the Survey Statistician
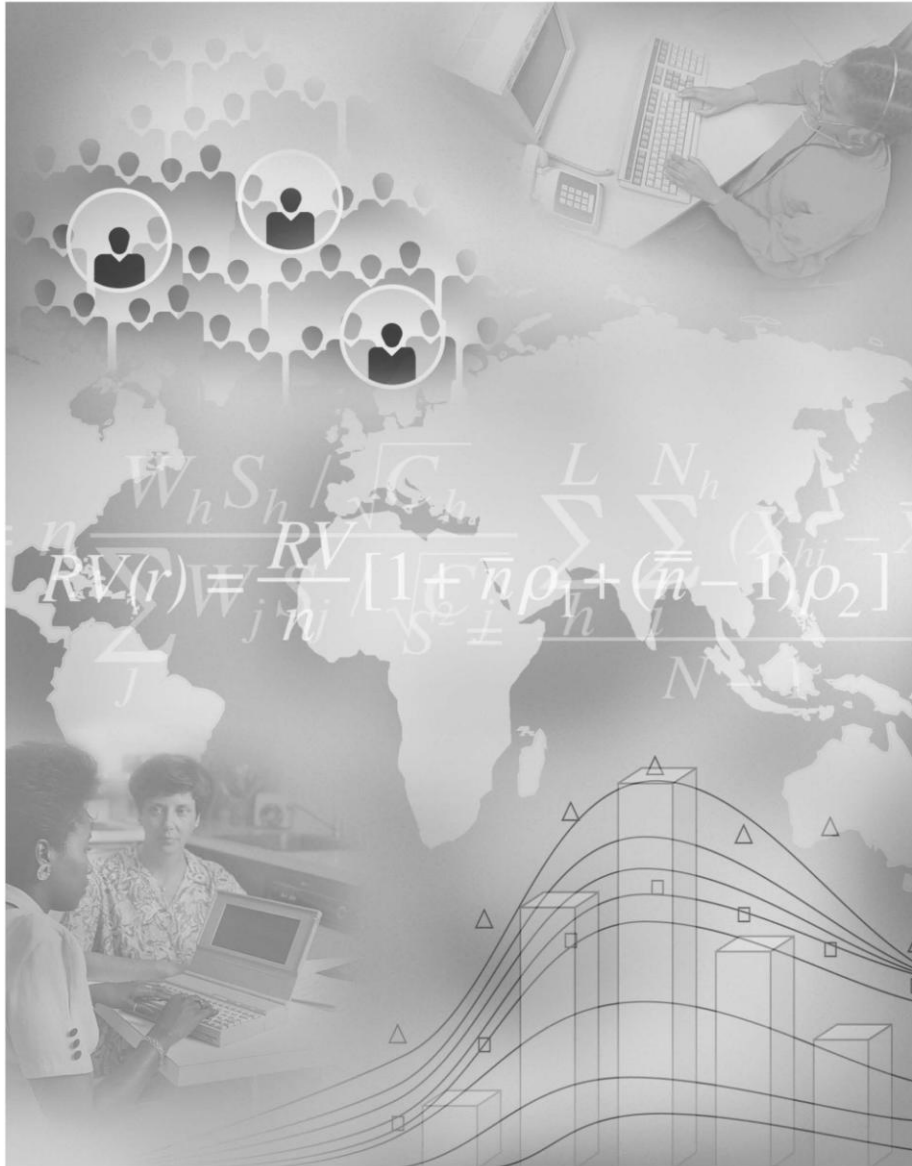
The Newsletter of the International Association of Survey Statisticians

International Statistical Institute

Institut International de Statistique

Editors
  Natalie Shlomo and Frank Yu

Section Editors
  Pierre Lavallée — Country Reports
  Robert Clark   — Ask the Experts
  Ineke Stoop   — New and Emerging Methods

Circulation/Production
  Henry Chiem
  Yovina Joymungul Poorun

In This Issue

Letter from the President

Dear Colleagues

It hardly seems like six months have passed since my letter in the previous issue of The Survey Statistician. But it has, and there are many things to inform you about as far as IASS is concerned.

The move of the IASS office from Libourne to ISI offices in The Hague is proceeding, though there are issues that still need to be dealt with by the Executive Director Catherine Meunier. A major obstacle to a smooth transition is that IASS is registered in France and hence subject to French law. This necessitates that the Association first be dissolved as a French association and then re-established as an association subject to Dutch law. To this end, Vice President Steve Heeringa has volunteered to redraft the statutes of the association to remove/amend the embedded references to the association as French and operating subject to French law. However, before these revised statutes can be accepted by members of the association, it will be necessary to dissolve the 'French version' of the IASS and create a 'Dutch version' to which these new statutes will apply. A suggested strategy and a timeline for carrying this out has now been prepared and is being considered by IASS Council. This strategy is necessary because Article 12 of the current statutes of the IASS is very precise about how the process of such an amendment to the statutes has to be implemented.

The development of the new IASS web site is still "a work in progress". A beta version of the website has been developed by Council member Olivier Dupriez and incoming Webmaster Justin Lokhorst, with the assistance of Mehmood Asghar, and can be found at http://www.ihsn.org/apps/iass/.

Growing the IASS membership base is a key objective, and Council members Eva Elvers and Mick Couper have prepared discussion papers addressing the related issues of increasing both individual and institutional membership of IASS. These are currently being considered by Council members. To a large extent increasing both individual and institutional membership of IASS will require addressing the "What's in it for me/us" issue. For individuals, this may mean going beyond the argument that membership of IASS is about a professional supporting his/her profession. The networking advantages of being a member of IASS, the access to IASS supported conferences and short courses and the methodological advice provided by articles in the *The Survey Statistician* all constitute direct benefits. More altruistic motives include IASS funding for survey statisticians from developing countries. For institutions, we have observed an increase in requests for reasons to continue being an IASS member given that the institution is already a member of ISI. Incoming President Danny Pfeffermann has suggested that IASS could set up a programme of short courses for member institutions, which could be presented by volunteer IASS members with only travel and accommodation costs charged. This clearly has considerable economic benefits for institutions, but depends on the goodwill and availability of IASS members who would have to present such courses.

From its inception, IASS has operated as a bilingual association, with both Anglophone and Francophone members, and with English and French versions of *The Survey Statistician* and its present web site. To a large extent, this has been possible because of bilingual administrative services provided by INSEE and French translation and proof-reading services provided by Statistics Canada. As you are aware, financial pressures meant that INSEE had to withdraw its administrative support last year. Unfortunately, similar financial pressures have

now meant that Statistics Canada will no longer be able to provide French translation services for *The Survey Statistician*, and this issue will be the last one that Statistics Canada will translate into French for distribution to the francophone members of the IASS. I have already expressed IASS's gratitude to Statistics Canada for its huge support in ensuring that a French language version of *The Survey Statistician* has been able to be published for so long. However, we now need to face the issue of how we can continue to provide our current level of French language resources for our francophone members. All suggestions are welcome!

With the closure of the IASS office in Libourne, the issue of what to do about the archived IASS material held there has arisen. This includes correspondence from a number of eminent survey statisticians (including Leslie Kish) at the time IASS was created, as well as spare copies of IASS publications. A start has been made on rationalising the materiel held in the archives, and this is now about 25% of what was previously held there. I am hopeful that a new home will be found for what is left.

Preparations for the ISI meeting in Hong Kong next year are proceeding quickly. I am very happy to inform you that Eric Rancourt, who chairs the IASS Program Committee for Hong Kong, has reported that of the 13 proposals for Invited Paper sessions made by IASS, 12 have been accepted. In addition, there is a proposal that an Invited Paper session of the IASC on new developments in survey computing will be co-sponsored by IASS. Eric is now working with the organisers for these sessions to ensure that everything goes smoothly next year, and at this stage I anticipate that the Hong Kong ISI meeting will be a very successful one for IASS. I am also happy to inform you that Council member Christine Bycroft has agreed to chair the IASS Program Committee for the Rio de Janeiro ISI meeting in 2015. The Rio meeting will be especially interesting to members of the IASS since past IASS President Pedro Silva has been nominated to take over as ISI President at this meeting.

On April 18, I represented IASS at the 100th Anniversary celebrations of the Polish Statistical Association in Poznan, Poland. The opportunity to this gave me great satisfaction since perhaps uniquely among the various branches of statistics in use today, statistical development originating in Poland lead to the paradigm that underlies much of the inferential ideas and methodology in use in survey statistics today. I refer of course to the use of randomization or design-based theory and to the fact that this theory was born out of the hugely influential work of Professor Jerzy Neyman and in particular his landmark paper "On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection" that was read before the Royal Statistical Society on June 19th, 1934, and subsequently published in the Journal of the Royal Statistical Society. The original ideas (and there were many) in this paper served to define the inferential 'pathway' for virtually all of survey sampling design and inference for the next 50 years, and still serve as the inferential basis for much official statistics methodology. I am sure that all of you join me in wishing the Polish Statistical Association another successful 100 years!

Finally, it is with considerable sadness that I have to inform you that David Binder, IASS member and one of the most eminent and influential survey statisticians of the last twenty five years, passed away in early June. To my mind, the following comment from Bob Rodriguez, President of the American Statistical Association, best expresses the feelings of many statisticians (not just survey statisticians) when they heard of David's death: "This is a sad day, not only for those of us who knew David, but also for our field. He will be remembered in many ways — as a brilliant researcher in survey data analysis, as a director general at Statistics Canada, and as

an international statistical leader. I will especially remember the gentle wisdom and unselfishness with which he helped colleagues and served our profession."

I am sure that you will join with me and the rest of the international statistical community in expressing our sincere condolences to his wife Marilyn and daughter Jill. Vale David Binder.

Ray Chambers,
President, IASS

Letter from the Editors

.
We have an exciting July 2012 issue of the IASS newsletter *The Survey Statistician* for your enjoyment.

In the *New and Emerging Methods* section (edited by Ineke Stoop), James Chipperfield and David Steel of Australia contribute an article on 'Split Questionnaire Designs'. The paper provides an overview of when these types of designs can be efficient. In *the Ask the Experts* section (edited by Robert Clark), Hyunshik Lee of the USA addresses the question of how one can find the contributions to the design effect (variance) made by each of the design components (stratification, clustering, weighting) under a complex sample design. This paper follows up on the recent discussion that was conducted over the SRMSNET mailing list where many persons contributed to the topic. We are also very pleased to have a contributing article for the *Book and Software Review* section by Matthias Templ and Valentin Todorov of Austria titled 'Official Statistics and Survey Methodology Meets R: An Overview of Corresponding Packages'. The list of comprehensive R-codes for the production and analysis of datasets arising from surveys is a welcome contribution given the growing use of the R package in survey methods and official statistics.

On behalf of the IASS community, we wish to thank all of the authors and editors of these sections for their important contributions. Please let Ineke Stoop (i.stoop@scp.nl) know if you would like to contribute to the New and Emerging Methods section in the future. If you have any questions which you would like to be answered by an expert, please send your questions to Robert Clark (rclark@uow.edu.au). If you are interested in writing a book review or software review, please get in touch with the editors: Frank Yu or Natalie Shlomo (frank.yu@abs.gov.au or n.shlomo@soton.ac.uk).

As in the past, this issue of *The Survey Statistician* includes a letter from the current President, Ray Chambers and updates from the Scientific Secretary, Ineke Stoop and the Chair of the Hong Kong 2013 Programme Committee, Eric Rancourt. In addition, we thank Marcel Vieira for putting together the list of upcoming conferences to advertise in the newsletter. Please send Marcel (marcel.vieira@ice.ufjf.br) any conference announcements that you would like advertised in the next newsletter to be issued in January 2013.

The *Country Report* section has always been a central feature of the IASS newsletter *The Survey Statistician* and we thank the editor of the section, Pierre Lavallée (pierre.lavallee@statcan.gc.ca) for his continuing efforts to get timely reports from

different countries. For this issue, we asked country representatives to contribute a report specifically related to the recent round of Censuses. We are pleased that we got many countries responding to this request with interesting reports on related Census topics. We wish to thank Pierre and all of the country representatives who contributed to the current issue of the newsletter. We ask all country representatives to continue submitting articles and sharing information on current activities, research and applications in survey methods in your countries.

The IASS Vice-president, Denise Silva, has taken on the responsibility of updating the list of country representatives. Please get in touch with her at denisebritz@gmail.com if you are currently serving as your country representative or alternatively can suggest a new representative for your country.

This issue of *The Survey Statistician* also includes the *Tables of Contents* from recent journals and we wish to thank the editorial team: Henry Chiem and Yovina Joymungul Poorun for putting this section together. We have also included an announcement for a new journal which will begin publishing in 2013: *The Journal of Survey Statistics and Methodology,* sponsored by AAPOR and the American Statistical Association.

Please send any advertisements, letters to the editors, or anything that you would like published in the IASS newsletter *The Survey Statistician* to the editors: frank.yu@abs.gov.au and n.shlomo@soton.ac.uk.

We would like to add a new section in the newsletter entitled 'Correspondence from Members'. In this section, we will include any correspondence from IASS members that is aimed to generate discussion and debate on methodological issues and matters of interest within the community. Please send your letters to the editor, Frank Yu (frank.yu@abs.gov.au).

We wish to send our warmest congratulations to Prof. J.N.K. Rao on his 75th birthday and a photo and brief description of the symposium that was held in his honour at Carlton University in Ottawa, Canada, May 30-June 1, 2012 is included in this newsletter.

On a different note, the entire IASS community is saddened at the early passing of our friend and colleague David Binder of Canada who passed away on June 3rd. We send our deepest condolences to David's family.

As always, we have many thanks for everyone working hard to put this newsletter together and in particular Henry Chiem and Yovina Joymungul Poorun of the Australian Bureau of Statistics for their invaluable assistance. We wish to thank the section editors, Ray Chambers for his continuing support, the IASS Executive Director and Rolande Charette for the translations into French. Statistics Canada has recently had to withdraw from the translation of the newsletter into French due to tight budgetary constraints. The editors wish to thank Statistics Canada who provided the French translation of the newsletter over the past many years. We hope we can find an alternative solution.

Finally, we would like to ask all IASS members to please take an active role in supporting the newsletter by volunteering to contribute articles, book/software reviews and country reports. We also ask IASS members to send in notifications about conferences and other important news items about their organizations or individual members. By taking an active role in contributing to the IASS newsletter,

we can ensure that it remains relevant to our membership and is interesting and enjoyable to read.

We hope you enjoy this July 2012 issue of *The Survey Statistician* and are happy to receive your feedback and comments on how we can make improvements.

The Survey Statistician is also available for downloading from the IASS website at http://isi.cbs.nl/iass/allUK.htm.

In Memorandum

David A. Binder

The IASS community sends our deepest sympathies to the family, friends and colleagues of David Binder who passed away on Sunday, June 3rd 2012 at the age of 62. The following excerpt is taken from the Statistical Society of Canada webpage at http://www.ssc.ca/en/award-winners/award-winners-2011 when David received his Distinguished Service Award in 2011:

Born in the United Kingdom to war refugees from Austria and Czechoslovakia, David and his family immigrated to Canada and settled in Toronto when he was two years old. David remained in Toronto until he graduated from the University of Toronto with a BSc, specializing in mathematics, statistics and economics. After taking some graduate courses in mathematical statistics as a part-time student at Carleton University, David moved to London, England where he earned a Ph.D. in statistics from Imperial College of Science & Technology in 1977. His Ph.D. research was supervised by David R. Cox and Ann F. S. Mitchell.

David's affiliation with Statistics Canada began in 1971 as a Survey Methodologist. Over the years, David moved up the ranks at Statistics Canada, culminating his career as the Director General of the Methodology Branch. This Branch had responsibility for providing statistical services to programs throughout the bureau, and for conducting research into state-of-the-art survey methods. He interrupted his career for three years (1973-6) to pursue post-graduate studies at the University of London in the United Kingdom, and later for two years (1979-81) to hold a teaching and research position at the University of Ottawa in the Faculty of Administration. During the four months prior to his retirement in 2004, David was a Senior Advisor to an Assistant Chief Statistician at Statistics Canada.

In spite of his demanding responsibilities at Statistics Canada, David found time to contribute generously to professional work, particularly for the Statistical Society of Canada. Over the years, David has served on a host of SSC committees and chaired many of the key ones, including the Elections, Awards, Membership, and Regional and Society Co-operation Committees. He also represented Ontario for four years on the Society's Board of Directors. In 2004, David was elected president of the SSC, serving as President-Elect (2004-05), President (2005-06) and Past-President (2006-

07). He was then appointed as the first Executive Director of the Society for the period 2007-10.

David's cumulative rich experience, his wide network of contacts, his farsightedness and ability to negotiate have all served the Society well over the years. He is a devoted member of the SSC who works hard either at the front lines, or in the background, to achieve the Society's mission and to advance its influence beyond traditional boundaries. For many years, David was the voice of the SSC before the American Statistical Association. As SSC Representative in the ASA Committee on Meetings, David was very effective in speaking for the SSC during the drafting of the new contract among the sponsoring societies of the Joint Statistical Meetings. In the ASA, he also served as Chair of the Survey Research Methods Section, and member of the Joint Statistical Meetings Advisory Committee, Council of Sections, ASA Census Advisory Committee of the US Bureau of the Census, and ASA Expert Panel to Review the Current Employment Statistics Survey of the Bureau of Labor Statistics. He has also served on panels to review statistical frameworks for the European Commission.

Tangible examples of David's deeds in the SSC abound. He led the SSC in challenging discussions on accreditation, on management of the SSC office and initiated discussions on how to engage young statisticians in the functioning of the Society. He was passionate about such engagement and viewed it as critical that the Society held the needs of young statisticians as prime concern. During his presidency, he implemented the AStat program and the accreditation of university courses, after much debate on the Board of Directors. David supervised the creation of the Probability Section and played a major role in the introduction of new awards, most notably the Award for Impact of Applied and Collaborative Work, developing the award's terms of reference and chairing its first selection committee. David worked closely with the Program Secretary and the Program Chair to regularize the SSC protocols for running Annual Meetings. He strongly supported collaborative research initiatives undertaken by NPCDS.

Also during David's presidency, an office transition took place, with the SSC contracting the services of the Canadian Mathematical Society's office. David was key in helping the staff understand the particulars of SSC operations, the philosophy of our Society, the management of our meetings, and the role of committees and the Executive. He was absolutely instrumental in streamlining the Society's day-to-day management and in providing needed stability at a time of major change. It then became apparent that David was going to be the most effective person in the SSC to assume the Executive Director's job. Those who have observed David's work in the SSC know that he is undaunted by difficult or challenging tasks and always rises to the occasion.

David A. Binder is an accomplished researcher who has made significant contributions in survey methodology and applications. He has published over 35 refereed research articles, over 40 publications in conference proceedings and over 15 reports and other publications. His articles have appeared in top journals, including Biometrika, Journal of the American Statistical Association, Survey Methodology and The Canadian Journal of Statistics. He served as Associate Editor for Survey Methodology and The Canadian Journal of Statistics, and has edited several conference proceedings. David served in NSERC's Statistical Sciences Grant Selection Committee. His statistical leadership and research contributions have been recognized with Elected Membership in the International Statistical Institute and Fellowship in the American Statistical Association.

The 2011 SSC Distinguished Service Award reads:

"To David A. Binder, for exceptionally devoted and highly effective service to the SSC, as representative internationally, as President and member of the Executive Committee, and Executive Director of the Society for the period 2004-2010; for career-long leadership in support of research and learning in statistical science in Canada."

Gérard Théodore

Gérard Théodore passed away on June 10th, 2012 in Paris. He was born in 1920, elected ISI member in 1971, one of the "founding father" of the IASS during the 39th ISI session held in Vienna in 1973 and IASS President from 1981 to 1983. Gérard Théodore worked in several services of the French Statistical System for nearly 40 years. He was "Compagnon de la Libération" since he joined the French Free Forces of the Général de Gaulle in London immediately after the Armistice between France and the Nazi Germany in June 1940 and lost his left leg during the battle of Bir-Hakeim in Libya in 1942.

Report from the Scientific Secretary

The IASS in the gap year

The World Statistics Conference in Dublin now seems a long time ago, and the WSC in Hong Kong is still a year away, so one could imagine these are quiet times for the IASS. Behind the screen, however, a great deal of work is being done. Firstly, the IASS office has been moved from the INSEE office at Libourne in France to the ISI office in The Hague. This also means that many issues, such as membership subscriptions and the organisation of short courses, are now integrated in the general ISI operating procedures. The institutional memory is still in place, however: Cathérine Meunier supports the IASS as Executive Director. One issue that remains as far as Libourne is concerned is the IASS archives there. In April, Catherine and the IASS President, Ray Chambers, started work on consolidating these archives, with the aim of eventually finding them a home elsewhere.

Preparation WSC 2013

The preparation of the WSC is in full swing now. The IASS council has proposed 13 Invited Paper Meetings (IPMs) and to our great delight all sessions have been accepted. Two IPMs will be combined, so the preparations will start for 12 IPMs on the following topics:
- Analytic Inference for Data from Complex Surveys
- New developments in multilevel model inference from complex sample survey data
- New developments in small area estimation and applications

- Rankings, use of ranks, and survey sampling
- Administrative and market data as tools for evaluating survey data and for research
- Split-questionnaire sampling and, more generally, matrix sampling are sampling designs in which different sets of data items are collected from different sample units.
- Response to natural disasters
- Administrative censuses: approaches when national population registers are not complete, or not available
- Issues related to major redesigns in National Statistical Offices
- Model-assisted approaches to combining information from different surveys (survey data integration)
- Non-random nonresponse and new developments in imputation
- The challenge from web panel surveys

The council is also working on a programme of short courses. We intend to keep a number of the successful courses from the past and introduce a number of new courses.

Sponsoring of courses

The objective of the IASS is to promote the study and development of the theory and practice of statistical censuses and surveys and associated subject throughout the world. To pursue this objective, the IASS is able to provide limited financial support to relevant conferences and workshops. In addition the IASS is keen to promote such activities and related outputs. The IASS has sponsored a number of courses in 2012 and has also received requests for sponsoring several courses and workshops in 2013. An overview will be given in December 2012.

Website and institutional membership

A number of practical issues have also been taken up by the council. There is work in progress on the modernisation of the IASS website, and the council is discussing how to increase institutional membership. One of the ideas to do this is to offer courses taught by volunteer IASS members at the location of the institutional members for free (which means no teaching fees, just travel and accommodation costs). More information on this in following editions of *The Survey Statistician*.

Contact Information for
IASS Country Representatives

Name	Country
Mr. Nacer-eddine Hammouda	Algeria
Ms. Alicia Masautis	Argentina
Mr. Paul Sutcliffe	Australia
Dr. Camille Vanderhoeft	Belgium
Mr. Walter Castillo Guerra	Bolivia
Mr. Moffat Malepa	Botswana
Mr. Cristiano Ferraz	Brazil
Mr. François Ilboudo	Burkina Faso
Mr. John Kovar	Canada
Ms. Maria de Lurdes Lopes	Cape Verde
Mr. Juan Eduardo Munoz	Chile
Mr. Huang Langhui	China
Mr. Mhadji Nailane	Comoros
Mr. Cakpo Benjamin Zanou	Côte d'Ivoire
Mr. Luis Carlos Silva	Cuba
Mr. Vaclav Cermak	Czech Republic
Mr. Peter Linde	Denmark
Dr. Imbi Traat	Estonia
Dr. Gayatri Vishwakarma	Ethiopia
Mr. Jean-Marc Museux	Eurostat (Europe)
Mr. Paavo Väisänen	Finland
Mr. Benoît Riandey	France
Mr. Jean-Pierre Zima Mefe	Gabon
Dr. Ralf Münnich	Germany
Mr. Anastassios Iliakopoulos	Greece
Mr. David Fitch	Guatemala
Ms. Fatoumata Danfaca	Guinea
Mr. Emmanuel Charles	Haiti
Mr. Zoltán Csereháti	Hungary
Dr. Gayatri Vishwakarma	India
Ms. Luisa Kadun Burck	Israel
Prof. Claudio Quintano	Italy
Dr. Ryozo Yoshino	Japan
Mr. Martins Liberts	Latvia
Prof. Bechara Hanna	Lebanon
Dr. Danutė Krapavickaitė	Lithuania
Mr. Antonio Baigorri Matamala	Luxembourg
Ms. Julia Rachel Ravelosoa	Madagascar
Dr. Suresh Chandra Babu	Malawi
Mr. Aziz Mohammad	Malaysia
Mr. Lamine Diop	Mali
Mr. Sidna Ould N'dah	Mauritania
Mr. Miguel Cervera	Mexico
Dr. Rudra Suwal	Nepal
Mr. Shyam Upadhyaya	Nepal
Mr. Jos de Ree	Netherlands
Mr. John Lopdell	New Zealand
Mr. Adetoun Aribiki Imolehin	Nigeria

Dr. Mark Griffin	Pacific Islands
Mr. Abdulhakeem A.H. Eideh	Palestine
Ms. Leonara Laguna	Peru
Mr. Gervacio G. Selda, Jr.	Philippines
Dr. Tomasz Zadlo	Poland
Mr. Paulo Jorge Gomes	Portugal
Mr. Alexis Lukaku Nzinga	D. R. Congo
Dr. Inho Park	Republic of Korea
Innocent Ngombe Bibemo	Central African Republic
Mr. Matar Gueye	Senegal
Ms. Dolores Lorca Lopez	Spain
Mrs. Anjeles Iztueta Azkue	Spain-Basque Region
Mr. Christina Prado Valle	Spain-Basque Region
Mr. Philippe Eichenberger	Switzerland
Mr. Peter Lundquist	Sweden
Dr. Ibrahim Ali	Syria
Mr. Vitalis Eustach Muba	Tanzania
M. Ouagadjio Bandoumal	Chad
Prof. Oztas Ayhan	Turkey
Dr. Olga Vasylyk	Ukraine
Prof. Peter Lynn	United Kingdom
Mr. Howard Hogan	United States
Mr. Federico Segui	Uruguay
Dr. Nguyen Quoc Anh	Vietnam

If you are interested in becoming your country representative or have any updates, amendments or changes to this list, please contact Denise Silva at denise.silva@ibge.gov.br.

News and Announcements

The Journal of Survey Statistics and Methodology (JSSM)

The JOURNAL OF SURVEY STATISTICS AND METHODOLOGY, sponsored by AAPOR and the American Statistical Association, will begin publishing in 2013. Its objective is to publish cutting edge scholarly articles on statistical and methodological issues for sample surveys, censuses, administrative record systems, and other related data. It aims to be the flagship journal for research on survey statistics and methodology.

Topics of interest include survey sample design, statistical inference, nonresponse, measurement error, the effects of modes of data collection, paradata and responsive survey design, combining data from multiple sources, record linkage, disclosure limitation, and other issues in survey statistics and methodology. The journal will publish both theoretical and applied papers, provided the theory is motivated by an important applied problem and the applied papers report on research that contributes generalizable knowledge to the field. Review papers are also welcomed. Papers on a broad range of surveys are encouraged, including (but not limited to) surveys concerning business, economics, marketing research, social science, environment, epidemiology, biostatistics and official statistics.

The journal will have three sections. The *Survey Statistics* section will present papers on innovative sampling procedures, imputation, weighting, measures of uncertainty, small area inference, new methods of analysis, and other statistical issues related to surveys. The *Survey Methodology* section will present papers that focus on methodological research, including methodological experiments, methods of data collection and use of paradata. The *Applications* section will contain papers involving innovative applications of methods and providing practical contributions and guidance, and/or significant new findings.

Joe Sedransk and Roger Tourangeau will be the editors of the new journal. Sedransk will be responsible for the Survey Statistics section and Tourangeau for the Methodology section; the two will jointly oversee the Applications papers.

Submitted papers will begin to be refereed in 2012, as soon as the editorial office is set up, the manuscript processing system is in place and editorial policies are finalized.

Look for the JSSM website.

75th Birthday of Professor J.N.K. Rao

Professor J.N.K. Rao celebrated his 75th birthday with over 40 friends, colleagues, former and current students at the Symposium on the Analysis of Survey Data and Small Area Estimation that was held between May 30-June 1, 2012 at the Field's Institute, Carleton University, Ottawa, Canada. Professor Rao is a distinguished Emeritus Research Professor at Carleton University, a long term Consultant to Statistics Canada and a Member of Statistics Canada's Advisory Committee on Methodology.

Professor Rao has been at the forefront of research in sampling theory and methods since 1960. He has made fundamental contributions to the design-based classical theory of sampling, to the foundations of sampling during the debates of the 1960s and 70s, to a variety of aspects of variance estimation, to the analysis of complex survey data and to small area estimation. Professor Rao is the author of the 2003 Wiley book on Small Area Estimation.

Professor Rao received the 2004 prestigious Waksberg Award for survey methodology and the 1994 Gold Medal of the Statistical Society of Canada. He is a Fellow of the American Statistical Association, Institute of Mathematical Statistics, and Royal Society of Canada and received Honorary Doctor of Mathematics in 2008 from the University of Waterloo.

The keynote address of the Symposium was provided by Professor Wayne Fuller on small area predictions using complex sample survey data. In addition, he presented anecdotes and photos from the early stages of Professor Rao's career. There were an additional 11 sessions at the Symposium, each with 3 invited papers covering a range of topics, particularly on the Analysis of Survey Data, Missing Data, Small Area Estimation, Reducing Non-sampling Errors and Design Based Inference (see: http://www.fields.utoronto.ca/programs/scientific/11-12/surveydata/ for the full programme). A conference banquet was held in honour of Professor Rao with a delicious Indian menu and many delightful tributes from colleagues and friends.

The entire IASS community wishes a very happy birthday to Professor J.N.K. Rao and many more years of happiness and creative activity.

CASS – Courses in Applied Social Surveys

How to Analyse Survey Data: Methods, Software and Applications

Venue: Highfield Campus, University of Southampton, UK
Presenters: Prof Danny Pfeffermann, Prof Patrick Sturgis, Dr Moshe Feder and Dr Dave Holmes
Dates of Course: Wednesday 19th - Friday 21st September 2012

To register for this course online visit the [online registration pages](#).

Summary of course:

Survey data are frequently used for analytic inference on statistical models holding for the population from which the sample is taken. Familiar examples include the analysis of labour market dynamics from labour force surveys, comparisons of pupils' achievements from educational surveys and estimation of causal relationships between risk factors and disease prevalence from health surveys. Survey data differ, however, from other data sets in four major aspects: 1- The samples are selected with known selection probabilities, which allows using the distribution over all possible sample selections as the basis for inference. 2- The selection probabilities may be related to the model outcome variable, in which case the model holding for the sample is different from the target population model. 3- Survey data are almost inevitably subject to nonresponse, which again may distort the population model if the response propensity is associated with the outcome. 4- The sample data are often clustered, implying that observations in the same cluster are correlated.

In this course we shall discuss and illustrate these problems and consider a large number of alternative approaches to address them. Available and new computer programs used to implement these approaches will be discussed, with examples using simulated and real data sets.

This course is based on the outcomes of a two-year research conducted by the four presenters, funded by an ESRC grant No. RES-062-23-2316

More information about the full short course CASS programme for 2012/13 can be found at: http://www.s3ri.soton.ac.uk/cass/programme.php

Ask the Experts

How should one find out the contributions to the design effect (variance) made by each of the design components (stratification, clustering, weighting) of a complex sample design?

Hyunshik Lee, Westat

In sample surveys, a sample is usually selected with a complex sample design other than the simple random sampling (SRS) method for convenience, necessity, enhancement of the sampling efficiency or any combination of these reasons. Then it is often of interest to compare the complex sample design with the SRS with the same sample size to see how much gain or loss is incurred by the design. This idea of comparing a complex design with the SRS was used early on by Cornfield (1951) but was popularized by Kish. He coined the name "design effect" in his book Kish (1965) and published subsequently a number of articles on the subject (Kish, 1965, 1987, and 1992). It has been proved to be useful in a number of ways. For example, to determine the sample size for a new survey or a redesigned survey, the simple SRS variance formula and an estimated or postulated design effect are used to determine the sample size for a given precision. If the design effect is known from the past survey with the same design, it can be used to estimate the variance for the current survey. Another important use is to adjust classical statistics valid under SRS and draw statistical inferences using the adjusted statistics (e.g., Rao and Scott, 1987).

A sample design becomes complex by imposing some structure before sample selection. Most commonly used design structures are stratification and clustering. We first discuss the effect of stratification, starting with the most basic stratified sample design, that is, the stratified SRS with H strata and sample allocation, n_h for stratum h, and the total sample size $n = \sum_{h=1}^{H} n_h$. The population mean $\bar{Y}$ of a y-variable is estimated by the stratified sample mean, namely,

$$\bar{y}_{st} = \sum_{h=1}^{H} \frac{N_h}{N} \bar{y}_h = \sum_{h=1}^{H} A_h \bar{y}_h \tag{1}$$

where N_h is the stratum population size, N is the total population size, $\bar{y}_h$ is the stratum sample mean, and $A_h = N_h/N$, assuming that N_h and N are known. By definition, the design effect for this stratified mean is given by:

$$\text{Deff}_s = \frac{V_{ST}(\bar{y}_{st})}{V_{SRS}(\bar{y})} = \frac{\sum_{h=1}^{H} A_h^2 \sigma_h^2 / n_h}{\sigma^2 / n} \tag{2}$$

where σ_h^2 is the stratum variance for stratum h and σ^2 is the population variance. We ignore the finite population correction and use n and n_h in the variance formulas instead of $(n-1)$ and $(n_h - 1)$ for ease of notation. We are more interested in a big picture idea rather than non-essential details. If stratum variances are the same as the population variance (i.e., $\sigma_h^2 = \sigma^2$ for all h), then the design effect is simply:

$$\text{Deff}_s = \sum_{h=1}^{H} A_h^2/a_h \qquad (3)$$

with $a_h = n_h/n$.

The minimum value of one by (3) is attained only when $A_h = a_h$, which is obtained by proportional allocation (i.e., $n_h = nA_h$). Equal stratum variance occurs when the stratification is done by random grouping. However, if the stratum variances are different, the stratified SRS design can be worse or better than the SRS design depending on the allocation method. The Neyman allocation which allocates the stratum sample size proportional to $N_h\sigma_h$ provides the smallest design effect.

Now let's turn to the clustering effect without stratification. We first consider an equal probability sample for a single-stage cluster design such as a self-weighting design, where the clusters are selected by probability proportional to size (PPS) with the measure of size (MOS) defined by the cluster size and a fixed number (b) of elements are selected with an equal probability sampling method within selected clusters. The clustering effect for the sample mean under such a design is measured by

$$\text{Deff}_c = 1 + \rho(b - 1) \qquad (4)$$

where ρ is the intra-cluster correlation and b is the cluster sample size. If the cluster sample size is not equal for all selected clusters, then b can be replaced by the simple average $\bar{b}$ or a weighted average in (4). Some authors discuss ways of improving the estimation of b (Lynn and Gabler, 2005). The formula (4) was originally due to Kish (1965, p. 162).

We assumed that the sampling weights are equal in the above discussion. However, even when an equal probability design was used, the final weight can be variable due to various reasons such as inaccurate MOS, nonresponse adjustment, post-stratification, etc. If the weight is not correlated (i.e., "haphazard") with the y-value, variable weights increase the variance for the weighted population mean estimate (the weighted sum of y-values divided by the sum of the weights), which is usually used to estimate the population mean, by the following factor:

$$\text{Deff}_w = 1 + c_w^2 \qquad (5)$$

where c_w^2 is the relative variance of the weights (i.e., $c_w^2 = \sum_{i=1}^{n}(w_i - \bar{w}^2)/(n\bar{w}^2)$). This is also due to Kish (1987). The factors in (4) and (5) are multiplicative, so the design effect for the weighted population mean estimate under a cluster sample design with haphazard weights is given by:

$$\text{Deff} = \text{Deff}_w \times \text{Deff}_c = (1 + c_w^2)\{1 + \rho(b^* - 1)\} \qquad (6)$$

where b^* is the simple average of cluster sample sizes or a more complex weighted average. Gabler, Häder, and Lahiri (1999) provided a model-based justification for the above formula.

Now let's consider a more general sample design we often encounter in reality, a stratified cluster design with unequal weights due to unequal probability sampling

and/or various weighting adjustments. The weighted population mean estimate is given by

$$\widehat{\bar{Y}} = \frac{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{b_{hi}} w_{hij} y_{hij}}{\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{b_{hi}} w_{hij}} \qquad (7)$$

where n_h is the number of clusters selected for stratum h, and b_{hi} is the cluster sample size for cluster hi.

Gabler, Häder, and Lynn (2006) developed the design effect formula for (7), which is expressed as a sum of stratum by stratum design effects under the assumption that stratum variances are the same (i.e., $\sigma_h^2 = \sigma^2$ for all h) but the intra-cluster correlation differs from stratum to stratum. They derived the formula for non-overlapping and exhaustive domains instead of strata but under their set-up, domains can be treated as strata. Their design effect formula is given by

$$\text{Deff} = \frac{\sum_{h=1}^{H} \left(\sum_{i=1}^{n_h} \sum_{j=1}^{b_{hi}} w_{hij} \right)^2}{\left(\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{b_{hi}} w_{hij} \right)^2} \times \frac{m}{m_h} \text{Deff}_h = \sum_{h=1}^{H} \frac{\widehat{M}_h^2}{\widehat{M}^2} \times \frac{m}{m_h} \text{Deff}_h \qquad (8)$$

where $m_h = \sum_{i=1}^{n_h} b_{hi}$ is the total element sample size for stratum h, m is the total element sample size for the entire sample, $\widehat{M}$ and $\widehat{M}_h$ are respectively estimates for the entire and stratum population sizes, and Deff_h is the design effect for stratum h, which is defined by:

$$\text{Deff}_h = \text{Deff}_{wh} \times \text{Deff}_{ch} \qquad (9)$$

The two terms on the right side of (9) are respectively the design effect due to variable weights similar to (5) and the design effect due to clustering similar to (4) for stratum h. The latter is specifically defined as

$$\text{Deff}_{ch} = 1 + \rho_h (b_h^* - 1) \qquad (10)$$

where ρ_h is the intra-cluster correlation for stratum h and

$$b_h^* = \frac{\sum_{i=1}^{n_h} \left(\sum_{j=1}^{b_{hi}} w_{hij} \right)^2}{\sum_{i=1}^{n_h} \sum_{j=1}^{b_{hi}} w_{hij}^2}. \qquad (11)$$

This is a sort of weighted average of the cluster sample size (i.e., b_{hi}) within stratum h. If the weights are benchmarked to the stratum population sizes, then the design effect formula in (8) becomes:

$$\text{Deff} = \sum_{h=1}^{H} \frac{M_h^2}{M^2} \times \frac{m}{m_h} \text{Deff}_h = \sum_{h=1}^{H} \left(A_h^2 / a_h \right) \text{Deff}_h \qquad (12)$$

where $A_h = M_h / M$ and $a_h = m_h / m$. The factor A_h^2 / a_h is the same as the one used in formula (3) for the stratification effect. If A_h is unknown, we can use $\widehat{A}_h = \widehat{M}_h / \widehat{M}$ as its estimate. If we further assume that stratum intra-cluster correlations are the same and the weights have the same relative variance across the strata, the overall design effect for the weighted population mean estimate can be written as:

$$\text{Deff} = \text{Deff}_s \times \text{Deff}_w \times \text{Deff}_c. \tag{13}$$

It is neat but this formula is applicable only when the assumptions behind of the formula are reasonably true. Those assumptions are reiterated below:

A1) Stratification is "haphazard" so that stratum variances are the same.

A2) Weights within stratum are "haphazard," meaning that they are not correlated with the survey variables.

A2*) Besides A2, the relative variances of stratum weights are equal across the strata.

A3) The stratum intra-cluster correlation differs from stratum to stratum.

A3*) The stratum intra-cluster correlation does not differ across the strata.

Under A1, A2, and A3, the design effect formula (8) or (12) is valid, whereas the formula (13) can be used under A1, A2*, and A3*.

These conditions for various design effect formulas can be met reasonably for some social surveys. However, business surveys are radically different from social surveys; large businesses are much more variables than small businesses in survey variables of main interest. To address this issue a business survey is designed to select large businesses with much larger probabilities (oftentimes with certainty) than those used for small businesses, resulting in highly variable weights and high correlations between the selection probabilities and the survey variables. This situation was considered by Spencer (2000) assuming a linear regression model between the y-value and the selection probability to derive the design effect for the weighted population total estimate. Park and Lee (2005) studied the design effect for the total in a social survey context. They highlighted that the design effect for the weighted population total estimate behaves very differently from the design effect for the weighted population mean estimate.

To address the radical difference in developing the design effect for business surveys, Lee et al. (2012) used a ratio model between the survey variable and the selection probability to study the sample allocation issue for a US business survey. The ratio model Lee et al. (2012) employed is given as follows:

$$
\begin{aligned}
& y_{hij} = \beta_h x_{hij} + e_{hij} \\
& V(y_{hij} \mid x_{hij}) = x_{hij}^2 \sigma^2 \\
& C(y_{hij}, y_{hi'j'} \mid x_{hij}, x_{hi'j'}) = \begin{cases} \rho_h x_{hij} x_{hi'j'} \sigma^2 & \text{if } i = i' \text{ and } j \neq j' \\ 0 & \text{if } i \neq i' \end{cases}
\end{aligned} \tag{14}
$$

If the sampling method is PPS with x-variable as the size measure and the ratio estimator is used to estimate the mean, then we can use a new variable $z_{hij} = y_{hij}/x_{hij}$ and the design effect formula (8) or (12) with appropriate modification.

Note that the model Gabler, Häder, and Lahiri (1999) used is the same as (14) with $x_{hij} = 1$. Henry (2012) studied the design effect for the calibration estimator under a more general set-up. Gambino (2009) discussed some pitfalls in applying the design effect. The design effect is a very useful statistic in sample surveys but the user should be careful and make sure that the particular formula being used is appropriate for the particular situation.

References

Cornfield, J. (1951). Modern methods in the sampling of human populations. *American Journal of Public Health*, 41, 654-661.

Gambino, J.G. (2009). Design Effect Caveats. *The American Statistician*, 63, 141-146.

Gbaler, S., Häder, S., and Lahiri, P. (1999). A model based justification of Kish's formula for design effect for weighting and clustering. *Survey Methodology*, 25, 105-106.

Gbaler, S., Häder, S., and Lynn, P. (2006). Design effect for multiple design samples. *Survey Methodology*, 32, 115-120.

Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.

Kish, L. (1987). Weighting in Deft. *The Survey Statistician*, June 1987.

Kish, L. (1992). Weighting for unequal p_i. *Journal of Official Statistics*, 8, 183-200.

Henry, K.A. (2012). Design effect for unequal weights. Presentation at the Washington Statistical Society Seminar, February 7, 2012, Washington, DC.

Lee, H., Li, T., Ferguson, G.R., Ponikowski, C.H., and Rhein, B. (2012). Sample allocation for the redesigned National Compensation Survey. To be presented at the Joint Statistical Meetings, San Diego, CA, August, 2012.

Lynn, P., and Gabler, S. (2005). Approximation to *b** in the prediction of design effects due to clustering. *Survey Methodology*, 25, 101-104.

Park, I., and Lee, H. (2005). Design effects for the weighted mean and total estimator under complex survey sampling. *Survey Methodology*, 30, 183-193.

Rao, J.N.K., and Scott, A.J. (1987). On simple adjustments to chi-square tests with sample survey data. *Annals of Statistics*, 15, 385-397.

Spencer, B.D. (2000). An approximate design effect for unequal weighting when measurements may correlate with selection probabilities. *Survey Methodology*, 26, 137-138.

Ask the Experts - Call for Questions

If you'd like to ask the experts a question, please contact Robert Clark at rclark@uow.edu.au.

New and Emerging Methods

Split Questionnaire Designs

James O. Chipperfield and David G. Steel

Abstract:

Survey designs typically impose the constraint that information on all characteristics (or data items) is collected from all units in the sample. A split questionnaire design (SQD) relaxes this constraint which means that a subset of the data items may be collected from any given unit in the sample. An SQD allows a high degree of flexibility in meeting the reliability requirements for a range of estimates while keeping survey costs and respondent load at reasonable levels. Recent work has considered estimation methods for a SQD and when an SQD is an efficient design choice when compared to the more traditional single phase and multi-phase designs. An SQD is also referred to as a matrix sampling design.

1 A Split Questionnaire Design

Consider a survey which collects information from respondents on K data items, denoted by **y** = $(y_1, y_2,..., y_K)'$. With few exceptions, survey designs collect information on all K data items from all units selected in the sample. Such designs are called single phase designs (SPDs) and lead to simplicity in the survey design and analysis and the requirement that only one questionnaire or collection instrument is developed, pilot tested and, perhaps, printed. An SPD can lead to sub-optimal designs and has little flexibility in managing respondent burden in situations where it may be desirable to collect only a subset of the data items from some or all of the respondents.

We will call a sample design that allows for different patterns, or sets, of information on data items to be collected from different sample units a Split Questionnaire Design (SQD). This can also be referred to as a Matrix Sampling Design (Thomas et al., 2006). In a survey that collects information on K data items, an SQD allows the use of all $J = \sum_{p=1}^{K} {}^{K}C_p = 2^K - 1$ different combinations in which information on the K different data items can be collected. The sample allocation for an SQD is defined by **n** $= (n^{(1)}, n^{(2)}, ..., n^{(j)})'$ where $n^{(j)}$ is the number of randomly selected units from which the j th pattern (or combination) of data items are collected. For example, when K=3 the entries in Table 1 show the 7 different patterns available to an SQD, where j = 1 indicates the pattern where only y_1 is collected from n_1 sample units. Some values of n_j may be zero in an SQD, which can have implications for analyses, as discussed in section 2.2

---
[1] James Chipperfield is Assistant Director, Methodology Divison, Australian Bureau of Statistics and and Senior Research Fellow of the Centre for Statistical and Survey Methodology (CSSM). David Steel is Director, CSSM, University of Wollongong, Australia.

Table 1: SQD Data Patterns for K = 3

Data pattern (j)	y_1	y_2	y_3	Sample size	Cost
1	X			$n^{(1)}$	$c^{(1)}$
2		X		$n^{(2)}$	$c^{(2)}$
3	X	X		$n^{(3)}$	$c^{(3)}$
4			X	$n^{(4)}$	$c^{(4)}$
5		X	X	$n^{(5)}$	$c^{(5)}$
6	X		X	$n^{(6)}$	$c^{(6)}$
7	X	X	X	$n^{(7)}$	$c^{(7)}$

The multi-phase design (MPD) is a special case of an SQD. The patterns available to an MPD are restricted to follow a monotone pattern: when information on y_k is collected, information on $y_{k-1}, y_{k-2}, \dots , y_1$ is always collected (e.g. patterns 1, 3, and 7 in Table 1 together form a monotone pattern). An MPD allows the use of only K of the J patterns available to an SQD.

There are at least four ingredients to identifying the optimal allocation, or value for **n**, for an SQD. The first ingredient is an expression for the variance of the estimates that are of interest, which will be a function of **n**. Chipperfield and Steel (2009) considers this problem when estimating population totals with the Best Linear Unbiased Estimator, which exploits the correlation between the data items. Chipperfield and Steel (2011) considers this problem within a likelihood framework when estimating means, regression coefficients and parameters in the multinomial distribution. Renssen and Nieuwenbroek (1997) and Merkouris (2004) suggested methods of estimation from independent surveys that have some data items in common; these authors noted the application of their methods to SQDs, where the data collected from each independent survey represents a different data item pattern. Wretman (1994), Raghunathan and Grizzle (1995) and Gelman, King, and Liu (1998) consider an SQD to manage respondent burden focussing on estimation issues.

The second ingredient is the values of a number of unknown design parameters, which are need to determine the optimal SQD allocation. For example, the optimal allocation problem for regression coefficients requires assuming that the covariance matrix for **y** is known.

The third ingredient is the cost for given allocation, **n**. Cost can be defined in terms of payments incurred by the statistical organisation. There may be several elements to the cost:

- c_f : the fixed cost for the survey. This cost is independent of the sample size.
- c_0: the fixed cost per unit. This incorporates the per unit costs that are incurred before any information is collected from the sample unit (e.g. time taken to make contact with a person in a dwelling).
- $c^{(j)}$ the marginal cost of collecting the data items from pattern j.

The total cost, *C*, for an SQD could be expressed as

$$C = c_f + c_0 n + \sum_j c^{(j)} n^{(j)} \tag{1}$$

The final ingredients are the practical constraints on the allocation, such as those imposed by the need to manage respondent burden.

The optimal allocation involves a trade-off between the accuracy of the estimates that are of interest to the design and the cost, subject to practical constraints that may be imposed.

2 Advantages and Implications of an SQD

2.1 Advantages

SQDs have three efficiency-based advantages over an SPD. Firstly, they allow information on data items with relatively high enumeration cost (lower variance requirement) to be collected from fewer units than data items with relatively low cost (or higher variance requirement). Secondly, the correlation between data items can be exploited to minimise the information loss due to not collecting all data items from all units in the sample. Thirdly, allowing some data items, or sets of data items, to be collected from more units than other data items allows maximum flexibility to meet the accuracy requirements on estimates that are important to the design. MPDs also have these three advantages but to a lesser extent, due to the restriction that the pattern of missing data must be monotone. Chipperfield and Steel (2009) and Chipperfield and Steel (2011) show that an SQD can be significantly more efficient when designing for means and regression coefficients than an SPD and MPD.

Another benefit of an SQD is its potential to increase the number of data items collected. Consider the case when the design objective is to estimate the means for *K* data items but, because of response burden constraints, we can only collect information on a maximum of *T* data items from any sample unit, where *T < K*. This situation can arise when a limit is placed on the total interview time. The Australian Bureau of Statistics recently considered an SQD for its Census of Population of Housing, whereby every household would receive two of three possible modules (see Bell & Whiting, 2007). An SQD can accommodate such constraints, unlike an SPD and an MPD. Shoemaker (1973) and Munger and Lloyd (1988) consider an SQD when estimating differences between groups in situations where an SPD is impractical or would result in concerns about the quality of responses, say due to respondent fatigue.

2.2 Implications of an SQD

An SQD does have important implications for analysis involving several or all of the data items. First, to ensure that estimates of parameters characterising relationships between all *K* data items, such as regression coefficients, are identifiable all *K* data items must be collected from at least some of the units in the sample.

Second, not collecting some data items from a unit means that a number of interactions are not observed on the unit. In particular, if *L* of *K* variables are not collected from a sample unit, where *0 < L < K*, then the *K*-way to (*K − L + 1*)-way interactions are not observed on the unit and $I_P = \Pi_{x=K-P-L+1}^{K-P-1} x$ of the *P*-way

interactions are not observed on the unit. For example, if one of the *K* variables is not collected from a sample unit, the *K*-way interaction and the *K - 1* second order interactions are not observed on the unit. Clearly an SQD potentially loses a significant amount of information about interactions. Analysts have a range of different models (and hence model parameters) in which they are interested and accounting for all possible models in the design will often be impractical.

Third, managing the survey process for an SQD could be more complicated than for an SPD (e.g. testing of multiple questionnaires may be required). With the replacement of pen-and-paper interviewing by computer-assisted interviewing by many survey organisations SQDs are becoming a practical option. Nevertheless, allowing all *J* data patterns to be considered in an SQD will make the sample design, questionnaire design and analysis more complicated. Many of the *J* data patterns can be removed from consideration at the design stage. To inform the decision of which patterns to remove, Chipperfield and Steel (2009) and Chipperfield and Steel (2011) suggest ranking the patterns in terms of their effectiveness in meeting the design objectives. Furthermore, the authors illustrated in an empirical study when designing for estimating regression coefficients and means with *K* = 6, that many of the optimal allocations for an SQD required only four or fewer data patterns.

Also, the analysis of data collected by an SQD will be more complex than analysis of data collected by an SPD. This is perhaps not so much of a problem given there is a range of statistical packages available to analyse data missing completely at random.

Fourth, the optimal SQD allocation problem requires more information on design parameters than an SQD. In particular, when means are of interest to the design, the optimal allocation problem for an SQD requires the covariance matrix of **y**.

Finally, it is well known that different routings through a questionnaire can sometimes affect response values. This is particularly the case for sensitive questions (for discussion see Lyberg et al., 1997 chapter 5). Care should be taken to ensure this issue is addressed when considering an SQD.

3 Conclusions

An SPD is likely to continue to be a popular design because of the simplicity of the survey design, data collection and analysis. However, in the following situations an SQD may be worthwhile:
- There is a wide variation in either the enumeration costs or variance requirements of the data items (e.g. some data items require the use of expensive medical equipment). An optimal SQD allocation would tend to collect inexpensive data items that could be used as auxiliary variables in estimation.
- There is specific multivariate estimate (e.g. interaction) that is of interest to the design.
- Multivariate analysis is not important. SQDs are particularly efficient for estimating means.

With the replacement of pen-and-paper data collection by computer-assisted data collection by many survey organisations comes the potential for sophisticated SQDs. The choice of which variables not to collect could be randomly made by the computer prior to contacting the respondent or could be made based on responses during the data collection.

A futuristic view of SQDs is that an individual in sample will be assigned a pattern during the interview, while maintaining a probability sampling approach or ensuring the data items are missing at random. For example, a young person may be less likely to be given a health module but perhaps more likely to be given an employment module; if a person is retired they may be asked questions about how they funded their retirement. At some point the design and estimation issues will be too complex to manage. However, given the pressure on statistical organisations to reduce the cost of survey-based collections it may be worthwhile considering the potential cost savings that SQDs could provide.

References

Bell, P., & Whiting, J. (2007). Imputation and estimation for a thematic form census. Research Paper, Australian Bureau of Statistics, 1352.0.55.092 .

Chipperfield, J. O., & Steel, D. G. (2009). Design and estimation for split questionnaire surveys. Journal of Official Statistics, 25 , 227-244.

Chipperfield, J. O., & Steel, D. G. (2011). Efficiency of split questionnaire surveys. Journal of Statistical Planning and Inference, 141 , 19251932.

Gelman, A., King, G., & Liu, C. (1998). Not asked and not answered: Multiple imputation for multiple surveys. Journal of the American Statistical Association, 93 , 846-857.

Lyberg, L., Biemer, P., Collins, M., Leeuw, E. de, Dippo, C., Schwarz, N., et al. (1997). Survey measurement and process control. John Wiley and Sons.

Merkouris, T. (2004). Combining independent regression estimators from multiple surveys. Journal of the American Statistical Association, 99 , 1131-1139.

Munger, G., & Lloyd, B. H. (1988). The use of multiple matrix sampling for survey research. Journal of Experimental Education(56), 187-191.

Raghunathan, T. E., & Grizzle, J. E. (1995). A split questionnaire design. Journal of the American Statistical Association(90), 54-63.

Renssen, R. H., & Nieuwenbroek, N. J. (1997). Aligning estimates for common variables in two or more surveys. Journal of the American Statistical Association(92), 368-374.

Shoemaker, D. M. (1973). Principles and procedures of multiple matrix sampling. Ballinger, USA.

Thomas, Raghunathan, Schenker, Katzoff, and Johnson (2006). An Evaluation of Matrix Sampling Methods Using Data from the NHANES. Survey Methodology, 32, 217-231.

Wretman, J. (1994). Estimation in sample surveys with split questionnaires. Research Report, University of Stockholm, 3 , 1-11

New and Emerging Methods – Call for Volunteers

If you're interested in contributing an article to the "New and Emerging Methods" section of a future edition of *The Survey Statistician*, please contact Ineke Stoop at i.stoop@scp.nl.

Book and Software Review

Official Statistics and Survey Methodology Meets R: An Overview of Corresponding Packages

M. Templ[1,2,3] V. Todorov[4]

Abstract: The free and open-source programming language and software environment, R, is currently the most widely used and the most popular software for statistics and data analysis. In this contribution, the usefulness of R in the daily work in a statistical office is given followed by an overview of the important packages in the above areas of official statistics and survey methodology

Keywords: R in the Statistical Office, R as Mediator, Survey Methodology with R

1 Some Facts About R

The R language (Ihaka and Gentleman, 2009, R Development Core Team, 2012) is a freely available environment for statistical computing and graphics and is available at the Comprehensive R Archive Network (CRAN) as Free Software under the terms of the Free Software Foundation's GNU General Public License (GPL) in source code and binary form.

It can be used for processing data and performing statistical calculations. In addition, R includes an extremely large number of methods for data analysis, and features excellent graphical capabilities (especially useful in official statistics for panel representations see, Sarkar, 2008, Wickham, 2009) included in a straightforward programming language (see also Todorov, 2010, Todorov and Templ, 2012). R is widely used for statistical analysis in a variety of fields and is backed up by a large number of add-on packages that extend the system. The R Core Team (http://www.r-project.org/) defines R as an environment rather than a statistical system or a programming language. R is an integrated suite of software facilities for data manipulation, calculation and graphical display which includes:

---

[1] Department of Statistics and Probability Theory, Vienna University of Technology Methods Unit, Statistics Austria
E-mail: templ@tuwien.ac.at

[2] Methods Unit, Statistics Austria

[3] Data-analysis OG
E-mail: office@data-analysis.at

[4] UNIDO
E-mail: v.todorov@unido.org

- a suite of operators for calculations on arrays, mostly written in C and integrated in R,
- the most comprehensive, coherent and integrated collection of methods for data analysis,
- graphical facilities for data analysis and can display either on-screen or on hardcopy,
- a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities,
- a flexible object-oriented system facilitating the code reuse,
- high performance computing with interfaces to compiled code and facilities for parallel and grid computing, and
- an environment that allows it to communicate with many other software tools.

2 R in the Statistical Office

For data processing and data analysis in a national or international statistical organization, at least several well established statistical software packages stay ready for selection (see also Todorov and Templ, 2012):

i. SAS, because of its traditional position in these organisations (if the necessary funds are available), its ability to handle extremely large data sets and its availability on any platform including mainframes (Lowman, 2009)
ii. SPSS is considered user friendly because of its point-and-click interface (although still providing the so called syntax);
iii. STATA should be the most cost-effective among the three and by design should be especially suitable for working with data generated from complex survey designs (as is done in many NSOs).

But if one is looking for flexibility in reading, manipulating, and writing data, availability of the most recent statistical methodology, versatile presentation capabilities for generating of tables and graphics which can readily be used in text processing systems such as LATEX (or Word), creating dynamical reports e.g. using **Sweave** [Leisch, 2003] or the **brew** package, and last but not least, if one is looking for a really economical solution, one arrives at R [see also Todorov and Templ, 2012]. In addition, several tools stay ready that allow software to integrate all kind of modern tools for scientific computing, programming and management of data and files into one environment, i.e. to combine the capacities of R with editors that also allow syntax highlighting and code completion, the use of modern subversion systems for code and file management, the use of modern text scripting languages as LATEX, and C, C++, Java programming features integration as well as easy-to-use (automatic) connections to powerful servers. Editors like eclipse and its R[5], C++, Java, LATEX, subversion and server-connection facilities integration or RStudio[6], are well designed and are available for providing a modern scientific computing environment.

Although R has powerful statistical analysis features, data manipulation tools, and versatile presentation capabilities, it has not yet become an everyday statistical

---

[5] See http://www.walware.de/goto/statet
[6] See http://support.rstudio.org

package in the national and international statistical offices. This is mainly due to the widespread opinion that R is hard to learn and has a very steep learning curve, which is true for learners without any programming skills. However, point-and-click graphical user interfaces, that also show the underlying code, are available [Fox, 2005] and interfaces to all popular toolkits for programming of graphical user interfaces are available.

3 R as a Mediator

In daily work in a national or international statistical organization, like for example Statistics Austria or UNIDO, many different systems for data analysis and data processing are used. Data exchange between statistical systems like **SAS**, **SPSS**, **EViews**, **Stata**, **Excel**, **R**, data-base systems **Access**, **MySql**, **DB2** or presentation **html xml** is often required and the importance of the statistical data and metadata exchange format SDMX is continuously growing.

In these cases R offers either by its base installation or through add-on packages available from CRAN, the most flexible import and export interfaces. The connection to all major database systems is easily accomplished with the packages **ROracle**, **RMySQL**, **RSQLite**, **RmSQL**, **RPgSQL**, **RODBC**, and **DBI**. The integration of other statistical systems is made possible through packages like **RWeka**, **X12**, **R2wd**, **R2ppt**, **RExcel** while data exchange among such systems is facilitated by the package **foreign**. More specific applications like web page generation or data and metadata exchange between organizations can be used the packages **R2HTML**, **sdmxer** or **RSDMX**.

If the data does not fit in the RAM of the computer when very large data files have to be handled, interfaces to RDBMS could be of help. Another way to solve the large data set problem is to use one of the packages **filehash**, **LaF**, **ff** or **bigmemory** or to connect to powerful servers with **rserve**[7] (also integrated in the RStudio editor).

R is designed for parallel computing and **parallel** is the recommended package for this task. There are also several other packages which could be used for parallel computing like, for example, **snow** or **multicore**. R also comes with nice interfaces for integrating compiled code (for example, package **Rccp**, **inline**, **rJava**) and with features for code profiling (package **profr** and **proftools**). In addition, R offers interfaces to almost all commercial and open-source linear program solvers that are often needed for special tasks in survey methodology.

4 R for Survey Statistics

R includes several methods which are useful for data processing and survey methodology in statistical offices and organizations where dealing with complex data sets from finite populations is a usual task.

The CRAN task view on *Official Statistics and Survey Methodology* (http://cran.rproject.org/view=OfficialStatistics) contains a list of packages which include methods typically used in official statistics and survey methodology. In the following presentation we list those packages and briefly outline their functionalities.

---

[7] See http://www.rforge.net/Rserve/

4.1 Complex Survey Designs

To draw survey samples, the package `sampling` includes various different algorithms (Brewer, Midzuno, pps, systematic, Sampford, balanced (cluster or stratified) sampling via the cube method, etc.) but also functionality to calibrate the design weights is available [see also Tillé, 2006].

For estimation purposes and to work with already drawn survey samples, the package `survey` [Lumley, 2010] - the standard package for that task - can be used once the given survey design is specified (stratified sampling design, cluster sampling, multistage sampling and pps sampling with or without replacement). The resulting object can be used to estimate (Horvitz-Thompson) totals, means, ratios and quantiles for domains or the whole survey sample, and to apply regression models. Variance estimation for means, totals and ratios can be done either by Taylor linearization or resampling (BRR, jack-knife, bootstrap or user-defined).
As an "add-on" package, `ReGenesees`[8] uses facilities from and extends the survey package. This package also offers graphical user interface. Package `EVER` provides variance estimation for complex designs by delete-a-group jack-knife replication for (Horvitz-Thompson) totals, means, absolute and relative frequency distributions, contingency tables, ratios, quantiles and regression coefficients even for domains.

The `laeken` package implements functions to estimate certain social inclusion indicators (at-risk-of-poverty rate, quintile share ratio, relative median risk-of-poverty gap, Gini coefficient) including their variance for domains and strata based on (calibrated) bootstrap resampling.

To perform simulation studies in official statistics, the package `simFrame` [Alfons et al., 2010] provides a framework for comparing different point and variance estimators under different survey designs as well as different conditions regarding missing values, representative and non-representative outliers.

Various other packages are available to select samples with specific designs: `pps`, `sampfling` and `SamplingStrata`.

4.2 Calibration

To calibrate the sampling weights to fit exact population characteristics and/or to calibrate for unit non-responses the package `survey` allows for post-stratification, generalized raking/calibration, GREG estimation and trimming of weights.
The `EVER` package provides facilities (function `kottcalibrate`()) for calibrating on a total number of units in the population, on marginal distributions or joint distributions of categorical variables, or on totals of quantitative variables.
The `calib`() function in package `sampling` allows to calibrate for nonresponse (with response homogeneity groups) for stratified samples. Similar but possibly faster is the implementation in `laeken` (function `calibWeights`()).

4.3 Editing and Visual Inspection of Microdata

For editing, the package `editrules` converts readable linear (in)equalities into matrix form, which then can be applied for editing the given data set. The package

---

[8] Not on CRAN, but available on
http://www.istat.it/it/strumenti/metodi-e-software/software/regenesees.

deducorrect depends on package **editrules** and applies deductive correction of simple rounding, typing and sign errors based on balanced edits. Values are changed so that the given balanced edits are fulfilled.

The package **rrcovNA** [Todorov et al., 2011] provides robust location and scatter estimation and robust principal component analysis with high breakdown point for incomplete data. It is thus applicable to find representative and non-representative outliers.

For visual inspection of microdata, the package **VIM** [Templ et al., 2012a] can be used.

With **VIM**, one can visualize missing values using suitable plot methods and analyse the structure of missing values in microdata using univariate, bivariate, multiple and multivariate plots. Missing values from specified variables are highlighted in selected variables. It also allows evaluating imputations visually. Moreover, the package comes also with a graphical user interface.

Package **tabplot** provides the tableplot visualization method, which is used to profile or explore large statistical data sets. Up to a dozen variables are shown column-wise as bar charts (numeric variables) or stacked bar charts (factors).

4.4 Imputation

A distinction between iterative model-based methods, k-nearest neighbor methods and miscellaneous methods is made. However, often the criteria for using a method depend on the scale of the data, which in official statistics are typically a mixture of continuous, semi-continuous, binary, categorical and count variables. Note that only few imputation methods can deal with mixed types of variables and semi-continuous ones and only the methods in package **VIM** account for robustness issues.

EM-based imputation methods are offered by package **mi** [Yu-Sung et al.], **mice** [van Buuren and Groothuis-Oudshoorn, 2011], **Amelia** [Honaker et al., 2011], **VIM** and **mix** [Schafer, 1997]. Package **mi** provides iterative EM-based multiple Bayesian regression imputation of missing values and model checking of the regression models used whereas the regression models for each variable can also be defined by the user. The data set may consist of continuous, semi-continuous, binary, and categorical and/or count variables.

The package **mice** [van Buuren and Groothuis-Oudshoorn, 2011] provides iterative EM-based multiple regression imputation as well and the dataset may consist of continuous, binary, categorical and/or count variables. Multiple imputation where first bootstrap samples are drawn for EM-based imputation can be carried out with **Amelia** [Honaker et al., 2011]. It is also possible to impute longitudinal data. Package **VIM** offers EM-based multiple imputation (function **irmi**()) using robust estimations [Templ et al., 2011b], which allows R to adequately deal with data including outliers. It can handle data consisting of continuous, semi-continuous, binary, and categorical and/or count variables.

Nearest neighbor imputation methods come along with the package **VIM**. It provides an implementation of the popular sequential and random (within a domain) hot-deck algorithm, but also a fast k-nearest neighbor (knn) algorithm which can be used for large data sets. It uses a modification of the *Gower Distance* for dealing with a

mixture of numerical, categorical, ordered, continuous and semi-continuous variables.

4.5 Statistical Disclosure Control

In its raw form, data from statistical agencies and other institutions are mostly confidential. Data providers must ensure confidentiality by modifying the original data so that no statistical unit can be re-identified and at the same time should guaranty a minimum amount of information loss.

For microdata perturbation, the package **sdcMicro** [Templ, 2008, Templ and Meindl, 2010] can be used for the generation of confidential (micro)data, i.e. for the generation of public and scientific-use files. All methods are implemented in C++ to allow fast computations. The package also provides a graphical user interface [Templ, 2008].

To simulate synthetic data, the package **simPopulation** [Alfons et al., 2011] offers methods for simulation of synthetic, confidential, close-to-reality populations for surveys based on sample data. Such population data can then be used for extensive simulation studies in official statistics using, for example, package **simFrame**.

For tabular data, the package **sdcTable** [Templ and Meindl, 2010] can be used to provide confidential (hierarchical) tabular data. It includes the HITAS and the HYPER-CUBE technique and offers interfaces to various commercial and open-source linear program solvers.

4.6 Time Series Analysis and Seasonal Adjustment

For general time series methodology we refer to TimeSeries task view *TimeSeries* on CRAN. Specifically, for survey methodology, decomposition of time series can be done with the function **decompose**(), or more advanced by using the function **stl**() or with the **StructTS**() function. Many powerful tools for seasonal adjustment can be accessed via package **x12**. It provides a wrapper function and GUI for the X12 binaries, which have to be installed first.

4.7 Statistical Record Matching

The package **StatMatch** [D'Orazio et al., 2006] provides functions to perform statistical matching between two data sources sharing a number of common variables. It creates a synthetic data set after matching of two data sources via a likelihood approach or hot-deck.

For record linkage methodology, the package **RecordLinkage** provides functions for linking and de-duplicating data sets and package **MatchIt** allows nearest neighbor matching, exact matching, optimal matching and full matching among other matching methods.

4.8 Indices and Indicators

The most comprehensive collection of indicator methodology is included in the package **laeken** which helps to estimate popular risk-of-poverty and inequality indicators (at risk-of-poverty rate, quintile share ratio, relative median risk-of-poverty gap, Gini coefficient, gender pay gap). In addition, standard and robust methods for

tail modelling of Pareto distributions are provided for semi-parametric estimation of indicators from continuous univariate distributions such as income variables.

Various indicators are included in package **ineq.** However **ineq** is not designed to deal with sampling weights directly (these can only be emulated approximately via **rep(x, weights)**).

For the calculation of indices, function **priceIndex**() from package **micEcon** can be used.

For visualization purposes, the package **sparkTable** offers tools to produce scalable sparklines for web sites and reports. It also contains visualization tools for presenting indicators in checker plot - grid-based representation of possible complex indicators, bar plots and time series in thematic maps [Templ et al., 2012b].

5 Summary and Conclusions

There is an increasing demand for statistical tools which combine the ease of use of the traditional software packages with the availability of the newest analytical methods as provided by the statistical programming language R. In this contribution, we briefly described its usefulness in the daily work in statistical offices.

We also listed and briefly described the most popular packages in R for survey methodology. The development of R packages in some particular areas of survey statistics is growing rapidly. For example, R provides many more methods for statistical disclosure control than in any other software package. The situation is similar in indicator methodology and survey statistics using robust methods. The development of the latter two was strongly supported by projects like the AMELI project [Münnich et al., 2011, Templ et al., 2011a].

References

A. Alfons, M. Templ, and P. Filzmoser. An object-oriented framework for statistical simulation: The R package **simFrame**. *Journal of Statistical Software*, 37(3):1-36, 2010.

A. Alfons, S. Kraft, M. Templ, and P. Filzmoser. Simulation of close-to-reality population data for household surveys with application to EU-SILC. *Statistical Methods & Applications*, 20(3):383-407, 2011. URL http://dx.doi.org/10.1007/s10260-011-0163-2.

M. D'Orazio, M. Di Zio, and M. Scanu. *Statistical Matching: Theory and Practice.* Wiley
Series in Survey Methodology. John Wiley & Sons, 2006. ISBN 9780470023549. URL
http://books.google.at/books?id=U75yCqrdXW4C.

J. Fox. The R Commander: A basic statistics graphical user interface to R. *Journal of Statistical Software*, 14(9):1-42, 2005. URL http://www.jstatsoft.org/v14/i09.

J. Honaker, G. King, and M. Blackwell. Amelia II: A program for missing data. *Journal of Statistical Software*, 45(7):1-47, 2011. URL http://www.jstatsoft.org/v45/i07/.

R. Ihaka and R. Gentleman. R: A language for data analysis and graphics. *Journal of

Computational and Graphical Statistics, 5(3):299-314, 2009.

F. Leisch. `Sweave`, part II: Package vignettes. *R News*, 3(2):21-24, 2003.

T. Lumley. *Complex Surveys: A Guide to Analysis Using R*. Wiley, 2010.

R. Münnich, A. Alfons, C. Bruch, P. Filzmoser, M. Graf, B. Hulliger, J-P. Kolb, R. Lehtonen, D. Lussmann, A. Meraner, R. Münnich, D. Nedyalkova, T. Schoch, M. Templ, M. Valaste, A. Veijanen, and S. Zins. Policy recommendations and methodological report. Research Project Report WP10 - D10.1/D10.2, FP7-SSH-2007-217322 AMELI, 2011. URL http://ameli.surveystatistics.net.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL http://www. R-project.org/. ISBN 3-900051-07-0.

D. Sarkar. *Lattice: Multivariate Data Visualization with* R. Springer, New York, 2008. ISBN 978-0-387-75968-5.

J.L. Schafer. *Analysis of Incomplete Multivariate Data*. Chapman & Hall, London, 1997.

M. Templ. Statistical disclosure control for microdata using the R-package sdcMicro. *Transactions on Data Privacy*, 1(2):67-85, 2008. URL http://www.tdp.cat/issues/abs.a004a08.php.

M. Templ and B. Meindl. Practical applications in statistical disclosure control using R. In J. Nin and J. Herranz, editors, *Privacy and Anonymity in Information Management Systems,* Advanced Information and Knowledge Processing, pages 31-62. Springer London, 2010. ISBN 978-1-84996-238-4. URL http://dx.doi.org/10.1007/978-1-84996-238-4_3. 10.1007/978-1-84996-238-4 3.

M. Templ, A. Alfons, P. Filzmoser, M. Graf, J. Holzer, B. Hulliger, A. Kowarik, S. Kraft, R. Lehtonen, D. Nedyalkova, T. Schoch, A. Veijanen, and S. Zins. *R packages plus manual.* Research Project Report WP10 - D10.3, FP7-SSH-2007-217322 AMELI, 2011a. URL http://ameli.surveystatistics.net.

M. Templ, A. Kowarik, and P. Filzmoser. Iterative stepwise regression imputation using standard and robust methods. *Comput Stat Data Anal*, 55(10):2793-2806, 2011b.

M. Templ, A. Alfons, and P. Filzmoser. Exploring incomplete data using visualization techniques. *Advances in Data Analysis and Classification*, 6(1):29-47, 2012a. doi: DOI: 10.1007/s11634-011-0102-y.

M. Templ, B. Hulliger, A. Kowarik, and K. Fürst. Combining geographical information and traditional plots: The checkerplot. *International Journal of Geographical Information Science*, 2012b. in press.

Y. Tillé . *Sampling Algorithms*. Springer Series in Statistics. Springer, 2006. ISBN 9780387308142. URL http://books.google.at/books?id=2auW1rVAwGMC.

V. Todorov. R in the statistical office: The UNIDO experience. Working Paper 03/2010 1, United Nations Industrial Development, 2010. URL http://www.unido.org/fileadmin/user_media/Services/Research_and_

Statistics/statistics/WP/WP_2010_03.pdf.

V. Todorov and M. Templ. R in the statistical office: Part 2. Working paper. in press., United Nations Industrial Development, 2012.

V. Todorov, M. Templ, and P. Filzmoser. Detection of multivariate outliers in business survey data with incomplete information. *Adv Data Anal Classif*, 5(1):37-56, 2011.

S. van Buuren and K. Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3):1-67, 2011. URL http://www. jstatsoft.org/v45/i03/.

H. Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009. ISBN 978-0-387-98140-6. URL http://had.co.nz/ggplot2/book.

S. Yu-Sung, A. Gelman, J. Hill, and M. Yajima. Multiple imputation with diagnostics (mi) in R: Opening windows into the black box. *Journal of Statistical Software*, 45.

We are interested in fostering review of books and software in the area of survey methods. This would include standard review of individual books or software packages. This may also include broader reviews of groups of text and monographs in specific sub-areas; or similarly broad reviews of available software. Of particular interest are some of the new R libraries that have been developed recently for survey methods. If you are able to write a) review for this section, please contact Natalie Shlomo (n.shlomo@soton.ac.uk).

ALGERIA

Algeria has conducted several types of censuses since its independence in 1962.

Five general censuses of population and housing (RGPH) (1966, 1977, 1987, 1998 and 2008)

Those censuses were carried out by the national commission for the population census (CNRP), which became the national census and statistical survey commission (CNRES) and later the National Office of Statistics (ONS). Preparations for a census of population and housing begin about two years in advance. Mapping delineation is the most time-consuming part of the process, since it is done from scratch for each census. To date, geographic information systems (GISs) have not been used, which makes it difficult to measure change at the sub-municipal level. Three types of questionnaires are administered: ordinary and collective household population, nomadic population and separately enumerated population. They have been used for various purposes: preliminary results, national sampling, wilaya (province) sampling (variable sampling fractions depending on the wilaya size, since 1987) and comprehensive. Each census was followed by a postcensal verification survey, but the results were not adjusted with any kind of coverage rate. Only missing households are taken into account in the adjustment process. There have been no major innovations to speak of in the five censuses. The questionnaires have not changed much, except in 2008, when additional demographic variables were introduced (international migration, fertility, mortality). There have been no detailed publications, except on fertility (municipal data). Coverage of labour force activity was very poor in the last two censuses (1998, 2008), which is why very few data on the subject have been published. In particular, none of the five censuses looked at income. Except in the 1987 census of population and housing, the ONS has not used specific applications to automatically adjust for inconsistencies detected by the data capture monitoring program. While questionnaire coding and capture are still done manually, the length of time they take has been reduced significantly. The initial results are published a few months after the end of the census. The use of electronic media to disseminate the results is not yet widespread. In our view, the censuses of population and housing are still underexploited despite their exorbitant costs. Secondary products are always possible (tabulations only), but only for the last census and only on request to the ONS. There are no public samples extracted from the censuses.

Two general censuses of agriculture (RGA) (1973 and 2001)

These were carried out by the ministry of agriculture. The next one is scheduled for 2012-2013 and will probably be conducted in conjunction with the ONS. Very few data were released.

An economic census (RE) conducted in 2011 by the ONS

The first operation of its kind in Algeria, phase 1 of the RE consisted in an enumeration of all economic establishments (except agriculture), regardless of industry and legal status, and all government institutions. It will be followed by a structure survey on a large sample during 2012. Three questionnaires were administered: economic sector, government entities and an opinion questionnaire. With regard to innovations in this economic census, questionnaires were captured, for the first time in Algeria, with a new technology: document scanning. The preliminary results are available online at www.ons.dz.

Collections Statistiques N° 168/ 2012
Série E: Economie N° 67
Premier recensement économique – 2011 – Résultats préliminaires de la première phase
[First economic census – 2011 – Phase 1 preliminary results]

AUSTRALIA

Australian Census Data Enhancement

In 2005, the ABS embarked on a project to enhance the value of Census data by bringing it together with other datasets, both ABS and non-ABS, to leverage more information from the combination of individual datasets than is available from the individual datasets separately. This project will be continued for the 2011 Census.

The Census Data Enhancement (CDE) project involves integrating unit record data from the Census of Population and Housing with other ABS and non-ABS datasets to create new datasets for statistical and research purposes. The project will also add value to data from the Census of Population and Housing by bringing it together with data from future Censuses.
The CDE project delivers significant public benefits without compromising the privacy of individuals or the confidentiality of their data. The project facilitates:
  - improved information to support good government policy making, program evaluation and service delivery; and
  - an improved and expanded range of official statistics.

The Census Data Enhancement (CDE) project will encompass a number of components, listed below. Of these components, 2 and 5 are new for the 2011 Census.
  1. Bringing together 2011 Census data with a small number of predetermined datasets during Census processing using name and address, for quality studies.
  2. Bringing together 2011 Census data with a small number of predetermined datasets during Census processing using name and address, to create statistical outputs.

3. Wave 2 of a 5% Statistical Longitudinal Census Dataset (SLCD), bringing together data from the 2006 and 2011 Censuses.
4. Bringing together the SLCD with other datasets without using name and address for statistical and research purposes.
5. Bringing together 2011 Census data with other datasets without using name and address after Census processing.

Wave 1 of the SLCD was created from the 2006 Census dataset by selecting a random sample of 5% of persons in the 2006 Census of Population and Housing. Wave 2 of the SLCD will endeavour to bring together the wave 1 records with their corresponding records in the 2011 Census.

Subsequent waves will be created with each new Census, providing a longitudinal dataset of information about 5% of the Australian population.

More information about the CDE project is available at http://www.abs.gov.au/ausstats/abs@.nsf/mf/2062.0 or from Graeme Thompson (graeme.thompson@abs.gov.au), Director Census Data Enhancement, Australian Bureau of Statistics, Canberra.

Australia: efforts to improve Aboriginal and Torres Strait Islander counts

One focus of the 2011 Australian Census was improved counts of Aboriginal and Torres Strait Islander people. The 2011 Census introduced a Northern Australia enumeration strategy and collection management unit, with additional funding and specific responsibility for enumeration of the Northern Territory and the difficult-to-collect northern areas of Queensland, Western Australia and South Australia.

Alongside this effort, the Post-Enumeration Survey (PES) was funded to increase the sample from areas expected to have a high density of Aboriginal and Torres Strait Islander people, including discrete Indigenous communities. This has approximately doubled the PES sample of these people, and will provide improved sampling error for estimates of this important sub-population.

Estimates from the PES at state and territory level will still have high standard errors for many states. Bayesian small-area estimation will be used to stabilise these estimates. The underlying model will be more sophisticated than was used in 2006. A key improvement will be use of regression against a measure of the difficulty of enumerating Aboriginal and Torres Strait Islander people within each Census collector's area. Logistic regression based on the 2006 PES was used in developing this difficulty measure, a key component of which is the proportion of Aboriginal and Torres Strait Islander people counted in the area.

For further information please contact Phillip Bell, Australian Bureau of Statistics Branch via email: philip.bell@abs.gov.au

BRAZIL

Cristiano Ferraz

The 2010 Brazilian Demographic Census
(with the assistance of Sonia Albieri and Zélia Bianchini)

The IBGE - Instituto Brasileiro de Geografia e Estatística is the official government agency in charge of the census in Brazil. The Brazilian Census of 2010 used innovative forms of data collection as it adopted hand held computers equipped with GPS systems, and allowed for questionnaires to be answered over the internet. Since 1960, the Brazilian demographic census research model is based on two questionnaires: a larger one, applied to a sample of households, and called the sample questionnaire; and a shorter one, applied to the households not selected for the sample, called the basic questionnaire. The 2010 census used five different sample fractions for the households, depending on the size of the municipalities given by its 2009 estimated population. The census official results started to be released in 2010, and included a series of publications in 2011. Several publications are still expected for 2012, and a volume on the census methodology description should be available by 2013. More information can be found at the IBGE web site:

http://www.ibge.gov.br/home/estatistica/populacao/censo2010/default.shtm

Third Brazilian School on Sampling and Survey Methodology - ESAMP, and Second International Workshop on Surveys for Policy Evaluation
(with the assistance of Marcel de Toledo Vieira)

The third edition of ESAMP and the second edition of the International Workshop on Surveys for Policy Evaluation took place in Juiz de Fora, Minas Gerais, between 22nd and 25th of November, 2011. The events congregated 5 international invited speakers (United Kingdom, Sweden, Canada and Colombia), and during the four days, 12 conferences, 6 special thematic sessions and two short courses were delivered, among activities involving oral and poster presentations of participants. Both meetings welcomed registrations from 13 different Brazilian states, from Colombia and United States. The ESAMP and the International Workshop are jointly sponsored by the ENCE - National School of Statistical Sciences (which is part of IBGE), the IBGE - Intituto Brasileiro de Geografia e Estatística, and the Joaquim Nabuco Foundation. The events are supported by the Brazilian Statistical Association. The International Workshop is also supported by the International Association of Survey Statisticians (IASS). The Fourth ESAMP and the Third International Workshop are scheduled to be held in November, 2013, in Brasília. General information on the meetings can be found at www.fundaj.gov.br/3esamp

CANADA

Every five years Statistics Canada takes a Census of Population, the last census was on May 10, 2011. The Census of Population produces population counts that feed into the population estimates program, which are needed to determine the distribution of seats in parliament, the distribution of federal transfer payments, and

the transfer and allocation of funds among regional and municipal governments, school boards and other local agencies within the provinces and territories. The 2011 Census contained 10 questions (basic demographic and language questions), as approved by the government. For the first time, Statistics Canada conducted the National Household Survey (NHS), a new voluntary survey that contained the content that had been previously collected on the long form census. This report will not cover the NHS.

The introduction of a new census collection approach, called the "Wave Methodology", was extremely successful in terms of increasing returns by Internet without negatively impacting the non-response follow-up workload. The Wave Methodology provided a streamlined approach to communicate with respondents and invited them using various methods to complete their census questionnaire (such as mail reminders, telephone voice broadcasts, field enumerators and paid advertising) at specific times throughout the collection period. With this new collection approach, the reliance on a large and decentralized workforce for the operations in the field was reduced. Communication between field staff, their supervisors and Head Office was significantly improved with the introduction of a Field Management System (FMS), which is a web-based application that was accessible by all field staff. With the FMS, timeliness of notifying the field that a dwelling had completed a census questionnaire by mail-back, on-line or through the Census Help Line was significantly improved. Mail-out to dwellings was increased from 73% achieved in 2006 to 79% in 2011. Internet responses increased from the 18% achieved in 2006 to 54.4% in 2011. The preliminary response rate achieved to the 2011 Census is 98.1%. The Census population counts were released on February 8, 2012. The 2011 Census counted 33,476,688 people in Canada.

For additional information, contact Normand Laniel at Normand.Laniel@statcan.gc.ca.

INDIA

Dr. Gayatri Vishwakarma

First time in India the concept of unique identity for a common man is being introduced. The **Unique Identification Authority of India** (**UIDAI**) is an agency of the Government of India implementing a unique identification project called the ***AADHAAR*** scheme. It was established in February 2009, and will own and operate the Unique Identification Number database. The authority aims to provide a unique ID number to all Indian citizens. The authority will maintain a database of residents containing biometric and other data.

AADHAAR is a 12-digit unique number for all residents in India. The number will be stored in a centralized database and linked to the basic demographics and biometric information – photograph, ten fingerprints and iris – of each individual. It is unique and robust enough to eliminate the large number of duplicate and fake identities in government and private databases. The random number generated will be devoid of any classification based on caste, creed, religion and geography.

It is believed that Unique National IDs (***AADHAAR***) will help address the rigged state elections and widespread embezzlement that affects subsidies and poverty alleviation programs. UIDAI does not use any existing databases citing problems of

fraud and duplicate/ghost beneficiaries in the existing databases. Instead, it will enroll the entire population using its multi-registrar enrollment model using verification processes prescribed by the UIDAI. This will ensure that the data collected is clean right from the beginning of the program.

The official estimates for the project is ₹18,000 crore (US$ 3.59 billion). The total numbers of *AADHAAR's* issued as of 07-Feb-2012 is 13.29 crores (US$ 133 million). This is about 11% of the total population of India.

Future prospects of *AADHAAR*: It will become the single source of identity verification. Residents would be spared the hassle of repeatedly providing supporting identity documents each time they wish to access services such as obtaining a bank account, passport, driving license and so on. By providing a clear proof of identity, *AADHAAR* will also facilitate entry for poor and underprivileged residents into the formal banking system and the opportunity to avail services provided by the government and the private sector. *AADHAAR* is giving migrants mobility of identity, financial inclusion with deeper penetration of banks, insurance and easy distribution of benefits of government schemes.

References:
1. "Rs.100 crore for Unique Identification Project", *The Hindu* (Chennai, India), 17 February 2009, http://www.hindu.com/2009/02/17/stories/2009021756751000.htm, retrieved 2009-06-26, "*... The Unique Identification Authority of India is being established under the aegis of the Planning Commission ….*"
2. "Nilekani to give numbers, ministries to issue cards", *The Economic Times*, 16 July 2009, http://economictimes.indiatimes.com/Nilekani-to-give-numbers-ministries-to-issue cards/articleshow/4782505.cms, retrieved 2009-07-18.
3. "Nilekani takes charge, says first set of IDs in 12-18 months". *The Times Of India*. 2009-07-24. http://timesofindia.indiatimes.com/articleshow/4812763.cms.
4. http://uidai.gov.in/what-is-aadhaar-number.html
5. http://uidai.gov.in/why-aadhaar.html

NEW ZEALAND

New Zealand 2011 Census

The New Zealand 2011 Census was called off 11 days before census day on 8 March 2011 as a result of the 22 February earthquake in Christchurch. The census was fully operational with 140 field offices set up, 7,500 field staff employed, a full advertising campaign underway and 25% of forms delivered. The 2011 Census could not be completed due to the national state of emergency and the probable impact on census results.

The collection model for the 2011 Census involved collectors personally delivering internet access codes and paper forms to all dwellings. Respondents would then have the choice of returning forms online, by mail or collector pick-up. The online response target was 30% of all forms.
The next New Zealand census will be held on 5 March 2013, seven years after the last completed census in 2006.

The goal of New Zealand's 2013 Census programme is to produce the best possible count of the population and dwellings. The 'best possible count' was defined in terms of Key Performance Indicators (KPIs) and it was necessary to have systems in place to monitor the collection progress against set measures and to develop a strategy for achieving the measures.

Statistics New Zealand developed a methodology for non-response follow-up optimisation based on historic non-response data, a Difficulty Index developed for the 2011 Census and delivery and collection information from the Field Collection system.

The primary aim of this project was to find out which geographic enumeration areas should be targeted for extra resource to ensure the KPIs are met. The KPI targets for response rates are 95% at the national level, 94% for population subgroups (Maori, Pacific Peoples, Asian and 15-29 year olds) and over 90% at the lowest census enumeration area level.

The methodology optimises the use of resources (collector hours) for the follow-up on non-response which takes place after the official collection phase ends. Usually, the follow-up is undertaken by selected collectors and their supervisors. The project was extended to include a decision-making tool and maps for census management.
For further information please refer to the 12 March 2012 Census Advisory on http://www.stats.govt.nz/tools_and_services/services/newsletters/census-advisory-newsletter.aspx

We have further useful information at www.stats.govt.nz. Join us on www.facebook.com/StatisticsNZ .

PHILIPPINES

Gervacio G. Selda, Jr.

Conduct of Pilot Census of the 2012 Census of Agriculture and Fisheries

The Census of Agriculture and Fisheries (CAF) is a large-scale operation geared towards the collection and compilation of statistics on the nation's agriculture and fishery sectors. For the Philippines, the Census of Agriculture and Fisheries (2012 CAF) will be conducted in February 2013 with January to December 2012 as its reference period. The conduct of said activity will be spearheaded by the National Statistics Office (NSO). The 2012 CAF is the sixth in a series of decennial agriculture census and the fifth in the decennial census of fisheries in the country. It will adopt a modular approach patterned after the World Census Programme for Agriculture 2010 (WCA 2010), as recommended by the Food and Agriculture Organization (FAO) of the United Nations. The modular approach consists of a core module to be carried out on a complete enumeration to provide key structural data; and one or more sample-based census supplementary modules - intended to provide more in-depth data. The new approach is within the framework of the system of integrated agricultural and fisheries censuses and surveys, allowing for better organization of statistical activities.

Because of the complexity of its operation, a pilot census will be conducted in March 2012. Its conduct provides an opportunity to test all aspects of preparation for the census operation and come-up with better procedures recommended for use in the

actual census. Preparations for this census started five years ago with the Statistical Research and Training Center (SRTC) - the research and training arm of the PSS - coordinating the preparatory works.

The core module of the 2012 CAF will cover a limited range of key data items on agriculture, aquaculture, and fishing. The supplementary modules, to be conducted after the core census module has been carried out, shall provide a more in-depth structural data on specific concerns such as irrigation and cropping pattern, livestock and poultry, aquaculture, fisheries, and capital formation in agriculture and fisheries.

Along with the core census module, which will cover all households and establishments engaged in agriculture, aquaculture, and fisheries, a community-based module will also be administered in all barangays – the lowest political and geographical classification of areas in the Philippines. The community-based module shall provide information on facilities and services related to agriculture, aquaculture, and fisheries that are present/available in the barangay. The dominant terrain of the barangay and the destructive natural calamities experienced in the barangay during the past five years will also be gathered.

Proposal to Reorganize the Philippine Statistical System

A bill to reorganize the Philippine Statistical System (PSS) is now pending with the legislative body of the country - the Philippine Congress. The bill once approved would put together the main statistical gathering and compilation agencies into one body that will be called the Philippine Statistical Authority (PSA). Currently, several agencies gather statistics as per given mandates to them. These are the National Statistics Office (NSO), Bureau of Agricultural Statistics (BAS), Bureau of Labor Employment and Statistics (BLES), and the National Statistical Coordination Board (NSCB). The conduct of statistical research and training activities solely fall under the mandate of the Statistical Research and Training Center (SRTC). The bill is one of the priority bills that the current administration in the Philippines has endorsed to the legislative body - Senate and House of Representatives - for immediate enactment.

The bill will make the PSA the central statistical authority of the government on primary data collection. It will be headed by a National Statistician having the rank of undersecretary, and attached to the National Economic and Development Authority (NEDA) – the premier planning body in the Philippines. According to the proposed bill, the PSA Board, to be headed by the National Statistician and composed of representatives of several department, agencies and private sector, would replace the NSCB as the highest policy-making body on statistical matters. The bill also proposes the creation of the Philippine Statistical Research and Training Institute, which upgrades the present SRTC, the research and training arm of the Philippine statistical system.

For further information, please contact, Gervacio G. Selda, Jr., Statistical Research and Training Center at ggseldajr@srtc.gov.ph.

POLAND

Tomasz Żądło

The national censuses in Poland were conducted in 2010 (the National Agricultural Census) and 2011 (the National Population and Housing Census). To obtain census data four sources were used:

- 28 administrative registers (data from different registers were combined based on the System of Personal Identification Numbers)
- the internet – self-enumeration (CAII)
- telephone interviews (CATI)
- a census with the participation of enumerators with the use of hand-held terminals (CAPI)

During these censuses only electronic forms were used. The GIS (Geographic Information Systems) tools where used to support the enumeration process. Because of the use of administrative registers, the internet self-enumeration and other electronic solutions, the number of enumerators was reduced c.a. 10 times comparing with 2002 census and it was possible to reduce the cost of the census by c.a. 50 million Euro.

During the National Population and Housing Census were asked questions concerning: first name and surname, Personal Identification Number (PESEL), sex, date of birth, address, migration, marital status, country of birth, citizenship, nationality, language. The full form is available at http://www.stat.gov.pl/e-census/dok/personal_form.pdf. The form for a sample survey in 20% of dwellings was larger and includes questions divided into the following categories: dwellings, building, education, economic activity, commuting o work, source of maintains, disability, fertility, production plans, country of birth and citizenship, internal and international migration, nationality and language (see http://www.stat.gov.pl/e-census/dok/List_of_questions_to_the_sample_survey.pdf).

The above description of the last censuses in Poland is based on the following website (in English): http://www.stat.gov.pl/e-census/index.html.

UNITED KINGDOM

Peter Lynn

The 2011 Population Census was the first in England and Wales to include online data collection. All households were sent a traditional paper Census form but were at the same time informed that they could alternatively complete the form online. Overall, 16% of returns were made online and the Office for National Statistics deemed the innovation to be a success. Online completion rates were highest amongst 26-35 year-olds and for large households with seven or more residents. Evaluation of various aspects of the Census is available online at http://tinyurl.com/census-update and further reports will be posted here as the Census data continue to be processed.

Meanwhile, The "Beyond 2011 Programme" is working towards the objective of ensuring that the 2011 Census will be the last one to use traditional Census forms. The programme is investigating and assessing alternative options for producing the required population and socio-demographic data. Options under consideration include the use of a continuous survey, a rolling census, and use of administrative data at aggregate and/or individual level. In comparing options, ONS has identified the key trade-offs as being between accuracy, frequency and geography. An outline of the programme can be found at http://tinyurl.com/census-beyond-2011.

UNITED STATES

Kathy Downey

Consumer Expenditure Survey

The Consumer Expenditure Survey (CE) program consists of two surveys—the quarterly Interview survey and the Diary survey—that provide information on the spending habits of American consumers, including data on their expenditures, income, and consumer unit (families and single consumers) characteristics. The overall mission of the CE is to collect, produce, and disseminate information that presents a statistical picture of consumer spending for use by government agencies and private data users. In particular, consumer expenditure data supplied by CE are a critical component in the calculation of the Consumer Price Index (CPI) and are used to estimate the cost weights for the CPI's consumer goods and services classification structure.

The CE program initiated the Gemini Project, a multiyear survey redesign effort, in early 2009. The mission of the project is to improve expenditure estimates in the CE by reducing measurement error. During the course of the project, the CE program will evaluate and implement survey design changes with the goals of improving overall data quality, increasing the analytic value of the data to users, and supporting greater operational flexibility to respond to changes in the data collection environment. Papers, presentations, and other project materials are available from the Gemini website at http://www.bls.gov/cex/geminimaterials.htm.

Motivation for redesign

The CE faces a number of issues that affect the quality of the data collected. These issues include evidence of measurement error in the survey data, changes in consumer spending behaviors, and the need for greater flexibility in the mode and method of data collection. All redesign changes proposed will be subject to budgetary constraints, and implementation decisions will be considered in terms of priorities and trade-offs. To allow for an unpredictable budgetary environment, both a complete redesign of the CE as well as more limited modifications to the current design are being investigated.

Even at the outset of the survey redesign process, a number of challenges were evident. Reconciling diverse user needs into a cohesive set of survey requirements is a non-trivial task. Despite identifying the needs of varied users and reconciling competing interests, there are certain to be some users with unmet needs. Additionally, it is a challenge to gather, respond, and act on stakeholder concerns and suggestions while maintaining forward project progress. Finally, because of time constraints and uncertainty over the direction and nature of forthcoming redesign recommendations, the CE program must a priori and concurrently investigate several redesign topics; research will be used to support or reject some forthcoming proposals. Ultimately, however, the redesign process is constrained by two overriding factors: the final survey design must fulfill the survey objectives while maintaining operational survey costs within specified budget levels.

Completed activities

In 2009 through 2011, CE completed a number of activities under the Gemini Project. These included a report which outlined definitions of data quality for CE and the creation of a research tracking system for the many CE methodological and cognitive issues tests informing plans for the redesign. CE has also contracted with survey

research companies to perform research, particularly on respondent record keeping ("U.S. Consumer Expenditure Records Study," http://www.bls.gov/cex/cesrvymethsgeisen1.pdf) and data capture technologies ("2011 Data Capture Technologies and Financial Software for Collecting Consumer Expenditure Data," http://www.bls.gov/cex/ceother2011westat.pdf).

In addition, several events have been held to collect information about key issues related to the redesign: a Survey Redesign Panel Discussion (January 2010), a Data Capture Technology Forum (March 2010), and a Data Users' Needs Forum (June 2010). In December 2010, CE and the Council of Professional Associations on Federal Statistics (COPAFS) held a CE Methods Workshop where five key topics central to the redesign were discussed: global questions, interview structure, proxy reporting, recall period, and split questionnaire designs.

Finally, CE has contracted with the Committee on National Statistics (CNSTAT) to assemble a Consensus Expert Panel, coordinate several research events, and produce a report with redesign recommendations based on the event discussions and other outside independent proposals. These research events included a First Panel Meeting (February, 2011) to gather background material on the CE, a Household Survey Producers Workshop (June, 2011), and a Redesign Options Workshop (October, 2011). CNSTAT is finalizing their recommendations and will deliver their final report to CE in August, 2012.

Current and future activities

Through 2012, the Gemini Project will continue to synthesize findings from information gathering, summarize results from completed research, and plan future research studies. Current research projects or special analyses include: a Bounding Interview Project, the CE Data Quality Profile, a project on Combining Split Questionnaire Files, Diary-to-Interview Imputation Methods, Evaluation of Financial Application Software, Expenditure Grouping Study, Exploratory Burden Index, a project on Survey Length and Telephone Questions, Records Information and Feasibility of Use, and a Web Diary Feasibility Test. Planned CE research studies include an analysis of the Business and Income Screener, Expenditure App Prototype and Feasibility Study, Global Grocery Expenditures Allocation, an Individual Diaries Field Study, and Telephone Administration in Interview 3 and 4.

A Redesign Proposal will be delivered in Spring 2013. In Fall 2013, CE plans to have an approved roadmap in place for the development and implementation of a redesigned CE. The development, testing, and implementation of a redesigned survey will occur in 2013 and beyond.

For more information, please see the Gemini project web page (http://www.bls.gov/cex/geminimaterials.htm) or contact the Gemini Project Manager, Kathy Downey (downey.kathy@bls.gov).

Statistics: Growing to Serve a Data-Dependent Society
July 28 — August 2, 2012

The Joint Statistical Meetings 2012

Date: July 28 – August 2 2012
Location: San Diego Convention Center, San Diego, California
Website: http://www.amstat.org/meetings/jsm/2012/index.cfm

JSM (the Joint Statistical Meetings) is the largest gathering of statisticians held in North America. It is held jointly with the American Statistical Association, the International Biometric Society (ENAR and WNAR), the Institute of Mathematical Statistics, the Statistical Society of Canada, and the International Chinese Statistical Association, and the International Indian Statistical Association. Attended by more than 6,000 people, meeting activities include oral presentations, panel sessions, poster presentations, continuing education courses, an exhibit hall (with state-of-the-art statistical products and opportunities), career placement services, society and section business meetings, committee meetings, social activities and networking opportunities.

San Diego, California, the host city for JSM 2012, offers a wide range of options for sharing time with friends and colleagues or sightseeing with family. For information, contact meetings@amstat.org.

The 2012 Joint Statistical Meetings will be held at the San Diego Convention Center, located at 111 West Harbor Drive, San Diego, CA 92101.

The 6th International Total Survey Error Workshop (ITSEW 2012) will be held in the Netherlands, September 2-4, 2012

The theme of the 6th International Total Survey Error Workshop is "Total Survey Error: Past, Present, and Future".

The workshop seeks presentations of research that explore or analyze the trade-offs of two or more sources of survey error within the context of relatively recent methodological or statistical approaches. However, abstracts that deal with the general topic of non-sampling error are welcome. The workshop format is informal to encourage open discussion and sharing of ideas. Thus, presentations discussing research still in progress or results of completed, progress or results of completed, unpublished research are quite appropriate.

Topics in the scope of the 6th ITSEW are:
- o Analysis of relation between nonresponse and measurement error (interviewer effects, contact or refusal conversion strategies)
- o Design of questionnaires for multiple survey modes
- o Decomposing mode effects
- o Mixed-mode data collection design including Internet web surveys
- o Quality of longitudinal data from Internet panels
- o Quality of paradata
- o Adaptive/responsive survey designs accounting for multiple errors.

Schedule:

September 2 (Sunday afternoon/evening): Arrival/Get together

September 3: Presentations; Evening program Reception/Poster session

September 4: Presentations; Evening Open mike discussion; Farewell

For more general information about the ITSEW, visit the website at www.cbs.nl/itsew

Professor Dr. Edith D. de Leeuw

Plantage Doklaan 40, NL-1018 CN Amsterdam

Tel + 31 20 622 34 38
Fax + 31 20 330 25 97
E-mail edithl@xs4all.nl

RSS Conference 2012
The annual conference of the Royal Statistical Society

ROYAL STATISTICAL SOCIETY

3-6 September, 2012, Telford UK
http://www.rssconference.org.uk/

The Royal Statistical Society's annual international conference will take place in September 2012 with its familiar mix of high-profile international plenary speakers, invited and contributed presentations, poster sessions, short courses and workshops, and a busy social and networking programme.

Added to this mix will be an expanded programme of professional development tutorials and workshops which should appeal to a broad spectrum of professional statisticians, researchers and the wider user community.

The full conference programme is now available.

Confirmed plenary speakers include **Hal Varian** (Google), **Martine Durand** (OECD), **David Hand** (Imperial College) and **Anthony Davison** (EPFL). **Andrew Dilnot**, new Chair of the UK Statistics Authority, will also give a keynote address.

Short courses will take place on Monday 3 September with the main conference opening on the morning of Tuesday 4 September. The conference will end on the afternoon of Thursday 6 September.

The venue will be the Telford International Centre in Shropshire.

Please contact Paul Gentry (p.gentry@rss.org.uk) for further information.

ENBIS-12
9 – 13 September 2012
University of Ljubljana, Faculty of Economics, Slovenia
www.enbis.org

The 12th Annual Conference of the European Network for Business and Industrial Statistics will take place at the University of Ljubljana, Faculty of Economics in Slovenia from 9-13 September, 2012.

It aims to build on the success of previous annual conferences by bringing together statistical practitioners in all branches of our profession. The conference offers plenty of opportunities to share ideas and learn about current best practices.

Contributions consist of review papers, state of the art advances in scientific research, overview and introductory papers, presentation of an area of expertise to the uninitiated or communication of best practice and practitioner experience.

We are delighted to announce keynote presentations by Marko Grobelnik of Jožef Stefan Institute, Slovenia; Rob Kozinets, Chair of Marketing, York University, Canada; Steve MacFeely, Director of Business Statistics & Methodology, Central Statistics Office, Ireland; John Pullinger, Librarian and Director General, Information Services at the House of Commons, London, President-Elect of RSS; and Grant Reinman, Senior Fellow, Statistics and Design For Variation at Pratt & Whitney, USA.

There will be many special sessions including these organised around the following themes: software, statistics education, French ENBIS, US ENBIS, ISBIS, DEINDE and Hidden Gems aimed at putting Central Europe back on the map of Quality Improvement.

Confirmed pre-conference and post-conference courses/workshops include "To explain or to predict?" led by Galit Shmueli and Ron Kenett on 9th September as well as "Statistics for Business Improvement" led by Roland Caulcutt on 13th September.

For more information, please contact Irena Ograjenšek at
irena.ograjensek@ef.uni-lj.si .

The 4rd International Workshop on Internet Survey Methods
September 12-13, 2012/Daejeon, Korea

The Fourth International Workshop on Internet Survey Methods

Date: September 12- 13, 2012

Venue: Statistical Center of Statistics Korea, Deajeon, South Korea

Organizer: International Statistical Cooperation Division, Statistics Korea

Theme: Quality Improvement of Official Statistics and the Role of Internet Survey Method

Languages: English and Korean (simultaneous interpretation)

Website: http://soc.kuleuven.be/web/home/11/81/eng

Workshop Description

The provisional theme of the 4th International Workshop on Internet Survey Methods is **"Quality Improvement of Official Statistics and the Role of Internet Survey Methods".** The workshop will focus on a wealth of major issues concerned with cutting-edge statistical methodologies as well as an abundance of innovative efforts and research performances which National Statistical Offices (NSOs) and other data producers have been doing to solve their difficulties and challenges.

The tentative discussion topics are:

- (Special Lecture) Advantages and Disadvantages of Utilizing Internet Survey Methods for Official Statistics
- Desing & Implementation Issues in Web Surveys
- Cases of Statistical Production Utilizing Internet Survey Methods
- Data Quality Issues in Web Surveys

The official language of the Workshop is English. The simultaneous interpretation from English to Korea will be provided for all the sessions.

Date 12 – 14 September, 2012

Location The State Enterprise «National Center for Business and Cultural Cooperation
"Ukrainian House"», Kyiv, Ukraine

Website http://iaos2012.ukrstat.gov.ua/

Contact iaos2012@ukrstat.gov.ua

The IAOS 2012 conference will focus on how statistical offices are ensuring that the key messages in their statistics are getting through to users and are adding value to decision making. In particular it will cover:

- Standards for statistical reporting and commentary writing: using plain and simple language
- Statistical and graphical presentation techniques including the use of trend analysis and data visualisation
- Developments in the use of IT: the web, providing on-line analysis facilities for users, using metadata to improve searches; iPods, wikis and blogs
- How third parties can add value
- Marketing our statistics
- Building relationships with customers, monitoring feedback and setting priorities
- Statistical literacy

H2R|2012

International Conference on Methods for
Surveying and Enumerating Hard–to-Reach Populations

Date:
October 31 – November 3, 2012

Venue:
Marriott New Orleans at the Convention Center, New Orleans, Louisiana, U.S.A.

Website:
www.amstat.org/meetings/h2r/2012

The conference will bring together survey methodologists, sociologists, statisticians, demographers, ethnographers, policy analysts and other professionals from around the world to present new and innovative concepts and techniques for surveying hard to reach populations. The conference will address both the statistical and survey design aspects of including hard to reach groups. Researchers will report findings from censuses and surveys and other research related to the identification, definition, measurement, and methodologies for surveying and enumerating undercounted populations. The conference will serve as a venue to network, and to share research and experiments designed to advance our understanding of the topic.

For further information contact: H2R2012@amstat.org

7ᴇ COLLOQUE FRANCOPHONE
SUR LES SONDAGES

5 au 7 novembre 2012

http://sondages2012.ensai.fr/

Ensai
Campus Ker Lann
Rue Blaise Pascal – BP 37203
35172 BRUZ cedex
FRANCE

La Société Française de Statistique (SFdS) confie à l'Ecole Nationale de la Statistique et de l'Analyse de l'Information (Ensai) l'organisation du 7e Colloque francophone sur les sondages du 5 au 7 novembre 2012.

L'objet du colloque est de faire le point sur l'état des pratiques et de la recherche dans les divers domaines de la méthodologie des enquêtes. Il réunira des chercheurs et des praticiens issus des instituts publics ou privés qui réalisent ou utilisent des enquêtes. Les thèmes abordés seront variés, et on trouvera notamment des applications en sciences politiques et sociales, marketing, en santé et sciences de la vie.

Conférenciers invités

Conférence d'ouverture : **Anne-Marie Dussaix |** professeur honoraire à l'Essec
Conférence de clôture : **Jean-Claude Deville |** Ensai/Crest
Pascal Ardilly | Insee
Daniel Bachelet | Commission des sondages
Yves Berger | Université de Southampton
Jay Breidt | Colorado State University
Yannick Carriou | Ipsos France
Daniel Defays | Eurostat
Camelia Goga | Université de Bourgogne
Beat Hulliger | University of Applied Sciences Northwestern Switzerland
Pierre Lavallée | Statistique Canada
Abdellatif Lfarakh | Haut-commissariat au plan du Maroc
Carl-Erik Särndal | Statistics Sweden, Université d'Örebro
Yves Tillé | Université de Neuchâtel
Aurélie Vanheuverzwyn | Médiamétrie
Josiane Warszawski | Inserm
Wesley Yung | Statistique Canada

CONTACT :
Président du comité d'organisation :
Éric Lesage : (sondages2012@ensai.fr)

NTTS 2013

New Techniques and Technologies for Statistics

Date:
5 – 7 March, 2013

Venue:
The Charlemagne building, Wetstraat 170 Rue de la Loi, 1040 Brussels, Belgium

Website:
www.ntts2013.eu

Contact:
ESTAT-NTTS@ec.europa.eu

NTTS (New Techniques and Technologies for Statistics) 2013 is an international scientific conference on the impact of new technologies on statistical collection, production and dissemination systems.

Present and share the outcomes of recent research activities in statistics in general and in official statistics including Eurostat ESSnet projects and VIP (Vision Infrastructure Projects). Promote new research methodological and technological development for use by Official Statistics.

Discuss future needs and developments of research in statistics, new paradigms for data use, access and retrieval (open data, big data, and organic data) and ICT developments and infrastructures for use by Official Statistics.

NTTS 2013 addresses research and development aspects related to innovative methods and techniques for official statistics with a particular emphasis on automatic and ICT-based methods. Papers are accepted in the following areas:

- New ways of collecting, accessing and using big amount of data.
- Integration, consolidation, combination of multiple data sources.
- Analysing data.
- Distributing, presenting and accessing data and microdata.
- Support for evidence-based policymaking.
- Use of standards for Official statistics.

ITACOSM2013

The Third Italian Conference on Survey Methodology (ITACOSM2013) will be held in Milan, 26-28 June 2013 at the University of Milan-Bicocca (http://www.unimib.it/go/102/Home/English).

This bi-annual international meeting of the Survey Sampling Group of the Italian Statistical Society is intended as a forum of scientific discussion for the developments of theory and application of survey sampling methodologies in human and natural sciences.

The conference program will include keynote lectures, specialized, contributed and poster sections.

Visit regularly http://www.statistica.unimib.it/itacosm13/ for key dates and updated info.

European Survey Research Association (ESRA) Fifth Conference, July 2013

Details to be announced…

http://surveymethodology.eu/home/

2013 Joint Statistical Meetings

August 3 - 8, 2013
Montréal, Québec, Canada, Palais de congrès de Montréal

http://www.amstat.org/meetings/jsm.cfm

The 59th World Statistics Congress
Hong Kong, China
25-30 August 2013

Date
25 – 30 August, 2013

Location
The Hong Kong Convention and Exhibition Centre (HKCEC), Hong Kong, China

Website
http://www.isi2013.hk/en/index.php

Contact
WSC2013-Enquiry@censtatd.gov.hk

The 59th World Statistics Congress (WSC) will be held in Hong Kong during 25-30 August 2013. The meeting venue will be the Hong Kong Convention and Exhibition Centre (HKCEC), which is a magnificent, multi-purpose venue located right in the heart of Hong Kong on the Victoria Harbour.

The 59th WSC provides a platform for the international statistical community to share and present the latest knowledge and innovation in statistics. The scientific programme encompasses a wide range of topics facilitating professional exchanges and sharing amongst experts and practitioners in various statistical spheres. Featuring the unique theme "Youth", a series of sessions will be organised on the "Theme Day" of the 59th WSC to address, from various statistical perspectives, topics surrounding the "Youth". Delegates are welcome to plan tailor-made events from a wide spectrum of activities including scientific programme, social programme, satellite meetings and short courses.

Third International Workshop on Surveys for Policy Evaluation and 4th Brazilian School on Sampling and Survey Methodology – ESAMP4

Organized by: IBGE and University of Brasilia (UNB)
Where: Brasilia, Brazil
When: 05.11.2013 to 08.11.2013
Website: to be confirmed.

SMALL AREA ESTIMATION 2014

We would like to inform that the next Small Area Estimation Conference will be organized by the Poznan University of Economics in the second half of June in 2014 in   Poland.

This conference will probably take place in the town of Wągrowiec situated 60 km from   Poznań   in   the   Pietrak   hotel   located   at   a   beutiful   lake (http://pietrak.pl/wagrowiec/galeria_en.html).

All details (final date, place, cost, accommodation etc.) will be posted on EWORSAE's website soon. Information about the website of SAE2014 will also be posted on a special web page, which is being developed by Poland.

Readers will be informed as soon as it is ready.

Survey Practice
Practical Information for Survey Researcher

www.surveypractice.org

April 2012

Introduction to Special Issue of Survey Practice on Item Nonresponse
D. A. Dillman

Determinants of Item Nonresponse to Web and Mail Respondents in Three Address-Based Mixed-Mode Surveys of the General Public
B. L. Messer, M. L. Edwards, D. A. Dillman

Comparing Item Nonresponse across Different Delivery Modes in General Population Surveys
V. M. Lesser, L. A. Newton, D. Yang

Item Nonresponse in a Client Survey of the General Public
G. D. Israel, A. J. Lamm

Do Mail and Internet Surveys Produce Different Item Nonresponse Rates? An Experiment Using Random Mode Assignment
M. M. Millar, D. A. Dillman,

Survey Practice Book List 2012: Recent Books and Journals in Public Opinion, Survey Methods, and Survey Statistics
M. Callegaro

SRM

Survey Research Methods

Vol 6, No 1 (2012)

http://w4.ub.uni-konstanz.de/srm/issue/current

Introduction to the Special Issue on the Theory of Human Values
E. Davidov, P. Schmidt, S. Schwartz

Value Systems of Fathers, Mothers and Adolescents: Do Parents and their Children Construe Basic Values in the Same Way?
D. Barni, A. Knafo

Congruence and Performance of Value Concepts in Social Research
T. Beckers, P. Siegers, A. Kuntz

Testing the discriminant validity of Schwartz' Portrait Value Questionnaire items – A replication and extension of Knoppen and Saris (2009)
C. Beierlein, E. Davidov, P. Schmidt, S. H. Schwartz, B. Rammstedt

A comparison of the invariance properties of the PVQ-40 and the PVQ-21 to measure human values across German and Polish Samples
J.Cieciuch, E. Davidov

Human Values and Trust in Institutions across Countries: A Multilevel Test of Schwartz's Hypothesis of Structural Equivalence
D. Morselli, D. Spini, T. Devos

Testing the Circular Structure of Human Values: A Meta-Analytical Structural Equation Modelling Approach
H. Steinmetz, R. Isidor, N. Baeuerle

Current Issue: March 28, Vol. 28 No. 1, 2012

A Multiplicative Masking Method for Preserving the Skewness of the Original Micro-records
N. Ruiz

Cell Bounds in k-way Tables Given Conditional Frequencies
B. J. Smucker, A. Slavković, X. Zhu

Book Reviews

In Other Journals
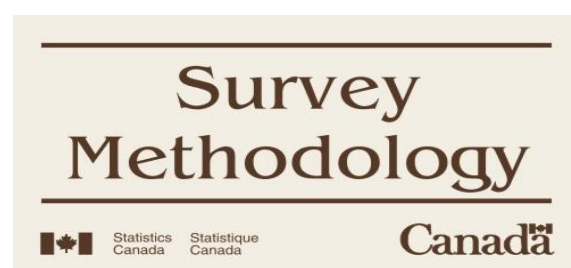
Survey Methodology
A Journal Published by Statistics Canada

Volume 37, Number 2, December 21

Modelling of complex survey data: Why model? Why is it a problem? How can we approach it?
D. Pfeffermann

A Bayesian analysis of small area probabilities under a constraint
B. Nandram, H. Sayit

On bias-robust mean squared error estimation for pseudo-linear small area estimators
R. Chambers, H. Chandra, N. Tzavidis

Variance estimation under composite imputation: The methodology behind SEVANI
J. F. Beaumont, J.Bissonnette

Alternative demographic sample designs being explored at the U.S. Census Bureau
P. E. Flanagan, R. A. Killion

Adaptive network and spatial sampling
S. Thompson

Alternative survey sample designs: Sampling with multiple overlapping frames
S. L. Lohr

Ten years of balanced sampling with the cube method: An appraisal
Y. Tillé

Innovations in survey sampling design: Discussion of three contributions presented at the U.S. Census Bureau
J. Opsomer

http://iospress.metapress.com/content/r32572329205/?p=d83645b238914d958a970f2342deb11f&pi=0

Reshaping health statistics: A new framework
J. Madans, C. Abou-Zahr, A. Bercovich, T. Boerma, D. Carlton, L. Castro, M. Desmedt, E. Domingo, J. Kahimbaara, M. Marquardt, H. Nviiri, E. Norgaard, E. Vassenden, M. Wolfson

Measuring health in population surveys
G. Kalton

Protecting confidentiality in a data enclave
P. S. Meyer, E. S. Robinson, J. Madans

The French health information system
M. Goldberg, E. Jougla, M. Fassa, R. Padieu, C. Quantin

Measuring health care costs of individuals with employer-sponsored health insurance in the U.S.: A comparison of survey and claims data
A. Aizcorbe, E. Liebman, S. Pack, D. M. Cutler, M. E. Chernew, A. B. Rosen

Combining health register data and biobank data
M. Gissler, H. M. Surcel

Health statistics in the Netherlands, review 1995–2009, preview 2010–2015
J. Kardaun, A. de Bruin, V. van Polanen Petel, S. van der Aart, J. van den Berg, O. van Hilten

Evaluation of the interest of using discharge abstract databases to estimate breast cancer incidence in two French departments
C. Quantin, E. Benzenine, M. Fassa, M. Hägi, E. Fournier, J. Gentil, D. Compain, E. Monnet, P. Arveux, A. Danzon

Volume 80, Issue 1, April 2012
http://onlinelibrary.wiley.com/doi/10.1111/insr.2012.80.issue-1/issuetoc

Volume 13, Number 1, March 2012
http://www.stat.gov.pl/pts/15_ENG_HTML.htm

Journal of Privacy and Confidentiality

Current Issue: Volume 3, Issue 2 (2012)

Editorial: In This Issue
S. E. Fienberg

How Uncertainty about Privacy and Confidentiality is Hampering Efforts to More Effectively Use Administrative Records in Producing U.S. National Statistics
G. W. Gates

Why It Matters to Distinguish Between Privacy & Confidentiality
K. Prewitt

A Dissenting View from Julia Lane
J. I. Lane

Use of Administrative Records and the Privacy-Confidentiality Trade-off
J. H. Madans

Commentary: Future U.S. National Statistics Use of Administrative Data
G. Duncan
Toward a Reconceptualization of Confidentiality Protection in the Context of Linkages with Administrative Records
S. E. Fienberg
Comment on Article by Gates
J. P. Reiter

Trust but Pre-Verify?
F. Scheuren

Privacy and the Statistician: What Do We Need to Know to Certify Nondisclosure?
A. M. Zaslavsky

Rejoinder
G. W. Gates

Volume 5, Issue 1, April 2012
http://www.tdp.cat/issues11/vol05n01.php

k-Concealment: An Alternative Model of k-Type Anonymity
T. Tassa, A. Mazza, A. Gionis

Utility-guided Clustering-based Transaction Data Anonymization
A. Gkoulalas-Divanis, G. Loukides

Distributing Data for Secure Database Services
V. Ganapathy, D. Thomas, T. Feder, H. Garcia-Molina, R. Motwani

A Practical Differentially Private Random Decision Tree Classifier
G.Jagannathan, K. Pillaipakkamnatt, R. N. Wright

Efficient Privacy Preserving Protocols for Similarity Join
B. Hawashin, F. Fotouhi, T. M. Truta, W. Grosky

Mobile Systems Privacy: 'MobiPriv' A Robust System for Snapshot or Continuous Querying Location Based Mobile Systems
L. Stenneth, P. S. Yu

BIOMETRIKA

Biometrika Vol. 99, Issue 2, June 2012
http://biomet.oxfordjournals.org/content/current

Dependence modelling for spatial extremes
Wadsworth J. L., Tawn J. A.

Stochastic blockmodels with a growing number of classes
D. S. Choi, P. J. Wolfe, E. M. Airoldi

Doubly misspecified models
N. X. Lin, J. Q. Shi, R. Henderson

Componentwise classification and clustering of functional data
A. Delaigle, P. Hall, N. Bathia

Global optimality of nonconvex penalized estimators
Y. Kim, S. Kwon

Pointwise nonparametric maximum likelihood estimator of stochastically ordered survivor functions
Y. Park, J. M. G. Taylor, J. D. Kalbfleisch

Analysing bivariate survival data with interval sampling and application to cancer epidemiology
H. Zhu, M.-C. Wang

Likelihood approaches for the invariant density ratio model with biased-sampling data
Y. Shen, J. Ning, J. Qin

Efficient estimation for the Cox model with varying coefficients
K. Chen, H. Lin, Y. Zhou

Nonparametric inference for assessing treatment efficacy in randomized clinical trials with a time-to-event outcome and all-or-none compliance
R. M. Elashoff, G. Li, and Y. Zhou

Corrected-loss estimation for quantile regression with covariate measurement errors
H. J. Wang, L. A. Stefanski, Z. Zhu

Multiple imputation in quantile regression
Y. Wei, Y. Ma, R. J. Carroll

Improved double-robust estimation in missing data and causal inference models
A. Rotnitzky, Q. Lei, M. Sued, J. M. Robins

Empirical bootstrap bias correction and estimation of prediction mean square error in small area estimation
D. Pfeffermann, S. Correa

Miscellanea:
A new residual for ordinal outcomes
C. Li, B. E. Shepherd
Structuring shrinkage: some correlated priors for regression
J. E. Griffin, P. J. Brown

Information dynamics and optimal sampling in capture-recapture
T. Schweder, D. Sadykova

A generalized Dunnett test for multi-arm multi-stage clinical studies with treatment selection
D. Magirr, T. Jaki, J. Whitehead

Inference for additive interaction under exposure misclassification
T. J. Vanderweele

Amendments and Corrections:
'On measuring the variability of small area estimators under a basic area level model'
G. S. Datta, J. N. K. Rao, D. D. Smith

Editorial: Publishing research—for whose benefit?
S. Day

The evolving Society: united we stand
V. Isham

Vignettes and health systems responsiveness in cross-country comparative analyses
N. Rice, S. Robone, P.C. Smith

Quantifying the weight of evidence from a forensic fingerprint comparison: a new paradigm
C. Neumann, I. W. Evett, J. Skerrett

Joint modelling of longitudinal outcome and interval-censored competing risk dropout in a schizophrenia clinical trial
R. Gueorguieva, R. Rosenheck, H. Lin

Victorian probability and Lewis Carroll
E. Seneta

Gender-specific differences and the impact of family integration on time trends in age-stratified Swiss suicide rates
A. Riebler, L. Held, H.Rue, M. Bopp

Stochastic population forecasts based on conditional expert opinions
F. C. Billari, R. Graziani, E. Melilli

Analysing intensive longitudinal data after summarization at landmarks: an application to daily pain evaluation in a clinical trial
P. Bunouf, J.-M. Grouin, G. Molenberghs, G. Koch

Finding dynamic treatment effects under anticipation: the effects of spanking on behaviour
M.-J. Lee, F. Huang

Making predictions from complex longitudinal data, with application to planning monitoring intervals in a national screening programme
M. J. Sweeting, S. G. Thompson

Evaluating continuous training programmes by using the generalized propensity score
J. Kluve, H. Schneider, A. Uhlendorff, Z. Zhao

Modelling and predicting temporal frequency of foot-and-mouth disease cases in countries with endemic foot-and-mouth disease
K. C. Young, W. O. Johnson, G.Jones, A. Perez, M. C. Thurmond

A Bayesian non-linear model for forecasting insurance loss payments
Y. Zhang, Vanja Dukic, J. Guszcza

Obituaries: Peter A. W. Lewis, 1932–2011; Wyndham Somerville Paige, 1927-2011
T. Springall

Journal of the American Statistical Association

http://amstat.tandfonline.com/toc/uasa20/current

Applications and Case Studies

Nonparametric Covariate-Adjusted Association Tests Based on the Generalized Kendall's Tau
W.Zhu, Y. Jiang, H. Zhang

Intrinsic Regression Models for Medial Representation of Subcortical Structures
X. Shi, H. Zhu, J. G. Ibrahim, F. Liang, J. Lieberman, M. Styner

Modeling Waves of Extreme Temperature: The Changing Tails of Four Cities
D. J. Dupuis

Adjustment for Missing Confounders Using External Validation Data and Propensity Scores
L. C. McCandless, S. Richardson, N. Best

Partially Hidden Markov Model for Time-Varying Principal Stratification in HIV Prevention Trials
J. Y. Dai, P. B. Gilbert, B. R. Mâsse

Using Conditional Kernel Density Estimation for Wind Power Density Forecasting
J. Jeon, J. W. Taylor

Bayesian Inference for Dynamic Treatment Regimes: Mobility, Equity, and Efficiency in Student Tracking
T. Zajonc

Bayesian Estimation and Prediction for Inhomogeneous Spatiotemporal Log-Gaussian Cox Processes Using Low-Rank Models, With Application to Criminal Surveillance
A. Rodrigues, P. J. Diggle

MRI Tissue Classification Using High-Resolution Bayesian Hidden Markov Normal Mixture Models
D. Feng, L. Tierney, V. Magnotta

Bias-Corrected Hierarchical Bayesian Classification With a Selected Subset of High-Dimensional Features
L. Li

Theory and Methods .Cross-Dimensional Inference of Dependent High-Dimensional Data
K. H. Desai, J. D. Storey

Sparse Estimation of Conditional Graphical Models With Application to Gene Networks
B. Li, H. Chun, H. Zhao

A Semiparametric Approach to Dimension Reduction
Y. Ma, L. Zhu

A Valid Matérn Class of Cross-Covariance Functions for Multivariate Random Fields With Any Number of Components
T. V. Apanasovich, M. G. Genton, Y. Sun

Multiple Imputation for M-Regression With Censored Covariates
H. J. Wang, X. Feng

Information Ratio Test for Model Misspecification in Quasi-Likelihood Inference
Q. M. Zhou, P. X.-K. Song, M. E. Thompson

Quantile Regression for Analyzing Heterogeneity in Ultra-High Dimension
L. Wang, Y. Wu, R. Li

Likelihood-Based Selection and Sharp Parameter Estimation
X. Shen, W. Pan, Y. Zhu

Block Bootstraps for Time Series With Fixed Regressors
D. J. Nordman, S. N. Lahiri

Semiparametric Double Balancing Score Estimation for Incomplete Data With Ignorable Missingness
Z. Hu, D. A. Follmann, J. Qin

One-Sided and Two-Sided Tolerance Intervals in General Mixed and Random Effects Models Using Small-Sample Asymptotics
G. Sharma, T. Mathew

Estimating Space and Space-Time Covariance Functions for Large Data Sets: A Weighted Composite Likelihood Approach
M. Bevilacqua, C. Gaetan, J. Mateu, E. Porcu

Modeling Nonstationary Processes Through Dimension Expansion
L. Bornn, G. Shaddick, J. V. Zidek

Estimation of Copula Models With Discrete Margins via Bayesian Data Augmentation
M.S. Smith, M. A. Khaled

A Heckman Selection-t Model
Y. V. Marchenko, M. G. Genton

Estimating Regression Parameters in an Extended Proportional Odds Model
Y. Qing Chen, N. Hu, S.-C. Cheng, P. Musoke, L. P. Zhao

Recursively Imputed Survival Trees
R. Zhu, M. R. Kosorok

Interim Design Modifications in Time-to-Event Studies
S. Irle, H. Schäfer

On Fractile Transformation of Covariates in Regression
B. Sen, P. Chaudhuri

Simplex Factor Models for Multivariate Unordered Categorical Data
A. Bhattacharya, D. B. Dunson

Bootstrapping for Significance of Compact Clusters in Multidimensional Datasets
R. Maitra, V. Melnykov, S. N. Lahiri

Sliced Latin Hypercube Designs
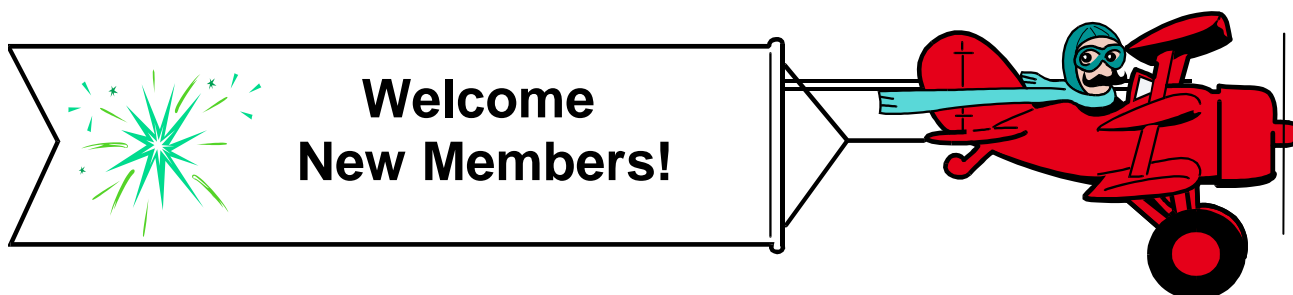P. Z. G. Qian

Optimal Designs for Rational Function Regression
D. Papp

Vast Volatility Matrix Estimation Using High-Frequency Data for Portfolio Selection
J. Fan, Y. Li, Ke Yu

Book Reviews

Welcome
New Members!

We are very pleased to welcome the following new members!

Country	Surname	Name
France	Randrianasolo	Toky
New Zealand	Dunnet	Gary
India	Chandra	Hukum
Uruguay	Galmes	Miguel Angel
Uganda	Okello	Patrick
United States	Lin	Yu-Chieh
United States	Bailey	Jeffrey T.
New Zealand	Seyb	Allyson Joan
Netherlands	Torres van Grinsven	Vanessa

IASS Officers and Council Members

Executive Officers

President (2011-2013):	Raymond Chambers (Australia)	ray.chambers@internode.on.net
President-elect:	Danny Pfeffermann (Israel)	d.pfeffermann@soton.ac.uk
Vice-Presidents (2011-2013):	Steve Heeringa (USA)	sheering@isr.umich.edu
	Denise Britz do N. Silva (Brazil)	denise.silva@ibge.gov.br
Scientific Secretary (2011-2013):	Ineke Stoop (The Netherlands)	i.stoop@scp.nl
Council Members (2009-2013):	Mike Hidiroglou (Canada)	mike.hidiroglou@statcan.gc.ca
	Edith D. de Leeuw (The Netherlands)	edithl@xs4all.nl
	Monica Pratesi (Italy)	m.pratesi@ec.unipi.it
	Mick P. Couper (USA)	mcouper@umich.edu
	Eva Elvers (Sweden)	eva.elvers@scb.se
	Yves Tille (Belgium)	yves.tille@unine.ch
Council Members (2011-2015):	Christine Bycroft (New Zealand)	christine.bycroft@stats.govt.nz
	Ka-Lin Karen Chan (China)	klchan@censtatd.gov.hk
	Olivier Dupriez (Belgium/USA)	odupriez@worldbank.org
	Natalie Shlomo (UK)	n.shlomo@soton.ac.uk
	Marcel de Toledo Vieira (Brazil)	marcel.vieira@ufjf.edu.br
	Alvaro Gonzalez Villalobos (Argentina)	alvarun@fibertel.com.ar

Committee Chairs

Hong Kong 2013 Programme Committee	Eric Rancourt (Canada)	eric.rancourt@statcan.gc.ca
Char of the 2013 Election Nominations Committee	Keith Rust (USA)	keithrust@westat.com
Chair of the jury of the 2013 Cochran-Hansen Prize	Yves Tille (Belgium)	yves.tille@unine.ch
Chair of the Rio 2015 Programme Committee	Christine Bycroft (New Zealand)	christine.bycroft@stats.govt.nz

The Secretariat

Executive Director:	Catherine Meunier (France)	catherine.meunier@insee.fr
Treasurer:	Ada van Krimpen (The Netherlands)	an.vankrimpen@cbs.nl
Webmaster:	Justin Lokhorst (Australia)	justin.lokhorst@abs.gov.au
IASS Secretariat Membership Officer	Margaret de Ruiter-Molloy (The Netherlands)	m.deruitermolloy@cbs.nl

Institutional Members

2 International Organizations

AFRISTAT
EUROSTAT

15 Bureaus of Statistics

AUSTRALIA – AUSTRALIAN BUREAU OF STATISTICS
BRAZIL – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA
CANADA – STATISTICS CANADA
CHINA – DIRECCAO DOS SERVICOS DE ESTATISTICA
DENMARK –DANMARKS STATISTIK
FINLAND – STATISTICS FINLAND
GERMANY –STATISTICHE BUNDESAMT
ITALY –INSTITUTO NAZIONALE DI STATISTICA
KOREA, REPUBLIC OF – STATISTICS KOREA
MEXICO –INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA (INEGI)
MAURITIUS – STATISTICS MAURITIUS
NEW ZEALAND – STATISTICS NEW ZEALAND
NORWAY – STATISTICS NORWAY
PORTUGAL –INSTITUTO NACIONAL DE ESTADÍSTICA (INE)
SWEDEN – STATISTISKA CENTRALBYRÅN

5 Universities, Research Centers, Private Statistics Firms

USA – CENTERS FOR DISEASE CONTROL AND PREVENTION
USA – RESEARCH TRIANGLE INSTITUTE
USA – SURVEY RESEARCH CENTER, UNIVERSITY OF MICHIGAN
USA – U.S. DEPARTMENT OF AGRICULTURE
USA – WESTAT

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 Membership Officer
Margaret de Ruiter-Molloy
International Statistical Institute
P.O. Box 24070, 2490 AB The Hague,
The Netherlands

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** ☐

Read
The Survey Statistician
online!

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