

# the Survey Statistician

The Newsletter of the International Association of Survey Statisticians

No. 74

July 2016







#### **Editors**

**Natalie Shlomo and Eric Rancourt**

#### **Section Editors**

**Pierre Lavallée** - Country Reports

**Ken Copeland** - Ask the Experts

**Denise Silva** - New and Emerging Methods

#### **Circulation /Production**

**Lori Young**

**Courtney Williamson**

**Olivier Dupriez**

*The Survey Statistician is published twice a year by the International Association of Survey Statisticians and distributed to all its members. The Survey Statistician is also available on the IASS website at <http://isi.cbs.nl/iass/alluk.htm>.*

Enquiries for membership in the Association or change of address for current members should be addressed to:

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

Comments on the contents or suggestions for articles in *The Survey Statistician* should be sent via e-mail to the editors, Natalie Shlomo ([Natalie.Shlomo@manchester.ac.uk](mailto:Natalie.Shlomo@manchester.ac.uk)) or Eric Rancourt ([Eric.Rancourt@Canada.ca](mailto:Eric.Rancourt@Canada.ca)).

**ISSN 2521-991X**

#### **In this Issue:**

- 3 Letter from the Editors**
- 5 Letter from the President**
- 7 Report from the Scientific Secretary**
- 9 News and Announcements**
- 13 Ask the Experts**  
Why the resurgence in interest in address-based sampling? By Rachel Harter
- 15 New and Emerging Methods**  
Indirect Sampling for Hard-to-Reach Populations by Pierre Lavallée
- 25 Country Reports**  
Algeria  
Argentina  
Canada  
Latvia  
New Zealand
- 32 Upcoming Conferences and Workshops**
- 50 In Other Journals**
- 65 Welcome New Members**
- 66 IASS Officers and Council Members**
- 67 Institutional Members**
- 68 Change of Address Form**





## Letter from the Editors

The July 2016 issue contains articles of interest and important information regarding upcoming conferences, contents of relevant journals, updates from the IASS Executive and more. We hope you enjoy this issue, and we would be happy to receive your feedback and comments on how we can make improvements.

In the *New and Emerging Methods* Section (edited by the Scientific Secretary Denise Silva), Pierre Lavallée from Statistics Canada has contributed an article titled: Indirect Sampling for Hard-to-Reach Populations. In the article, Pierre summarises the state-of-the-art of indirect sampling which together with the generalised weight share method for estimation, provides a unified mathematical framework for sampling hard-to-reach population.

In the *Ask the Experts* Section (edited by Ken Copeland), Rachel Harter from RTI International, has provided a response to the question: Why the resurgence in interest in address-based sampling? The author discusses the challenges of using address lists as sampling frames.

Please contact Denise Silva ([denise.silva@ibge.gov.br](mailto:denise.silva@ibge.gov.br)) if you would like to contribute an article to the *New and Emerging Methods* Section. If you have any questions which you would like to be answered by an expert, please send them to Ken Copeland ([copeland-kennon@norc.org](mailto:copeland-kennon@norc.org)). If you are interested in writing a book or software review, please contact Natalie Shlomo ([Natalie.Shlomo@manchester.ac.uk](mailto:Natalie.Shlomo@manchester.ac.uk)). Finally, if you would like to contribute brief articles or editorials to the newsletter, please send them directly to the editors of the newsletter, Eric Rancourt and Natalie Shlomo.

The *Country Report* Section has always been a central feature of the IASS *The Survey Statistician* and we thank all the country representatives for their contribution and coordination of the reports. We also thank the editor of the section, Pierre Lavallée ([Pierre.Lavallee@Canada.ca](mailto:Pierre.Lavallee@Canada.ca)) for his continuing efforts to obtain timely reports from the different countries. We ask all country representatives to please share information on your country's current activities, applications, research and developments in survey methods.

This issue of *The Survey Statistician* includes a letter and updates from our IASS President, Steven Heeringa and from our Scientific Secretary, Denise Silva. In addition, in the *News and Announcement* section we have highlights from the recent ICES-V conference that was held in Geneva in June 2016. In addition, we congratulate Jack Gambino who is the newly elected President of the Canadian Statistical Society. Finally, we present new ethical guidelines from the American Statistical Association.

We thank Lori Young from Statistics Canada for collating the advertisements of upcoming conferences and for preparing the tables of contents in the *In Other Journals* section. This is a very time-consuming and detailed task but the information she gathers is deeply appreciated by the members. We also thank Lori for her hard work in collating all the articles into this newsletter that you see before you.



Please take an active role in supporting the IASS newsletter by volunteering to contribute articles, book/software reviews and country reports and/or by making it known to friends and colleagues. We also ask IASS members to send in notifications about conferences and other important news items about their organisations or individual members.

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

Eric Rancourt [Eric.Rancourt@Canada.ca](mailto:Eric.Rancourt@Canada.ca)

Natalie Shlomo [Natalie.Shlomo@manchester](mailto:Natalie.Shlomo@manchester)



## INTERNATIONAL ASSOCIATION OF SURVEY STATISTICIANS (IASS)

Dear IASS colleagues,

Summer has arrived in the northern hemisphere and the Copa America and Euro 2016 are history (congratulations Chile and Portugal). In Brazil, our colleagues in Rio who helped to host the 2015 World Statistics Congress are awaiting the start of the 2016 "Summer" Olympics. Here in Ann Arbor, we are wrapping up our 69<sup>th</sup> Summer Institute in Survey Research Techniques. Regardless of where you are in the world, I hope that you will have the time to relax and read this issue of *The Survey Statistician* including the Report of our Scientific Secretary (Denise Silva), the reports of the Country Representatives and all the other regular content and special articles that the editors, Natalie Shlomo and Eric Rancourt, have worked very hard to assemble.

The 2017 World Statistics Congress will be held July 16-21, 2017 in Marrakech, Morocco and I expect the IASS and its members will be well-represented as participants in the Scientific Program and as instructors in the ISI Short Course Program. Marcel Viera ([marcel.viera@ice.ufjf.br](mailto:marcel.viera@ice.ufjf.br)) is serving as the IASS representative to the WSC 2017 Scientific Program Committee (SPC) and over the past ten months has worked tirelessly in this role on behalf of the IASS. The SPC has finalized its selections for the Invited Paper Sessions (IPS) and as in the past IASS members who attend the 2017 WSC can expect many invited paper sessions on topics of relevance to survey statisticians, methodologists and practitioners. The submission period for WSC 2017 Special Topic Sessions (STS) is currently open but final session proposals must be submitted to the Scientific Program Committee by August 15. If you plan to submit an abstract for a Contributed Paper Session (CPS), the submission window will open September 15, 2016 and close February 15, 2017. More information on the Scientific Program for the 2017 WSC and regular updates as the meeting dates draw closer can be found at the official web site, <http://www.isi2017.org/>

In late January, ISI President, Pedro Silva, formed the WSC 2017 Short Course Committee. Ralf Muennich ([muennich@uni-trier.de](mailto:muennich@uni-trier.de)) generously agreed to serve as the IASS representative to the ISI Committee that will review and rate short course proposals and make the final selection of program offerings. The short courses will be offered immediately prior to the WSC 2017 (July 13-15, 2017). The call for short course proposals was distributed by the ISI permanent office on June 2, 2016. The deadline for the submission of proposals is August 4. Since the mid-1970s, IASS members have been instrumental in the development and instruction of the biennial short course programs. I encourage all of you to consider this opportunity to assist in our training mission through offering a one or two day course on a topic in survey design, methodology or statistics. For details on how and where to submit a short course proposal, please see the ISI announcement or feel free to e-mail Ralf or the Scientific Secretary, Denise Silva ([denisebritz@gmail.com](mailto:denisebritz@gmail.com)).

Prior to the WSC 2017, the IASS 2017 Cochran-Hansen Prize (<http://isi-iass.org/home/cochran-hansen-prize/>) will again be awarded to a young



statistician from a developing or transition country. This award is based on a paper competition that will be refereed by a committee of IASS members appointed by the IASS officers. IASS Vice-President, Monica Pratesi, has agreed to Chair the 2017 Cochran-Hansen Prize. If you are considering competing for the C-H Prize or are an academic advisor, professional supervisor or mentor to a young statistician that you believe should submit their written work to this competition, I encourage you start planning the submission now. Additional details on the C-H prize competition will be circulated early this fall. If you are planning to compete for the Cochran-Hansen Prize, please anticipate that papers will need to be submitted to the committee between December 1, 2016 and February 15, 2017.

Outside the biennial meetings of the World Statistics Congress, the IASS has rarely played the primary role as organizer of a major conference devoted to survey methodology and statistics. In large part, we are fortunate each year to have access to a large and diverse set of conferences, symposia, workshops and training activities devoted to survey methodology and survey statistics. The IASS has chosen each year to provide modest financial support to approximately 4 to 5 regional and international meetings that are specifically focused on survey statistics and survey methodology. In addition, in early January of this year, the IASS officers submitted an application to the ISI World Bank Trust Fund for Statistical Capacity Building and IASS and as a result was awarded a grant to support travel and expenses for three young statisticians from developing countries to attend the 9<sup>th</sup> French Colloquium on Survey Sampling, <http://isi-iass.org/home/events/9th-french-colloquium-on-survey-sampling/>. The Report of the IASS Scientific Secretary included in this issue of TSS provides more specific information on this special initiative and the 2016 conferences and meetings that the officers have selected to receive IASS support. If you are organizing a conference and would like to submit a request for IASS financial support please feel free to submit the proposal to the Scientific Secretary ([denisebritz@gmail.com](mailto:denisebritz@gmail.com)) or to me ([sheering@umich.edu](mailto:sheering@umich.edu)).

I will close this letter with an apology and a suggestion to our members.

The apology is for not making the progress that I promised (January 2016) toward completing a draft of a Strategic Plan for our Association. The IASS Strategic Planning committee will return to this important initiative early this fall and we will keep you informed as we reach milestones in revising the draft document and are able to share it with you for comment.

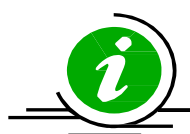
The suggestion relates to one key element of our strategic planning—engaging young survey statisticians and methodologists and encouraging them to join the global network that IASS membership should and does provide. There are a number of approaches that IASS can take toward rejuvenation of its membership. One approach that I know can be effective is for senior, established members of the IASS network to sponsor the membership of students and more junior statisticians, encouraging them to apply for IASS membership and paying their initial year's membership dues. Leslie Kish, my mentor when I came to this profession in 1975, insisted that I join IASS as a student. Not all of us can be as persuasive as Leslie was.



But if each of us committed annually to introduce a student or junior colleague in our work place to the IASS (and in cases of financial need paying their membership fee) we would quickly begin to bring much-needed “youth” to our international association. As always, I would appreciate hearing your personal thoughts on this suggestion or, more generally, on ideas that you may have to promote our membership and improve our ability to meet the collective aims of our organization.

Yours in the science and practice of survey research,

Steve Heeringa  
IASS President



## Report from the Scientific Secretary

---

---

During this first semester of the year, IASS has been involved with preparations for the 2017 World Statistics Congress (WSC). The 2017 WSC Scientific Programme Committee (SPC) met in Marrakesh last May to take decisions about the Invited Paper Session proposals. We are pleased to report that 16 out of 30 sessions submitted by the IASS community were accepted, and others were recommended as Special Topic Sessions. Some sessions were endorsed by more than one association, particularly by both IASS and IAOS. The 2017 WSC programme will also include two lectures: one for the IASS Cochran-Hansen Award and the IASS President's Invited Lecture.

We like to thank the IASS members for making an active contribution to the development of the WSC scientific programme. Although no IPS will have an association label for the WSC, it is good to know that these invited paper sessions will constitute great opportunities to place survey statistics subjects in the spotlight of the 2017 congress. We now call the IASS colleagues to promote 2017 WSC, submitting proposals for short courses and special topic sessions. There is still time to do it. Please check the congress website <http://www.isi2017.org/>.

IASS has co-sponsored the Fifth International Conference on Establishment Surveys (ICES V) that was held for the first time in Europe, from 20-23 June 2016 in Geneva, Switzerland. As reported by Polly Phipps and Paul Smith, “the conference programme included two keynote addresses, 16 invited sessions, 12 topic contributed sessions, and 20 contributed sessions, with a total of 189 presentations.

In addition, the conference held a student contest, six short courses, seven introductory overview lectures, and four software demonstration sessions. ICES-V was well attended, with 324 persons attending from 43 countries, including many IASS members who took advantage of IASS sponsorship and received a discount on registration”. The conference website is still available with more information about the event.



As a sponsor of the 9<sup>th</sup> French Colloquium in Survey Sampling and with the assistance of the ISI World Bank Trust Fund for Statistical Capacity Building (WBTFSCB), IASS has set up a competitive award to financially support young statisticians from developing countries to cover expenses to attend the event. Based on 9 submissions, the conference scientific committee has selected three candidates (two young statisticians from Burkina Faso and one from the Republic of Cameroon). The Colloquium will take place in Canada, on 11 - 14 October 2016. For more information, have a look at (<http://sondages2016.sfds.asso.fr/>). The awardees and their papers are:

- Daouda Ilboudo (Burkina Faso) - Evaluation de l'impact d'un programme/projet sur les performances des entreprises d'une localité: plan de sondage, mesure et significativité de l'impact (Evaluation of the impact of a program/project on the performances of enterprises of a village: sample design, measurement and significativeness of its impact);
- Arouna Sow (Burkina Faso) - Estimation provinciale du taux de chômage au Burkina Faso: une illustration de l'application des techniques d'estimation sur petits domaines (Provincial estimation of the unemployment rate in Burkina Faso: an example of the application of small area estimation techniques);
- Belmondo Voufo Tanakem (Cameroon) - Estimation sur petit domaine: application à la construction de la carte de pauvreté du Cameroun (Small Area estimation: application to the construction of the poverty map of Cameroon).

In addition, IASS is also committed to support the Baltic-Nordic-Ukrainian (BNU) Network Summer School on Survey Statistics that will take place on 22-26 August 2016 in Kyiv, Ukraine. (<http://wiki.helsinki.fi/display/BNU/Events>).

I take this opportunity to inform that the IASS President represented ISI at the June 25-26 ESRI National GIS Executive Forum (<http://www.esri.com/events/ngef>). The Forum was attended by leaders and senior staff of the geography and geo-spatial statistical division of around 120 countries. The discussions were focused on the application of GIS systems and geo-spatial statistics as well as their integration into national statistical systems.

All members are invited to contribute to our newsletter by volunteering to send articles, book/software reviews and country reports to the next issue of *The Survey Statistician* to be published in January 2017. If you would like to suggest a subject or write an article to the *New and Emerging Methods* section, please get in touch with me at [denisebritz@gmail.com](mailto:denisebritz@gmail.com) or [denise.silva@ibge.gov.br](mailto:denise.silva@ibge.gov.br).

Very best wishes

Denise Silva



---

---



## News and Announcements

---

---

### ICES-V Highlights

By

**Polly Phipps (Bureau of Labor Statistics) and Paul Smith (University of Southampton)**

The Fifth International Conference on Establishment Surveys (ICES-V) was held for the first time in Europe, from 20-23 June 2016 in Geneva, Switzerland. The programme included two keynote addresses, 16 invited sessions, 12 topic contributed sessions, and 20 contributed sessions, with a total of 189 presentations. In addition, the conference held a student contest, six short courses, seven introductory overview lectures, and four software demonstration sessions. ICES-V was well attended, with 324 persons attending from 43 countries, including many IASS members who took advantage of IASS sponsorship and received a discount on registration.

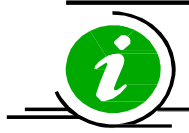
With so many sessions, it is difficult to summarize the programme, but here are a few highlights. The two keynote speakers, Peter van de Ven of the Organisation for Economic Co-operation and Development and Connie Citro of the U.S. National Academies of Sciences, Engineering and Medicine, spoke on globalisation, defining the statistical unit of measurement, the integration of administrative and survey data, and the use of multiple data sources – themes repeated in many of the conference sessions. As establishment survey and statistical measurement moves into the future, the trend towards statistics on an international level led to discussions on statistical units, the structure of complex units, integrating statistics on a global scale, and breaking down stovepipes within and across national statistical institutes, countries, and continents. Another issue of great relevance covered multi-source and mixed-mode designs, including moving from fixed samples to efficient designs to obtain a representative response distribution, and targeting data collection efforts to strata that need more attention – with the added complexity of changing and heterogeneous business populations that can be difficult to reach. Respondent burden continued to be a key area of concern at ICES-V – covering a wide range of topics -- the use of multiple data sources, the need for more auxiliary variables, imputation and estimation issues, as well as a focus on respondent relations and experiences.



The ICES-V introductory overview lectures complemented conference discussions with speakers on adaptive design, big data, administrative data, imputation, disclosure, and economic classification. Six short courses were held and were well attended, covering sampling, business survey data collection, record linkage, and estimation. Another highlight was the 15 software demonstrations, where participants could sit down with developers and view and discuss their work. They covered a wide range of common tasks, including sampling, questionnaire design, data collection and management, processing, estimation, modeling, and geographic information systems. Importantly, thinking about the next generation of ICES contributors, the Programme Committee developed and carried out a student contest to help create interest in establishment survey work and test technical skills and creativity.

The benefits of ICES include opportunities to catch up with colleagues, make new contacts, and learn about and discuss current topics and issues in the field of establishment surveys and statistics. New and experimental ideas were presented, as well as further development and progress on topics presented in past conferences. We encourage you to attend ICES-VI, the location and dates are not yet certain, but the conference is likely to be held in 2020. We thank the many sponsors for their support, as well as the Swiss Federal Statistical Office for hosting the conference.





## **Revising the American Statistical Association's Ethical Guidelines for Statistical Practice**

---

---

**By**

**Howard Hogan, Chair, ASA Committee on Professional Ethics**

[This note reports the work of the ASA Committee on Professional Ethics. Any views expressed are those of the Author and not necessarily those of the U.S. Census Bureau.]

In 1981, the ASA Board of Directors approved its first Code of Conduct. A few years later, the Committee on Professional Ethics was formed. Within its duties is to “maintain and promulgate, subject to Board review and approval, the set of ASA Ethical Guidelines that describes the general view of ethics in statistical practice....”

The most recent Guidelines had been approved by the ASA Board in 1999.

A principal motivation for the revision was the passage of time; good practice required that the Guidelines be reviewed at least once a decade. Were there developments in statistics that might require changing the Guidelines? Many people mentioned “Big Data.” A second motivation was to make the Guidelines more accessible. Could we make the Guidelines more readable on Tablets and Smart-Phones, for example?

In terms of content, the Guidelines now cover:

- A. Professional Integrity and Accountability
- B. Integrity of data and methods
- C. Responsibilities to Science, Public and Client
- D. Responsibilities to Research Subjects
- E. Responsibilities to Research Team Colleagues
- F. Responsibilities to Other Statisticians
- G. Responsibilities Regarding Allegations of Misconduct
- H. Responsibilities of Employers

They are guided by the following philosophy:

Good statistical practice is fundamentally based on transparent assumptions, reproducible results, and valid interpretations. Above all, professionalism in statistical practice presumes the goal of advancing knowledge while avoiding harm; using statistics in pursuit of unethical ends is inherently unethical. The principles expressed here should guide both those whose primary occupation is statistics and those in all other disciplines who use statistical methods in their professional work.

“Big Data” is not explicitly addressed, but the Committee feels that the ethical questions are well covered by:

D 5: Considers whether appropriate research-subject approvals were obtained before participating in a study involving human beings or organizations, before analyzing data from such a study, and while reviewing manuscripts for publication or internal use.



The statistician considers the treatment of research subjects ... when evaluating the appropriateness of the data source(s).

D 6: In contemplating whether to participate in an analysis of data from a particular source, refuses to do so if participating in the analysis could reasonably be interpreted by individuals who provided information as sanctioning a violation of their rights.

In March, the Committee submitted its proposal to the ASA board, which adopted the Guidelines at its April meeting.

The revised Guidelines will soon be posted on the [amstat.org](http://amstat.org) website.

---

## **President of the Statistical Society of Canada: Jack Gambino**

**By**

**Eric Rancourt, Statistics Canada**

---

On July 1, 2016, Jack Gambino, from Statistics Canada, became president of the



Statistical Society of Canada (SSC). Jack is the sixth of an elite group of Statistics Canada employees to have held this position in the 38 years of the SSC. He received BSc and MSc degrees in mathematics and statistics from McGill University in the late 1970s, and went on to complete a PhD in statistics at the University of Toronto in 1982. He has worked as a survey statistician at Statistics Canada since then, gradually moving into managerial and executive positions. He is currently director of the Household Survey Methods Division, where he manages a large team responsible for the design, development, implementation, and maintenance of Statistics Canada's household surveys. His focus has been towards survey sampling theory and methods, but he always kept an interest in Bayesian inference. Jack

has been involved in the SSC in various capacities over the years, including local representative and associate editor of *Liaison*. He has also been active in the Ottawa chapter, serving as president for one year. He has been an associate editor of the journal *Survey Methodology* for many years. Jack is an elected member of the ISI and a member of the IASS.





## Ask the Experts

---

---

### **“Why the resurgence in interest in address-based sampling?”**

**By**

**Rachel Harter, RTI International**

Address-based sampling (ABS) is a valuable tool for survey designers who can take advantage of its strengths and adaptability. The survey industry in the U.S. is currently learning to access and use high quality address lists that were inaccessible to the public in the past. Address lists have become available through direct marketing companies at the same time that traditional ways of selecting households and collecting data have suffered from the challenges of declining response rates and rising costs. The time is right to explore the use of address lists as the basis for sampling frames to help solve these challenges.

Historically, there was no publicly available sampling frame of all U.S. households. While the U.S. Postal Service (USPS) does not make its database of mail addresses available to anyone except the U.S. Census Bureau, qualified direct marketing companies with their own mailing lists have taken advantage of USPS services to clean and update their lists. Thus mass mailers have the best mailing addresses, and the USPS operates more efficiently. Some direct marketing companies, in turn, lease copies of their address lists to organisations that can sample from them, either directly or through sampling vendors. This commercial availability of more than 145 million residential addresses has generated interest and created alternative ways of conducting population and household surveys. The address frames can be used to enhance various modes of data collection and survey designs. A few examples of the use of address frames illustrate their benefits.

First, in-person surveys have often relied on geographic multistage designs with field enumeration to build frames of housing units in selected geographic areas. The cost of building such frames is prohibitively expensive for most surveys. Address lists provide a viable, low-cost frame alternative in most areas of the country.

Second, as more and more households give up their landline telephones in favor of mobile phones, the coverage of landline-based random-digit dial (RDD) surveys has suffered. Mobile phones in the U.S. may be dialed for surveys, but by law an auto-dialer may not be used, increasing the costs of telephone surveys. For surveys that must be conducted by phone, a dual-frame (landline and cell) RDD sample may be best; however, mail surveys hold great promise as an alternative to RDD. With a well-



designed questionnaire and protocol that reduces the need for interviewers, response rates by mail from a sample of addresses can rival or exceed those of RDD surveys, at far lower cost.

Third, address lists provide a relatively inexpensive way to contact households when the web is the preferred mode for survey responses. Contacting households by mail may be the starting point for a mixed mode survey.

Fourth, some organisations use address lists for two-phase sampling, using the first phase results to stratify, screen, and/or subsample for the second phase.

One advantage of address lists is the ability to determine the locations of the homes associated with the addresses. In general, having address locations enables much more precise geographic targeting for local surveys than RDD (especially cell phones) can offer. In addition, assigning addresses to census geographies enables area-level data from the decennial census or American Community Survey to be appended to the frame for improved stratification, weighting, or estimation. Other auxiliary data, matched on addresses, can be appended, as well.

Address lists have some coverage issues. Many vacant homes and other addresses not receiving mail are not on the USPS primary mail delivery file, and may or may not be on a vendor's file. Some addresses cannot be geocoded, which poses a coverage problem for in-person interviewing and for local surveys where the ability to identify the household's location is key. For example, some households have mail delivered to a post office box rather than their home. Some addresses are shared by multiple households, and some households have multiple addresses. Nevertheless, coverage of U.S. households by address lists is considered to be high, especially in urban and suburban areas. Furthermore, coverage issues are not necessarily a problem unless there is coverage bias.

Survey designers considering address frames should expect a learning curve to understand the address file characteristics and to develop methodologies for reducing errors. After all, the address lists were developed for mail delivery, not for sampling frames. The industry has generated numerous publications as it learns how to use this valuable resource.

In the U.S., address lists are currently the best available source for sampling frames of housing units as proxies for households. Recently the American Association for Public Opinion Research (AAPOR) commissioned a panel of experts to compile an ABS report that was the source for this brief article. For more information, please see the report and references cited there:

American Association for Public Opinion Research (2016). *Address-based Sampling*. Report prepared for AAPOR Council by the Task Force on Address-based Sampling; R. Harter, Chair. Available at:

[http://www.aapor.org/AAPOR\\_Main/media/MainSiteFiles/AAPOR\\_Report\\_1\\_7\\_16\\_CLEAN-COPY-FINAL.pdf](http://www.aapor.org/AAPOR_Main/media/MainSiteFiles/AAPOR_Report_1_7_16_CLEAN-COPY-FINAL.pdf) .





# New and Emerging Methods

## Indirect Sampling for Hard-to-Reach Populations

By  
Pierre Lavallée<sup>1</sup>

**Note:** This paper is a condensed version of Chapter 21 of the book *Hard-to-Survey Populations* (Tourangeau, Edwads, Johnson, Wolter, Bates, Editors; Cambridge University Press, 2014).

### 1. Introduction

In survey sampling, some populations are hard to reach because they happen to be hard to survey. Their relative rareness and the absence of a suitable sampling frame are two main reasons for this. As mentioned by Kalton and Anderson (1986), “*The initial consideration in designing a sample for a rare population is whether there exists a separate frame for that population. If a separate frame exists, is available for sample selection and is deemed adequate, the sample may be selected from it using standard methods and no problems arise*”. When no sampling frame is available for the desired target population, one might then choose a sampling frame that is indirectly related to the targeted rare population. We can then speak of two populations  $U^A$  and  $U^B$  that are related to one another. We wish to produce an estimate for the population  $U^B$  by selecting a sample from the population  $U^A$  for which a sampling frame is available and using the existing links between the two populations. This sampling process is referred to as *Indirect Sampling* (Lavallée, 2002, 2007). Producing estimates in the context of Indirect Sampling can be difficult to achieve if the links between  $U^A$  and  $U^B$  are not one-to-one. A solution for this is to use the *Generalised Weight Share Method* (GWSM). The population  $U^B$  can be the hard-to-reach population itself or it can be a population that contains it as a subpopulation. Fortunately, for the statistician, it turns out that hard-to-reach populations can often be found by surveying clusters. This is the case, for example, with infectious diseases (Thompson, 1992). In this paper then, we will assume that the population  $U^B$  is partitioned into clusters. Selection of clusters will then be performed through the Indirect Sampling process. Since sampling clusters rather than individual units might allow tracking more easily units that are part of the hard-to-reach populations, we can foresee considerable reductions in costs since a large part of the costs are related to the identification of the hard-to-reach populations. As well, cluster sampling allows for the production of results at the cluster level itself, in addition to the units. Sampling hard-to-reach populations through selection of clusters can be illustrated by the following three examples that we will develop further in Section 3.

<sup>1</sup> Pierre Lavallée, Statistics Canada, Ottawa, Ontario, K1A 0T6, CANADA, pierre.lavallee@canada.ca.



*Example 1: Surveying a rare ethnic group.* Suppose that we want to survey a rare ethnic group within a city. The members of this ethnic group might be living in small communities that can be viewed as clusters. By surveying clusters, entire communities are surveyed. Therefore, the yield of the sample in terms of people from the targeted ethnic group is likely to be larger than by surveying individuals directly, i.e., ignoring the clustered aspect of the target population.

*Example 2: Surveying smokers.* Suppose that we want to measure the health issues of smokers in a given country. Smokers can be viewed as a hard-to-reach population because there is no list available of smokers. Sometimes such a list is available (for example, from local health centres because of cancer diagnostics) but it offers only a partial coverage of smokers. If no list is available, sample selection of smokers can then be done through families (including, for example, aunts, cousins, etc.), which can be viewed as clusters. By sampling entire families, rather than individuals, we are more likely to find some smokers.

*Example 3: Surveying homeless people.* Suppose that we want to measure health and living conditions of homeless people in a big city. Ardilly and LeBlanc (2001) proceeded by reaching the homeless persons through services that are provided to them: overnight accommodation and meals. These two types of services can be seen as being part of two different sampling frames. During a certain period of time, one can assume that each homeless person will receive a given set of services. By sampling services, we get indirectly a sample of homeless persons to be surveyed.

Using Indirect Sampling as briefly described above is one way for surveying hard-to-reach populations. Other approaches exist, such as *Network Sampling*, *Adaptive Cluster Sampling*, *Snowball Sampling*, *Respondent-Driven Sampling*, and the use of *multiple frames*. With appropriate modifications, these approaches can all be put into the context of Indirect Sampling (Lavallée, 2014). One can use the theory and developments surrounding Indirect Sampling and the GWSM to obtain a unified mathematical framework for the above approaches. See Lavallée (2002, 2007) for further details on Indirect Sampling and the GWSM.

## 2. Indirect Sampling

With Indirect Sampling, we wish to produce an estimate for a population  $U^B$  using a sample selected from the sampling frame  $U^A$  and the existing links between the two populations. We proceed by selecting a sample  $s^A$  containing  $m^A$  units from the population  $U^A$  containing  $M^A$  units according to some sample design chosen by the survey statistician. Let  $\pi_j^A$  represent the selection probability of unit  $j$  in  $U^A$ . We assume that  $\pi_j^A > 0$  for all  $j \in U^A$ . On the other hand, the population  $U^B$  contains  $M^B$  units. It is divided into  $N^B$  clusters, where cluster  $i$  contains  $M_i^B$  units.

We assume there exists a relationship between a unit  $j$  of population  $U^A$  and some unit  $k$  of cluster  $i$  in  $U^B$ , the latter noted as unit  $ik$ . This relationship is described by an indicator variable  $l_{j,ik}$ , where  $l_{j,ik} = 1$  if there is a link between unit  $j \in U^A$  and unit  $ik \in U^B$ , and 0 otherwise. Note that there might be some cases where no link exists



between a unit  $j$  of population  $U^A$  and any unit  $k$  from cluster  $i$  of population  $U^B$ , which amounts to  $L_j^A = \sum_{i=1}^{N^B} \sum_{k=1}^{M_i^B} l_{j,ik} = 0$ . In such a case, unit  $j$  is of no use to help find units of  $U^B$ , which compromises the efficiency of the sampling strategy but does not introduce any bias.

To produce unbiased estimates, each unit  $k$  of the population  $U^B$  must have a non-zero probability of being surveyed through  $U^A$ . This translates into the following constraint: Each cluster  $i$  of  $U^B$  must have at least one link with a unit  $j$  of  $U^A$ , i.e.,  $L_i^B = \sum_{k=1}^{M_i^B} \sum_{j=1}^{M^A} l_{j,ik} > 0$ .

The Indirect Sampling process is as follows:

1. We select a sample  $s^A$  of  $m^A$  units from the frame  $U^A$ .
2. For each unit  $j$  selected in  $s^A$ , we identify the units  $ik$  of  $U^B$  that are linked to  $j$ , i.e.,  $l_{j,ik} = 1$ .
3. For each unit  $ik$  identified, we assume that we can set up the list of  $M_i^B$  units of cluster  $i$  containing this unit. Each cluster  $i$  then represents, within itself, a population  $U_i^B$  where
4.  $U^B = \bigcup_{i=1}^{N^B} U_i^B$ . Let  $s^B$  be the set of  $n^B$  clusters identified by the units  $j \in s^A$ , i.e.,  $s^B = \{i \in U^B \mid \exists j \in s^A \text{ and } L_{j,i} > 0\}$  where  $L_{j,i} = \sum_{k=1}^{M_i^B} l_{j,ik}$ .
5. We survey **all** units  $k$  of the cluster  $i \in s^B$  and measure a variable of interest  $y_{ik}$  as well as the number of links  $L_{ik}^B = \sum_{j=1}^{M^A} l_{j,ik}$  between unit  $ik$  of  $U^B$  and the frame  $U^A$ .

The last step is an important aspect to which the survey process (or measurement) is subjected. If a unit is selected in the sample then every unit of the cluster containing the selected unit will be surveyed.

For the population  $U^B$ , we seek to estimate the total  $Y^B = \sum_{i=1}^{N^B} \sum_{k=1}^{M_i^B} y_{ik}$ . The estimation of the total  $Y^B$  indirectly through the sample  $s^A$  selected from the frame  $U^A$  can be a major challenge, in particular if the links between the units of  $U^A$  and  $U^B$  are not one-to-one. Indeed, with multiple ways to reach a given unit of  $U^B$  comes an issue of double counting, which goes unaddressed if inferences are made using the inclusion probabilities or weights arising from  $U^A$ . A correct procedure must somehow factor in the multiple entry points into  $U^B$ . A solution to this estimation problem is the use of the GWSM (Lavallée, 1995, 2002, 2007).

## 2.1 The Generalised Weight Share Method (GWSM)

The GWSM produces an estimation weight for each surveyed unit from the population  $U^B$  by transforming the sampling weights inherited from  $U^A$ . This estimation weight is basically an average of the sampling weights of the selected units from population  $U^A$ . Lavallée (1995) presented for the first time the GWSM within the context of the problem of cross-sectional weighting for longitudinal household surveys. The GWSM is a



generalisation of the weight share method described by Ernst (1989).

By using the GWSM, we want to assign an *estimation weight*  $w_{ik}$  to each unit  $k$  of a surveyed cluster  $i$ . To estimate the total  $Y^B$  belonging to the population  $U^B$ , we can then use the estimator

$$\hat{Y}^B = \sum_{i=1}^{n^B} \sum_{k=1}^{M_i^B} w_{ik} y_{ik} \quad (1)$$

where  $n^B$  is the number of surveyed clusters and  $w_{ik}$ , the weight assigned to unit  $k$  of cluster  $i$ . With the GWSM, the estimation method is based on the sample  $s^A$ , together with the existing links between  $U^A$  and  $U^B$ . The links are in fact used as a bridge to go between the populations  $U^A$  and  $U^B$ .

The steps of the GWSM are as follows:

1. For each unit  $k$  of cluster  $i$  of  $s^B$ , we calculate an initial weight  $w'_{ik}$ , as follows:

$$w'_{ik} = \sum_{j=1}^{M^A} l_{j,ik} \frac{t_j}{\pi_j^A} \quad (2)$$

where  $t_j = 1$  if  $j \in s^A$ , and 0 otherwise. In this step, we assess the amount of over-representation for  $ik$  in the sample drawn  $s^B$  due to the multiple links leading to it from  $U^A$ . Note that a unit  $ik$  having no link with any unit  $j$  of  $U^A$  automatically has an initial weight of zero. Observe also that a unit  $ik$  exhibiting just one link with units in  $U^A$  gets an initial weight equal to its design weight.

2. For each unit  $k$  of cluster  $i$  of  $s^B$ , we get the total number of links  $L_{ik}^B$ :

$$L_{ik}^B = \sum_{j=1}^{M^A} l_{j,ik} \quad (3)$$

We assess here the maximum possible number of contributors to the over-representation of  $ik$ .

3. We calculate the final weight  $w_i$ :

$$w_i = \frac{1}{L_i^B} \sum_{k=1}^{M_i^B} w'_{ik} \quad (4)$$

where  $L_i^B = \sum_{k=1}^{M_i^B} L_{ik}^B$ .

4. We assign  $w_{ik} = w_i$  for all  $k \in U_i^B$ .

Putting all together, we have for a unit  $ik$  of the population  $U^B$  that

$$w_{ik} = \sum_{j=1}^{M^A} \frac{t_j}{\pi_j^A} \frac{L_{j,i}}{L_i^B} \quad (5)$$

To estimate the total  $Y^B$ , we use equation (1). Because the estimation weights coming from the GWSM are the same for the set of  $M_i^B$  units of each cluster  $i$ , the estimator (1) can be written as a function of clusters only. Thus, we have

$$\hat{Y}^B = \sum_{i=1}^{n^B} \sum_{k=1}^{M_i^B} w_{ik} y_{ik} = \sum_{i=1}^{n^B} w_i \sum_{k=1}^{M_i^B} y_{ik} = \sum_{i=1}^{n^B} w_i Y_i \quad (6)$$



The estimator  $\hat{Y}^B$ , given by (1), can then also be written as

$$\hat{Y}^B = \sum_{j=1}^{M^A} \frac{t_j}{\pi_j^A} Z_j \quad (7)$$

where  $Z_j = \sum_{i=1}^{N^B} Y_i L_{j,i} / L_i^B$ . Equations (6) and (7) reveal the advantage there is in addressing duplication through the weights rather than inclusion probabilities: the linearity of the transformation used in the GWSM results in an estimator that can just as easily be interpreted under  $U^A$  than under  $U^B$ . We can switch from one form to the other as we see fit. It turns out that form (6) is natural for point estimation and (7) better suited for variance estimation.

The GWSM offers some clear advantages over other estimation methods. First, the GWSM exploits selection probabilities  $\pi_j^A$  only for the selected units  $j$  in the sample  $s^A$ . This is a major simplification compared to other weighting methods such as the one based on the exact calculation of selection probabilities of surveyed units. Second, in a given context, the links between the population  $U^A$  from which the sample is selected and the population  $U^B$  can be very complex and yet, the GWSM can be applied relatively easily. Third, since we are surveying the full set of units from the selected clusters, it can happen that we must calculate an estimation weight for a unit of  $U^B$  that is surveyed, but that has no link with population  $U^A$  from which the sample is selected. The problem is then to get an estimation weight for these units so that we can produce unbiased estimates. The GWSM allows for the calculation of estimation weights for these units. Finally, in simple problems related to classical sampling (or direct sampling) theory where no duplication exists, the GWSM gives generally the same results as the classical theory.

### 3. Indirect Sampling for Hard-to-Reach Populations

Suppose that it is desired to produce estimates for a hard-to-reach population  $U_d^B$  that is contained in a larger population  $U^B$ , i.e.  $U_d^B \subseteq U^B$ . This target population can be, for example, a rare ethnic group (Example 1 above), smokers (Example 2), or homeless people (Example 3). Let this target population  $U_d^B$  contain  $M_d^B$  units. Note that the quantity  $M_d^B$  is often unknown itself or it can even be a parameter of interest to be estimated. As above, the population  $U^B$  is divided into  $N^B$  clusters, where cluster  $i$  contains  $M_i^B$  units.

Although, by definition, we do not have a sampling frame for the hard-to-survey population  $U_d^B$ , we might have access to a frame  $U^A$  that is at least related to population  $U^B$ . Contrary to populations  $U_d^B$  and  $U^B$ , we assume that we have some information on each of the  $M^A$  units of  $U^A$ . Considering Example 1, we might decide to survey the desired ethnic group by simply using an area frame. The frame  $U^A$  is then simply a set of addresses (cities, streets, etc.) from which we can select a sample  $s^A$  of  $m^A$  units with some sample design. Looking at Example 2, because no list of smokers is available,



sample selection of smokers can then be done through the individuals from families, which can be viewed as clusters. In this case, the frame  $U^A$  is one of individuals where, unfortunately, we are not able to know, without surveying, whether a given person is a smoker or not. Lastly, considering Example 3, by constructing a sampling frame  $U^A$  from the set of services (overnight accommodation and meals) provided to homeless people, this target group can then be reached.

An appropriate sampling frame  $U^A$  for the hard-to-survey population  $U_d^B$  should have well-defined links  $l_{j,ik}$  between the units  $j$  of population  $U^A$  and the units  $k$  of clusters  $i$  of the population  $U^B$ , if not  $U_d^B$  itself. If  $L_j^A = \sum_{i=1}^{N^B} \sum_{k=1}^{M_i^B} l_{j,ik} = 0$  for a unit  $j$  of  $s^A$ , then there are simply no units of  $U^B$  identified with this unit  $j$ . In Example, 3, this would mean that a given service (e.g., a meal) has not been delivered to any homeless person. Because of the selection process, if a unit is selected in the sample, then all units of the cluster containing the selected unit will be surveyed. For hard-to-reach populations, this turns out to be beneficial because it helps the identification of units of the hard-to-reach population. In example 2, surveying all units in selected clusters means that every family member of the selected person would be submitted to the search for smokers. We now describe in detail the three above examples in the context of Indirect Sampling with the application of the GWSM.

### 3.1 Example 1: Surveying a Rare Ethnic Group

Recall that we want to survey a rare ethnic group within a city. The members of this ethnic group are assumed to be living in households that are part of small communities, which can be viewed as clusters. Typical examples of small communities in cities are Chinatown and Little Italy, which can be viewed as clusters of their respective ethnic groups. For some other ethnic groups, the communities can be the “family” in the broad sense, which can be defined as grand-parents with their children, grand-children and spouses.

We are interested in households containing at least one member of the ethnic group of interest. For this survey, it is desirable to use an area frame. The frame  $U^A$  is then a set of addresses from which we can select a sample  $s^A$  of  $m^A$  addresses according to some sample design.

The population  $U^B$  corresponds to the complete population of the city, in which we have communities (clusters) that contain households (units). The hard-to-reach population  $U_d^B$  (i.e., the ethnic group of interest) can be viewed as a domain of interest  $d$  of  $U^B$ .

We seek to estimate the total  $Y_d^B = \sum_{i=1}^{N^B} \sum_{k=1}^{M_i^B} y_{d,ik}$ , where  $y_{d,ik} = y_{ik}$  if unit (household)  $k$  of cluster (community)  $i$  belongs to the domain  $d$ , and 0 otherwise.  $Y_d^B$  corresponds to the total of the variable  $y$  for the ethnic group of interest (domain  $d$ ). We can also write

$Y_d^B = \sum_{i=1}^{N^B} Y_{d,i}$  where  $Y_{d,i} = \sum_{k=1}^{M_i^B} y_{d,ik}$ . Note that because each community of domain  $d$  contains only households of the ethnic group of interest, we can write  $Y_d^B = \sum_{i=1}^{N_d^B} Y_i$  where  $N_d^B$  is the total number of communities of the ethnic group in the population.



Using Indirect Sampling, we proceed as follows: We select a sample  $s^A$  of addresses from the frame  $U^A$ . For each address  $j$  selected in  $s^A$ , we identify the households  $ik$  of  $U^B$  that have  $l_{j,ik} = 1$ , i.e., that are living at address  $j$ . For each household  $ik$  belonging to the ethnic group of interest, we survey **all** the  $M_i^B$  households of the related community  $i$ . We then measure the variable of interest  $y_{ik}$  and the number of links  $L_{ik}^B = \sum_{j=1}^{M^A} l_{j,ik}$  between each surveyed household  $ik$  of  $U^B$  and the frame of addresses  $U^A$ . Note that in general, a household belongs to a single address, which implies that  $L_{ik}^B = 1$ . We have  $L_{ik}^B = 0$  whenever the address is not part of the frame  $U^A$ . If the household  $ik$  identified by some address  $j$  in  $s^A$  does not belong to the ethnic group of interest, we stop surveying the cluster  $i$  to which this household belongs. Because this cluster  $i$  is not part of the ethnic group of interest  $U_d^B$ , there is no point in measuring the variable of interest  $y_{ik}$  and we then set  $y_{ik} = 0$  for all  $k \in i$ .

To estimate  $Y_d^B$ , we use the estimator

$$\hat{Y}_d^B = \sum_{i=1}^{n^B} \sum_{k=1}^{M_i^B} w_{ik} y_{d,ik} = \sum_{i=1}^{n^B} w_i Y_{d,i} \quad (8)$$

where the weights  $w_{ik}$  (or  $w_i$ ) obtained using the GWSM are given by (5). Using (8), we can then easily estimate totals for the ethnic group of interest.

### 3.2 Example 2: Surveying Smokers

We want to measure the health issues of smokers in a country. As mentioned earlier, smokers can be viewed as a hard-to-reach population because no list of smokers is available. However, sample selection of smokers can be done through families, which can be viewed as clusters.

The frame  $U^A$  can be a list of the persons living in the country. This list could be obtained, for example, from the last census or from administrative records. We select a sample  $s^A$  containing  $m^A$  persons from  $U^A$  containing  $M^A$  persons according to some sample design. The population  $U^B$  contains the same persons as in  $U^A$ , and thus,  $M^B = M^A$ . It is divided into  $N^B$  families (clusters), where family  $i$  contains  $M_i^B$  persons. Since populations  $U^A$  and  $U^B$  are the same, there is a one-to-one relationship between person  $j$  of population  $U^A$  and person  $k$  of clusters  $i$  of the population  $U^B$ , which is reflected into the indicator variable  $l_{j,ik}$ , where  $l_{j,ik} = 1$  if  $j$  and  $k$  refers to the same person, and 0 otherwise. This implies that the indices  $j$  and  $k$  are exchangeable and, because of this, only the index  $j$  will thereafter be used to identify the persons of  $U^A$  or  $U^B$ . Since the population  $U^B$  is the list of persons living in the country, the hard-to-reach population  $U_d^B$  (i.e., the smokers) can be viewed as a domain of interest  $d$ , where  $U_d^B \subseteq U^B$ .

To survey the population of smokers using Indirect Sampling, we can proceed as follows: We select a sample  $s^A$  of  $m^A$  persons (smoker or not) from the frame  $U^A$ . For each person  $j$  selected in  $s^A$ , we set up the list of the  $M_i^B$  members of its family (cluster)  $i$  and



identify **all** smokers in the family by interviewing them, for example. Let  $s^B$  be the set of  $n^B$  families identified by the persons of  $s^A$ . We survey **all** the smokers contained in each family  $i \in s^B$  and measure a variable of interest  $y_{ik}$ .

We want to estimate the total  $Y_d^B = \sum_{i=1}^{n^B} \sum_{j=1}^{M_i^B} y_{d,ij}$ , where  $y_{d,ij} = y_{ij}$  if person  $j$  of family  $i$  is a smoker, and 0 otherwise. We can also write  $Y_d^B = \sum_{i=1}^{n^B} Y_{d,i}$  where  $Y_{d,i} = \sum_{j=1}^{M_i^B} y_{d,ij}$ .

To estimate  $Y_d^B$ , we use the estimator

$$\hat{Y}_d^B = \sum_{i=1}^{n^B} \sum_{j=1}^{M_i^B} w_{ij} y_{d,ij} \quad (9)$$

where the weights  $w_{ij}$  (or  $w_i$  since  $w_{ij} = w_i$  for  $j \in i$ ) obtained using the GWSM are given by (5). Now, because of the one-to-one relationship between  $U^A$  and  $U^B$ , we have  $L_{j,i} = 1$  (i.e., each person belongs to only one family) and  $L_i^B = M_i^B$ . Therefore, equation (5) reduces to

$$w_i = \sum_{j=1}^{M^A} \frac{t_j}{\pi_j^A} \frac{1}{M_i^B} \quad (10)$$

for  $j \in i$ . Using (9) and (10), we can then easily estimate totals for smokers.

### 3.3 Example 3: Surveying Homeless People

We want to measure the health and living conditions of homeless people in a big city. Following Ardilly and LeBlanc (2001), we proceed by reaching homeless persons through services provided to them: overnight accommodations and meals. During a certain period of time, one can assume that each homeless person will receive a given set of services. By sampling services, we get ultimately a sample of homeless persons to be surveyed.

The frame  $U^A$  of  $M^A$  services (overnight accommodations and meals) is built by considering a given reference period, e.g., a particular week of the year. For the meals, we can consider all individual plates served at breakfast, lunch and dinner during the full week. For overnight accommodations, we can consider all beds offered in the social centres of the city during the same week. With Indirect Sampling, we then select a sample  $s^A$  of  $m^A$  services from  $U^A$ . The population  $U^B$  (which is the same as the target population) contains  $M^B$  homeless people. Contrary to Examples 1 and 2,  $U^B$  is not divided here into clusters, which is the same as saying that each cluster  $i$  contains  $M_i^B = 1$  homeless person. We can then omit index  $i$  for the rest of the section.

There is a relationship between the homeless person  $k$  of  $U^B$  and service  $j$  of  $U^A$  if this homeless person used this service (overnight accommodation or meal). That is,  $l_{j,k} = 1$  if the service  $j \in U^A$  has been used by the homeless person  $k \in U^B$ , and 0 otherwise. To produce unbiased estimates, each homeless person  $k$  of  $U^B$  must have used at least one service. This means that we must have  $L_k^B = \sum_{j=1}^{M^A} l_{j,k} > 0$ .



In practice, this might not be respected if some homeless people happen not to use any services provided by the city during the reference week chosen for sampling.

For homeless people, the Indirect Sampling process is as follows: We first select a sample  $s^A$  of  $m^A$  services from the frame  $U^A$ . For each service  $j$  selected in  $s^A$ , we identify the homeless person  $k$  of  $U^B$  that has used this service, i.e., with  $l_{j,k}=1$ . For each homeless person  $k$  identified, we assume that we can set up the list of all services used by this person. In practice, this can be done by asking the social centres if the surveyed homeless person  $k$  has used other services during the reference week.

Let  $s^B$  be the set of  $m^B$  homeless people identified by the units  $j \in s^A$ . For the (target) population  $U^B$ , we seek to estimate the total  $Y^B = \sum_{k=1}^{M^B} y_k$  using

$$\hat{Y}^B = \sum_{k=1}^{m^B} w_k y_k \quad (11)$$

From (5), we have

$$w_k = \sum_{j=1}^{M^A} \frac{t_j}{\pi_j^A} \frac{l_{j,k}}{L_k^B} \quad (12)$$

where  $L_k^B = \sum_{j=1}^{M^A} l_{j,k}$  is the total number of services used by homeless person  $j$  during the reference week. As we can see from (12), estimator (11) takes into account duplicate selections of homeless persons by dividing the weights  $w_k$  by the total number of services  $L_k^B$  used by the homeless persons  $k$ . That is, even if a given homeless person has been selected in  $s^B$  through more than one service, no overestimation occurs because the weights (12) take into account this multiplicity. Note that this holds in general with the weights (5) obtained from the GWSM. Using (11) and (12), we can then easily estimate totals for homeless people.

#### 4. Conclusion

In this paper, we looked at the possibility of using Indirect sampling for hard-to-reach populations. We exploited the fact that hard-to-reach populations are often found in clusters. The selection of clusters is then performed through the Indirect Sampling process, and the estimates are produced by using the GWSM. By sampling clusters rather than individual units, we can track more easily units that are part of the hard-to-reach populations.

Approaches other than Indirect Sampling exist for surveying hard-to-reach populations. Network Sampling, Adaptive Cluster Sampling, Snowball Sampling, Respondent-Driven Sampling and the use of multiple frames are a subset of such methods. As mentioned by Lavallée (2014), these approaches can be put into the context of Indirect Sampling. One can then use the theory and developments surrounding Indirect Sampling and the GWSM to obtain a unified mathematical framework for the above approaches. For instance, calibration, non-response adjustment and the use of weighted links (in particular, the ones that minimise the variance of estimates) can all be used for improving the efficiency of the above methods, as described in Lavallée (2002, 2007).



## Bibliography

Ardilly, P., & LeBlanc, D. (2001). Sampling and weighting a survey of homeless persons: A French example. *Survey Methodology*, Vol. 27, No. 1, June 2001, pp. 109-118.

Ernst, L. (1989). Weighting issues for longitudinal household and family estimates. In Kasprzyk, D., Duncan, G., Kalton, G., Singh, M.P. (Editors), *Panel Surveys*, John Wiley and Sons, New York, pp. 135-159.

Kalton, G., & Anderson, D.W. (1986). Sampling rare populations. *Journal of the Royal Statistical Society, A*, Vol. 149, Part 1, pp. 65-82.

Lavallée, P. (1995). Cross-sectional weighting of longitudinal surveys of individuals and households using the weight share method. *Survey Methodology*, Vol. 21, No. 1, pp. 25-32.

Lavallée, P. (2002). *Le sondage indirect, ou la méthode généralisée du partage des poids*. Éditions de l'Université de Bruxelles (Belgique) et Éditions Ellipses (France), 215 pages.

Lavallée, P. (2007). *Indirect sampling*. Springer, New York, 245 pages.

Lavallée, P. (2014). Indirect sampling for hard-to-reach populations. In *Hard-to-Survey Populations* (Tourangeau, Edwards, Johnson, Wolter, Bates, Editors), Cambridge University Press, pp. 445-467.

Thompson, S.K. (1992). *Sampling*. John Wiley and Sons, New York, 343 pages.





---

## **ALGERIA**

---

**Reporting: M. Nacer-eddine Hammouda**

Algeria's statistical information system is experiencing a certain inertia. Changes, either methodological or thematic, are rare. One important thing to note is the first survey on time use (ENET) that was conducted in 2012. This survey is part of the AI-Insaf program "promoting gender equality and women's empowerment." This program was implemented in partnership between the Algerian government and United Nations agencies and was signed in February 2009. The survey was conducted by the Office national des statistiques (ONS) with the assistance of UN Women and initial results are available (<http://www.ons.dz/-Enquete-sur-l-emploi-du-temps-ENET-.html>). This survey was supposed to give a better assessment of female participation in the labour force. In this respect, it was inconclusive,<sup>2</sup> since the estimate of female economic activity was lower than the ILO's indirect estimates. This leads us once again to question the methodological approach adopted and the choice of the duration of the survey in the field (from May 6 to July 10, 2012).

Regarding employment (<http://www.ons.dz/-Emploi-et-chomage-.html>), the ONS went from a quasi-annual survey to a semi-annual survey in 2014, conducted in April and September. The sample was partially renewed, unlike in the past where samples were completely renewed, other than in a few inconclusive attempts at partial renewal. This survey was conducted only once in 2015 due to budget constraints.

The Ministry of Health, Population and Hospital Reform took charge of conducting an MICS survey in 2012/2013. This is the fourth such survey conducted by Algeria. The preliminary report was published in 2015.

There are plans to conduct a national disability survey in 2015/2016 on a considerable sample.

Another survey on health risk factors (STEPS) is also planned for 2016 with the assistance of the WHO, targeting the population aged 18 to 69. A pilot survey was conducted in 2004 on a sample of 4,000 (2 X 2,000) households in two wilayas (Sétif and Mostaganem) but on persons aged 25-64. Another national survey,

---

<sup>2</sup> This is also the case for other countries.



under the TAHINA project, was conducted the same year and focused on risk factors in the population 35 years and older.

CREAD (the applied economics research centre for development) conducted a survey of young people aged 15 to 29 as part of the SAHWA project ([www.sahwa.eu](http://www.sahwa.eu)), funded by the European Union. This survey covers five countries in the MENA region (Algeria, Egypt, Lebanon, Morocco and Tunisia) and examines a number of aspects (living conditions, education, employment, migration, culture and politics). The data are now being processed and will be widely disseminated.

For more information, please contact: M. Nacer-eddine Hammouda, [nacereddine.hammouda@ensae.org](mailto:nacereddine.hammouda@ensae.org)

---

## **ARGENTINA**

---

**Reporting: Veronica Beritich**

### **The government is rebuilding INDEC, its discredited statistics institute**

On 30th December 2015 the new president, Mauricio Macri, declared a “national statistical emergency”. A decree that gives Jorge Todesca, who has been named to clean up and repair the government’s statistics institute, a free hand to appoint new directors and allows INDEC to suspend publication of data on GDP, inflation, poverty and unemployment until the end of 2016.

On 13th February 2016 at The Economist’s interview, Mr. Todesca says: “Teams were decimated”. That is because statisticians, whose findings displeased the former government from 2007 to 2015, were sacked or just resigned or were banished to back rooms without equipment.

In 2013 Argentina became the first country to be censured by the IMF for misreporting GDP and prices. With poverty rising, in part because of high inflation, INDEC simply stopped reporting the poverty rate in 2014.

INDEC has rehired experts who were ousted. Some of them will resume responsibility for poverty and employment data. By May they intend to put together a “basic food basket” and publish its price, a step towards calculating the level of extreme poverty.

They will also create a broader measure of poverty, using a bigger basket of goods and services, including transport. And they hope to have a “multidimensional poverty index”, which is likely to include such things as access to health care and education, by early 2017.

Others are overseeing the construction of a new consumer-price index. Putting together a new consumer-price index is expected to take until September, even though it will probably be based on the widely-used methodology of the



International Labour Organisation and, much like the series used until 2014, on prices just in Buenos Aires and its suburbs. Responding to criticisms, on January 2014, former INDEC has devised a new index, the IPCNu, which monitored prices nationally but new authorities believe that the national samples used in the IPCNu may not be reliable enough.

Until a new consumer-price index comes out, INDEC advises Argentines to consult two in which it has some confidence: those published by the city of Buenos Aires and by the province of San Luis. Data on GDP and employment will take longer. The government plans to submit to economic monitoring by the IMF. Argentina has refused it since 2006. Mr Todesca hopes the IMF will soon lift its statistical censure. “Argentina was once a pioneer in Latin America” in publishing data, he points out. Now, just being one of the crowd would be an achievement.

Further information on this report can be found in the interview to Mr Todesca “An Augean stable. The government is rebuilding its discredited statistics institute”, February 13, 2016 at [www.economist.com](http://www.economist.com).

General information can be found at [www.indec.gov.ar](http://www.indec.gov.ar).

---

## **CANADA**

---

**Reporting: Claude Turmelle**

### **The 2015 Redesign of the Canadian Labour Force Survey**

The Canadian Labor Force Survey provides estimates of employment and unemployment at the regional and national levels which are among the most timely and important measures of performance of the Canadian economy. With the release of the survey results only 10 days after the completion of data collection, the LFS estimates are the first of the major monthly economic data series to be released.

Every ten years, after the decennial census of population, the LFS undergoes a sample redesign to account for the evolution of the population and labor market characteristics, to adjust to the current and expected needs of data users, and to update the geographical information used to carry out the survey. The 2015 redesign introduced a number of major changes to the methodology of the survey. These changes were introduced to reduce survey costs, use modern collection methods, and allow data users to perform state-of-the-art analyses.

In this survey redesign, the Primary Sampling Units (PSUs) were constructed from the Dissemination Areas defined for the 2011 Census. In addition to streamlining the work involved with the sample redesign, this change makes the LFS geography



more standard, which helps in the comparison of estimates across surveys and in analysis involving multi-level modeling.

The use of existing lists of addresses has been expanded significantly in the 2015 design. The Dwelling Universe File is an extraction of addresses from Statistics Canada's residential address register (AR). It is now being used to produce the list of addresses for over 90% of the PSUs in the LFS sample. This reduces the work of field interviewers who would otherwise have to create the list of addresses by directly observing the neighborhoods / PSUs in the LFS sample. The frame service also supplies telephone numbers that will help interviewers establish contact with sampled households.

A web-based electronic questionnaire (EQ) pilot test was conducted in 2013 to determine the impact that the introduction of EQ collection would have on the survey. It was demonstrated that response rates would not be affected and mode effects were not observed. Therefore, the LFS has introduced the third collection method in 2015. Eligible respondents, who accept the offer at the end of their first month's interview, can now complete their questionnaire online for their remaining 5 months in the survey. Currently, about 17% of all LFS questionnaires are completed online.

Last but not least, a significant change was made in 2015 to variance estimation. Starting in January 2015, the bootstrap method has replaced jackknife as the variance estimation technique for the LFS. This allows users to compute and report design-based variance estimates for state-of-the-art analyses on their own.

For more information, contact Claude Turmelle, Chief methodologist, Household Survey Methods Division, Statistics Canada, at [Claude.Turmelle@Canada.ca](mailto:Claude.Turmelle@Canada.ca).

---

## **LATVIA**

---

### **Projection of Demographic Indicators Using Cohort Method**

Central Statistical Bureau of Latvia has done the first cohort analyses of married couples using the data from administrative register and population census. This is an example how administrative register data can be used to produce new statistics.

Couples who concluded marriage in 2001 were included in the research cohort (9212 couple in total). The impact of various demographic and social factors (age of spouses, education, children) – on divorce were analysed. The research was carried out using OCMA (Office of Citizenship and Migration Affairs) register data and population census data 2011.

After 15 years 60.4% of couples are still married, 32.6% are divorced, while 7.0% marriages ended as one of the spouses has passed away. The projection what will



happen with couples in future is necessary to answer the question – *how many couples will stay together forever?* Following data were used for the projection:

- The annual number of marriages since 1970,
- The number of divorces in 2014 depending on the length of the marriage,
- Number of couples married in 2001 and divorced till 2016.

The possibility to divorce after a certain number of years of cohabitation is calculated by dividing the number of divorced couples (with certain cohabiting length) with the number of marriages registered in the specified year:

$$p_1(n) = \frac{d(2014, n)}{m(2014 - n)} * 100\% \quad p_2(n) = \frac{d(2001 + n, n)}{m(2001)} * 100\%$$

where:  $m(i)$  – the number of marriages concluded in the year  $i$ ;  
 $d(i, j)$  – the number of divorces in the year  $i$  with the length of cohabitation  $j$  years;  
 $p(j)$  – probability (estimated as ratio) to divorce after  $j$  years of cohabitation ( $p_1$  – using the data of divorced couples in 2014,  $p_2$  – using data of spouses in 2001 divorced until 2016).

According to the projection 14.4% of spouses married in 2001 will divorce in 2016 or later. If we add together, 32.6% of couples have already divorced during 15 years of cohabitation and 14.4% will divorce in future, it makes that 47.0% of spouses married in 2001 will divorce, but 53.0% will be married all their life.

For more information on the research visit the website of the Central Statistical Bureau of Latvia (<http://www.csb.gov.lv/en/notikumi/csb-experts-presented-research-changes-occurred-during-last-15-years-demographic-and-social>) or contact Kristīne Lece ([Kristine.Lece@csb.gov.lv](mailto:Kristine.Lece@csb.gov.lv)) and Dana Bukša ([Dana.Buksa@csb.gov.lv](mailto:Dana.Buksa@csb.gov.lv)).

---

## **NEW ZEALAND**

---

**Reporting: Felibel Zabala**

### **Small domain estimates for cultural measures of New Zealand's indigenous population**

Statistics New Zealand has published small domain estimates for iwi on four cultural measures. Iwi are tribes of New Zealand's indigenous (Māori) population. The estimates use census and survey data. We used the statistical software package Stan to produce the estimates, and published the results using an interactive R Shiny application.

This is the first time we've released estimates for a household survey that use small domain estimation methods, though we've also used small domain estimates to



produce period life tables for 2012–14 and intercensal estimates of New Zealand's estimated resident population.

See [Iwi statistics using small domain estimation](#) on [www.stats.govt.nz](http://www.stats.govt.nz), or contact Gareth Minshall at [gareth.minshall@stats.govt.nz](mailto:gareth.minshall@stats.govt.nz) for more details on the iwi estimates.

See [New Zealand period life tables: Methodology for 2012–14](#) on [www.stats.govt.nz](http://www.stats.govt.nz), or contact John Bryant at [john.bryant@stats.govt.nz](mailto:john.bryant@stats.govt.nz) for more details on the period life tables.

## Testing a modernised collection model for the census

Statistics NZ has been field testing a modernized collection model that we're introducing at the next New Zealand Census of Population and Dwellings in 2018. We want to understand how the new collection model will affect coverage and response rates, and see how well our systems and processes work.

We mailed the test to around 22,000 households in several regions on 22 March, and expect to have results by the end of June 2016.

Our new 'digital-first' collection model will have an altered mix of delivery modes. Forms will be mailed out to dwellings, and the internet will replace paper as the primary response mode.

For the census test we used an experimental design that will enable us to optimize self-response, and maximize online response. We want to learn more about:

- the best timing for mailing out the initial call-to-action contact material to respondents
- the effect of using different messaging options on the 'postcard' reminder we send
- the overall mix of components we should use in our respondent reminder strategy.

See [Census transformation in New Zealand](#) on [www.stats.govt.nz](http://www.stats.govt.nz) or contact Charlie Dohrman at [charlie.dohrman@stats.govt.nz](mailto:charlie.dohrman@stats.govt.nz) for more details.

## Developments in environmental statistics: *Environment Aotearoa 2015*

Statistics NZ and the Ministry for the Environment released *Environment Aotearoa 2015* in October 2015, the first 'whole of environment' report for New Zealand since 2007.

The release consists of four parts, each designed to address different stakeholder and customer needs. An infographic presents the high level key findings, which are discussed more fully in the summary report written in plain English. The report is presented in a 'pressure-state-impact' format for each of the five environmental domains.



The indicator website presents statistical findings, definitions, and methodologies for 122 indicators, with raw datasets (including metadata) available from the Ministry for the Environment's Data Service. We conducted statistical analysis of environmental trends where data were of sufficient quality, and the processes we followed are detailed in a technical report.

We compiled *Environment Aotearoa 2015* from existing data. Crown research institutes, universities, and central and local government organizations were key data suppliers. We selected indicators, in consultation with experts, based on their relevance to the topics and data quality. We assessed data quality against the six criteria from the Principles and Protocols for Producers of Tier 1 Statistics: relevance, accuracy, timeliness, consistency/coherency, accessibility, and interpretability.

Given the variability in data quality, we had to carefully consider how to convey data quality issues while making more information available. The use of data quality icons and metadata helped us to do this. Most indicators relate to environmental state, while pressures and impact indicators are relatively undeveloped. We will consider how the integration of indicators produced under the System of Environmental-Economic Accounting framework may enhance our future analysis of economic pressures and impacts.

## **Environmental Reporting Act 2015**

New Zealand's national-level environmental reporting framework is set out under the Environmental Reporting Act 2015. The Act prescribes a pressure-state-impact framework covering the five environmental domains: air, atmosphere and climate, fresh water, land, and marine. Biodiversity information is presented as a cross-cutting theme. Pressures include both human and natural pressures. Impact categories are defined in the Act and cover: ecological integrity, public health, the economy, te ao Māori (the Māori world view), and culture and recreation.

See [Environmental Reporting Series](#) on [www.stats.govt.nz](http://www.stats.govt.nz) or contact [adam.tipper@stats.govt.nz](mailto:adam.tipper@stats.govt.nz), [lilian.morrison@stats.govt.nz](mailto:lilian.morrison@stats.govt.nz), or [nafees.anwar@stats.govt.nz](mailto:nafees.anwar@stats.govt.nz) for more information.





## Upcoming Conferences and Workshops

---

---



---

### **Small Area Estimation Conference 2016**

**Where:** Maastricht, The Netherlands

**When:** August 17-19, 2016

**Homepage:** <http://www.sae2016.nl/>

Welcome to the website of the Small Area Estimation 2016 conference.

This conference is organized by Maastricht University School of Business and Economics and Statistics Netherlands.

The conference will be held in Maastricht in the faculty building of Maastricht University School of Business and Economics on **August 17-19, 2016**.

The following speakers already confirmed to deliver a presentation.

See also the [provisional program](#) for more details.

#### **Keynote speakers:**

- Prof. Dr. Thomas Louis, Department of Biostatistics, John Hopkins Bloomberg School of Public Health
- Prof. Dr. Jiming Jiang, Department of Statistics, University of California

#### **Invited and special topic speakers:**

- Dr. W. Bell, Census Bureau
- Dr. H.J. Boonstra, Statistics Netherlands
- Dr. H. Chandra, ICAR - Indian Agricultural Statistics Research Institute, New-Delhi



- Prof. Dr. G.S. Datta, Department of Statistics, University of Georgia
- Dr. S. Falorsi, Italian National Statistical Institute, Rome
- Prof. Dr. M. Ghosh, Department of Statistics, University of Florida
- Prof. Dr. S. Holan, Department of Statistics, University of Missouri, Colombia
- Prof. Dr. P. Lahiri, Joint program in Survey Methodology, University of Maryland, College Park, USA
- Prof. Dr. I. Molina, Department of Statistics, Universidad Carlos III de Madrid
- Prof. Dr. J. Opsomer, Department of Statistics, Colorado State University
- Prof. Dr. J.N.K. Rao, School of Mathematics and Statistics, Carleton University, Ottawa
- Prof. Dr. J. Sunil Rao, Department of Public Health Sciences, University of Miami
- Dr. R. Steorts, Department of Statistical Science, Duke University



### **Baltic-Nordic-Ukrainian (BNU) Network Summer School on Survey Statistics**

**Organized by:** Baltic-Nordic-Ukrainian (BNU) Network institutions

**Where:** Kyiv, Ukraine

**When:** August 22-26, 2016

**Homepage:** <http://probability.univ.kiev.ua/school16/>

### **Objectives:**

The Baltic-Nordic-Ukrainian Summer School on Survey Statistics aims to promote interest and skill in the theory, methodology and practice of survey statistics, and to provide an opportunity for university teachers, research students and survey practitioners to discuss their problems and to learn from the experiences in other countries.

The Baltic-Nordic-Ukrainian Network on Survey Statistics is a non-governmental group of enthusiastic survey statisticians in Belarus, Estonia, Finland, Latvia, Lithuania, Sweden and Ukraine and their home institutions, most of them being a university or a national statistical agency. A Baltic-Nordic network for co-operation on education and research in survey statistics grew continuously since 1996. Since then the network has expanded to include partner universities from Ukraine and further, from Belarus. It has arranged three summer schools, twelve workshops and four conferences. It has also arranged exchange visits for university teachers, research students and survey practitioners to universities within the Network with a total duration of more than 100 months. The website of the Network is <http://wiki.helsinki.fi/display/BNU/Home>

Participants are personally invited by the organizers.

Registration Fees:



- EUR 150 for regular participants from EU countries
- EUR 75 for student participants from EU countries
- EUR 50 for accompanying persons from EU countries
- UAH 600 for participants from Ukraine and Belarus

Registration fees will cover:

- School materials;
  - Coffee breaks;
  - Lunches on working days;
  - An excursion to National Museum of Ukrainian Architecture and Culture (Pyrogovo Museum) with lunch;
  - Welcome and farewell parties.
- A discounted rate of 5% of registration fee is offered as an option for current IASS members.

Registration fees are paid on arrival.

Institute of Continuing Education of Kyiv National University  
<http://univ.kiev.ua/en/dep/pislja-dup/>. Address: Kyiv, Vasylykivska str.



### **2016 International Population Data Linkage Conference**

**Organized by:** International Population Data Linkage Network (<http://www.ipdln.org/>)

**Where:** Swansea, United Kingdom

**When:** August 22-26, 2016

**Homepage:** <http://www.ipdlnconference2016.org/>

### **Linking Data - Improving Lives**

‘Big Data’ and its potential to improve health and society is much discussed across the world. Linking disparate population-scaled datasets together is a key technique in the production of powerful new data resources that can illuminate, evaluate and steer policy and practice right across society.



However, with this work comes the responsibility to build trustworthy and well-governed approaches that safeguard privacy at all times. Opportunities for new infrastructures and architectures present themselves, new analytical techniques are needed, and a cadre of people equipped with a new blend of skills is required.

This unique conference is designed to provide researchers, policy makers, practitioners, administrators, regulators, and data guardian's opportunities to learn more about the cutting edge work that is underway across the world, together with a chance to showcase their own achievements, and listen to talks from international leaders in the field.

It is a great pleasure to be able to host this important international event in Wales this year, as part of its bi-annual journey around the world. We hope you will take the opportunity to join us either by submitting an abstract of your work, or by attending as a delegate.

### **Conference Sub-Themes**

- *Analytical approaches to distributed data*
- *Data and linkage quality*
- *Advanced analytics*
- *Privacy, regulation & governance*
- *Applied projects*
- *Linking to emerging data types*
- *Public engagement*
- *Capacity building*
- *Delivering and measuring impact*

### **Who Should Attend?**

Data Analysts who use large linked datasets to help understand society, human health, and well-being, such as:

- Researchers and analysts in Universities, the private sector, non-for profit and community organisations
- Public policy makers in local and national governments

Those who enable the use of linked data, including:

- Technical experts working on linking, merging and preparing large complex datasets
- Technologists and statisticians developing new analytic methods
- Privacy, Information Governance and regulatory experts

Users of the results of linked data analyses, including:

- Leaders of organisations interested in social policy, health and human development
- Leaders in industry that develop products and knowledge in these areas





### **Third Global International Conference on Big Data for Official Statistics**

**Organized by:** Central Statistical Office Ireland and the United Nations Statistics Division

**Where:** Dublin, Ireland

**When:** August 30 - September 1, 2016

**Homepage:** <http://unstats.un.org/unsd/bigdata/conferences/2016>

#### **Overview**

The daily use of mobile phones and social media, as well as the routine checking of equipment, such as cars or home appliances, generate continuous streams of electronic data. These data sources, commonly referred to as Big Data, can potentially be used in the compilation of official statistics for the purpose of evidence-based decision making. The UN Statistical Commission therefore established a Global Working Group (GWG) in 2014, which was mandated to provide strategic vision, direction and coordination on the use of Big Data for Official Statistics. The GWG promotes the practical use of Big Data, capacity building and sharing experiences, while finding solutions for the associated challenges.

The GWG organized its first global conference on Big Data in Beijing in October 2014 and showcased a variety of examples of Big Data projects for official statistics, such as Mobile phone data for daytime population statistics, Satellite imagery data for agriculture statistics and Social media data for consumer confidence indicators. The second global conference took place in Abu Dhabi in October 2015, where the GWG tried to more systematically progress in the use of Big Data in the production process of official statistics.

In line with the priorities of work and with the urgent needs for support of SDG indicators, the third Global Conference on Big Data for Official Statistics wants to take the next steps in the utilization of Big Data in the production of official statistics, namely by (1) providing guidance for access to proprietary data and for successful partnerships with data owners, by (2) providing training courses on topics such as methodology, IT tools and project management related to Big Data, and by (3) supporting projects, which demonstrate the use of Big Data for official statistics, and especially for the compilation of SDG indicators.



---

### **Royal Statistical Society 2016 International Conference**

**Organized by:** The Royal Statistical Society

**Where:** University Place, Manchester, United Kingdom

**When:** September 5 - 8, 2016

**Homepage:**

[http://www.rss.org.uk/RSS/Events/RSS\\_Conference/RSS\\_2016\\_International\\_Conference/RSS/Events/Conference/RSS\\_2016\\_International\\_Conference.aspx?hkey=44c3b1e0-42b0-42e1-867f-6a78b09121de](http://www.rss.org.uk/RSS/Events/RSS_Conference/RSS_2016_International_Conference/RSS/Events/Conference/RSS_2016_International_Conference.aspx?hkey=44c3b1e0-42b0-42e1-867f-6a78b09121de)

The RSS 2016 Conference will take place in Manchester – the European City of Science 2016.

Because of the range of topics presented and discussed and the breadth of the audience, this conference is the only one in the UK where all statisticians and users of data gather together. So it's the best forum to share information, network and learn from one another.

Now in its 24th year, the RSS conference has gained prestige for its focus on current statistical issues, how it fosters the exchange of ideas and information and the quality of its speakers. Plenary speakers in 2016 will include Christl Donnelly (Imperial College London) and Xiao-Li Meng (Harvard University).

The main conference program will be preceded by one-day training courses on Monday 5 September and in addition professional development workshops will run throughout the event.



---

# PRIVACY IN STATISTICAL DATABASES

Dubrovnik, Croatia, Sep. 14-16, 2016



## **Privacy in Statistical Databases 2016**

**Organized by:** UNESCO Chair in Data Privacy

**Where:** Dubrovnik, Croatia

**When:** September 14-16, 2016

**Homepage:** <http://unescoprivacychair.urv.cat/psd2016/>

Privacy in statistical databases is about finding tradeoffs to the tension between the increasing societal and economical demand for accurate information and the legal and ethical obligation to protect the privacy of individuals and enterprises which are the respondents providing the statistical data. In the case of statistical databases, the motivation for respondent privacy is one of survival: statistical agencies or survey institutes cannot expect to collect accurate information from individual or corporate respondents unless these feel the privacy of their responses is guaranteed.

Beyond respondent privacy, there are two additional privacy dimensions to be considered: privacy for the data owners (organizations owning or gathering the data, who would not like to share the data they have collected at great expense) and privacy for the users (those who submit queries to the database and would like their analyses to stay private). Co-utility has shown to provide sustainable solutions to user privacy.

"Privacy in Statistical Databases 2016" (PSD 2016) is a conference sponsored and organized by the [UNESCO Chair in Data Privacy](#) with proceedings published by Springer-Verlag in Lecture Notes in Computer Science. Generous support by the Templeton World Charity Foundation is acknowledged. The purpose of PSD 2016 is to attract world-wide, high-level research in statistical database privacy.

PSD 2016 is a successor to [PSD 2014](#) (Eivissa, Sep. 17-19, 2014), [PSD 2012](#) (Palermo, Sep. 26-28, 2012), [PSD 2010](#) (Corfu, Sep. 22-24, 2010), [PSD 2008](#) (Istanbul, Sep. 24-26, 2008), [PSD 2006](#) (Rome, Dec. 13-15, 2006) and [PSD 2004](#) (Barcelona, June 9-11, 2004), all with proceedings published by Springer in LNCS 8744, LNCS 7556, LNCS 6344, LNCS 5262, LNCS 4302 and LNCS 3050, respectively. Those seven PSD conferences follow a tradition of high-quality technical conferences on SDC which started with "Statistical Data Protection- SDP'98", held in Lisbon in 1998 and with proceedings published by OPOCE, and continued with the AMRADS project SDC Workshop, held in Luxemburg in 2001 and with proceedings published in Springer LNCS 2316.

Like the aforementioned preceding conferences, PSD 2016 originates in Europe, but wishes to stay a worldwide event in database privacy and SDC. Thus, contributions and attendees from overseas are welcome.

---





### **ITSEW2016 – International Total Survey Error Workshop**

**Organized by:** Australian Bureau of Statistics

**Where:** Sydney, Australia

**When:** October 9-12, 2016

**Homepage:** <https://consol.eventsair.com/QuickEventWebsitePortal/itsew2016/itsew16>

The theme of the 2016 International Total Survey Error Workshop is “*Will Total Survey Error Save Survey Science?*”. Consistent with this theme, presentations are encouraged on the following topics: inferences from nonprobability samples, questionnaire designs to minimize Total Survey Error (TSE), collection designs to mitigate errors and methods for reducing the impacts of error on estimation and analysis.

Some of the additional topics that previous workshop presentation have covered and which continue to be of interest in 2016 include the following:

- Simultaneous evaluations of the contributions to total survey error from two or more error sources; example, from nonresponse and response error.
- Interactions between error sources; example, the interplay between item nonresponse error and editing error or interviewer and respondent error.
- Methods for simulating the effects of multiple error sources on the Total Survey Error (TSE).
- The conceptualization and historical development of TSE.
- Reviews of the literature and meta-analyses of TSE.
- The impact of the TSE concept on survey methodology.
- Survey models and other representations of TSE.
- New methods for estimating the components of the mean squared error.
- Methodologies for comparing the TSE for two or more modes of data collection.
- Reports on efforts to evaluate TSE in complex surveys.
- Uses of quality profiles for understanding and minimizing TSE.
- Uses of prior estimates of non-sampling error in the design of new surveys or for allocating resources for survey error reduction.
- Exemplary studies of non-sampling error components, either planned or in progress.
- Clients' and other users' perceptions of the TSE concept.
- Continuations of work presented in earlier ITSEWs.





### **9th French Colloquium on Survey Sampling**

**Organized by:** Société Française de Statistique

**Where:** l'Université du Québec en Outaouais, Québec, Canada

**When:** October 11-14, 2016

**Homepage:** <http://sondages2016.sfds.asso.fr/>

The Ninth French Colloquium on Survey Sampling (Colloque francophone sur les sondages) will take place on **October 12-14, 2016**, on the main campus of the Université du Québec en Outaouais (UQO), in Gatineau (Canada). It will be preceded by training workshops on **October 11, 2016**, also on the main campus of UQO. This ninth Colloquium is organized by the Société Française de Statistique (SFdS) and its Enquêtes, Modèles et Applications group, and by UQO. UQO is considered a university in which the human aspect promotes learning, thought and creation, UQO is strong as a part of the Université du Québec network and can rely on the educational resources and shared services of the largest university network in Canada.

The Gatineau Colloquium will capitalize on two synergies, namely the synergy from the meeting of several continents and the synergy from the meeting of specialists from various communities and disciplines: statisticians and statistics users (for example, sociologists, demographers and political scientists) from academia, governments and the private sector.

Looking forward to see you!



### **Conference of European Statistics Stakeholders 2016**

**Organized by:** Eurostat, the European Central Bank, the European Statistical Advisory Committee, the FENStatS, the Hungarian Central Statistical Office and the Hungarian Statistical Association

**Where:** Budapest, Hungary

**When:** October 20–21, 2016

**Homepage:** <http://www.ksh.hu/cess2016/>

Eurostat, the European Central Bank, the European Statistical Advisory Committee, the Federation of European National Statistical Societies and the Hungarian Central Statistical Office, with the involvement of the Hungarian Statistical Association is organising the second Conference of European Statistics Stakeholders in Budapest on 20–21 October 2016.

The aim of the Conference is to bring together European methodologists, producers, and users of statistics to discuss user needs, to share best practices in the production of official statistics, to present innovative ways of visualising and communicating statistics, and to advance new methodological ideas for collecting and analysing data.

We would like to draw your attention to the publication opportunities for CESS 2016 authors: All presenters at CESS 2016 will have the opportunity to submit their extended papers for publication in a Eurostat e-book (subject to a peer review by members of the Scientific Programme Committee). Moreover, authors of selected papers will be invited to extend their papers for publication in the ECB Statistics Paper Series ([SPS](#)).

We are confident that the topics of the Conference will be appealing to a broad range of participants who will also enjoy the attractions of Budapest and Hungary.

For further information, please contact the organising team: [cess2016@ksh.hu](mailto:cess2016@ksh.hu)







---

### **ICAS VII The International Conference on Agricultural Statistics 2016**

**Organized by:** Italian National Institute of Statistics (Istat) and Food and Agriculture Organization (FAO/UN)

**Where:** Rome, Italy

**When:** October 26-28, 2016

**Homepage:** <http://icas2016.istat.it/>

### **Modernization of Agricultural Statistics in Support of the Sustainable Development Agenda**

The Seventh International Conference on Agricultural Statistics (ICAS VII) will be held in Rome – Italy on 26-28 October 2016. ICAS VII is organized by the Italian National Institute of Statistics, in close collaboration with the Food and Agriculture Organization of the UN (FAO). The Conference focuses on bringing together research and best practices in the field of agriculture statistics, in response to the changing needs and opportunities for agricultural statistics.

ICAS VII convenes senior agricultural statisticians from all over the world. Most of them represent national statistical offices and ministries of agriculture, but the Conference is open to all producers, suppliers, trainers and users of agricultural statistics, such as economists, statisticians, agronomists, researchers, analysts and decision-makers from government entities, academia, development partners and international organizations





## **International Conference on Questionnaire Design, Development, Evaluation, and Testing 2016 QDET2**

**Organized by:** The American Statistical Association and endorsed by several organizations

**Where:** Miami, United States

**When:** November 9-13, 2016

**Homepage:** <http://www.amstat.org/meetings/qdet2/>

### **Welcome to the 2016 International Conference on Questionnaire Design, Development, Evaluation, and Testing (QDET2)**

QDET2 will provide a creative environment for researchers from across the world working in government, academia, and the private sector to share new solutions and fresh approaches to survey measurement, questionnaire design, and evaluation. The conference builds on the resounding success of the first QDET, held in 2002, and will be the first international conference in more than a decade devoted to both the design of survey questionnaires and the methods used for their development, evaluation, and testing.

*Student and early career awards, as well as travel support for students and those from developing nations are available! Check out the QDET2 [awards](#) page for details and deadlines.*

#### **QDET2 is endorsed by these organizations:**

- Center for Survey Research at the University of Massachusetts Boston
- European Social Survey European Research Infrastructure Consortium
- The European Survey Research Association
- The Federal Statistical Office of Germany (Destatis)
- The Joint Program in Survey Methodology, University of Maryland, College Park
- The Michigan Program in Survey Methodology
- The National Statistics Institute (INE) of Spain
- Office for National Statistics U.K.
- The Royal Statistical Society
- Statistics Canada
- Statistics Netherlands
- Statistics New Zealand
- Survey of Health, Ageing and Retirement in Europe (SHARE) European Research Infrastructure Consortium.





**23rd Australian Statistical Conference 2016, 14th Australasian Data Mining Conference (AusDM), 9th Australian Conference on Teaching Statistics (OZCOTS)**

**Organized by:** Big Data: Mining, Analyzing and Teaching

**Where:** Canberra, Australia

**When:** December 5 – 9, 2016

**Homepage:** <http://asc2016.com.au>

On behalf of the Statistical Society of Australia (SSA), the Organising Committee invites you to attend the 23<sup>rd</sup> Australian Statistical Conference which is to be held in conjunction with the 14<sup>th</sup> Australasian Data Mining Conference (AusDM) and the 9<sup>th</sup> Australian Conference on Teaching Statistics (OZCOTS). This conference will provide unique insight into statistics, data mining and statistics teaching.

The major focus of the conference will be Big Data: Mining, Analysing and Teaching. It aims to bring together the statistical approaches to data analysis with the techniques of data mining and their use in teaching statistics. Application to Big Data will be of particular interest.

The SSAI Program of the conference will cover a wide range of classical and Bayesian methods in mathematical statistics, their application to environmental, social, biological and official statistics and international engagement. The programs for [AusDM](#) and [OZCOTS](#) can be found on their respective host pages and the program outline is on the Program [Page](#).

Attending the ASC 2016 will give each delegate a unique opportunity to meet, discuss and learn with meet like-minded individuals.

On behalf of the Organising Committee we look forward to welcoming you to Canberra in 2016 and encourage you to attend this amazing occasion!





---

### **The Spirit of Official Statistics: Partnership and continuous innovation**

**Organized by:** International Association of Official Statistics

**Where:** Abu Dhabi, United Arab Emirates

**When:** December 6-8, 2016

**Homepage:** <http://www.iaos2016.ae/IAOS-2016-conference.php>

### **About IAOS 2016 Conference**

Hosted by the Emirate of Abu Dhabi from 6 to 8 December 2016, the 15th edition of IAOS Conference will provide a unique platform for the international community to share their knowledge and present their insights on innovations in statistics. We are committed to delivering a remarkable international event, which will surpass all expectations and boost the statistical knowledge in the region and the world.

The conference will be held at the multi-award winning venue, Abu Dhabi National Exhibition Centre (ADNEC), which offers a unique experience to conference participants.



The theme of the IAOS2016 conference is: **“The Spirit of Official Statistics: Partnership and continuous innovation”**. During the conference, we shall explore what is and what will be the value of official statistics in light of emerging data



To register or for more detailed information on the conference themes, please see the IAOS 2016 web site at (<http://www.iaos2016.ae>) or contact the programme chair Eric Rancourt ([Eric.Rancourt@Canada.ca](mailto:Eric.Rancourt@Canada.ca)) directly.



<http://ec.europa.eu/eurostat/cros/content/ntts-2017>

The NTTS 2017 conference will take place from **14 to 16 March 2017** in **Brussels**.

The [NTTS 2017 Scientific Committee](#) welcomes the submission of abstracts within all [areas](#) of relevance to the conference. The deadline for [submission of abstracts](#) is 28 October 2016.

Authors of NTTS papers have the opportunity to submit their paper for consideration for a [special NTTS issue of the Journal of Official Statistics](#), as well as a [special NTTS section of the Eurostat Review on National Accounts and Macroeconomic Indicators](#).





The 7th Conference of the European Survey Research Association (ESRA) will be held in Lisbon, Portugal, from 17th to 21st July 2017. The conference will be hosted by the University of Lisbon, jointly by the School of Economics and Management (ISEG), the Centre for Research in Social Sciences and Management (CSG) and the Institute for Social Sciences (ICS).

### **Call for sessions now open**

The scientific committee is now inviting researchers who are active in the field of survey research and survey methodology to **submit proposals to organise sessions** at the conference.

For your session to be considered, please [submit an abstract](#) (max. 400 words) together with 3-5 keywords by **15<sup>th</sup> September 2016**. To submit a proposal you must login to your ESRA account (or create a new account if you do not already have one) and then follow the instructions provided.

Session proposals are invited in any area of survey methodology, or in substantive areas of survey research. We encourage proposals from researchers with a variety of backgrounds, including academic research, national statistics and market research.

The following are examples of topics that are of particular interest: Sample designs, coverage, and sampling

- Fieldwork processes
- Unit and item nonresponse
- Weighting and imputation
- Questionnaire development, testing and piloting
- Interviewers and interviewer effects
- Mixing modes and mode effects
- Online survey methods and surveys on mobile devices
- Linking survey data to auxiliary data sources
- Using paradata to evaluate survey quality
- Methods for cross-national and cross-cultural surveys
- Longitudinal surveys and longitudinal analysis techniques
- Analysing, monitoring and reducing Total Survey Error
- Experiments in general population surveys
- Data documentation, archiving and data access
- Survey analysis techniques
- Election polling and public opinion
- Social indicators
- Substantive applications of survey research



As a session organiser you are expected to encourage paper submissions to your session, review paper proposals with regards to their suitability in terms of focus and quality, liaise with presenters and the conference organisers and chair your session during the conference. If the conference committee accepts your proposal, your session will be advertised in an open call for papers that will be issued in October 2016.

*The organizing committee - Annelies Blom, Caroline Roberts, Alice Ramos, João Peixoto, Monica Fraga, Margarida Piteira, Eldad Davidov and Bart Meuleman*

For all questions regarding transportation, the conference venue, accommodation, payments and other logistics, please contact [info@europeansurveyresearch.org](mailto:info@europeansurveyresearch.org).

For questions regarding the scientific program, contact [conference@europeansurveyresearch.org](mailto:conference@europeansurveyresearch.org).



### **61st World Statistics Congress of the International Statistical Institute**

**Organized by:** The High Commission for Planning - HCP

**Where:** Marrakech, Morocco

**When:** July 16 – 21, 2017

**Homepage:** <http://www.isi2017.org>

We are immensely pleased to invite you to attend the 61st ISI World Statistics Congress (WSC) to be held in Marrakech, Morocco from 16 to 21 July 2017.

The biennial WSC is the flagship conference of the International Statistical Institute (ISI) and its seven associations. It brings together eminent statisticians and members of the statistical community from the five continents to present, discuss, promote and disseminate research and best practices in every field of Statistics and its applications.



ISI2017 will feature a rich scientific programme focusing on the latest knowledge and innovation in Statistics. It will also be an excellent opportunity to benefit from additional scientific activities such as satellite meetings and short courses.

The High Commission for Planning (HCP), the main producer of official statistics in Morocco, is pleased to host and organize ISI2017 in Marrakech. Also called the Red City, Marrakech, offers a wide variety of attractions, from historical palaces, tombs, mosques, gardens, and beautiful nature, to endless shopping, dining and entertainment places. Marrakech, the realm of well-being, art, culture and gastronomy, has been elected as the travelers' top destination in the world for the year 2015.

To ensure an enjoyable and fruitful ambiance for sharing and networking, an attractive and diversified social programme is being designed. It shall give the participants an opportunity to experience the rich culture of Marrakech and its magical smells and colors.

The venue of the ISI2017 is the Mansour Eddahbi Hotel & Palais des Congrès, a great palace located at the heart of the empirical city of Marrakech, within walking distance from the city's major attractions. Find out more about the venue at [www.mansoureddahbi.com](http://www.mansoureddahbi.com).

We have no doubt that taking part in ISI2017 will be a great and enjoyable opportunity for sharing experiences and networking in the magical and warm ambiance of the Red City – Marrakech.

Come and join us there for ISI2017!





## In Other Journals

---

### Journal of Survey Statistics and Methodology

---

**VOLUME 4 / NUMBER 2 / JUNE 2016**

<https://jssam.oxfordjournals.org/content/4/2.author-index>

#### **Survey Statistics**

**Multiple Imputation in Two-stage Cluster Samples Using the Weighted Finite Population Bayesian Bootstrap\***

*Hanzhi Zhou, Michael R. Elliott, and Trivellore E. Raghunathan*

**A Note on Optimum Allocation for Multiobjective Sample Surveys Using Fuzzy Programming\***

*A. K. P. C. Swain*

**Haplotype-based Statistical Inference for Population-based Case-control and Cross-Sectional Studies with Complex Sample Designs\***

*Daoying Lin, Lingxiao Wang, and Yan Li*

**A Propensity-score-adjustment Method for Nonignorable Nonresponse\***

*Minsun Kim Riddles, Jae Kwang Kim, and Jongho Im*

#### **Survey Methodology**

**Reducing Nonresponse and Nonresponse Error in a Telephone Survey: An Informative Case Study**

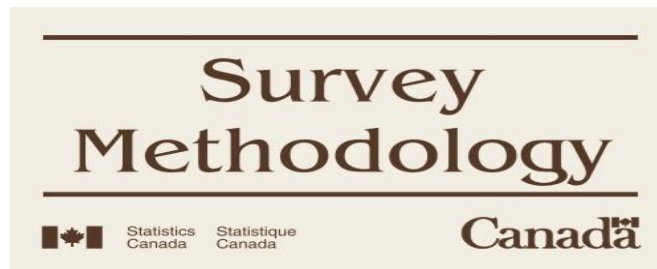
*Floyd J. Fowler, Anthony M. Roman, Rumel Mahmood, and Carol A. Cosenza*

**Biases in Random Route Surveys**

*Johannes J. Bauer*

\*Edited by Joe Sedransk, whose term as editor ended 31 December 2015.





**VOL 42, NO 1 (JUNE 2016)**

<http://www.statcan.gc.ca/pub/12-001-x/12-001-x2016001-eng.htm>

**A generalized Fellegi-Holt paradigm for automatic error localization**

*Sander Scholtus*

**Statistical matching using fractional imputation**

*Jae Kwang Kim, Emily Berg and Taesung Park*

**Comparison of unit level and area level small area estimators**

*Michael A. Hidioglou and Yong You*

**Comparison of some positive variance estimators for the Fay-Herriot small area model**

*Susana Rubin-Bleuer and Yong You*

**A comparison between nonparametric estimators for finite population distribution functions**

*Leo Pasquazzi and Lucio de Capitani*

**A note on regression estimation with unknown population size**

*by Michael A. Hidioglou, Jae Kwang Kim and Christian Olivier Nambeu*

**Register-based sampling for household panels**

*Jan A. van den Brakel*

**Nonresponse adjustments with misspecified models in stratified designs**

*by Ismael Flores Cervantes and J. Michael Brick*

**A short note on quantile and expectile estimation in unequal probability samples**

*Linda Schulze Waltrup and Göran Kauermann*

### **Addendum**

**Model-assisted optimal allocation for planned domains using composite estimation**

*Wilford B. Molefe and Robert Graham Clark*





## Journal of Official Statistics

---

### Volume 32, Issue 2 (June 2016)

<http://www.degruyter.com/view/j/jos.2016.32.issue-2/issue-files/jos.2016.32.issue-2.xml>

#### **On a Modular Approach to the Design of Integrated Social Surveys**

*Ioannidis, Evangelos / Merkouris, Takis / Zhang, Li-Chun / Karlberg, Martin / Petrakos, Michalis / Reis, Fernando / Stavropoulos, Photis*

#### **Discussion**

*Chipperfield, James O.*

#### **Discussion**

*Dolson, David*

#### **Discussion**

*Gonzalez, Jeffrey M. / Eltinge, John L.*

#### **Rejoinder**

*Ioannidis, Evangelos / Merkouris, Takis / Zhang, Li-Chun / Karlberg, Martin / Petrakos, Michalis / Reis, Fernando / Stavropoulos, Photis*

#### **The Impact of Question Format, Context, and Content on Survey Answers in Early and Late Adolescence**

*Diersch, Nadine / Walther, Eva*

#### **End User Licence to Open Government Data? A Simulated Penetration Attack on Two Social Survey Datasets**

*Elliot, Mark / Mackey, Elaine / O'Shea, Susan / Tudor, Caroline / Spicer, Keith*

#### **Interviewer Effects on a Network-Size Filter Question**

*Josten, Michael / Trappmann, Mark*

#### **The FEWS Index: Fixed Effects with a Window Splice**

*Krsinich, Frances*

#### **“Do the Germans Really Work Six Weeks More than the French?” – Measuring Working Time with the Labour Force Survey in France and Germany**

*Körner, Thomas / Wolff, Loup*

#### **Random Walks on Directed Networks: Inference and Respondent-Driven Sampling**

*Malmros, Jens / Masuda, Naoki / Britton, Tom*



**Modernizing a Major Federal Government Survey: A Review of the Redesign of the Current Population Survey Health Insurance Questions**

*Pascale, Joanne*

**Misspecification Effects in the Analysis of Panel Data**

*Vieira, Marcel de Toledo / Smith, Peter W.F. / Salgueiro, Maria de Fátima*

**Weight Smoothing for Generalized Linear Models Using a Laplace Prior**

*Xia, Xi / Elliott, Michael R.*

**Book Review**

*Phipps, Polly A. / Toth, Daniell*

**Book Review**

*Beatty, Paul C.*



---

**VOL 9, NO 3 (2016)**

<http://www.surveypractice.org/index.php/SurveyPractice/index>

**Trends in Telephone Outcomes, 2008 - 2015**

*David Dutwin, Paul Lavrakas*

**Post-stratification or non-response adjustment?**

*Stas Kolenikov*

**Gridlocked: The impact of adapting survey grids for smartphones**

*Ashley Richards, Rebecca Powell, Joe Murphy, Mai Nguyen, Shengchao Yu*

**Random Digit Dialing versus Address-Based Sampling using Telephone Data Collection**

*Bonnie E Shook-Sa, Lauren Klein Warren, David Roe, Douglas Currivan*

**Attrition when dropping CAPI from a CATI/CAPI panel survey**

*Oliver Lipps, Florence Lebert*

**Ranking: Perceptions of Tied Ranks and Equal Intervals on a Modified Visual Analog Scale**

*Jerry Timbrook, William F. Moroney*

**The Use of Vignettes in Evaluating Asian Language Questionnaire Items**

*Mandy Sha*



**Vol. 10, No. 1 (2016)**

<https://ojs.ub.uni-konstanz.de/srm/index>

### **Table of Contents**

**Enhancing Social Surveys Through the Postal Collection of Shed Milk Teeth: an Example of a Large-Scale Cost-Effective Collection on a Longitudinal Study**  
*Samantha Jayne Parsons, Lucinda Platt*

**Two of a Kind. Similarities Between Ranking and Rating Data in Measuring Values.**

*Guy Moors, Ingrid Vriens, John P.T.M. Gelissen, Jeroen K. Vermunt*

**Adaptive Survey Designs for Nonresponse and Measurement Error in Multi-Purpose Surveys**

*Melania Calinescu, Barry Schouten*

**Helping Respondents Provide Good Answers in Web Surveys**

*Mick Couper, Chan Zhang*



## **Statistical Journal of the IAOS: Journal of the International Association for Official Statistics**

**VOL 32, NO 2 (2016)**

<http://content.iospress.com/journals/statistical-journal-of-the-iaos/32/2>

**Interview with Stephen Penneck**

*West, Kirsten*

**A conversation with Dr. Peter Agre about Nobel prizes and science diplomacy**

*West, Kirsten*

**Plans and innovations for the 2020 decennial census of the United States**

*Blumerman, Lisa M. | Bishop, Deirdre Dalpiaz | Dinwiddie, James L.*

**Adaptive design research for the 2020 Census**

*Konicki, Scott | Adams, Tamara*



**An approach for using administrative records to reduce contacts in the 2020 Decennial Census**

*Morris, Darcy Steeg | Keller, Andrew | Clark, Brian*

**Imputation research for the 2020 Census**

*Keller, Andrew*

**Towards a political economy of statistics**

*Heine, Klaus | Olmanns, Erich*

**Official statistical leadership at the crossroads again: An information age perspective**

*Ramasamy, Ramachandran*

***The Māori statistics framework: A tool for indigenous people's development***

*Coutts, Karen | Morris, John | Jones, Ngareta*

**Obtaining numerically consistent estimates from a mix of administrative data and surveys**

*Waal, Ton de*

**Collecting and managing fuzzy data in statistical relational databases**

*Hudec, Miroslav | Praženka, Dušan*

**Iterative method for reducing the impact of outlying data points: Ensuring data quality**

*Jesilevska, Svetlana*

**Comments on four papers on synthetic data in Volume 32 Issue 1 the Statistical Journal of the IAOS**

*Raab, Gillian M.*

***Discussion of the synthetic data papers published in the previous issue***

*Drechsler, Jörg*



**VOL 84, ISSUE 1 (April 2016)**

<http://onlinelibrary.wiley.com/doi/10.1111/insr.v84.1/issuetoc>

**Nalini Ravishanker joins Raymond Chambers as co-Editor-in-Chief of the *International Statistical Review***

*Raymond Chambers*



**A Conversation with Jack Kalbfleisch and Jerry Lawless**

*Chris J. Wild*

**On Invariant Post-randomization for Statistical Disclosure Control**

*Tapan K. Nayak<sup>1,\*</sup> and Samson A. Adeshiyan<sup>2</sup>*

**Alternative Indicators for the Risk of Non-response Bias: A Simulation Study**

*Raphael Nishimura<sup>1,\*</sup>, James Wagner<sup>2</sup> and Michael Elliott<sup>2</sup>*

**Treatment and Control Groups in a Dynamic Setting**

*Götz Rohwer<sup>\*</sup>*

**Calibration Weighting Methods for Complex Surveys**

*Changbao Wu<sup>1,\*</sup> and Wilson W. Lu<sup>2</sup>*

**Reliable Robust Regression Diagnostics**

*Silvia Salini<sup>1,\*</sup>, Andrea Cerioli<sup>2</sup>, Fabrizio Laurin<sup>2</sup> and Marco Rian<sup>2</sup>*

**A Review of Modern Computational Algorithms for Bayesian Optimal Design**

*Elizabeth G. Ryan<sup>1,2,\*</sup>, Christopher C. Drovandi<sup>1,3</sup>, James M. McGree<sup>1,3</sup> and Anthony N. Pettitt<sup>1,3</sup>*

**Genomic Clinical Trials and Predictive Medicine**

*Matthew Schipper<sup>\*</sup>*

**Statistical Learning with Sparsity: The Lasso and Generalizations**

*Carl M. O'Brien<sup>\*</sup>*

**Bayesian Nonparametric Data Analysis**

*Diego Andres Perez Ruiz<sup>\*</sup>*

**Current Trends in Bayesian Methodology with Applications**

*Lili Zhao<sup>\*</sup>*

**Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction**

*Carol Joyce Blumberg<sup>\*</sup>*

**Bayesian and Frequentist Regression Methods**

*Sanjib Basu<sup>\*</sup>*

**John Napier: Life, Logarithms and Legacy**

*David Bellhouse<sup>\*</sup>*

**Analysis of Categorical Data with R**

*Shuangzhe Liu<sup>\*</sup>*

**Meta-Analysis: A Structural Equation Modeling Approach**

*Reijo Sund<sup>\*</sup>*



---

# TRANSACTIONS ON DATA PRIVACY

Foundations and Technologies

<http://www.tdp.cat>

---

**Volume 9, Issue 1, April 2016**

<http://www.tdp.cat/issues16/vol09n01.php>

## TABLE OF CONTENTS

### **Genomic Data Privacy Protection using Compressed Sensing**

*Aminmohammad Roozgard, Nafise Barzigar, Pramode Verma, Samuel Cheng*

### **Differential Privacy Models for Location-Based Services**

*Ehab ElSalamouny, Sebastien Gambs*

### **Evaluating the impact of k-anonymization on the inference of interaction networks**

*Pedro Rijo, Alexandre P. Francisco, Mário J. Silva*

### **On Differentially Private Longest Increasing Subsequence Computation in Data Stream**

*Luca Bonomi, Li Xiong*



**Vol 179 Issue 3 (June 2016)**

<http://onlinelibrary.wiley.com/doi/10.1111/rssa.2016.179.issue-3/issuetoc>

**‘Big data’ and data sharing**

*David J. Hand*

**Political socialization in flux: linking family non-intactness during childhood to adult civic engagement**

*Timo Hener<sup>1</sup>, Helmut Rainer<sup>2</sup> and Thomas Siedler*

**Behind and beyond the (head count) employment rate**

*Andrea Brandolini and Eliana Viviano<sup>\*</sup>*

**A comparison of multiple-imputation methods for handling missing data in repeated measurements observational studies**

*Oya Kalaycioglu<sup>1,2,\*</sup>, Andrew Copas<sup>3</sup>, Michael King<sup>1</sup> and Rumana Z. Omar*

**Mothers' long-run career patterns after first birth**

*Sylvia Frühwirth-Schnatter<sup>1</sup>, Christoph Pamminger<sup>2</sup>, Andrea Weber<sup>3,4</sup> and Rudolf Winter-Ebmer<sup>2,5,\*</sup>*

**Does more balanced survey response imply less non-response bias?**

*Barry Schouten<sup>1,2,\*</sup>, Fannie Cobben<sup>1</sup>, Peter Lundquist<sup>3</sup> and James Wagner<sup>4</sup>*

**Olley–Pakes productivity decomposition: computation and inference**

*Ari Hyytinen<sup>1,\*</sup>, Pekka Ilmakunnas<sup>2</sup> and Mika Maliranta<sup>3</sup>*

**Small area estimation with state space common factor models for rotating panels**

*Jan A. van den Brakel<sup>1,2,\*</sup> and Sabine Krieg*

**Sequence analysis of call record data: exploring the role of different cost settings**

*Mark Hanly<sup>1,\*</sup>, Paul Clarke<sup>2</sup> and Fiona Steele<sup>3</sup>*

**Born leaders: political selection and the relative age effect in the US Congress**

*D. Muller and L. Page, Daniel Muller<sup>1,\*</sup> and Lionel Page<sup>2</sup>*



**The effect of private police on crime: evidence from a geographic regression discontinuity design**

*John M. MacDonald\*, Jonathan Klick and Ben Grunwald*

**J-divergence measurements of economic inequality**

*Nicholas Rohde\**

**Books for review:**

**Data Analysis for Network Cyber-security**

*Sebastian Dietz*

**A Certain Uncertainty: Nature's Random Ways**

*Kuldeep Kumar*

**Probability and Statistics by Example 1: Basic Probability and Statistics**

*Shalabh*

**Graphical Data Analysis with R**

*Andrey Kostenko*

**Exposure–Response Modeling Methods and Practical Implementation**

*Darryl Neil Penenberg*





## Journal of the American Statistical Association



---

**Volume 111, Issue 513 (March 2016)**

<http://www.tandfonline.com/toc/uasa20/110/511>

**A Functional Approach to Deconvolve Dynamic Neuroimaging Data**

*Ci-Ren Jiang, John A. D. Aston & Jane-Ling Wang*

**A Bayesian Partial Identification Approach to Inferring the Prevalence of Accounting Misconduct**

*P. Richard Hahn, Jared S. Murray & Ioanna Manolopoulou*

**Meta-Analytic Framework for Sparse K-Means to Identify Disease Subtypes in Multiple Transcriptomic Studies**

*Zhiguang Huo, Ying Ding, Silvia Liu, Steffi Oesterreich & George Tseng*

**Collective Estimation of Multiple Bivariate Density Functions With Application to Angular-Sampling-Based Protein Loop Modeling**

*Mehdi Maadooliat, Lan Zhou, Seyed Morteza Najibi, Xin Gao & Jianhua Z. Huang*

**Calibrating an Ice Sheet Model Using High-Dimensional Binary Spatial Data**

*Won Chang, Murali Haran, Patrick Applegate & David Pollard*

**Perturbation Detection Through Modeling of Gene Expression on a Latent Biological Pathway Network: A Bayesian Hierarchical Approach**

*Lisa M. Pham, Luis Carvalho, Scott Schaus & Eric D. Kolaczyk*

**A Model-Based Approach to Climate Reconstruction Using Tree-Ring Data**

*Matthew R. Schofield, Richard J. Barker, Andrew Gelman, Edward R. Cook & Keith R. Briffa*

**Theory and Methods:**

**Constrained Maximum Likelihood Estimation for Model Calibration Using Summary-Level Information From External Big Data Sources**

*Nilanjan Chatterjee, Yi-Hau Chen, Paige Maas & Raymond J. Carroll*

**Comment**

*Peisong Han & Jerald F. Lawless*



**Comment**

*Sebastien Haneuse & Claudia Rivera*

**Comment**

*Thomas A. Louis & Niels Keiding*

**Comment: A Human Genetics Perspective**

*Joel A. Mefford, Noah A. Zaitlen & John S. Witte*

**Comment: Addressing the Need for Portability in Big Data Model Building and Calibration**

*Chirag J. Patel & Francesca Dominici*

**Rejoinder**

*Nilanjan Chatterjee, Yi-Hau Chen, Paige Maas & Raymond J. Carroll*

**Instrumental Variables Estimation with Some Invalid Instruments and its Application to Mendelian Randomization**

*Hyunseung Kang, Anru Zhang, T. Tony Cai & Dylan S. Small*

**Generalizing Quantile Regression for Counting Processes with Applications to Recurrent Events**

*Xiaoyan Sun, Limin Peng, Yijian Huang & HuiChuan J. Lai*

**A Potential Tale of Two-by-Two Tables From Completely Randomized Experiments**

*Peng Ding & Tirthankar Dasgupta*

**Ultrahigh-Dimensional Multiclass Linear Discriminant Analysis by Pairwise Sure Independence Screening**

*Rui Pan, Hansheng Wang & Runze Li*

**Constrained Estimators and Consistency of a Regression Model on a Lexis Diagram**

*Wenjiang Fu*

**Statistical Inference in Hidden Markov Models Using  $k$ -Segment Constraints**

*Michalis K. Titsias, Christopher C. Holmes & Christopher Yau*

**Pairwise Likelihood Inference for Nested Hidden Markov Chain Models for Multilevel Longitudinal Data**

*Francesco Bartolucci & Monia Lupparelli*

**Large-Scale Multiple Testing of Correlations**

*T. Tony Cai & Weidong Liu*

**Personalized Prediction and Sparsity Pursuit in Latent Factor Models**

*Yunzhang Zhu, Xiaotong Shen & Changqing Ye*



**Minimax and Adaptive Estimation of Covariance Operator for Random Variables Observed on a Lattice Graph**

*T. Tony Cai & Ming Yuan*

**Estimation and Testing of Varying Coefficients in Quantile Regression**

*Xingdong Feng & Liping Zhu*

**Feature Augmentation via Nonparametrics and Selection (FANS) in High-Dimensional Classification**

*Jianqing Fan, Yang Feng, Jiancheng Jiang & Xin Tong*

**Spline-Lasso in High-Dimensional Linear Regression**

*Jianhua Guo, Jianchang Hu, Bing-Yi Jing & Zhen Zhang*

**Efficient Sequential Monte Carlo with Multiple Proposals and Control Variates**

*Wentao Li, Rong Chen & Zhiqiang Tan*

**Stepwise Signal Extraction via Marginal Likelihood**

*Chao Du, Chu-Lan Michael Kao & S. C. Kou*

**Hierarchical Testing in the High-Dimensional Setting With Correlated Variables**

*Jacopo Mandozzi & Peter Bühlmann*

**Quantile Regression in the Secondary Analysis of Case–Control Data**

*Ying Wei, Xiaoyu Song, Mengling Liu, Iuliana Ionita-Laza & Joan Reibman*

**Variable Selection with Prior Information for Generalized Linear Models via the Prior LASSO Method**

*Yuan Jiang, Yunxiao He & Heping Zhang*

**An Adaptive Exchange Algorithm for Sampling From Distributions With Intractable Normalizing Constants**

*Faming Liang, Ick Hoon Jin, Qifan Song & Jun S. Liu*

**Asymptotically Normal and Efficient Estimation of Covariate-Adjusted Gaussian Graphical Model**

*Mengjie Chen, Zhao Ren, Hongyu Zhao & Harrison Zhou*

**Estimating Variance Components in Functional Linear Models With Applications to Genetic Heritability**

*Matthew Reimherr & Dan Nicolae*

**Extended MaxT Tests of One-Sided Hypotheses**

*Zeng-Hua Lu*



**Volume 103, Issue 2 (June 2016)**

<http://biomet.oxfordjournals.org/content/current>

**Maximum likelihood estimation for semiparametric transformation models with interval-censored data**

*Donglin Zeng, Lu Mao, and D. Y. Lin*

**Nonparametric maximum likelihood estimation for the multi sample Wicksell corpuscle problem**

*Kwun Chuen Gary Chan and Jing Qin*

**Regression analysis of networked data**

*Yan Zhou and Peter X.-K. Song*

**Exact simulation of max-stable processes**

*Clément Dombry, Sebastian Engelke, and Marco Oesting*

**Data augmentation for models based on rejection sampling**

*Lizhen Lin, and David B. Dunson*

**Exponential tilting in Bayesian asymptotics**

*S. A. Kharroubi and T. J. Sweeting*

**Partial least squares for dependent data**

*Marco Singer, Tatyana Krivobokova, Axel Munk, and Bert de Groot*

**Skew-normal antedependence models for skewed longitudinal data**

*Shu-Ching Chang and Dale L. Zimmerman*

**A pairwise interaction model for multivariate functional and longitudinal data**

*Jeng-Min Chiou and Hans-Georg Müller*

**On exchangeable multinomial distributions**

*E. Olusegun George, Kyeongmi Cheon, Yilian Yuan, and Aniko Szabo*

**A pairwise likelihood-based approach for changepoint detection in multivariate time series models**

*Ting Fung Ma and Chun Yip Yau*

**Nonparametric identification and maximum likelihood estimation for hidden Markov models**

*G. Alexandrovich, H. Holzmänn, and A. Leister*



### **Quantile-based classifiers**

*C. Hennig and C. Viroli*

### **Modelling complex survey data with population level information: an empirical likelihood approach**

*M. Oguz-Alper and Y. G. Berger*

### **Calibrated propensity score method for survey nonresponse in cluster sampling**

*Jae Kwang Kim, Yongchan Kwon, and Myunghee Cho Paik*

### **Miscellanea:**

#### **On varieties of doubly robust estimators under missingness not at random with a shadow variable**

*Wang Miao and Eric J. Tchetgen Tchetgen*

#### **Sharp sensitivity bounds for mediation under unmeasured mediator-outcome confounding**

*Peng Ding and Tyler J. Vanderweele*

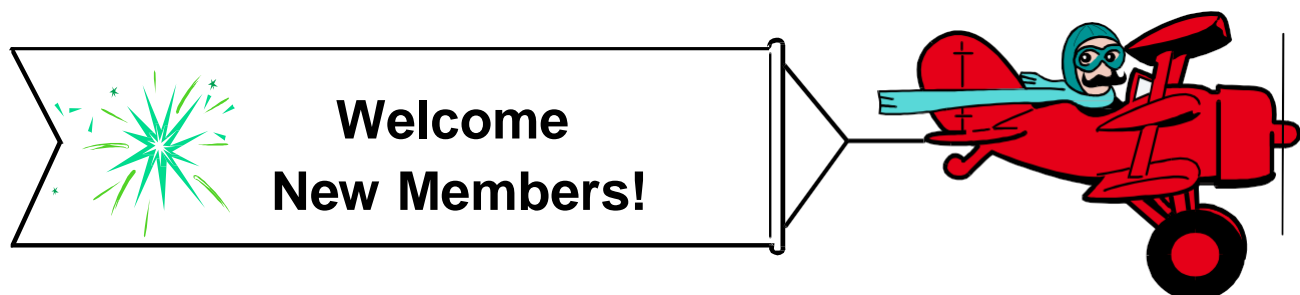
### **Amendments and Corrections:**

#### **‘Clarifying missing at random and related definitions, and implications when coupled with exchangeability’**

*Fabrizia Mealli and Donald B. Rubin*

**Retraction Notice:** Paul R. Rosenbaum  
MEALLI, F. and RUBIN, D. B.





**We are very pleased to welcome the following new members!**

<b>Title</b>	<b>First name</b>	<b>Surname</b>	<b>Country</b>
MR.	Raymond	Cyr	Canada
DR.	Joerg	Drechsler	Germany
DR.	Imoro Ibn Saeed	Bashiru	Ghana
DR.	Smart Asomaning	Sarpong	Ghana
MR.	Ayodele Dauda	Lamidi	Nigeria
DR.	Maciej Eryk	Beresewicz	Poland
MRS	Loveness Nyaradzo	Dzikiti	South Africa
MR.	Achraf	Mrabet	Tunisia
MS	Natalia	Rozora	Ukraine
MR.	Paul	Burton	United States
MR.	Wen	Chang	United States
DR.	Stephanie	Eckman	United States
MR.	Daniel	Guzman Herrera	United States
MS	Lisa	Holland	United States
MR.	Frost	Hubbard	United States
MR.	Andrew	Hupp	United States
DR.	Arthur	Kennickell	United States
DR.	Chris	Moriarity	United States
PROF	Bruce David	Spencer	United States
MR.	Daniel	Zahs	United States
PROF	Zhengyuan	Zhu	United States



# IASS Officers and Council Members

---

## Executive Officers

<b>President:</b>	Steve Heeringa (USA)	<a href="mailto:sheering@isr.umich.edu">sheering@isr.umich.edu</a>
<b>President-elect:</b>	Peter Lynn (UK)	<a href="mailto:PLynn@essex.ac.uk">PLynn@essex.ac.uk</a>
<b>Vice-Presidents:</b>	Monica Pratesi (Italy) Imbi Traat (Estonic)	<a href="mailto:Monica.pratesi@unipi.it">Monica.pratesi@unipi.it</a> <a href="mailto:Imbi.traat@ut.ee">Imbi.traat@ut.ee</a>
<b>Scientific Secretary:</b>	Denise Silva (Brazil)	<a href="mailto:denisebritz@gmail.com">denisebritz@gmail.com</a>
<b>Council Members (2013-2017)</b>	Michael Brick (USA) Daniela Cocchi (Italy) Jack Gambino (Canada) Risto Lehtonen (Finland) Ralf Münnich (Germany) Jean Opsomer (USA)	<a href="mailto:mikebrick@westat.com">mikebrick@westat.com</a> <a href="mailto:daniela.cocchi@unibo.it">daniela.cocchi@unibo.it</a> <a href="mailto:jack.gambino@Canada.ca">jack.gambino@Canada.ca</a> <a href="mailto:risto.lehtonen@helsinki.fi">risto.lehtonen@helsinki.fi</a> <a href="mailto:muennich@uni-trier.de">muennich@uni-trier.de</a> <a href="mailto:jopsomer@stat.colostate.edu">jopsomer@stat.colostate.edu</a>
<b>(2015-2019)</b>	Hukum Chandra (India) Maria Giovanna Ranalli (Italy) Timo Schmid (Germany) David Steel (Australia) Ineke Stoop (The Netherlands) Niko Tzavidis (UK)	<a href="mailto:Hchandra12@gmail.com">Hchandra12@gmail.com</a> <a href="mailto:Giovanna.ranalli@stat.unipg.it">Giovanna.ranalli@stat.unipg.it</a> <a href="mailto:Timo.schmid@gmx.org">Timo.schmid@gmx.org</a> <a href="mailto:dstell@uow.edu.au">dstell@uow.edu.au</a> <a href="mailto:i.stoop@scp.nl">i.stoop@scp.nl</a> <a href="mailto:n.tzavidis@soton.ac.uk">n.tzavidis@soton.ac.uk</a>
<b>Chair of the Morocco IASS Programme Committee 2017</b>	Marcel Vieira (Brazil)	<a href="mailto:Marcel.vieira@ice.ufjf.br">Marcel.vieira@ice.ufjf.br</a>
<b>Chair of the committee for the Cochran Hansen prize award</b>	Risto Lehtonen (Finland)	<a href="mailto:risto.lehtonen@helsinki.fi">risto.lehtonen@helsinki.fi</a>

## Ex Officio Members

<b>ISI Associate Director, Operation Manager / IASS Liaison</b>	Shabani Mehta	<a href="mailto:s.mehta@cbs.nl">s.mehta@cbs.nl</a>
<b>Transition Executive Director:</b>	Catherine Meunier (France)	<a href="mailto:katherine.meunier@orange.fr">katherine.meunier@orange.fr</a>
<b>Director, Executive/Treasurer of the ISI</b>	Ada van Krimpen (The Netherlands)	<a href="mailto:an.vankrimpen@cbs.nl">an.vankrimpen@cbs.nl</a>
<b>Finance Manager:</b>	Michael Leeuwe	
<b>IASS Webmaster:</b>	Mehmood Asghar and Olivier Dupriez	<a href="mailto:odupriez@worldbank.org">odupriez@worldbank.org</a>
<b>ISI Membership Officer</b>	Margaret de Ruiter-Molloy (The Netherlands)	<a href="mailto:m.deruitermolloy@cbs.nl">m.deruitermolloy@cbs.nl</a>



## 2 International Organisations

AFRISTAT  
EUROSTAT

## 23 Bureaus of Statistics

**AUSTRALIA** – AUSTRALIAN BUREAU OF STATISTICS  
**BRAZIL** – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE)  
**CANADA** – STATISTICS CANADA  
**DENMARK** – DANMARKS STATISTIK  
**FINLAND** – STATISTICS FINLAND  
**GERMANY** – STATISTISCHE BUNDESAMT  
**ITALY** – INSTITUTO NAZIONALE DI STATISTICA (ISTAT)  
**KOREA, REPUBLIC OF** – STATISTICS KOREA  
**LUXEMBOURG** – EUROPEAN COMMISSION – EUROSTAT  
**MACAO, SAR China** – DIRECÇÃO DOS SERVIÇOS DE  
 ESTATÍSTICA E CENSOS  
**MALI** – AFRISTAT  
**MAURITIUS** – STATISTICS MAURITIUS  
**MEXICO** – INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA (INEGI)  
**MEXICO** – NUMÉRIKA-MEDICION Y ANALISIS ESTAD. AVANZADO, SC  
**NEW ZEALAND** – STATISTICS NEW ZEALAND  
**NORWAY** – STATISTICS NORWAY  
**PORTUGAL** – INSTITUTO NACIONAL DE ESTADÍSTICA (INE)  
**SWEDEN** – STATISTICS SWEDEN  
**UNITED STATES** – NASS, RESEARCH &  
 DEVELOPMENT DIVISION  
**UNITED STATES** – NATIONAL CENTER FOR HEALTH  
 STATISTICS  
**UNITED STATES** – RESEARCH TRIANGLE INSTITUTE  
**UNITED STATES** - SURVEY RESEARCH CENTER  
**UNITED STATES** – WESTAT INC.



# 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-iass.org/home/services/the-survey-statistician/>