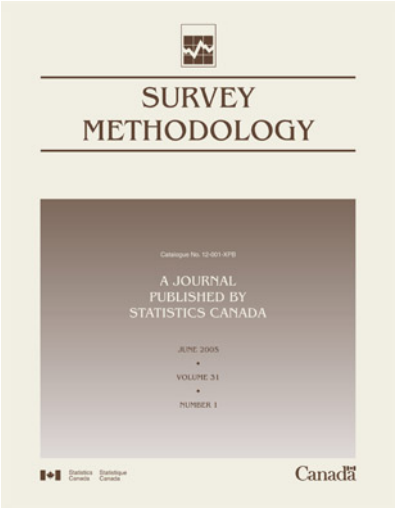
Catalogue no. 12-001-XIE

Survey Methodology

2005



Statistics Canada Statistique Canada

Canada

How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Business Survey Methods Division, Statistics Canada, Ottawa, Ontario, K1A 0T6 (telephone: 1 800 263-1136).

For information on the wide range of data available from Statistics Canada, you can contact us by calling one of our toll-free numbers. You can also contact us by e-mail or by visiting our website.

National inquiries line	1 800 263-1136
National telecommunications device for the hearing impaired	1 800 363-7629
Depository Services Program inquiries	1 800 700-1033
Fax line for Depository Services Program	1 800 889-9734
E-mail inquiries	infostats@statcan.ca
Website	www.statcan.ca

Information to access the product

This product, catalogue no. 12-001-XIE, is available for free. To obtain a single issue, visit our website at www.statcan.ca and select Our Products and Services.

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner and in the official language of their choice. To this end, the Agency has developed standards of service that its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1 800 263-1136. The service standards are also published on www.statcan.ca under About Statistics Canada > Providing services to Canadians.

Statistics Canada
Business Survey Methods Division

Survey Methodology

2005

Published by authority of the Minister responsible for Statistics Canada

November 2005

Catalogue no. 12-001-XIE
ISSN 1492-0921

Frequency: semi-annual

Ottawa

Cette publication est disponible en français sur demande (nº 12-001-XIF au catalogue).

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Vol. 29, No. 1, pp. 1-2
Statistics Canada, Catalogue No. 12-001

In This Issue

This issue of *Survey Methodology* contains the third in an annual invited paper series in honour of Joseph Waksberg. A brief description of the series and a short biography of Joseph Waksberg were given in the June 2001 issue of the journal. I would like to thank the members of the Award Selection Committee, Chris Skinner (Chair), David Binder, Paul Biemer and Mike Brick for having chosen Tim Holt, who has had a very distinguished career in both academia and in official statistics, as the author of the third paper in the Waksberg Invited Paper Series.

In his paper entitled "Methodological Issues in the Development and Use of Statistical Indicators for International Comparisons", Holt first describes the wide range of national statistical indicators suggested by various United Nations committees to monitor and compare development in such areas as demography, health, economics and employment, and he considers how these can be prioritized for implementation. He then discusses the need for sound statistical infrastructure in each country, and the importance of base population estimates, administrative sources of data, and good meta-data for indicators that are produced. Holt goes on to discuss several methodological issues related to the implementation of such indicators, and interpretation of international comparisons.

The next six papers in this issue form a special section on small area estimation. The first three papers present general methodology, while the last three discuss small area estimation methods in more specific contexts.

Meeden presents a new Bayesian approach to small area estimation. Instead of using the usual Bayesian approach that implicitly links one area to another area, Meeden instead uses a noninformative or objective Bayesian approach. It applies a Polya posterior idea to obtain model-based estimates of small area parameters, all without introducing a model or a prior explicitly. One advantage of this approach is that population parameters other than means can be estimated with sensible estimates of their precision.

You, Rao and Gambino approach the problem of estimating unemployment in small domains by using an extension of the well-known Fay-Herriot model by borrowing strength across both areas and time. The authors use the structure of the Canadian Labour Force Survey to produce some interesting variations on this model. They use the short period – 6 months – that rotation groups are in the sample to produce efficient Hierarchical Bayes estimates which neatly avoids the seasonality problem common to designs with longer time periods. The result of this method is large reduction of the coefficient of variation especially in the smaller areas.

In their paper, Lehtonen, Särndal and Veijanen examine the effect of model choice for different types of estimators of domain totals. They point out that earlier literature on small domain estimation has not emphasized enough the distinction between the types of estimators and the model choice. They show analytically and empirically that model improvement has different effects on different estimator types. One of their main results is that, under some conditions, model improvement leads to a larger decrease in mean squared error in smaller domains for the generalized regression estimator. The opposite holds for the synthetic estimator. Also, model improvement is in general more beneficial to the synthetic estimator than to the generalized regression estimator since the former can have a large bias.

Chung, Lee and Kim consider small area estimation using the Korean Economically Active Population Survey. They compare synthetic estimation, a composite estimator that combines the synthetic and direct estimators, and a hierarchical Bayes estimation method based on multi-level modelling. They describe the estimators and the model selection for the hierarchical Bayesian approach. They find that all of these approaches improve significantly over direct estimates for unplanned small areas; however, the composite estimator was best overall.

Di Consiglio, Falorsi, Falorsi and Russo empirically compare several small area estimators using data from the Italian Labour Force Survey to estimate numbers of employed, unemployed, and persons looking for jobs within Local Labour Market Areas. Auxiliary data and target parameters are based on census data. Comparisons are done both conditionally on realized sample sizes within a small area and unconditionally. Several types of small area estimator – expansion, post-stratified ratio, synthetic, composite, sample size dependent, and empirical best linear unbiased predictors – are compared. They conclude that the best estimators overall are a composite estimator and a sample size dependent estimator.

In the final paper of the special section, Harter, Macaluso and Wolter present a case-study of small domain estimation techniques to estimate employment at the county/industry division level using data from the U.S. Current Employment Statistics program and lagged administrative data on employment. They discuss such issues as the availability, quality and choice of auxiliary data, problems in micro-matching of survey and administrative data, and regular monitoring of the entire process in order to build in the quality needed to support small area estimation.

The paper by de Waal deals with the error localization problem: the identification of erroneous fields in erroneous data. A well known method to solve this problem in numerical data is based on vertex generation, in particular the Chernikova algorithm. De Waal extends this approach to identify errors in a mix of categorical and numerical data. The paper shows that many results for numerical data also hold true for a mix of categorical and numerical data. This paper provides a nice readable introduction to Error Localization and its implementation.

Haziza and Rao discuss the problem of unweighted imputation for missing survey data. They show that unweighted imputation, unlike weighted imputation, generally leads to biased estimators under the design-based approach (*i.e.*, uniform response). They propose a bias-adjusted estimator which is simple to obtain and has the desirable property that it is approximately unbiased under both the design-based and the model-based approaches. They also derive linearization variance estimators for the proposed estimators. A simulation shows the good performance of the bias-adjusted estimator, especially when the correlation between the variable of interest and the inclusion probability is high.

The paper by Johnson and Deely develops optimal and approximately optimal fixed cost sampling allocations for simultaneous estimation in multiple independent Poisson processes based on the Bayes risk and the Bayes estimator under two different loss functions. The results from this approach are straightforward, interesting and are connected to the classical stratified random sampling allocations. Techniques for finding "representative" conjugate priors, under more general hierarchical models for allocation purposes are also presented.

In the last paper of this issue, Tracey, Singh and Arnab investigate calibrating to the second order moment of a auxiliary variable, when available, to improve the efficiency of estimators. They show that this new estimator can be more efficient than the combined regression estimator in stratified sampling and provide a variance estimator for the new estimator. Finally, they extend the method to double sampling and conclude with some limited simulation results.

Finally, we note that a paper from the December 2002 issue of this journal has just won an award. The paper by Balgobin Nandram, Geunshik Han, and Jai Won Choi, entitled "A Hierarchical Bayesian Nonignorable Nonresponse Model for Multinomial Data from Small Areas", has received the Statistical Science Award as the best paper of the year in applied statistics, awarded by the Statistical Awards Ceremony Committee of the Centers for Disease Control and Prevention and the Agency for Toxic Substances and Disease Registry. Congratulations to Drs. Nandram, Han and Choi!

M.P. Singh