

## CLOSING REMARKS

**Michael Hidirolou, Statistics Canada**

It is my pleasure to bring to a close this year's symposium on "Modelling Survey Data for Social and Economic Research".

The symposium started out on Wednesday with two workshops concerning the analysis of survey data collected from complex surveys. The workshop given by Barry Graubart, entitled "Analysis of Complex Survey Data with Applications to Health Data" covered a number of issues related to the analysis of survey data. The second workshop given by Roland Thomas and Bruno Zumbo entitled "Item Response Theory with Application to Complex Sample Surveys" covered issues linked to basic measurement concepts, and the analysis of data from complex surveys. Both workshops were very well attended with over 150 registrants.

Our keynote speaker, Sir David Cox, opened the symposium on Thursday. We would like to thank Professor Cox for sharing with us some interesting insights on the analysis of data using graphical chain models to explain possible causal interpretation. Symposium 2002 was quite well attended. Over 400 participants from a wide number of countries contributed to its success. Forty-three papers were presented in 12 sessions.

Data analysis methodology for samples drawn from infinite populations has evolved rapidly during the last forty years. The proper analysis of survey data, however, started with H.O. Hartley's paper in 1959 where he compared domain means. During the sixties, seventies, and eighties many studies examined the effect of sampling design on the common types of statistical analyses. Some of early papers dealing with regression analysis are Konijn (1962), Kish and Frankel (1974), and Fuller (1975). The effect of design effect on chi-square was studied by Nathan (1969,1975), Holt, Scott and Ewing (1980), Fellegi (1980), and Rao and Scott in a series of papers written in the eighties. Fairly general methodology for computing variances of complex estimates from survey designs was given by Binder in 1983. The computerisation of some of these theoretical developments started in the seventies. Some of the computer software dedicated to the analysis of survey data included: CLUSTERS, OSIRIS, SUPER CARP, SUDAAN, and NASVAR.

This brings us to 2002. We have covered quite a bit of ground in these last three days on the topic of modelling survey data for social and economic research. We have heard a wide variety of topics concerning the modelling of survey data. The papers can be grouped under three main headings, namely (i) Theoretical developments; (ii) The application of the existing theory to a number of social and economic surveys; and (iii) The current status of survey analysis software. In terms of theoretical developments, mentioning only a few topics, we have seen how to account for survey design when small area data are analysed, how survey statisticians have started to account for the impact of imputation on data analysis, the adaptation of well-established longitudinal analysis techniques to survey data, as well as trying to undo the complexity of survey data. We have also seen a number of applications of the theory to complex surveys such as social and economic data. We have also heard several fine examples of how survey analysis software is evolving with the evolution of the methodology for the analysis of survey data. This software included STATA, the inclusion of the SURVEYLOGISTIC Procedure in SAS, WESVAR, and SEVANI.

As mentioned by David Binder in his opening remarks, the symposium fulfilled our expectations that there had been a significant number of developments in the area of data modelling and analysis.

I would now like to thank the organizers of this symposium for having such a highly successful meeting. In total, there are probably around 30 people who were involved in the organization in one way or another in this symposium. The main organizing committee, headed by Mary March, included Jean-Francois Bastien, Joseph Duggan, Annette Everett, Milorad Kovacevic, Denis Lemire, Sylvain Perron, Georgia Roberts, Adam Wronski, and Xuelin Zhang. I would like to thank all of the numerous volunteers who worked hard on this event. I would like to recognise the useful assistance of the support staff in the Statistics Canada Methodology Branch, as well as the

various services at Statistics Canada, Michel Beaulne and his very capable systems group, and the Fairmont Chateau Laurier Hotel banquet staff. I would like to thank our interpreters for two days of solid work. Next, thanks to our international visitors (some 100) who took the time to come and share their experiences with us. Last but not least, thanks to all our presenters for all the hard work spent for preparing the material and making this symposium possible.

Next year's symposium is entitled, Challenges in Survey Taking for the Next Decade. Douglas Yeo is the Chair of the Organising Committee for that conference.