out according to 5 economic categories and 6 employment classes. The overall sample size has been fixed in 3,000 survey units, assigned to strata according to the Neyman optimal allocation.

The data collection has been organised using self-completion questionnaires sent by mail. The arrival (by mail or fax) of responses is continuously monitored and reminders are sent to non respondents by mail; telephone contacts are used in a second stage. The response rate has been, in average, close to 70%, with a turnover coverage of up to 70%. The data editing process is developed in four steps: 1) first automatic check of data coherence; 2) questionnaires' revision; 3) non response imputation by a regression model, using as explicative variable turnover in the same quarter of the previous year; 4) detection and treatment of outlier observations, based on macro-editing techniques. Quarterly elementary indexes are calculated, in each stratum, multiplying the average change occurred from each quarter and the same quarter of the previous year by the index (with base 2000) referring to the same quarter of the previous year. Higher order indexes are calculated by a Laspeyres scheme applied to indexes by stratum and using weights derived from the business register. In April 2005, indexes referring to the period I-2001 up to IV-2004 have been published in a specific press release. Starting from data referred to the first quarter of 2005, quarterly indexes are published in a press release including other service activities as well. At the moment, the average delay from the end of the reference period is 90 days, with the aim of a reduction to 60 days by the middle of 2006. The latest published data refer to the second quarter of 2005.

For further information on the survey methodology or data products, please contact Alfredo Cirianni at cirianni@istat.it

New Zealand

Vina Cullum

In a previous IASS country report (Sept 2004) Statistics New Zealand introduced our **Business model Transformation Strategy** (BmTS), which aims to improve, streamline and standardise our business processes in order to optimise the Statistics New Zealand business model from end to end perspective.

We are now almost two years through the initial three-year implementation phase of the BmTS. In that time, initial versions of a number of BmTS components, or elements of these, have been successfully designed. These are now being implemented in six strategic programmes of work, such as National Accounts and the Programme of Social Statistics, and will be available for future programmes and projects to re-use or move forward an iteration. The components designed and implemented to date include:

- workflow framework and dashboard - that allow users to specify, configure and run the business processes they need to produce statistical information for customers;
- an input data environment, along with a supporting metadata store;
- generic modules for statistical transformations such as editing and imputation and seasonal adjustment, along with a library of such transformations.

At the end of the initial three-years, Statistics New Zealand will have designed and built a new platform on which future programmes and projects undertaken by Statistics New Zealand will be founded. The following six years will involve migrating other Statistics New Zealand collections and outputs to this new platform. Eventually, BmTS will become business as usual for Statistics New Zealand.

The Statistical & Methodological Services (SMS) Group within Statistics New Zealand plays a key role in the BmTS, via the setting of statistical and methodological standards for use across the whole organisation. The SMS group currently has a number of research projects underway to research, evaluate and select methods and tools to support such processes as sampling, coding, editing and imputation, weighting and estimation, and statistical disclosure control.

The selected methods and tools will then be included in the transformations component of the BmTS, and implemented via one or more of the six strategic programmes mentioned above.

For more information on the BmTS, please contact Gary Dunnet:
gary.dunnet@stats.govt.nz

For more information on research within the Statistical & Methodological Services Group, please contact Walter Davis:
walter.davis@stats.govt.nz

Statistics New Zealand currently provides **access to microdata** for researchers through a data laboratory with facilities in three cities, and a small but growing number of Confidentialised Unit Record Files (CURFs) that are produced on CD-ROM for licensed users. A new development is the completion of a pilot run of a remote access system using software that has been created by the Australian Bureau of Statistics. This pilot work was carried out on the New Zealand Income Survey 2003 dataset, and work is now being undertaken to create further datasets for remote access in order to have a remote access system launched later this year.

Further information about microdata access at Statistics New Zealand can be found at:

http://www.stats.govt.nz/curf-programme/default.htm
http://www.stats.govt.nz/products-and-services/datalab.htm and
http://www.stats.govt.nz/about-us/policies-and-guidelines/general/microdata-access-protocols.htm

The Official Statistics Research and Data Archive Centre (OSRDAC), set-up in 2004 as a result of a review of the New Zealand Official Statistics System, is preparing to launch a new data archive later this year. The data archive, to be called Source, will hold important unit record dataset created by

Statistics New Zealand and other government agencies. Users will be able to locate datasets through an online catalogue, and then request access to them through the Statistics New Zealand Data Laboratory, Confidentialised Unit Record File programme or the Remote Access Data Laboratory service that is being developed. Each dataset will be archived along with high-quality metadata and documentation to ensure that it can be understood in the future.

For more information contact Hamish James, Manager, OSRDAC:
hamish.james@stats.govt.nz

New Zealand held its **32nd Census** in March, 2006. Statistics New Zealand's 6,500 collectors delivered roughly 1.5 Million Dwelling Forms and 4.1 Million Individual Forms to every household in New Zealand.

This was the first time that Statistics New Zealand provided the option to complete the census online. Approximately 104,000 Dwelling Forms and 290,000 Individual Forms were submitted online. The online forms, just like the paper forms, were available in English and Mäori.

Collectors did not need to return to households where all census forms had been completed online because of a new mobile phone text message communication system. This involved collectors receiving text messages notifying them when census forms in their area had been submitted online, as well as to advise of requests for additional forms.

The Post Enumeration Survey, which will measure coverage of the census, commenced its field work phase in late March. Separate evaluation studies will also be conducted to understand the factors behind the take-up of the Online Form and the Mäori/English forms.

For more information see
www.stats.census.govt.nz or contact
Nancy McBeth, GM Census 2006:
nancy.mcbeth@stats.govt.nz

Spain

Dolores Lorca

In 2005 the National Statistical Institute (INE) has conducted a survey to study the homeless population of Spain, the **Homeless People Survey 2005**. It has been the first experience of its type in the Institute. The scope of this survey is the population of people not having a fixed residence, so this population is missed in the traditional household surveys. We have followed the INSEE-France approach of accessing the persons through a sample of the services they get (accommodation and meals) in the centres that the social institutions and the Administration have to assist them. The base of this survey has been a previous survey conducted in 2003 to study the centres that provide services to the homeless persons, the Homeless People Survey (Centres) 2003. The frame of the centres used in the survey of centres of 2003, updated in 2005, has been the frame used in the new survey targeted to homeless persons. Only centres located in municipalities with 20,000 or more inhabitants have been taken into account. The survey was conducted during the month of February of 2005.

The sample design has followed the INSEE-France methodology, so we have actually drawn a sample of services and then we have studied the persons these services were provided to. The sample of services has followed a two-stage sampling design with stratification of the primary sampling units. The primary sampling units have been the social centres and the ultimate ones the services (accommodation and meals) provided by them. First we stratified the population of centres according to two criteria, firstly the type of centre, dividing the centres in those devoted only to meals and the others that also provide accommodation and secondly the size of the centre measured by the number of services provided in an average day.

In order to obtain the desired accuracy level and in line with international experiences the theoretical sample size was planned at 4,000 services. The allocation of the sample of centres has been done proportional to the

number of services they provide. The selection of the sample of centres has been performed with probability proportional to its size (number of services provided) in each stratum. Depending on the size and type of the centre a number of interviews have been planned in each of the selected centres also taking into account that an interviewer can only conduct three interviews per day.

The weighting procedure has been performed following the weight-share method as described in Ardilly and Le Blanc (1). The selection process of the persons through the services has had as a consequence that the sample of persons has included repeated elements. The weight assigned to each person is the sum of the weights of the sample services that were provided to this person divided by the number of services (sample or not) that were provided to this person during the sample period.

For further information: results, methodology, questionnaire, please visit www.ine.es/inebase/menu3_soc_en.htm or contact Pedro Ruiz at pruizsal@ine.es or Alejandro Salinero at salinero@ine.es

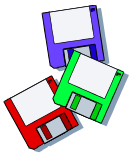
References:

(1) Ardilly, Pascal ; Le Blanc, David. *Sampling and Weighting a Survey of Homeless Persons: A French Example.* Survey Methodology, 2001, vol 27, 1, pp 109-118.

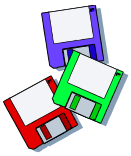
Country Representatives
Updated: June 5, 2006

Country (Region)	**Name**	
Algeria	Mme Nacer-Eddine Hammouda	A
Argentina	Mme Alicia Masautis	A
Australia	Mr. Paul Sutcliffe	A
Bahrain (Arabian Gulf)	Mr. Mohamed F. Abulata	
Belgium	Dr. (Mr.) Camille Vanderhoeft	A
Bolivia	Mr. Walter Castillo Guerra	
Botswana	Mrs. Ndiye Nko	A
Brazil	Dr. Pedro Louis do Nascimento Silva	A
Burkina Faso	M. François Ilboudo	
Cameroon	Mme Règine Zebaze	
Canada	Mr. John Kovar	A
Cape Verde	Mme Maria de Lurdes Fernandes Lopes	A
Chad	M. Ouagadjio Bandoumal	
Chile	Mr. Juan Eduardo Munoz	
China	Mr. Huang Langhui	A
Comoros	M. Mhadji Nailane	A
Congo	M. Alexis Lukaku Nzinga	A
Cuba	Mr. Luis Carlos Silva	
Czech Republic	Mr. Vaclav Cermak	
Denmark	Mr. Peter Linde	
Estonia	Dr. Imbi Traat	A
Finland	Mr. Paavo Väisänen	A
France	M. Benoît Riandey	A
Gabon	M. Jean-Pierre Zima Mefe	
Germany	Prof. Rolf Wiegert	
Ghana	Mrs. Elisabeth A. Allotey	
Greece	Mr. Anastassios Iliakopoulos	

Guatemala	Mr. David Fitch	A
Guinea	Mme Fatoumata Danfaca	A
Haiti	M. Emmanuel Charles	
Hungary	Mr. Laszlo Mihalyffy	A
India	*Vacant*	
Indonesia	Mrs. Sri Budianti Sukmadi	
Israel	Luisa Kadun Burck	A
Italy	Prof. Claudio Quintano	A
Ivory Coast	M. Cakpo Benjamin Zanou	A
Japan	*Vacant*	
Kenya	Mr. Philip Gachuki	
Korea	Prof. Hae-Young Lee	
Latvia	Dr. Janis Lapins	A
Lebanon	Prof. Bechara Hanna	
Libya	Mr. Farouk El Bishiti	
Lithuania	Dr. Danute Krapavickaite	A
Luxembourg	Mr. Antonio Baigorri Matamala	
Madagascar	Mme Julia Rachel Ravelosoa	A
Malawi	Dr. Suresh Chandra Babu	
Mali	*Vacant*	
Mauritania	M. Sidna Ould N'Dah	
Mexico	Mr. Miguel Cervera	
Morocco	*Vacant*	
Nepal	Mr. Shyam Upadhyaya	
Netherlands	Mr. Jos de Ree	
New Zealand	Ms. Vina Cullum	A
Nigeria	Mr. O.O. Ajayi	
Pakistan	Mr. Mohammad Ishaq (S.O.)	
Palestine	Mr. Faisal Awartani	
Papua New Guinea	Mr. John Shadlow	
Peru	Ms. Leonara Laguna	A
Philippines	Mr. Gervacio G. Selda, Jr. (Sun)	A
Poland	Mr. Janusz Wywial	A
Portugal	Mr. Paulo Jorge Gomes	
Russia	Mrs. Galina N. Sotnikova	
Senegal	M. Matar Gueye	A
Spain	Ms. Dolores Lorca Lopez	A
Spain (Basque Region)	Mrs. Anjeles Iztueta Azkue	A
Sweden	Mr. Peter Lindquist	A
Switzerland	Mr. Philippe Eichenberger	
Syria	Mr. Ibrahim Ali	A
Tanzania	Mr. Vitalis Eustach Muba	
Togo	M. Ousman Koriko	
Trinidad & Tobago	*Vacant*	
Tunisia	M. Abdellatif Sellami	
Turkey	Mr. Oztas Ayhan	A
United Kingdom	Prof. Peter Lynn	A
United States	Dr. Howard Hogar	
Uruguay	Mr. Daniel la Buonora	
Vietnam	Dr. Nguyen Quoc Anh	A



SOFTWARE REVIEW



Software reviews over the last several years have presented summaries of the features of a number of systems or components of systems that deal directly with survey estimation issues, particularly sampling error and weighting procedures. They have addressed such systems as EpiInfo, Wesvar, SUDAAN, SAS sampling error PROCedures, Stata, IVEware, and software for generalized regression weighting. Some of this software is available for free over the internet, and others require payment of a licensing fee.

We are considering additional reviews, and we would like to get reader advice about software that has not been reviewed to date, or for updated reviews on some software. We also would like to hear from you about whether there are software systems that address other survey sampling issues that you'd like us to review. Please keep in mind that we review only software that is available for purchase or for free download. We do not review proprietary or other software that our readers cannot purchase or access through a download.

Please send your advice and ideas to Jim Lepkowski, the review editor, at jimlep@umich.edu.

ASK THE EXPERTS

Anders Christianson

Q. What is meant by blurring in the context of statistical disclosure control?

A. Eric Schulte Nordholt, The Netherlands

Blurring in the context of statistical disclosure control has not a negative meaning as in the usual context. It means that a reported value is replaced by an average. There are many possible ways to implement blurring. Groups of records for averaging may be formed by matching on other variables or by sorting on the variable of interest. The number of records in a group (whose data will be averaged) may be fixed or random. The average associated with a particular group may be assigned to all members of a group, or to the "middle" member (as in a moving average). It may be performed on more than one variable with different groupings for each variable. More information on statistical disclosure control can be found in the glossary that has been published under the knob GLOSSARY at the home page of the CASC project on the internet: http://neon.vb.cbs.nl/casc.

Q. We made a survey last year in the housing sector, and we are going to repeat it next year in order to measure changes. What do we need to think of? Should we make an independent sample?

A. Anders Christianson

An independent sample is probably not a good idea. The reason is that you could benefit largely from the positive correlations that occur for many housing survey variables, between the two years that you compare, by using (a large part of) last year's sample again, and basing your inference on the differences for individual households. Such a sampling procedure can create drastic gains in sampling variances compared to the one based on two independent samples, because the positive correlations are very favorable for the same sample used at both occasions with an estimator based on differences. In Wallis & Roberts (1965) an example is given showing that 25 units, measured twice, were comparable in precision to two samples of 2,222 units, measured once each! As the authors put it: "This illustrates the potential importance of proper statistical planning before collecting data."

Another issue that is important to keep in mind is that the estimators really will estimate the true change. If you make changes to the methodology of the survey, you will have to exclude those as explanation of the difference that is obtained. So, the same mode of data collection, and the same questionnaire at the two occasions, would probably be a wise design decision.

Wallis W A & Roberts H V: Statistics: A New Approach, Twelfth printing 1965.

Q. What can you tell about survey quality awareness?

A. Anders Christianson

Quality awareness among users and producers is, of course, of great importance for statistical surveys as well as in any other branch. For statistical surveys, however, quality has many facets, some of which can be evaluated by means of a so-called customer satisfaction survey. For instance, the user as well as the producer of a statistical survey easily obtains the timeliness accuracy. For other aspects of survey quality, more sophisticated methods must be used to assess the quality. Most important is the accuracy of the survey.

Accuracy is an indicator of the degree to which the user can rely on the results of a

survey. The accuracy should meet the needs of the users and the inevitable inaccuracies should be properly reported, thus enabling the users to make their own judgments as to whether the quality of the data supports intended uses or not.

Perfect survey quality is merely a theoretical ideal. Inaccuracies come from errors in the survey. Sources of errors are the sampling procedure (sampling errors) and other steps of the survey (nonsampling errors), particularly nonresponse and measurement errors.

It is crucial that the producer obtains a fair assessment of the accuracy of the survey by means of evaluation efforts like the computing of confidence intervals and nonresponse rates, to mention the two most common measures. Thus, the producer's perceived accuracy is obtained. It is equally important that this is communicated to the users to affect their perceived accuracy. This is done by means of a quality declaration.

Generally, the assessment of quality measures of random errors, i.e. sampling quality, is readily obtained by the computation of confidence intervals (or point estimates together with their coefficients of variation, CV) that measure the very uncertainty associated with the sampling error, taking advantage of the sampling theory. Unfortunately, there is no unified theory that evaluates response quality in a similar way. Actually the quality awareness varies with the different sources of errors. This may be described by the quality awareness ladder, introducing the following four steps of quality awareness:

Step 4: Quantitative measure of inaccuracy (confidence interval) available
Step 3: Quantitative measure of indicator of issue (nonresponse rate) available
Step 2: Vague awareness of quality issue occurs (as often for measurement error)
Step 1: No awareness of quality issue (sometimes detected "by accident")

Obviously, the quality awareness varies from one issue to another. Sampling errors represent the highest step of quality awareness, whereas indicators or judgments normally evaluate nonsampling errors. There is no single measure that brings together all the uncertainty of a statistical survey. However, much work in the realm of survey methods development today is devoted to bringing the awareness of quality issues up the ladder, and much of this work has been successful, though far from completed in terms of a single "total error quality indicator". Especially, works on measurement error models and variance due to imputation make it possible to provide quality information beyond that of rates.

Statistical agencies have developed, and are using, quality guidelines and/or policies to help staff have quality in their mind during all steps of surveys and to provide a framework for correctly informing the users of data quality.

Christianson A and Polfeldt T (1995): Evaluation and Improvement of Response Quality at Statistics Sweden – A Research Approach. In Statistics Sweden: European Harmonization, National Decentralization and Quality. Proceedings of the 1st International Conference on Methodological Issues in Official Statistics. Stockholm, June 12-13 1995.

Christianson A and Tortora R D (1995): Issues in Surveying Businesses: An International Survey. Chapter 14 of Cox B G, Binder D A, Nanjamma Chinnappa B, Christianson A, Colledge M J, and Kott P S. Business Survey methods. Wiley ISBN 0-471-59852-6.

Question: What is a tolerable nonresponse rate?

A. Anders Christianson

This is a somewhat controversial question. Historically, the view upon what is a tolerable nonresponse rate, was stricter, say 40 years ago, than it is today. Some agencies even applied so called minimum performance standards, implying that survey results were suppressed when standards were not met. Improved weighting methods and a hardening survey climate have made the view of what is a tolerable nonresponse rate more liberal.

However, there is still a common agreement in two respects:

First, it depends on the purpose of the survey if the nonresponse rate should be considered acceptable or not. Some decisions demand a higher degree of accuracy than others do. It is the survey user's responsibility to take the uncertainty of nonresponse into account when she or he makes decisions.

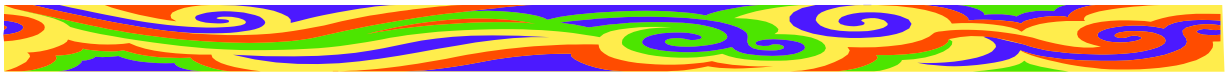
Thus, it is the survey producer's (if other than the user) obligation to communicate a fair account of this uncertainty to the users. There are also limits as to when a probability sample still is a probability sample when the non-response increases, thus jeopardizing the basis of inference from a sample to the population from which it is drawn.

So, there is no straightforward answer to your question in terms of a specific percentage. An interesting discussion on this, giving different points of view, is to be found in the July 2002 issue of the Survey Statistician in the "discussion corner" starting with an article entitled "Avoid the Need to Impute."

CHANGE OF ADDRESS

Members are encouraged to inform the IASS Secretariat of changes of address as soon as possible. Mailings of the proceedings of the IASS papers presented at the ISI sessions, and "The Survey Statistician" will be delayed and may be lost if the Secretariat does not have your correct address.

You may notify Ms. Claude Olivier of your change of address by completing and mailing the Change of Address form given at the end of this newsletter. Alternatively, you can provide the same information to Ms. Olivier by email to claude.olivier@insee.fr.



Articles

In Memoriam

Joseph Waksberg

Joseph Waksberg (known universally as "Joe") died suddenly on January 10, 2006. He had celebrated his 90th birthday in September. At the time of his death, he was the Chair of the Board of Directors of Westat, a statistical survey research organization which he had joined in 1973 upon his retirement from the Census Bureau. Throughout a career that spanned more than 60 years, he made many important contributions to methods of survey sampling, developed innovative applications of the methods, and conducted research in a broad array of survey methodology issues.

Many statisticians, of course, have made major contributions to the field, but what set Joe apart was his ability to blend the theoretical with the pragmatic, and to understand how best to blend the statistical requirements of the task with the operational aspects, whether because of cost limitations, or a possible clash with field or processing requirements. Invariably, when faced with such conflicts, Joe was able to find the best and most appropriate compromise that both satisfied the needs of adherence to statistical standards and also met time and cost constraints. He also had that rare ability to think quickly. For example, during the 1970 Decennial Census, it became apparent that interviewers were overstating the number of vacant units. Within a week or so, Joe proposed, developed, and implemented an effective and efficient national sample design to revisit a sample of reported vacant units in each area to determine the extent of overstatement, along with a statistical technique to be applied to each local area to correct for the overstatement.

Solutions to what had seemed insoluble problems seemed to come to him almost without effort. Sometime during the early 1950s, for example, some Census Bureau staff met in an attempt to come to grips with the problem of "new construction". At that time, area segments were selected with probabilities proportional to estimated size, and the housing units within them were listed by interviewers for subsequent sample selection. Far too often, and not surprisingly, interviewers found many times the number of housing units expected because of large, newly constructed developments. The unexpected growth was causing serious problems of many types and, although ad hoc solutions had been developed, the problem was far from solved and only getting worse. It was Joe who suddenly sat up in his chair and uttered the equivalent of "Eureka". He proposed using the Bureau's ongoing construction statistics program as the basis for developing an independent count of permits filed for new construction, and using it as the source of "new construction" sampling for household surveys, an approach still in use today.

While working on a survey to determine the expenditures by homeowners of alterations and repairs, Joe raised questions about their ability to recall costs and the associated timings and the effect such recall problems might have on the data. Together with John

Neter, Joe subsequently developed a major study to investigate such problems. The results of their research are widely cited for the light they shed on the magnitude of various types of recall problems.

And, then, there is his work on models to estimate election night results. First, Joe developed a very elegant theory of sampling using procedures that Warren Mitofsky had discovered to improve the efficiency of sampling telephone numbers. The Mitofsky-Waksberg method quickly became the standard for all telephone sampling for decades. Subsequently, he adapted the methodology for use in identifying controls for case-control studies, extending the utility of the method into a very different application. As a result of Joe's contributions in these areas, Westat become a leader in telephone survey research. .

Another of Joe's facets was that he didn't hesitate to learn from others. As soon as he'd read a paper or heard a talk about a new approach to some problem or other, Joe began the process of finding how and how best to adapt it for use in meeting present or future challenges. When technological advances resulted in a modification of the RDD approach he'd developed, he was a leader in the research to enhance this new method of sampling telephone numbers. His mind was always open to something new, even if it meant he had to do some adapting to fit his need, and his zeal in seeking it out spread far and wide among his colleagues.

In sum, can there be any doubt that much of Joe's research can be called "seminal"?

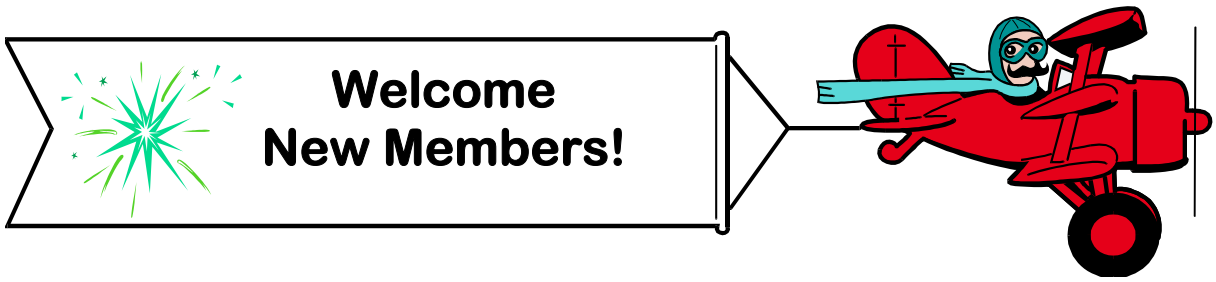
Joe was an exceptionally gifted statistician, but what really set him apart were his unique qualities as a human being. He was that very rare statistician who could explain a complicated statistical idea to the unsophisticated and untutored, using only basic English, and without the need of a blackboard covered with mathematical symbols and formulae. And if you didn't quite get it the first time, there was no hesitation, no impatience, in explaining it again until you understood. Many frequent beneficiaries of his generosity can attest both to that unique skill and to its importance in making those around him better equipped to do their jobs.

He also had that wonderful ability of never making anyone feel awkward or inferior, no matter what their question or their level of accomplishment. Another of his talents was his ability to write clearly and easily, no matter the subject, and he strove, constantly, to impart the need to do so to others. No matter how bad the product he was reviewing or how many comments filled the margins, however, he always found something to praise, and took great pains to explain in supportive words the why's and wherefore's of his comments. In dealing with others, whether close colleagues or clients, Joe never showed anger or said anything hurtful or demeaning; rather, he was always quite careful to ensure that, no matter how great the difference or the disagreement, it did not carry over beyond that moment, and that all parting was amicable, showing neither anger nor displeasure. As was correctly said by the MC at his retirement party, "Joe was the Sara Lee of statisticians; no one ever didn't like Sara Lee." Truly, no one ever didn't like Joe Waksberg!

A long-time member and Fellow of the ASA, Joe also served as President of WSS, as Chair of the Sections of Survey Research Methods and Social Statistics, and as member of the Nominating and Elections Committee and the Committee on ASA Fellows. He also delivered the 1966 Morris Hansen Lecture. He was an elected member of the ISI, and received numerous honors and awards, including the Gold and Silver Medal Awards of the U.S. Department of Commerce for exceptional service to the Federal Statistical System, and the Roger Herriot Award from the Washington Statistical Society. And in 2001, *Survey Methodology*, with support from Westat, established the annual Waksberg Invited Paper Series in his honor, with the paper also being presented at Statistics Canada's Annual Symposium.

Last, but not least, he was teacher and mentor to generations of statisticians.

Daniel Levine



We are very pleased to welcome the following new members:

Country	Last Name	First Name
Argentina	Wolman	Guido
Nigeria	Onyeka	Aloy
Spain	Alfaro Cortes	Esteban
	Alfaro Navarro	Jose Luis
	Cancelo de la Torre	Jose Ramon
	Fernandez Aviles Calderon	Gema
	Gamez Martinez	Matias
	Garcia Rubio	Noelia
	Larraz Iribas	Beatriz
	Mondejar Jimenez	Jose
	Mondejar Jimenez	Juan Antonio
	Montero Lorenzo	Jose Maria
	Prado	Cristina
	Vargas Vargas	Manuel
Sweden	Elvers	Eva
Tunisia	Mubila	Maurice
United Kingdom	Correa	Solange
	Trujillo	Leonardo
United States	Banks	David
	Balaji	Lakshmi

Dear New Member:
For questions or input regarding *The Survey Statistician*, please contact:
Steven Heeringa
Institute for Social Research
426 Thompson St., Rm 4046
Ann Arbor, MI 48104 USA
sheering@isr.umich.edu

Announcements

Third International Conference on Establishment Surveys (ICES-III): Survey Methods for Businesses, Farms, and Institutions

McGill University, Royal Victoria Hospital and Mount Royal

June 18-21, 2007
Hyatt Regency Montréal
Montréal, Québec, Canada

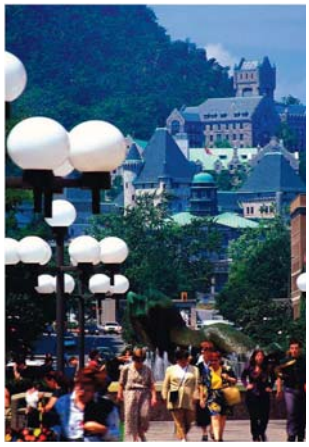


Photo Credit: © Tourisme Montréal, Stéphan Poulin

Co-sponsored by: American Statistical Association (ASA), ASA Section on Survey Research Methods, ASA Section on Government Statistics, International Association of Survey Statisticians, and Statistical Society of Canada.

The first International Conference on Establishment Surveys (ICES-I) in 1993 convened more than 400 experts in the area of surveys of businesses, farms, and institutions.

ICES-I set the stage by formally documenting the state of the art in 1993. In 2000, ICES-II took a forward look at methods for surveying businesses, farms, and institutions. ICES-III will build on the cross-national interdisciplinary research cooperation on the unique features of establishment surveys started at ICES-I in Buffalo, NY in 1993, and continued at ICES-II in 2000, also in Buffalo.

ICES-III is organizing a stimulating invited and contributed program. This is just a sampling of the invited sessions:

- Surveys of Environmental Protection: Experiences and Challenges
- Commodity Flow Surveys
- Use of Metadata for Establishment Surveys
- Collecting Data Electronically from Businesses: the Dream and the Reality

The contributed program will become final in September 2006.

The Conference also includes:

- Two full day and two half day short courses on Monday, June 18,
- Robert W. Edwards of the International Monetary Fund as the Keynote speaker on Monday evening,
- Eight introductory overview lectures,
- Poster sessions, and software demonstrations throughout the conference,
- A terrific social program to see and learn about Montréal,
- A celebration dinner on Thursday evening, June 21, and
- Following the Conference, a CD ROM of all presented papers will be sent to all Conference participants.

All Conference information can be accessed on the ICES-III Web Site www.amstat.org/meetings/ices/2007/index.cfm. Online registration opens on Monday, December 18, 2006. All Conference activities will be held at the Hyatt Regency Montréal. The Hyatt Regency Montréal web site is http://montreal.hyatt.com/. Hotel reservations open on March 1, 2007.

Montréal is a city rich in history, art, culture, and learning and sports. It combines the smarts of North America with the laid back *insouciance* of Europe, high tech and high style. The city's architecture too is an artful mix of old and new, with graceful historic structures lovingly preserved and merged with the cool lines of the 21st

century. The visual arts flourish here at the Montréal Museum of Fine Arts, the Musée D'art Contemporain, and hundreds of top-notch local galleries.

Robert Rutchik

* *

Call for Software Demonstrations

Since 2000, when the second ICES was held, many processing systems have been developed by statistical agencies or software organizations around the globe. Therefore, ICES III is setting aside a room for demonstrating software used in establishment surveys. These demonstrations should target live processing of data and be able to adapt to the interests of specific audiences. Inflexible slide shows or presentations are discouraged. Proposed software packages should be in use for one or more establishment surveys already and be designed to automate survey processes such as:

- Sample design and selection
- Data collection, instrument design, capture and coding
- Record linkage and matching
- Editing and imputation
- Survey data analysis
- Weighting, estimation, and tabulation
- Time series and seasonal adjustment
- Disclosure analysis
- Data dissemination and presentation

Schedule and Equipment: Demonstrations will take place during regular conference sessions from June 19–20. They will be split into four groups, with a dedicated half day for each. ICES III will provide participants with high-speed internet connections, 18" flat-panel monitors, tables, and chairs. Participants must provide their own computers with demonstration software set up.

How to send your proposal: Submit a 200-word abstract by December 1, 2006, and include a description of the software package, potential applications in other survey organizations, and special equipment required for the demonstration. A registration fee and completed form will be required later. Registration forms and details may be obtained at www.amstat.org/meetings/ices/2007. Proposals and questions should be sent to Dale Atkinson at Dale_Atkinson@nass.usda.gov or (703) 877-8000, ext. 130.

Visit our web site at www.amstat.org/meetings/ices/2007

Dale Atkinson
ICES-III Organizing Committee
USDA/NASS/RDD
Chief, Census and Survey Research Branch

New Journal: Survey Research Methods

Peter Lynn, Editor

A new peer-reviewed journal, *Survey Research Methods*, has been set up under the auspices of The European Survey Research Association (ESRA). The journal, edited by Peter Lynn of the University of Essex, UK, publishes articles in English which discuss methodological issues related to survey research. Two types of papers are in-scope:

1. Papers discussing methodological issues in substantive research using survey data;

2. Papers that discuss methodological issues that are more or less independent of the specific field of substantive research.

Topics of particular interest include survey design, sample design, question and questionnaire design, data collection, nonresponse, data capture, data processing, coding and editing, measurement errors, imputation, weighting and survey data analysis methods. The journal aims to be multidisciplinary. The journal will be published electronically and aims for a fast review and publication process, so that papers could be published within three months of initial submission.

Further information, including details of how to submit papers, can be found on the journal website http://esra.sqp.nl/esra/journal/ .

Survey Methodology on-line!

Survey Methodology, the journal published by Statistics Canada is now available on-line in fully searchable pdf format. Starting with the December 2005 issue, all articles of the journal are now being made available free of charge directly on the web site upon release. Further, plans are to also include previous issues. Currently, all articles of the latest seven issues have been posted and work is in progress to add those of the previous 10 years. Printed versions will still be produced and subscription to receive them remains in effect. Older issues can be obtained upon request in paper or pdf scanned formats. *Survey Methodology* can be accessed from Statistics Canada's web site at www.statcan.ca/bsolc/english/bsolc?catno=12-001-X. For submission of papers and inquiries about the journal, please contact the editor, John Kovar at kovar@statcan.ca.

Eric Rancourt

Cochran-Hansen Prize 2007: Competition for Young Survey Statisticians from Developing and Transition Countries

In celebration of its 25th anniversary the International Association of Survey Statisticians established the Cochran-Hansen Prize to be awarded every two years to the best paper on survey research methods submitted by a young statistician from a developing or transition country.

Participation in the competition for the Prize is open to nationals of developing or transition countries who are living in such countries and who were born in 1967 or later.

Papers submitted must be unpublished original works, They may include materials from the participant's university thesis. They should be in either English or French. Papers for consideration should be submitted to the IASS Secretariat at the address below to arrive by December 29, 2006. Each submission should be accompanied by a cover letter that gives the participant's year of birth, nationality, and country of residence.

The papers submitted will be examined by the Cochran-Hansen Prize Committee appointed by IASS. The decision of the Committee is final.

The winner of the prize will be invited to present his/her paper at the 56th Session of the International Statistical Institute to be held in Lisbon, Portugal, August 22-29, 2007, and the name of the winner will be announced at the ISI General Assembly in Lisbon.

The author of the winning paper will receive the Cochran-Hansen Prize in the form of books and journal subscriptions to the value of about 500 Euros and will have reasonable travel and living expenses paid in order to present the paper at the ISI Session in Lisbon.

For further information please write to:
Madame Claude OLIVIER
IASS Secretariat
International Association of Survey Statisticians
CEFIL-INSEE, 3 rue de la Cité, 33500 Libourne, France
Tel : +33 5 57 55 56 17
Fax : +33 5 57 55 56 20
E-mail : Claude.olivier@insee.fr

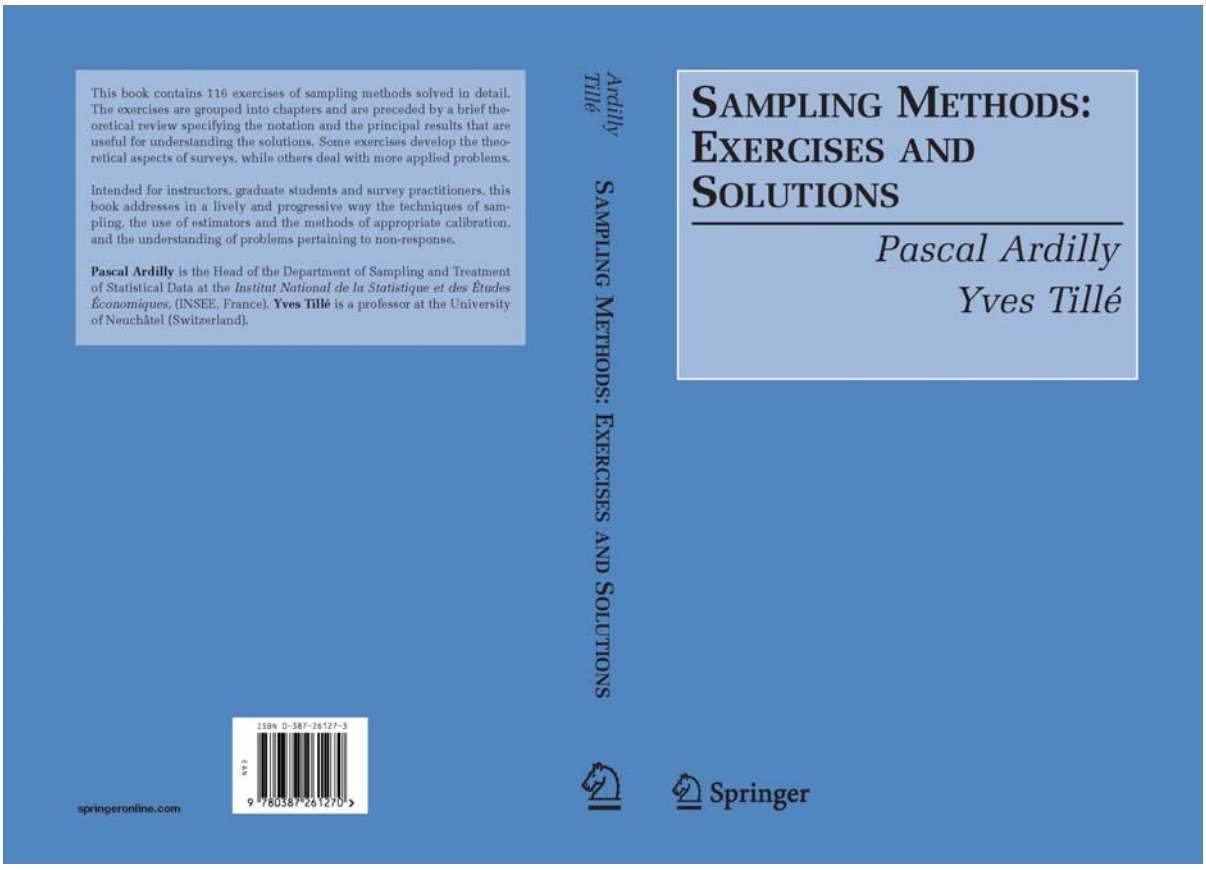
To All Members

- The IASS needs your contribution.
- Please do not forget to renew your membership.
- As of January 2002, French Francs are no longer accepted. As a consequence, the payment of dues and subscriptions must be made in either Euros or U.S. dollars.

Announcing the English Version of a New Book by IASS Members
Pascal Ardilly and Yves Tillé

This book, the original version of which was published by the authors in French in 2003, contains 116 exercises of sampling methods solved in detail. The exercises are grouped into chapters and are preceded by a brief theoretical review specifying the notation and the principal results that are useful for understanding the solutions. Some exercises develop the theoretical aspects of surveys, while others deal with more applied problems. Intended for instructors, graduate students and survey practitioners, this book addresses in a lively and progressive way the techniques of sampling, the use of estimators and the methods of appropriate calibration, and the understanding of problems pertaining to non-response.

Pascal Ardilly is the Head of the Department of Sampling and Treatment of Statistical Data at the *Institut National de la Statistique et des Études Économiques*, (INSEE, France). **Yves Tillé** is a professor at the University of Neuchâtel (Switzerland).
springeronline.com



The new glossary on Statistical Disclosure Control

At the fourth Joint ECE / Eurostat Work Session on Statistical Data Confidentiality (7-9 April 2003) in Luxembourg the idea for a glossary on Statistical Disclosure Control was launched. A team of five people was created to work on this glossary: Mark Elliot (University of Manchester), Anco Hundepool (Statistics Netherlands), Eric Schulte Nordholt (Statistics Netherlands), Jean-Louis Tambay (Statistics Canada) and Thomas Wende (Destatis, Germany). The five people who produced the glossary were present at that Work Session and also met on 18 August 2003 at the ISI Session in Berlin. In the period 2003-2005 preliminary versions have been presented so that a number of experts in the field from all over the world could comment on these versions. The authors acknowledge the input of many people during the drafting of this new glossary, especially: Paul Feuvrier (INSEE, France), Kingsley Purdam (University of Manchester), Barry Schouten (Statistics Netherlands), Duncan Smith (University of Manchester) and Peter-Paul de Wolf (Statistics Netherlands). The final version of this new glossary on Statistical Disclosure Control was presented as Working Paper 45 by Eric Schulte Nordholt at the fifth Joint ECE / Eurostat Work Session on Statistical Data Confidentiality (9-11 November, 2005) in Geneva. More information on this Work Session (including all presented papers) can be found at http://www.unece.org/stats/documents/2005.11.confidentiality.htm.

The aim of this glossary is twofold: firstly, it should help people who are new in the field to get acquainted with the terminology used in Statistical Disclosure Control and secondly it can be used in courses on Statistical Disclosure Control as a back-up facility. We hope that this glossary will be useful and the two aims will be reached. As we believe in the future a more flexible structure is better, Anco Hundepool has put the glossary on the internet pages of the CASC (Computational Aspects of Statistical Confidentiality) project: http://neon.vb.cbs.nl/casc. The glossary can be found under the knob GLOSSARY on the home page of the CASC website. If you have any comments or questions, please forward them to Eric Schulte Nordholt (e-mail: ESLE@CBS.NL) so that they can be used to improve the electronic glossary.

GLOSSARY ON STATISTICAL DISCLOSURE CONTROL

Mark Elliot (University of Manchester)
Anco Hundepool (Statistics Netherlands)
Eric Schulte Nordholt (Statistics Netherlands)
Jean-Louis Tambay (Statistics Canada)
Thomas Wende (Destatis, Germany)

Version September 2005

Introduction

At the Joint ECE / Eurostat Work Session on Statistical Data Confidentiality (7-9 April 2003) in Luxembourg the idea for a glossary on Statistical Disclosure Control was launched. The five people who produced this glossary were present at that Work Session and also met on 18 August 2003 at the ISI Session in Berlin. This new glossary on Statistical Disclosure Control will be presented at the next Joint ECE / Eurostat Work Session on Statistical Data Confidentiality (9-11 November, 2005) in Geneva. In the meantime preliminary versions have been presented so that experts in the field from all over the world could comment on these versions. The aim of this glossary is twofold: firstly, it should help people who are new in the field to get acquainted with the terminology used in Statistical Disclosure Control and secondly it can be used in courses on Statistical Disclosure Control as a back-up facility. We hope that this glossary will be useful and the two aims will be reached. If you have any comments or questions, please forward them to Eric Schulte Nordholt (e-mail: ESLE@CBS.NL) so that they can be taken into account for future versions.

Acknowledgements

The authors acknowledge the input of many people during the drafting of this document, especially: Paul Feuvrier (INSEE), Kingsley Purdam (University of Manchester), Barry Schouten (Statistics Netherlands), Duncan Smith (University of Manchester) and Peter-Paul de Wolf (Statistics Netherlands).

A

Ambiguity rule: Synonym of (p,q) rule.

Analysis server: A form of **remote data laboratory** designed to run analysis on data stored on a safe server. The user sees the results of their analysis but not the data.

Anonymised data: Data containing only anonymised records.

Anonymised record: A record from which direct identifiers have been removed.

Approximate disclosure: Approximate disclosure happens if a user is able to determine an estimate of a respondent value that is close to the real value. If the estimator is exactly the real value the disclosure is exact.

Argus: Two software packages for Statistical Disclosure Control are called Argus. μ-Argus is a specialized software tool for the protection of **microdata**. The two main techniques used for this are **global recoding** and **local suppression**. In the case of **global recoding** several categories of a variable are collapsed into a single one. The effect of local suppression is that one or more values in an unsafe combination are suppressed, i.e. replaced by a missing value. Both **global recoding** and **local suppression** lead to a loss of information, because either less detailed information is provided or some information is not given at all. τ-Argus is a specialized software tool for the protection of tabular data. τ-Argus is used to produce safe tables. τ-Argus uses the same two main techniques as μ-Argus: global recoding and local suppression. For τ-Argus the latter consists of suppression of cells in a table.

Attribute disclosure: Attribute disclosure is **attribution** independent of identification. This form of disclosure is of primary concern to **NSIs** involved in **tabular data** release and arises from the presence of empty cells either in a released table or linkable set of tables after any subtraction has taken place. Minimally, the presence of an empty cell within a table means that an intruder may infer from mere knowledge that a population unit is represented in the table and that the intruder does not possess the combination of attributes within the empty cell.

Attribution: Attribution is the association or disassociation of a particular attribute with a particular population unit.

B

Barnardisation: A method of disclosure control for tables of counts that involves randomly adding or subtracting 1 from some cells in the table.

Blurring: Blurring replaces a reported value by an average. There are many possible ways to implement blurring. Groups of records for averaging may be formed by matching on other variables or by sorting on the variable of interest. The number of records in a group (whose data will be averaged) may be fixed or random. The average associated with a particular group may be assigned to all members of a group, or to the "middle" member (as in a moving average). It may be performed on more than one variable with different groupings for each variable.

Bottom coding: See **top and bottom coding**.

Bounds: The range of possible values of a cell in a table of frequency counts where the cell value has been perturbed or suppressed. Where only margins of tables are released it is possible to infer bounds for the unreleased joint distribution. One method for inferring the bounds across a table is known as the **Shuttle algorithm**.

C

Calculated interval: The interval containing possible values for a suppressed cell in a table, given the table structure and the values published.

Cell suppression: In tabular data the cell suppression SDC method consists of **primary and complementary (secondary) suppression**. Primary suppression can be characterised as withholding the values of all risky cells from publication, which means that their value is not shown in the table but replaced by a symbol such as '×' to indicate the suppression. According to the definition of risky cells, in frequency count tables all cells containing small counts and in tables of magnitudes all cells containing small counts or presenting a case of **dominance** have to be primary suppressed. To reach the desired protection for risky cells, it is necessary to suppress additional non-risky cells, which is called **complementary (secondary) suppression**. The pattern of complementary suppressed cells has to be carefully chosen to provide the desired level of ambiguity for the risky cells with the least amount of suppressed information.

Complementary suppression: Synonym of **secondary suppression**.

Complete disclosure: Synonym of **exact disclosure**.

Concentration rule: Synonym of **(n,k) rule**.

Confidentiality edit: The confidentiality edit is a procedure developed by the U.S. Census Bureau to provide protection in data tables prepared from the 1990 Census. There are two different approaches: one was used for the regular Census data; the other was used for the long-form data, which were filled by a sample of the population. Both techniques apply statistical disclosure limitation techniques to the **microdata** files before they are used to prepare tables. The adjusted files themselves are not released; they are used only to prepare tables. For the regular Census microdata file, the confidentiality edit involves "data swapping" or "switching" of attributes between matched records from different geographical units. For small blocks, the Census Bureau increases the sampling fraction. After the microdata file has been treated in this way, it can be used directly to prepare tables and no further disclosure analysis is needed. For long form data, sampling provides sufficient confidentiality protection, except in small geographic regions. To provide additional protection in small geographic regions, one household is randomly selected and a sample of its data fields are blanked and replaced by imputed values.

Controlled rounding: To solve the additivity problem, a procedure called controlled rounding was developed. It is a form of **random rounding**, but it is constrained to have the sum of the published entries in each row and column equal to the appropriate published marginal totals. Linear programming methods are used to identify a controlled rounding pattern for a table.

Controlled Tabular Adjustment (CTA): A method to protect tabular data based on the selective adjustment of cell values. Sensitive cell values are replaced by either of their closest safe values and small adjustments are made to other cells to restore the table additivity. Controlled tabular adjustment has been developed as an alternative to **cell suppression**.

Conventional rounding: A disclosure control method for tables of counts. When using conventional rounding, each count is rounded to the nearest multiple of a fixed base. For example, using a base of 5, counts ending in 1 or 2 are rounded down and replaced by counts ending in 0 and counts ending in 3 or 4 are rounded up and replaced by counts ending in 5. Counts ending between 6 and 9 are treated similarly. Counts with a last digit of 0 or 5 are kept unchanged. When rounding to base 10, a count ending in 5 may always be rounded up, or it may be rounded up or down based on a rounding convention.

D

Data divergence: The sum of all differences between two datasets (data-data divergence) or between a single dataset and reality (data-world divergence). Sources of data divergence include: data ageing, response errors, coding or data entry errors, differences in coding and the effect of disclosure control.

Data intruder: A data user who attempts to disclose information about a population unit through **identification** or **attribution**.

Data intrusion detection. The detection of a **data intruder** through their behaviour. This is most likely to occur through analysis of a pattern of requests submitted to a **remote data laboratory**. At present this is only a theoretical possibility, but it is likely to become more relevant as **virtual safe settings** become more prevalent.

Data Intrusion Simulation (DIS). A method of estimating the probability that a **data intruder** who has matched an arbitrary population unit against a sample unique in a target microdata file has done so correctly.

Data protection: Data protection refers to the set of privacy-motivated laws, policies and procedures that aim to minimise intrusion into respondents' privacy caused by the collection, storage and dissemination of personal data.

Data swapping: A disclosure control method for microdata that involves swapping the values of variables for records that match on a representative key. In the literature this technique is also sometimes referred to as "multidimensional transformation". It is a transformation technique that guarantees (under certain conditions) the maintenance of a set of statistics, such as means, variances and univariate distributions.

Data utility: A summary term describing the value of a given data release as an analytical resource. This comprises the data's **analytical completeness** and its **analytical validity**. Disclosure control methods usually have an adverse effect on data utility. Ideally, the goal of any disclosure control regime should be to maximise data utility whilst minimising disclosure risk. In practice disclosure control decisions are a trade-off between utility and **disclosure risk**.

Deterministic rounding: Synonym of **conventional rounding**.

Direct identification: Identification of a statistical unit from its **formal identifiers**.

Disclosive cells: Synonym of **risky cells**.

Disclosure: Disclosure relates to the inappropriate attribution of information to a data subject, whether an individual or an organisation. Disclosure has two components: **identification** and **attribution**.

Disclosure by fishing: This is an attack method where an intruder identifies risky records within a target data set and then attempts to find population units corresponding to those records. It is the type of disclosure that can be assessed through a **special unique analysis**.

Disclosure by matching: Disclosure by the linking of records within an identification dataset with those in an anonymised dataset.

Disclosure by response knowledge: This is disclosure resulting from the knowledge that a person was participating in a particular survey. If an intruder knows that a specific individual has participated in the survey, and that consequently his or her data are in the data set, identification and disclosure can be accomplished more easily.

Disclosure by spontaneous recognition: This means the recognition of an individual within the dataset. This may occur by accident or because a data intruder is searching for a particular individual. This is more likely to be successful if the individual has a rare combination of characteristics which is known to the intruder.

Disclosure control methods: There are two main approaches to control the disclosure of confidential data. The first is to reduce the information content of the data provided to the external user. For the release of tabular data this type of technique is called **restriction-based disclosure control method** and for the release of microdata the expression disclosure control by data reduction is used. The second is to change the data before the dissemination in such a way that the disclosure risk for the confidential data is decreased, but the information content is retained as much as possible. These are called **perturbation based disclosure control methods**.

Disclosure from analytical outputs: The use of output to make attributions about individual population units. This situation might arise to users that can interrogate data but do not have direct access to them such as in a **remote data laboratory**. One particular concern is the publication of residuals.

Disclosure limitation methods: Synonym of **disclosure control methods**.

Disclosure risk: A disclosure risk occurs if an unacceptably narrow estimation of a respondent's confidential information is possible or if exact disclosure is possible with a high level of confidence.

Disclosure scenarios: Depending on the intention of the intruder, his or her type of a priori knowledge and the microdata available, three different types of disclosure or disclosure scenarios are possible for microdata: **disclosure by matching, disclosure by response knowledge** and **disclosure by spontaneous recognition**.

Dissemination: Supply of data in any form whatever: publications, access to databases, microfiches, telephone communications, etc.

Disturbing the data: This process involves changing the data in some systematic fashion, with the result that the figures are insufficiently precise to disclose information about individual cases.

Dominance rule: Synonym of **(n,k) rule**.

E

Exact disclosure: Exact disclosure occurs if a user is able to determine the exact attribute for an individual entity from released information.

F

Formal identifier: Any variable or set of variables which is structurally unique for every population unit, for example a population registration number. If the formal identifier is known to the intruder, identification of a target individual is directly possible for him or her, without the necessity to have additional knowledge before studying the microdata. Some combinations of variables such as name and address are pragmatic formal identifiers, where non-unique instances are empirically possible, but with negligible probability.

G

Global recoding: Problems of confidentiality can be tackled by changing the structure of data. Thus, rows or columns in tables can be combined into larger class intervals or new groupings of characteristics. This may be a simpler solution than the suppression of individual items, but it tends to reduce the descriptive and analytical value of the table. This protection technique may also be used to protect microdata.

H

HITAS: A heuristic approach to **cell suppression** in hierarchical tables.

I

Identification: Identification is the association of a particular record within a set of data with a particular population unit.

Identification dataset: A dataset that contains formal identifiers.

Identification data: Those personal data that allow direct identification of the data subject, and which are needed for the collection, checking and matching of the data, but are not subsequently used for drawing up statistical results.

Identification key: Synonym of **key**.

Identification risk: This risk is defined as the probability that an intruder identifies at least one respondent in the disseminated microdata. This identification may lead to the disclosure of (sensitive) information about the respondent. The risk of identification depends on the number and nature of **quasi-identifiers** in the microdata and in the a priori knowledge of the intruder.

Identifying variable: A variable that either is a formal identifier or forms part of a formal identifier.

Indirect identification: Inferring the identity of a population unit within a microdata release other than from **direct identification**.

Inferential disclosure: Inferential disclosure occurs when information can be inferred with high confidence from statistical properties of the released data. For example, the data may show a high correlation between income and purchase price of home. As the purchase price of a home is typically public information, a third party might use this information to infer the income of a data subject. In general, NSIs are not concerned with inferential disclosure for two reasons. First, a major purpose of statistical data is to enable users to infer and understand relationships between variables. If NSIs equated disclosure with inference, no data could be released. Second, inferences are designed to predict aggregate behaviour, not individual attributes, and thus often poor predictors of individual data values.

Informed consent: Basic ethical tenet of scientific research on human populations. Sociologists do not involve a human being as a subject in research without the informed consent of the subject or the subject's legally authorized representative, except as otherwise specified. Informed consent refers to a person's agreement to allow personal data to be provided for research and statistical purposes. Agreement is based on full exposure of the facts the person needs to make the decision intelligently, including awareness of any risks involved, of uses and users of the data, and of alternatives to providing the data.

Intruder: A data user who attempts to link a respondent to a microdata record or make attributions about particular population units from aggregate data. Intruders may be motivated by a wish to discredit or otherwise harm the NSI, the survey or the government in general, to gain notoriety or publicity, or to gain profitable knowledge about particular respondents.

J

K

Key: A set of **key variables**.

Key variable: A variable in common between two datasets, which may therefore be used for linking records between them. A key variable can either be a **formal identifier** or a **quasi-identifier**.

L

Licensing agreement: A permit, issued under certain conditions, for researchers to use confidential data for specific purposes and for specific periods of time. This agreement consists of contractual and ethical obligations, as well as penalties for improper disclosure or use of identifiable information. These penalties can vary from withdrawal of the license and denial of access to additional data sets to the forfeiting of a deposit paid prior to the release of a **microdata** file. A licensing agreement is almost always combined with the signing of a contract. This contract includes a number of requirements: specification of the intended use of the data; instruction not to release the **microdata** file to another recipient; prior review and approval by the releasing agency for all user outputs to be published or disseminated; terms and location of access and enforceable penalties.

Local recoding: A disclosure control technique for microdata where two (or more) different versions of a variable are used dependent on some other variable. The different versions will have different levels of coding. This will depend on the distribution of the first variable conditional on the second. A typical example occurs where the

distribution of a variable is heavily skewed in some geographical areas. In the areas where the distribution is skewed minor categories may be combined to produce a courser variable.

Local suppression: Protection technique that diminishes the risk of recognition of information about individuals or enterprises by suppressing individual scores on **identifying variables**.

Lower bound: The lowest possible value of a cell in a table of frequency counts where the cell value has been perturbed or suppressed.

M

Macrodata: Synonym of **tabular data**.

Microaggregation: Records are grouped based on a proximity measure of variables of interest, and the same small groups of records are used in calculating aggregates for those variables. The aggregates are released instead of the individual record values.

Microdata: A microdata set consists of a set of records containing information on individual respondents or on economic entities.

Minimal unique: A combination of variable values that are unique in the **microdata** set at hand and contain no proper subset with this property (so it is a minimal set with the uniqueness property).

N

NSI(s): Abbreviation for National Statistical Institute(s).

(n,k) rule: A cell is regarded as confidential, if the n largest units contribute more than k % to the cell total, e.g. n=2 and k=85 means that a cell is defined as risky if the two largest units contribute more than 85 % to the cell total. The n and k are given by the statistical authority. In some **NSIs** the values of n and k are confidential.

O

On-site facility: A facility that has been established on the premises of several NSIs. It is a place where external researchers can be permitted access to potentially disclosive data under contractual agreements which cover the maintenance of confidentiality, and which place strict controls on the uses to which the data can be put. The on-site facility can be seen as a 'safe setting' in which confidential data can be analysed. The on-site facility itself would consist of a secure hermetic working and data storage environment in which the confidentiality of the data for research can be ensured. Both the physical and the IT aspects of security would be considered here. The on-site facility also includes administrative and support facilities to external users, and ensures that the agreed conditions for access to the data were complied with.

Ordinary rounding: Synonym of **conventional rounding**.

Oversuppression: A situation that may occur during the application of the technique of cell suppression. This denotes the fact that more information has been suppressed than strictly necessary to maintain confidentiality.

P

Partial disclosure: Synonym of **approximate disclosure**.

Passive confidentiality: For foreign trade statistics, EU countries generally apply the principle of “passive confidentiality”, that is they take appropriate measures only at the request of importers or exporters who feel that their interests would be harmed by the dissemination of data.

Personal data: Any information relating to an identified or identifiable natural person (’data subject’). An identifiable person is one who can be identified, directly or indirectly. Where an individual is not identifiable, data are said to be anonymous.

Perturbation based disclosure control methods: Techniques for the release of data that change the data before the dissemination in such a way that the disclosure risk for the confidential data is decreased but the information content is retained as far as possible. Perturbation based methods falsify the data before publication by introducing an element of error purposely for confidentiality reasons. For example, an error can be inserted in the cell values after a table is created, which means that the error is introduced to the output of the data and will therefore be referred to as output perturbation. The error can also be inserted in the original data on the **microdata** level, which is the input of the tables one wants to create; the method will then be referred to as data perturbation - input perturbation being the better but uncommonly used expression. Possible perturbation methods are:

- rounding;
- perturbation, for example, by the addition of random noise or by the **Post Randomisation Method**;
- disclosure control methods for microdata applied to tabular data.

Population unique: A record within a dataset which is unique within the population on a given **key**.

P-percent rule: A **(p,q) rule** where q is 100 %, meaning that from general knowledge any respondent can estimate the contribution of another respondent to within 100 % (i.e., knows the value to be nonnegative and less than a certain value which can be up to twice the actual value).

(p,q) rule: It is assumed that out of publicly available information the contribution of one individual to the cell total can be estimated to within q per cent (q=error before publication); after the publication of the statistic the value can be estimated to within p percent (p=error after publication). In the (p,q) rule the ratio p/q represents the information gain through publication. If the information gain is unacceptable the cell is declared as confidential. The parameter values p and q are determined by the statistical authority and thus define the acceptable level of information gain. In some **NSIs** the values of p and q are confidential.

Post Randomisation Method (PRAM): Protection method for microdata in which the scores of a categorial variable are changed with certain probabilities into other scores. It is thus intentional misclassification with known misclassification probabilities.

Primary confidentiality: It concerns tabular cell data, whose dissemination would permit attribute disclosure. The two main reasons for declaring data to be primary confidential are:
- too few units in a cell;
- dominance of one or two units in a cell.

The limits of what constitutes "too few" or "dominance" vary between statistical domains.

Primary protection: Protection using disclosure control methods for all cells containing small counts or cases of dominance.

Primary suppression: This technique can be characterized as withholding all disclosive cells from publication, which means that their value is not shown in the table, but replaced by a symbol such as '×' to indicate the suppression. According to the definition of disclosive cells, in frequency count tables all cells containing small counts and in tables of magnitudes all cells containing small counts or representing cases of dominance have to be primary suppressed.

Prior-posterior rule: Synonym of the **(p,q) rule**.

Privacy: Privacy is a concept that applies to data subjects while confidentiality applies to data. The concept is defined as follows: "It is the status accorded to data which has been agreed upon between the person or organisation furnishing the data and the organisation receiving it and which describes the degree of protection which will be provided." There is a definite relationship between confidentiality and privacy. Breach of confidentiality can result in disclosure of data which harms the individual. This is an attack on privacy because it is an intrusion into a person's self-determination on the way his or her personal data are used. Informational privacy encompasses an individual's freedom from excessive intrusion in the quest for information and an individual's ability to choose the extent and circumstances under which his or her beliefs, behaviours, opinions and attitudes will be shared with or withheld from others.

Probability based disclosures (approximate or exact): Sometimes although a fact is not disclosed with certainty, the published data can be used to make a statement that has a high probability of being correct.

Q

Quasi-identifier: Variable values or combinations of variable values within a dataset that are not structural uniques but might be empirically unique and therefore in principle uniquely identify a population unit.

R

Randomized response: Randomized response is a technique used to collect sensitive information from individuals in such a way that survey interviewers and those who process the data do not know which of two alternative questions the respondent has answered.

Random perturbation: This is a disclosure control method according to which a noise, in the form of a random value is added to the true value or, in the case of categorical variables, where another value is randomly substituted for the true value.

Random rounding: In order to reduce the amount of data loss that occurs with suppression, alternative methods have been investigated to protect sensitive cells in tables of frequencies. Perturbation methods such as random rounding and controlled rounding are examples of such alternatives. In random rounding cell values are rounded, but instead of using standard rounding conventions a random decision is made as to whether they will be rounded up or down. The

rounding mechanism can be set up to produce unbiased rounded results.

Rank swapping: Rank swapping provides a way of using continuous variables to define pairs of records for swapping. Instead of insisting that variables match (agree exactly), they are defined to be close based on their proximity to each other on a list sorted on the continuous variable. Records which are close in rank on the sorted variable are designated as pairs for swapping. Frequently in rank swapping the variable used in the sort is the one that will be swapped.

Record linkage process: Process attempting to classify pairs of matches in a product space A×B from two files A and B into M, the set of true links, and U, the set of non-true links.

Record swapping: A special case of **data swapping,** where the geographical codes of records are swapped.

Remote access: On-line access to protected microdata.

Remote data laboratory: A virtual environment providing remote execution facilities.

Remote execution: Submitting scripts on-line for execution on disclosive microdata stored within an institute's protected network. If the results are regarded as **safe data**, they are sent to the submitter of the script. Otherwise, the submitter is informed that the request cannot be acquiesced. Remote execution may either work through submitting scripts for a particular statistical package such as SAS, SPSS or STATA which runs on the remote server or via a tailor made client system which sits on the user's desk top.

Residual disclosure: Disclosure that occurs by combining released information with previously released or publicly available information. For example, tables for nonoverlapping areas can be subtracted from a larger region, leaving confidential residual information for small areas.

Restricted access: Imposing conditions on access to the **microdata**. Users can either have access to the whole range of raw protected data and process individually the information they are interested in - which is the ideal situation for them - or their access to the protected data is restricted and they can only have a certain number of outputs (e.g. tables) or maybe only outputs of a certain structure. Restricted access is sometimes necessary to ensure that linkage between tables cannot happen.

Restricted data: Synonym of **safe data**.

Restriction based disclosure control method: Method for the release of **tabular data**, which consists in reducing access to the data provided to the external user. This method reduces the content of information provided to the user of the **tabular data**. This is implemented by not publishing all the figures derived from the collected data or by not publishing the information in as detailed a form as would be possible.

Risky cells: The cells of a table which are non-publishable due to the risk of statistical disclosure are referred to as risky cells. By definition there are three types of risky cells: small counts, dominance and complementary suppression cells.

Risky data: Data are considered to be disclosive when they allow statistical units to be identified, either directly or indirectly, thereby disclosing individual information. To determine whether a statistical unit is identifiable, account shall be taken of all the means that might reasonably be used by a third party to identify the said statistical unit.

Rounding: Rounding belongs to the group of disclosure control methods based on output-perturbation. It is used to protect small counts in **tabular data** against disclosure. The basic idea behind this disclosure control method is to round each count up or down either deterministically or probabilistically to the nearest integer multiple of a rounding base. The additive nature of the table is generally destroyed by this process. Rounding can also serve as a **recoding** method for microdata.

R-U map: A graphical representation of the trade off between disclosure risk and data utility.

S

Safe data: Microdata or **macrodata** that have been protected by suitable **Statistical Disclosure Control** methods.

Safe setting: An environment such as a **microdata** lab whereby access to a disclosive dataset can be controlled.

Safety interval: The minimal **calculated interval** that is required for the value of a cell that does not satisfy the primary suppression rule.

Sample unique: A record within a dataset which is unique within that dataset on a given **key**.

Sampling: In the context of disclosure control, this refers to releasing only a proportion of the original data records on a **microdata** file.

Sampling fraction: The proportion of the population contained within a data release. With simple random sampling, the sample fraction represents the proportion of population units that are selected in the sample. With more complex sampling methods, this is usually the ratio of the number of units in the sample to the number of units in the population from which the sample is selected.

Scenario analysis: A set of pseudo-criminological methods for analysing and classifying the plausible risk channels for a data intrusion. The methods are based around first delineating the means, motives and opportunity that an intruder may have for conducting the attack. The output of such an analysis is a specification of a set of **keys** likely to be held by **data intruders**.

Secondary data intrusion: After an attempt to match between identification and target datasets an intruder may discriminate between non-unique matches by further direct investigations using additional variables.

Secondary disclosure risk: It concerns data which is not primary disclosive, but whose dissemination, when combined with other data permits the identification of a microdata unit or the disclosure of a unit's attribute.

Secondary suppression: To reach the desired protection for risky cells, it is necessary to suppress additional non-risky cells, which is called secondary suppression or complementary suppression. The pattern of complementary suppressed cells has to be carefully chosen to provide the desired level of ambiguity for the disclosive cells at the highest level of information contained in the released statistics.

Security: An efficient disclosure control method provides protection against exact disclosure or unwanted narrow estimation of the attributes of an individual entity, in other words, a useful technique prevents exact or partial disclosure. The security level is accordingly high. In the case of disclosure control methods for the release of **microdata** this protection is ensured if the identification of a respondent is not possible, because the identification is the prerequisite for disclosure.

Sensitive cell: Cell for which knowledge of the value would permit an unduly accurate estimate of the contribution of an individual respondent. Sensitive cells are identified by the application of a dominance rule such as the (n,k) rule or the (p,q) rule to their microdata.

Sensitive variables: Variables contained in a data record apart from the key variables, that belong to the private domain of respondents who would not like them to be disclosed. There is no exact definition given for what a 'sensitive variable' is and therefore, the division into key and sensitive variables is somehow arbitrary. Some data are clearly sensitive such as the possession of a criminal record, one's medical condition or credit record, but there are other cases where the distinction depends on the circumstances, e.g. the income of a person might be regarded as a sensitive variable in some countries and as quasi-identifier in others, or in some societies the religion of an individual might count as a key and a sensitive variable at the same time. All variables that contain one or more sensitive categories, i.e. categories that contain sensitive information about an individual or enterprise, are called sensitive variables.

Shuttle algorithm: A method for finding lower and upper cell bounds by iterating through dependencies between cell counts. There exist many dependencies between individual counts and aggregations of counts in contingency tables. Where not all individual counts are known, but some aggregated counts are known, the dependencies can be used to make inferences about the missing counts. The Shuttle algorithm constructs a specific subset of the many possible dependencies and recursively iterates through them in order to find bounds on missing counts. As many dependencies will involve unknown counts, the dependencies need to be expressed in terms of inequalities involving lower and upper bounds, rather than simple equalities. The algorithm ends when a complete iteration fails to tighten the bounds on any cell counts.

Special uniques analysis: A method of analysing the per-record risk of **microdata**.

Statistical confidentiality: The protection of data that relate to single statistical units and are obtained directly for statistical purposes or indirectly from administrative or other sources against any breach of the right to confidentiality. It implies the prevention of unlawful disclosure.

Statistical Data Protection (**SDP):** Statistical Data Protection is a more general concept which takes into account all steps of production. SDP is multidisciplinary and draws on computer science (data security), statistics and operations research.

Statistical disclosure: Statistical disclosure is said to take place if the dissemination of a statistic enables the external user of the data to obtain a

better estimate for a confidential piece of information than would be possible without it.

Statistical Disclosure Control (SDC): Statistical Disclosure Control techniques can be defined as the set of methods to reduce the risk of disclosing information on individuals, businesses or other organisations. Such methods are only related to the dissemination step and are usually based on restricting the amount of or modifying the data released.

Statistical Disclosure Limitation (**SDL):** Synonym of **Statistical Disclosure Control**.

Subadditivity: One of the properties of the (n,k) rule or (p,q) rule that assists in the search for complementary cells. The property means that the sensitivity of a union of disjoint cells cannot be greater than the sum of the cells' individual sensitivities (triangle inequality). Subadditivity is an important property because it means that aggregates of cells that are not sensitive are not sensitive either and do not need to be tested.

Subtraction: The principle whereby an intruder may attack a table of population counts by removing known individuals from the table. If this leads to the presence of certain zeroes in the table then that table is vulnerable to **attribute disclosure**.

Suppression: One of the most commonly used ways of protecting sensitive cells in a table is via suppression. It is obvious that in a row or column with a suppressed sensitive cell, at least one additional cell must be suppressed, or the value in the sensitive cell could be calculated exactly by **subtraction** from the marginal total. For this reason, certain other cells must also be suppressed. These are referred to as **secondary suppressions**. While it is possible to select cells for secondary suppression manually, it is difficult to guarantee that the result provides adequate protection.

SUDA: A software system for conducting analyses on population uniques and special sample uniques. The **special uniques analysis** method implemented in SUDA for measuring and assessing disclosure risk is based on resampling methods and used by the ONS.

Swapping (or switching): Swapping (or switching) involves selecting a sample of the records, finding a match in the data base on a set of predetermined variables and swapping all or some of the other variables between the matched records. Swapping (or switching) was illustrated as part of the confidentiality edit for tables of frequency data.

Synthetic data: An approach to confidentiality where instead of disseminating real data, synthetic data that have been generated from one or more population models are released.

Synthetic substitution: See **Controlled Tabular Adjustment**.

T

Table server: A form of **remote data laboratory** designed to release safe tables.

Tables of frequency (count) data: These tables present the number of units of analysis in a cell. When data are from a sample, the cells may contain weighted counts, where weights are used to bring sample results to the population levels. Frequencies may also be represented as percentages.

Tables of magnitude data: Tables of magnitude data present the aggregate of a "quantity of interest" over all units of analysis in the cell. When data are from a sample, the cells may contain weighted aggregates, where quantities are multiplied by units' weights to bring sample results up to population levels. The data may be presented as averages by dividing the aggregates by the number of units in their cells.

Tabular data: Aggregate information on entities presented in tables.

Target dataset: An anonymised dataset in which an intruder attempts to identify particular population units.

Threshold rule: Usually, with the threshold rule, a cell in a table of frequencies is defined to be sensitive if the number of respondents is less than some specified number. Some agencies require at least five respondents in a cell, others require three. When thresholds are not respected, an agency may restructure tables and combine categories or use cell suppression, rounding or the confidentiality edit, or provide other additional protection in order to satisfy the rule.

Top and bottom coding: It consists in setting top-codes or bottom-codes on quantitative variables. A top-code for a variable is an upper limit on all published values of that variable. Any value greater than this upper limit is replaced by the upper limit or is not published on the **microdata** file at all. Similarly, a bottom-code is a lower limit on all published values for a variable. Different

limits may be used for different quantitative variables, or for different subpopulations.

U

Union unique A sample unique that is also population unique. The proportion of sample uniques that are union uniques is one measure of file level disclosure risk.

Uniqueness: The term is used to characterise the situation where an individual can be distinguished from all other members in a population or sample in terms of information available on **microdata** records (or within a given **key**). The existence of uniqueness is determined by the size of the population or sample and the degree to which it is segmented by geographic information and the number and detail of characteristics provided for each unit in the dataset (or within the key).

Upper bound: The highest possible value of a cell in a table of frequency counts where the cell value has been perturbed or suppressed.

V

Virtual safe setting: Synonym of **remote data laboratory**.

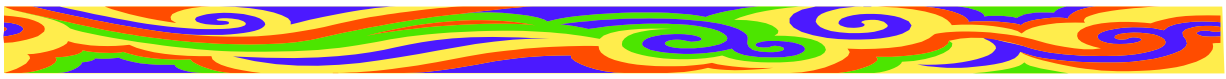
W

Waiver approach: Instead of suppressing tabular data, some agencies ask respondents for permission to publish cells even though doing so may cause these respondents' sensitive information to be estimated accurately. This is referred to as the waiver approach. Waivers are signed records of the respondents' granting permission to publish such cells. This method is most useful with small surveys or sets of tables involving only a few cases of dominance, where only a few waivers are needed. Of course, respondents must believe that their data are not particularly sensitive before they will sign waivers.

X

Y

Z



Visit the new and improved IASS web site and read *The Survey Statistician* on line!

http://www.cbs.nl/isi/iass/SurvStat.htm

Important Notices

- A PDF file of the newsletter is available on the IASS web site. Currently, a few members prefer to be notified only when a new issue is posted, instead of receiving the hard copy. At this point we do not have a process in place to support this option. A process will be developed when an adequate number of members choose the above. Until that time, all members will continue to receive hard copies of the newsletter. Please send an e-mail to sheering@isr.umich.edu if you would like to take advantage of this option.

- Members are encouraged to view the IASS website (www.isi-iass.org) and provide comments or suggestions to Eric Rancourt: eric.rancourt@statcan.ca.

-

In Other Journals

Journal of Official Statistics
An International Review Published by Statistics Sweden

JOS is a scholarly quarterly that specializes in statistical methodology and applications. Survey methodology and other issues pertinent to the production of statistics at national offices and other statistical organizations are emphasized. All manuscripts are rigorously reviewed by independent referees and members of the Editorial Board.

Contents of Issue: September, Volume 21, Number 3, 2005

Journal of Official Statistics

Contents of Issue December, Volume 21, Number 4, 2005

IASS Officers and Council Members

President (2005–2007):	Gordon Brackstone, gordon.brackstone@sympatico.ca
President-elect:	Pedro Luis do Nascimento Silva (Brazil), pedrosilva@ibge.gov.br)
Vice-Presidents :	John G. Kovar (Canada), Kovar@statcan.ca
	Jelke Bethlehem (Netherlands), jbtm@cbs.nl
Scientific Secretary:	Lilli Japec (Sweden), lilli.japec@scb.se
Council Members (2003–2007):	Pascal Ardilly (France), pascal.ardilly@insee.fr
	J-J. Droesbeke (Belgium), jjdroesb@ulb.ac.be
	Guido Ferrari (Italy), ferrari@ds.unifi.it
	Linda Hewitt (Trinidad and Tobago), linhew@trinidad.net
	Naman Keita (Mali), naman.keita@fao.org
	Hiek Som (Cambodia), hiek.som@fao.org
Council Members: (2005-2009)	Geoffrey Lee (Australia), geoff.lee@abs.gov.au
	Don Royce (Canada), Don.Royce@statcan.ca
	Giuliana Coccia (Italy), coccia@istat.it
	Leyla Mohadjer (United States), LeylaMohadjer@Westat.com
	Sarah M. Nusser (United States), nusser@iastate.edu
	Wei Yuan (China), wyuan@ruc.edu.cn
Committee Chairs:	
2005 Election Nominations Committee:	Claudio Quintano (Italy), claudio.quintano@uniparthenope.it
2009 Durban Programme Committee	Leyla Mohadjer (USA), LeylaMohadjer@Westat.com
The Secretariat:	
Executive Director:	Michel Péronnet (France), michel.peronnet@insee.fr
Treasurer:	Michel Kagan (France), michel.kagan@insee.fr
Executive Secretary:	Anna Maria Vespa-Leyder (France), vespa@cepii.fr
Secretary:	Claude Olivier (France), claude.olivier@insee.fr

INTERNATIONAL ASSOCIATION OF SURVEY STATISTICIANS

CHANGE OF ADDRESS FORM

If your home or business address has changed, please copy, complete, and mail this form to:

IASS Secretariat
c/o INSEE-CEFIL
Att. Ms. Claude Olivier
3, rue de la Cité
33500 Libourne – France

Name: Mr./Mrs./Miss/Ms. _____ **First name:** _____

E-mail address *(please just indicate one)*: _____

May we list your e-mail address on the IASS web site?

Yes ☐ **No** ☐

Home address

Street: _____
City: _____
State/Province: _____ **Zip/Postal code:** _____
Country: _____
Telephone number: _____
Fax number: _____

Business address

Company: _____
Street: _____
City: _____
State/Province: _____ **Zip/Postal code:** _____
Country: _____
Telephone number and extension: _____
Fax number: _____

Please specify address to which your IASS correspondence should be sent:

Home ☐ **Business** ☐

Institutional Members

4 International Organizations

AFRISTAT
CICRED
EUROSTAT
UNITED NATIONS STATISTICAL DIVISION

31 Bureaus of Statistics

ARGENTINA - INSTITUTO NACIONAL DE ESTADISTICA Y CENSOS - INDEC
AUSTRALIA - AUSTRALIAN BUREAU OF STATISTICS
BELGIUM - INSTITUT NATIONAL DE STATISTIQUE
BRAZIL - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA - IBGE
CANADA - STATISTICS CANADA
CHINA - GOVERNO DE MACAU
COTE D'IVOIRE - INSTITUT NATIONAL DE LA STATISTIQUE
CZECH REPUBLIC - CZECH STATISTICAL OFFICE
DENMARK - DANMARKS STATISTIK
FINLAND - STATISTICS FINLAND
FRANCE - INSTITUT NATIONAL DE STATISTIQUE ET D'ÉTUDES ÉCONOMIQUES - INSEE
GAMBIA - CENTRAL STATISTICS DEPARTMENT
GERMANY - STATISTICHE BUNDESAMT
GREECE - NATIONAL STATISTICAL SERVICE OF GREECE
IRAN - STATISTICAL CENTER OF IRAN
ITALY - INSTITUTO CENTRALE DI STATISTICA - ISTAT
MEXICO - INSTITUTO NACIONAL DE ESTADISTICA, GEOGRAFIA E INFORMATICA - INEGI
NETHERLANDS - CENTRAL BUREAU OF STATISTICS
NEW ZEALAND - STATISTICS NEW ZEALAND
NIGERIA - FEDERAL OFFICE OF STATISTICS
NORWAY - CENTRAL BUREAU OF STATISTICS
PORTUGAL - INSTITUTO NACIONAL DE ESTATISTICA - INE
REPUBLIC OF KOREA - NATIONAL STATISTICAL OFFICE - NSO
SPAIN - INSTITUTO NACIONAL DE ESTADISTICA
SWEDEN - STATISTICS SWEDEN
SWITZERLAND - OFFICE FEDERAL DE STATISTIQUE
TANZANIA - BUREAU OF STATISTICS
UNITED KINGDOM - OFFICE FOR NATIONAL STATISTICS
USA - BUREAU OF THE CENSUS
USA - DEPARTMENT OF EDUCATION
USA - DEPARTMENT OF HEALTH AND HUMAN SERVICES

6 Universities, Research Centers, Private Statistics Firms

ARGENTINA - UNIVERSITAD NACIONAL DE TRES DE FEBRERO
DENMARK - SFI
EUROPE - A.C. NIELSEN MANAGEMENT SERVICE
FRANCE - INSTITUT NATIONAL D'ÉTUDES DEMOGRAPHIQUES - INED
USA - RESEARCH TRIANGLE INSTITUTE
USA - Westat

Read

The Survey Statistician

on line!

http://isi.cbs.nl/iass/index.htm