International Association of Survey Statisticians

THE SURVEY STATISTICIAN

Association Internationale des Statisticiens D'Enquetes
(AISE)

No.30

May 1994

Contents

Introduction 1

General Assembly Council 1

Reports from Executive Director, President, IASS Programme Committee, IASS Nominations Committee, IASS Scientific Secretary 2-7

Special Articles

"Which Survey Should Be Conducted As A Post-Censal Survey - The Israel Experience" by Malka Kantorowitz 7

"The 1992-93 New Zealand Household Health Survey" by Denise Brown 10

"International Blaise Users Conference" by Tony Manners 12

"Measurement Errors and Survey Data Analysis" by Chris Skinner 15

"Using Survey Research Skills in Program Evaluation: A Study of the Quality of Student Financial Aid Programs" by Benjamin F King and Ronald S Fecso 19

Question/Answer Column 20

Country Reports 22

Introduction

I must begin this edition of the Survey Statistician with my profuse apologies for the long delay which was largely due to my workload. Initially the delay was caused by an acute shortage of material for inclusion in the Survey Statistician but colleagues (most notably Anders Christianson and Lilli Japec) at Statistics Sweden came to the rescue and persuaded a number of people to submit articles. This is very gratifying and I do hope that having been prompted in this way IASS members will get into the habit of submitting material (short articles, items for the diary, news, questions to be answered by the IASS team of experts etc.).

The copy dates are:

30 April and
31 October each year

Material should be sent to:

Denise Lievesley
ESRC Data Archive
University of Essex
Wivenhoe Park
Colchester
Essex C04 3SQ, UK

tel: 44 206 872009
fax: 44 206 872003
e-mail: denise@essex.ac.uk

It makes my life much easier if this can be sent in electronic form especially for longer contributions.

Please feel free to contact Denise Lievesley if you would like to discuss possible contributions.

****

Survey of Members

The IASS Council has decided to conduct a survey amongst all present and recent members to find out their views on the role and activities of the association. The survey will be administered by Gösta Forsman with student assistance from the Department of Statistics, University of Linköping, Sweden. Please participate in this survey and take the opportunity to inform the council of your views regarding the future of the IASS.

******

1993 ISI Session
Florence Italy

These are abbreviated versions of the minutes of the IASS General Assembly and Councils, and Reports of the various IASS offices. Full versions of these reports may be obtained from Anne Marie Vespa-Leyder at the IASS Office, c/o INSEE, 18 boulevard Adolphe-Pinard, 75675 Paris Cedex 14, France.

Summary of General Assembly Report

The outgoing President announced the results of the mail ballot for the election of the IASS Officers and Council members (these results were given in the last issue of the Survey Statistician). Bernard Grais has been replaced by Alain Charraud as Executive Director. He and Anne Marie Vespa-Leyder (who has left INSEE) work at the Statistical Division of the French Ministry of Labour.

INSEE and the Ministry of Labour were thanked for their support, and Graham Kalton and Bernard Grais were thanked for their valuable service to the Association.

The IASS membership is in decline so a number of ways of recruiting and promoting membership were considered. It has been agreed that the Association will make similar arrangements for reduced price subscriptions to the Journal of Official Statistics, as those which exist for Survey Methodology.

The IASS will co-sponsor the conference to be held in Bristol in April 1995 on Survey Measurement and Process Quality. It was decided to strengthen the relationships with IAOS, to conduct a membership survey (see earlier on this) and to strengthen the network of country representatives.

Extracts from the Reports of the Council Meetings

Appointments

The following appointments were made: Alain Charraud as Executive Director and Lars Lyberg and Kirk Wolter as auditors.

Workshops/conferences

Chris Skinner, the new incoming Scientific Secretary, will begin to plan workshops to be held alongside the next ISI meeting in Beijing. It is also hoped to assist the IAOS in their conference to be held in Zimbabwe in October

1994, and the possibility of organising an IASS conference in Hungary in 1996 was raised.

Publications

The changes to the Survey Statistician which have been reported in the previous issue were discussed. Kirk Wolter will continue to be responsible for the IASS information in the ISI Newsletter and it was agreed that we should try to fill the allocated two pages each issue.

Declining membership

The Council decided to form a committee that should produce a plan of action to increase the membership. The Council urged the Committee to work fast. The Committee members are:

Lars Lyberg (Chair)
Alain Charraud
Gösta Forsman
Fred Smith

Country Representatives

The Council appointed Lars Lyberg and Dennis Trewin to coordinate the network of Local Representatives.

Finances

The Council decided that the dues should remain the same, i.e. FF 130 (or FF 65 for members from developing countries) until the 1995 ISI meeting.

Statutes

The Council appointed a working group to look over the current statutes. The aim of this group was to make sure that the list of candidates and their nationalities are in accordance with articles 6.3 and 6.15 of the IASS statutes. Five council members were appointed to examine the statutes: F Smith (Chair), V Verma, D Trewin, A Marton, C Skinner. They produced a suggested change to the statutes which the Council accepted. The suggested changes will be sent to the IASS members before the next General Assembly.

Chairpersons

1997 Programme Committee Chair:
A Christianson, Sweden,
Nominations Committee Chair:
H Hogan, USA

Abbreviated Report of the Executive Director Bernard Grais

Activities of the Secretariat

This report begins by outlining the role and responsibilities of the IASS Secretariat. This comprised, in addition to Bernard Grais,

- Alain Charraud, who had acted as Deputy Executive Director with responsibility for managing the IASS business in the Paris office upon M. Grais' move to Luxembourg
- Mrs Anne Marie Vespa-Leyder who acted as Executive Secretary and was Editor of the bulletin
- Ms F Anselmi, the Secretary
- Mr F Fabre who manages the computerized file of IASS members.

Most of the time of these individuals is given voluntarily but INSEE makes the time of Ms Anselmi available to the Association as well as providing many free services.

Sincere thanks are due to Mr Champsaur, Director General of INSEE, Mr Franchet, Director General of EUROSTAT and Mr Seibel, Director of the Statistical Division of the French Ministry of Labour, for their support.

Publications

1. The Association's main publication is its official bulletin, the "Survey Statistician", published twice a year in English and French by the Secretariat. Translation is kindly provided as a free service by Statistics Canada who we wish to thank.

   In addition, the Spanish Institute of Statistics provides a free Spanish version of the journal, called the "Estadistico de Encuestas", sent to Spanish-speaking members of the Association, and Ms Carmen Arribas is in charge of this task. We thank this Institute and Ms Arribas for this assistance.

   Since the Cairo session, 4 issues of the bulletin have been published, issue No.26 containing the minutes of the Council and General Assembly meetings and the reports submitted to the Assembly.

   The Council, in the Paris meeting, had asked

Tim Holt to review the Survey Statistician and to identify ways in which it might be changed to serve better the IASS members. Following the discussion of his report in the Cairo meeting, a number of recommendations were made, among them a wish for a greater emphasis on the scientific content of the bulletin.

2. Since the Amsterdam Session (1985), Statistics Canada's bilingual review, "Techniques d'enquêtes/Survey Methodology" (2 issues a year) has been adopted as the Association's technical publication. Members are entitled to a preferential subscription rate which is reduced by half for members from developing countries, the difference being made up by the Association.

   A new procedure was adopted in agreement with Statistics Canada: the list of subscribers is sent, two or three times a year to Statistics Canada, together with the payment of the corresponding subscription fees. This procedure is satisfactory: it alleviates the burden on the Secretariat and it improves co-ordination.

3. The Secretariat sends the ISI Newsletter published 3 times a year, to all members of the Association who do not also belong to the ISI. In each issue a column is devoted to IASS. Thanks to the co-ordination carried out by Kirk Wolter, and following the recommendations of Tim Holt, the information content of this column has been greatly improved, at the same time that the burden on Ms Vespa-Leyder was alleviated.

4. As a follow up to the Cairo Session, the Summary Booklet of the invited papers presented at the IASS sessions or at the ISI/IASS joint sessions, was sent to the members.

5. A new edition of the Association's Membership Directory has been prepared.

Membership

As in mid June 1993, there were 1030 members registered, of which 264 were also members of the ISI, therefore entitled under present regulations to be members of IASS without payment of dues (unless belonging to more than one section and choosing to pay the IASS fee).

The geographical breakdown was as follows:

- Sub-Saharan Africa	104
- North Africa and Middle East	45
- Asia (except Middle East)	169
- Oceania and Pacific	49
- North America (excluding Mexico)	219
- Latin America and Caribbean	91
- Western Europe	334
- Eastern Europe	19
Total	1030

There were also 16 institutional members, most of them being national statistical institutes. These organizations deserve our appreciation for supporting the Association through the payment of dues amounting to FF 2000 (reduced to FF 800 for developing countries).

It is worth noting the excellent participation of Western Europe and North America; the lower participation of Africa and Latin America; and more particularly the very low figure for Eastern Europe.

A sharp decline in membership has been experienced since June 1991: the number of registered members on file is approximately 18% less than before the Cairo session. However, it is difficult to make an exact comparison, since a member is a person who has paid his dues at least once for the last 3 years. Those who are also members of ISI are considered to be IASS members without any payment of dues in some cases (in accordance with the statutes). Therefore, the apparent change in membership is affected by the dates on which the file has been purged for non payment of dues.

The sharp decrease in membership is confirmed by the figures on the receipts from dues, which have fallen by 14% in 1992 compared to 1990.

The reduction in membership is most significant for developing countries: between June 1990 and June 1993 it is 52% in Sub Saharan Africa, 34% in Latin America and Carribean, 20% in Asia and 8% in North Africa and Middle East. But, there is also a decrease in developed countries: 7% in North America and 5% in Western Europe. Only, Oceania and Pacific experience a big increase in membership (+ 49%) presumably due to the efforts of our Australian colleagues in recruiting new members. The number of members in Eastern Europe stays stable at a very low level.

This sharp decline in membership must be a warning for us: the membership figures are now lower than before the Paris session in 1989. This decline can be due to some extent to the world recession. But efforts to keep former members, to recruit new ones and to better meet their needs should be undertaken and pursued vigorously. A special attention should be paid to the enlargement of membership from developing countries and Eastern Europe.

Financial report

The accounts for 1991 and 1992 are the following. As for 1989 and 1990, in order to improve their understandability, we present the results and a balance sheet for each year:

Results for 1991 and 1992

Receipts	1991	1992
- Membership fees (including subscriptions to Survey Methodology)	103252.63	94400.53
- Resale of travellers cheques (Cairo)	19121.26	
- Royalties (proceedings of the International Conference on telephone Survey Methodology)	883.85	392.85
- Royalties (proceedings of the International Conference on Measurement Errors in Surveys)	-	1843.93
Total	123257.74	96637.31

EXPENSES	1991	1992
- Secretariat current expenditure	946.06	682.19
- Presidential expenditures	1897.22	
- Secretariat exceptional expenditures (Cairo)	1593.10	6353.52
- Travel and meetings (Cairo)	23393.00	9536.43
- Purchase of travellers cheques (Cairo)	21816.20	
- Subscription to Survey Methodology (1990,1991,1992)	9690.51	33701.51
- ISI Newsletter	20213.55	19019.12
- Reimbursement to the ISI of the Philippines Statistical Office contribution		2092.69
Total	79549.64	71385.46
RESULTS	+ 43708.10	+ 25251.85
- Purchase of securities	187073.52	
Excess of receipts over expenses	- 143365.42	25251.85

Balance sheet on December 31 1991

ASSETS		LIABILITIES	
Securities	428758.98	Reserve	476024.66
Bank	115043.44	Provision for Cairo debts	12000.00
Post Office Account	34337.16	Provision for Survey Methodology subscriptions	16000.00
		Results 1991	43708.10
		Appreciation on securities	30406.82
Total	578139.58		578139.58

Balance sheet on December 31 1992

ASSETS		LIABILITIES	
Securities	469066.72	Reserve	578129.58
Bank	130091.97		
Post Office Account	44530.48		
		Results 1992	25251.85
		Appreciation on securities	40307.74
Total	643689.17		643689.17

On the receipts side, the main issue is the sharp decline, already mentioned, in the income from dues, which decreased respectively by 6.4% and 8.6% in 1991 and 1992 compared to the previous year. The income coming from our participation in the surplus of conferences and royalties remain at a relatively low level.

In the membership fees are included the subscriptions to Survey Methodology. The members are entitled to a preferential rate (FF 150) which is reduced by half for developing countries, the difference being covered by the Association. For 1991 and 1992, the Association collected respectively FF 17325 and FF 13875 for the subscriptions to Survey Methodology. These sums have been transferred to Statistics Canada, and the IASS has paid in addition FF 4125 and FF 3825 as subsidies to members from developing countries for the subscription.

On the expenses side, the Secretariat exceptional expenditures are due to the participation in the ISI biennal Session in Cairo (1991).

The participation in the ISI Newsletter appears to be very expensive. The cost in 1991 and

1992 is twice the amount of 1989 and 1990. I suggest that this point be discussed with the ISI Permanent Office.

The policy of investment of the reserve fund in bonds and monetary funds has been pursued. The returns from these funds are automatically incorporated in the capital and do not appear in the receipts. They represent the major part of the appreciation on securities in the balance sheets. For the 2 years, the appreciation of the invested funds has been FF 70714.56. These returns, which are of the same order of magnitude as the results themselves, are expected to decrease in the future as the interest rates are falling in Europe.

******

General Report from The President
Graham Kalton

Introduction

During the past two years, the Association has made some progress in developing the services it provides to its members. A start has been made on expanding the Survey Statistician, the coverage of the IASS booklets of papers relating to survey research from the ISI sessions was expanded with the Cairo session to include a volume of contributed papers, a series of short courses has been organized in conjunction with the ISI Session in Florence, and the Association continues to have a healthy scientific program both at the ISI Session and at other conferences.

Survey Statistician

It was agreed at the ISI Session in Cairo that the Association should make changes to the Survey Statistician to expand its contents and improve its value to the membership. Denise Lievesley has kindly undertaken to act as Editor of the Survey Statistician to implement the changes, and Gordon Brackstone has kindly agreed to serve as Associate Editor for the Country Reports. Professor Lievesley has outlined the changes that are planned in Issue no.28. The development of the Survey Statistician will necessarily take place incrementally. Some of the changes have been introduced in Issue no.29.

Short Courses at the University of Padua

The Association organized a program of six short courses to be held at the University of Padua from August 21-25, 1993, just prior to the ISI Session in Florence. In the event, two of these courses had to be cancelled. Those given were:

- Workshop on Survey Sampling in Developing Countries
  Instructors: Graham Kalton, Colm O'Muircheartaigh, Nanjamma Chinnappa and Anis Maitra

- Variance Estimation in Complex Surveys
  Instructors: Wayne A Fuller, Michael Hidiroglou, Kirk Wolter

- Survey Measurement: Methods for Error Evaluation and Reduction
  Instructors: Paul P Biemer, Barbara H Forsyth and Judith T Lessler

- Environmental Sampling, Statistical Modeling and Observational Economy
  Instructors: S D Gore, G P Patil, A K Sinha, C Taillie, S K Thompson and G Lovison.

The Association is deeply indebted to its Scientific Secretary, Dr Kirk Wolter, for his unstinting efforts in setting up, publicizing and organizing these courses, and to Professor Luigi Fabbris of the University of Padua for making all the local arrangements. The Association also acknowledges the generous contributions made by the instructors in teaching these courses.

Conferences

The Association has participated in two successful conferences in the past two years:

- International Scientific Conference on Small Area Statistics and Survey Designs, Warsaw, September 30-October 3 1992. This conference was jointly organized by IASS, the Central Statistical Office of Poland and the Polish Statistical Association. Sponsorship for the conference was provided by EUROSTAT, the Central Statistical Office of Poland, and the Polish Statistical Association. Proceedings of the conference have been published by the Central Statistical Office of Poland.

- International Conference on Establishment Surveys, Buffalo, New York, June 27-30 1993. IASS was one of several co-sponsors for this conference. Other co-sponsors were the Sections on Survey Research Methods and on Business and Economic Statistics of the American Statistical Association, Statistics Canada, the Statistical Society of

Canada, the National Association of Business Economists and the American Agricultural Economists Association. A goal of the conference was to produce a monograph that provides a comprehensive review of the methodology for surveys of business, farms and institutions.

ISI Newsletter

The IASS uses its space in the ISI Newsletter to inform its members of conferences and other current activities that are likely to be of interest to them. The Association is grateful to the Scientific Secretary, Kirk Wolter, for providing the IASS material for the ISI Newsletter.

Committee Reports

Nanjamma Chinnappa, Chair of the 1993 Program Committee, and her Committee members are to be congratulated for organizing an excellent set of meetings for the Florence Session. Len Cook is Chair of the 1995 Program Committee. His committee is now actively engaged in organizing the program for the Beijing Session.

Susan Linacre, Chair of the Nominations Committee, and her committee members are to be thanked for their efforts in developing an excellent state of candidates to stand in the elections for officers of the Association.

IASS Secretariat

The IASS is extremely grateful to INSEE for providing a home for the IASS secretariat and for its other contributions to the Association. INSEE supports the IASS secretariat and the Survey Statistician, and produces the booklets of relevant papers from the ISI sessions which it distributes to IASS members. INSEE's generous support contributes greatly to the services that the Association provides to its members.

Bernard Grais, who has served the Association for the past four years as Executive Director, moved to take up a position in EUROSTAT at Luxembourg earlier in the year. As a result of his move, he is unable to continue as Executive Director after the Florence Session. His dedication and valuable contributions to the Association will be missed. We wish him well in his new position.

In view of Mr Grais' move, Alain Charraud was invited and has agreed to serve as Deputy Executive Director of the Association. Mr. Charraud is Associate Director in the Statistical Division of the Ministry of Labor.

We are grateful to him for stepping in at short notice to manage the IASS secretariat in the past few months.

Mrs Anne Marie Vespa-Leyder continues in her position as Executive Secretary. She has moved from INSEE and is now working in the Statistical Division of the Ministry of Labor with Mr Charraud.

Thanks are due to her for her dedicated service to the Association. She plays important roles in the distribution of the Survey Statistician, in setting up and distributing the booklets of ISI papers, in managing the IASS directory, and a variety of other IASS activities.

Conclusion

I should like to record my thanks to the many members of the Association who have helped me with IASS matters during my term of office. I should especially like to record my appreciation of the work of Kirk Wolter who, as Scientific Secretary, contributed a good deal of his time to the Association.

I should also like to thank the many members of IASS who have contributed to the Association in a wide variety of ways - through committees, serving as local representatives, contributing to the Survey Statistician, organizing conferences, etc. The success of the Association depends on the active participation of its members.

It has been an honor and privilege for me to have served as President of the IASS. I hope that I will be able to continue to serve the Association in the future.

******

Abbreviated version of the Report from the IASS Programme Committee
Nanjamma Chinnappa

This committee was formed in March 1990 with representation from developing and developed countries, academia, official and private organisations.

Suggestions for topics and organisers were invited from all IASS members, and a survey was conducted (with a response from 166 members) through country representatives. Following consultation with other sections of ISI about co-sponsoring topics and an analysis of

the contributions of members, a list of 30 highly ranked topics and 29 additional topics was drawn up. The programme committee evaluated this list based on the availability of organisers and speakers, popularity of the topic, whether there was new and useful work and whether it had been covered in recent meetings. A balance was sought between applications and theory and between topics of interest to developing and developed countries. This list was discussed at Cairo and a final list of 10 topics for invited paper meetings and 10 for organised contributed paper meetings was agreed. Two invited paper sessions were named in honour of the memory of Morris Hansen and M N Murthy to commemorate their contributions to the theory and practice of sample surveys.

******

Abbreviated version of the Report from the IASS Nominations Committee
Susan Linacre

The nominations committee was formed following the Cairo session. In establishing the committee an attempt was made to ensure good representation across countries and across academic and national and international statistical agencies. Nominations received from individual members of the IASS together with those of the nominations committee itself totalled

8 for President-elect
13 for Vice-President
8 for Scientific Secretary
55 for Council Members

The nominations committee ranked members and through this process shortlisted 2 nominations for President-elect, 4 for Vice-President, 2 for Scientific Secretary and 9 for Council Members. Despite a wide array of potential candidates being put forward the shortlist was disappointingly under representative of females and of some geographic areas. Continued efforts to improve the representation are required.

******

Abbreviated version of the Report of the IASS Scientific Secretary
Kirk Wolter

In view of the strong support of members for the proposition that IASS should do more to provide specialised training opportunities, this became a priority of the Scientific Secretary. Six courses were planned of which four took place in conjunction with the ISI session in Florence. The IASS is very grateful to Professor Luigi Fabbris for his assistance in arranging these courses at Padua. 65 people from a wide range of countries registered in advance for one or more courses.

The registration fees were designed to encourage participation by students, by statisticians from developing countries and by statisticians in more than one course. Funding was sought to assist those from developing countries and countries in transition to attend the courses and some was provided by EUROSTAT and by the ISI National Organising Committee.

******

Which Survey Should Be Conducted As A Post-Censal Survey - The Israel Experience

Malka Kantorowitz*
Israel Central Bureau of Statistics

Several types of surveys can be classified as Post-Censal Surveys (PCS), whose purpose is to investigate certain areas not covered in-depth in the census itself. For certain surveys, whose target population covers only a subgroup of the total population, a census is the only source that can be used as a screening tool to form a suitable sampling frame. For other surveys, even if relating to the whole population, a census can be used as a frame to achieve a more efficient sample, e.g. for disproportionate stratified sampling. Sometimes, individual data from the census can be used for a PCS to avoid asking the same questions again. It is necessary, of course, to design the census so as to provide suitable data for these purposes.

A number of PCS's were conducted in Israel in the past: Survey of Vacant Dwelling Units (1983); Surveys of Graduates of Higher Education (1961, 1974, 1984); Survey of Households with Persons Aged 60 and Over (1985); and Travelling Habits Surveys (1972/3 and 1984/5). The next census in Israel will take place in 1994 and a number of PCS's are under consideration. The decision of which surveys should be conducted as a PCS will be based on previous experience.

Conducting a large number of PCS's simultaneously, as close as possible to a

census, strains the executive capabilities of the organization (especially if it is a small one) and may harm the quality of the surveys. Thus, it is desirable to limit PCS's unless their advantages are significant. Census deficiencies and the time-gap between the census and the survey can affect the accuracy and efficiency of the various PCS's, depending on the extent of: (a) response, recording or coding errors in the census relating to the required data; (b) item non-response in the census for the relevant variables; (c) out-dated data, if they change over time, such as education; (d) initial under-coverage in the census; (e) under-coverage of certain groups, e.g. new immigrants, because of the time-gap.

Survey of Vacant Dwelling Units - In 1983, a survey was conducted two months after the census in order to obtain estimates of the extent and characteristics of vacant dwellings. The survey was conducted as a PCS since the only way of screening vacant dwellings, comprising 7% of all dwellings, was by using the census.

To ensure as complete coverage as possible, the sampling frame included all dwelling units for which no questionnaire had been received in the census, i.e. not only dwellings which were recorded in the census as vacant, but also "closed" dwelling and those occupied by non-residents (diplomats or tourists), or used for businesses. Despite this extension, which increased the rate of "out of scope" units in the sample, under-coverage still remained because of vacant dwellings not recorded at all in the census, or because occupancy of dwellings is dynamic, even within a short time. In this survey, the only stratification was geographical regions, so there was no mis-stratification problem.

To conduct surveys of this kind in future the only reasonable way is as a PCS and it is even recommended to consider incorporating the survey in the census itself, for the benefit of both the survey and the census.

Surveys of Graduates of Higher Education
Three surveys of this type, with some variations, were carried out in Israel as PCS's in 1961, 1974 and in 1984. For these surveys, it is necessary to use suitable screening variables from a census to form the sampling frame since no other frame is available. In 1984, for example, the survey population contained about 16% of all adults (aged 20 and over).

Misclassification relating to education characteristics due to census deficiencies could cause both biases and an increase in sampling variances. Kantorowitz & Nathan showed that response errors to questions on education are very common ("The Estimation of Response Error Micro-Effects from Repeated Surveys for Invariant Characteristics", Proceedings of the Third ARC, 1987, pp.359-390).

Since one of the primary aims of these surveys is to obtain estimates of the size of various groups, reducing under-coverage is especially important. Therefore the frame of the 1984 survey, for instance, included all persons with at least one positive indicator of higher education (by means of three questions included in the long questionnaire administered to the 20% census sample - highest degree obtained, last school attended, and number of years of study). Persons who reported an occupation that indicated possible higher education and also all non-respondents were included in the frame. To limit under-coverage arising from the time-gap between the census and the survey (15-23 months), persons reporting in the census to be studying in universities or in post-secondary schools were also included.

The sample was designed so that sampling probabilities were inversely proportional to the expected proportion of persons belonging to the population in each stratum, in order to reduce the number of "out of scope" cases in the sample. Marginal groups were sampled with relatively low probabilities, causing "migrants" into the boosted strata, thereby increasing sampling variances. This was the price paid for the successful attempts to reduce biases by covering most of the survey population, thus decreasing the Mean Square Error.

It was found that if the screening was based only on highest degree, coverage would be 82% for academic graduates and only 63% for post-secondary graduates. Adding last school attended would increase coverage by 9% and 12% respectively. The additional contribution of years of study would be very important for the post-secondary (15%) and less (only 2%) for the academic population. The "unknown" and those reporting on "other" schools or degrees, contributed 6% and 9% to the two groups.

Despite the extension of the frame, under-coverage still existed: (a) under-coverage in the census of the relevant population was probably higher than the overall estimated 2.6% since

the rate of temporary absence overseas is higher among the survey population; (b) the frame did not include immigrants and returning Israeli citizens, with a relatively large proportion of persons with higher education, assessed to be about 3% of the survey population;
(c) under-coverage arising from out-dated information on education was partially reduced by including those studying at the time of the census, however some graduates of post-secondary schools could have both commenced and completed their studies between the census and the survey. Altogether, the under-coverage in the survey was estimated to be at least 7% and, for certain sub-groups could be considerably higher.

Surveys similar to the Survey of Graduates of Higher Education must be conducted as PCS's, in the absence of an alternative frame. Screening deficiencies can be reduced only by improving the screening questions in the census and extending them, as will be done in the next census. Nevertheless, the inclusion of ineligible cases in the frame is inevitable, while more marginal groups should be sampled with decreased probabilities. Efforts should be devoted to include also immigrants and returning Israelis and to conduct the survey as close as possible to the census.

Survey of Households with Elderly Persons (Aged 60 and Over)
A comprehensive survey was conducted in 1985 as a PCS, the survey population constituting about 30% of all households. Other small scale surveys were carried out previously as an annex to the current Labour Force Survey. The use of the census for screening of the relevant households was reasonably good and cheap. An efficient sample could be designed as a non-proportional stratified sample to meet the aims of the survey, e.g. strata of households with persons in the higher age-groups were boosted.

Nevertheless, under-coverage still remained (even after using supporting information from the Population Register), estimated to be at least 5%. This was due to response and recording errors in census age, due to census under-coverage, and due to the time-gap between the census and the survey (21-29 months) leading to non-coverage of immigrants and Israelis returning after the census.

Mis-stratification due to response errors in stratification variables was relatively small. But the type of household changed considerably over time, mainly due to deaths and moving either to institutions or to other households. These are more common amongst the higher age-groups that were boosted in the sample, so mis-stratification led mainly to "migrants" from high probability strata to low probability ones and to an increase in the rate of ineligible cases in the sample. The accuracy of the survey estimates was not seriously affected.

Such surveys, if in-depth investigation is required, are preferable to conduct as PCS's: providing the time-gap is shorter and efforts are made to reduce under-coverage. Alternative less satisfactory frames can be considered for such surveys in the inter-censal period and for a limited investigation the current LFS can be used.

Travelling Habits Surveys - A number of such surveys were conducted in Israel and two were PCS's: the 1972/3 one was a nationwide survey with a sample of 55,000 households, and the 1984/5 survey was limited to the three large metropolitan areas, in each a sample of about 1,500 households. The investigation in the 1972/3 survey started shortly after the census and terminated after eight months, whereas in the 1984/5 survey there was a time-gap of 15 to 22 months.

One reason for conducting these surveys as PCS's was that the census 20% sample questionnaire could provide data for disproportionate stratified sampling, i.e. assigning higher selection probabilities to households according to the number of vehicles in use. However, this did not produce the expected gain in efficiency due to: (a) mis-stratification of dwelling units resulting from response errors in the census; (b) unoccupied dwellings in the census and census non-response that could not be classified into the strata; (c) changes of number of vehicles over time; and, (d) turnover of occupants between the census and the survey.

The frame did not cover certain dwellings, such as vacant dwellings, nor refusals, and absentees. Therefore, to avoid under-coverage, a complementary sample was selected to represent them. Although this introduced complications and resulted in more ineligible units in the sample, it contributed more than 7% of the survey population.

It was also planned to use certain individual data (such as income) from the census instead of collecting them in the survey. This was

unsuccessful, since the survey was conducted too long after the census and there were matching problems and changes over time relating to the data of the same individuals, or to the residents of the sampled dwellings.

Most of these disadvantages intensify when the time-lag between the census and the survey is longer. Therefore, large scale surveys similar to the one conducted following the 1972 census, if conducted immediately after the census and make use of the census staff and organization, are more suitable to be conducted as a PCS. Otherwise, Travelling Habits Surveys are not recommended to be conducted as PCS's.

---

*Editorial note: we are very grateful to Malka Kantorowitz for supplying this article at short notice.

******

The 1992-93 New Zealand Household Health Survey

Denise Brown
Social Policy Division,
Statistics New Zealand

Introduction

New Zealand's first major national health survey was conducted by Statistics New Zealand during the twelve month period April 1992 to March 1993. The 1992-93 Household Health Survey collected information from over 7,000 New Zealanders about the state of their health, their use of health services, and aspects of lifestyle and behaviour which impact on health.

This article describes methodological aspects of the 1992-93 Household Health Survey, including the survey design, data collection, questionnaire content, and processing and dissemination of the survey results. It begins with a brief discussion of the background to the survey.

Background to the survey

In New Zealand, as in other developed countries, there has been over time a gradual shift away from viewing health in terms of survival, through a phase of defining it in terms of freedom from disease, to the current emphasis on positive themes of social and emotional well-being and quality of life. This shift in focus is reflected in changing attitudes to health care where much greater emphasis is being placed on preventative measures to protect health, and on the identification and avoidance of health risk factors. More attention is also being given now to broader influences on health, such as income, education and employment.

These changing attitudes to health and health care have not been matched by parallel developments in health statistics. Most of the health information available in New Zealand at present consists of statistics about death, illness and hospitalisation. What is known about the prevalence of illness comes mainly from administrative sources. As data sources, they are often deficient in information about the demographic and socio-economic characteristics of individuals and their lifestyle, or particular behaviours which have directly affected their health. An important gap in health information has always been primary health care services such as those provided by general practitioners, pharmacists, community-based nurses and other health professionals who are usually the first level of contact with the health system. There are also serious gaps in data on health status, particularly with regard to disability and well-being, and factors which influence health status, such as smoking and exercise.

Late in 1991, a working party was established to undertake the development of a proposal for a New Zealand health survey. The working party comprised representatives from area health boards, Statistics New Zealand, the Ministry of Health and the Wellington Clinical School. Experts outside the working party were consulted on specific issues as required. The survey, which was conducted by Statistics New Zealand and sponsored by the Ministry of Health, went into the field on 6 April 1992.

Aims and objectives

The overall aim of the 1992-93 Household Health Survey was to provide information on patterns of health service use, satisfaction with service, and health status among different population groups to assist the development of policy and decision-making by funders, purchasers and providers of services.

The broad objectives of the survey were:

- to measure health service use not currently recorded in the national statistics, and factors which influence level of use;

- to obtain data on health service use according to socio-demographic variables;
- to measure individuals' satisfaction and perceptions of their health services;
- to measure individuals' health status and unmet needs for health services; and
- to provide information on selected health behaviour not at present available nationally.

Survey design

The survey was run as a supplement to the Household Labour Force Survey, since this provided an appropriate framework and cost-effective vehicle for the collection of national information. The Household Labour Force Survey collects information quarterly from a statistically representative random sample of 16,000 households from rural and urban areas in the North and South Islands of New Zealand. Each quarter, one-eighth of the sampled households are rotated out of the survey, after being in it for two years. The sample of households for the Household Health Survey consisted of those households in the final quarter of inclusion in the Household Labour Force Survey.

One respondent per household was selected for interview, using a procedure designed to avoid interviewer bias and ensure the best random selection of an individual within the household. The decision to select only one individual per household was based on the following considerations:

- the objective of the Health Survey was to provide information on the health of individuals rather than families or households;
- it is methodological preferable to have health data free from the clustering effects which result when multiple members of a household are interviewed. Such clustering effects are difficult to quantify and would add to the costs of analysis.
- multiple interviewing in a household increases the household respondent burden, making it more likely that there would be partial refusals leading to a higher non-response rate. Interviewers would have to make more than one call/visit to get all household members, and this would increase field costs.

Data collection

Nearly 9,000 respondents who were being interviewed for the Household Labour Force Survey, were invited to participate in the 1992-93 Household Health Survey. The achieved sample size was 7,065 making a response rate of 89 percent for the survey. The health interview was conducted by trained Statistics New Zealand interviewers at the conclusion of the interview for the Household Labour Force Survey. In the main, the interviews were conducted by telephone, but where this was not feasible a face-to-face interview was carried out.

A personal interview was conducted, where possible, with each selected adult respondent. In order to obtain an interview with appropriate respondents, interviewers made up to 10 telephone calls or three personal visits in addition to the calls for the Household Labour Force Survey. Children aged 0-14 years were not interviewed, but a parent or guardian was asked to answer on their behalf. This situation was the only one where a proxy interview was routinely called for.

Interviews averaged 30 minutes per respondent, and were spread evenly over a twelve month period to eliminate bias in the data caused by seasonal or periodically varying conditions.

Questionnaires

Two interviewer-administered questionnaires were used in the 1992-93 Household Health Survey. The first, The Household Control Document, was used solely for the purpose of randomly selecting an individual within the household for interview. This selection process involved ordering household members by age. On each household control document, there was a randomly generated table that gave, for each possible number of people present, one specified position in the age order. The person who was in that position in the age order was the selected respondent.

The second questionnaire, the Health Survey Questionnaire, was used to collect information from individuals about their health status, health-related behaviour and use of health services. The topics chosen for inclusion in the survey were those identified as being of highest priority in terms of their relevance to important areas of health status and health service use; their relevance to the development of health policy; the unavailability of the information from

other data sources; the prevalence of the characteristic throughout the community; and the ability to appropriately address the topic in a telephone survey.

A range of additional demographic and socio-economic information was obtained directly from Household Labour Force Survey data records held from previous interviews with the respondent.

In addition to the topic selection criteria noted earlier, a number of other factors were considered in designing the questionnaire. These included the use of easily understood words and concepts, the use of standard definitions and classifications wherever possible, comparability with overseas health surveys, the sensitivity of the topics and the overall length of the questionnaire.

The questionnaire was fully field tested.

Processing

A combination of clerical and computer-based systems was used to process the survey data. Clerical edits were initially applied by interviewers to ensure the completeness and consistency of the questionnaires before being returned to Statistics New Zealand. All questionnaires were again checked on receipt in the office to ensure interview workloads were fully accounted for and that all questionnaires for each respondent were completed.

Information on the questionnaires was converted into machine-readable format after the completion of the clerical checks. An extensive range of computer edits was then applied to each record on the file to check that logical sequences had been followed in the questionnaires, that specific values lay within valid ranges and that relationships between items were within limits deemed acceptable for the purposes of the survey. The edits were designed to detect errors which may have occurred (e.g. during data transcription or original recording of the information by the interviewer) and to identify cases which, although not necessarily errors, were sufficiently unusual to warrant examination. Listings of all records were produced and amendments made on the computer file as required.

There was no imputation of individual responses. Where a response to a question was omitted it was assigned to the 'not specified' category.

Estimates obtained from the survey were derived using a Horvitz-Thomson estimator. This procedure uses survey results to calculate approximate values for the entire population. A system of linear weighting was used to ensure that the age, sex and ethnic distribution within the sample conformed to an independently estimated distribution of the total population based on the Population Census.

Dissemination

Final results from the survey were provided to the sponsoring agency, the Ministry of Health, within two months of the completion of the fieldwork. The first official release of data was in October in the form of a published report, *A Picture of Health*. This report, which was jointly prepared by Statistics New Zealand and the Ministry of Health, presents an overview of the survey findings. The data are presented in a non-technical way, with commentary accompanied by graphics and tables. More detailed analysis of data from the survey will be contained in later publications.

In addition to the published survey results, special tabulations can be produced on request to meet the individual requirements of users.

Conclusion

The 1992-93 Household Health Survey collected a wide range of health-related information, from a representative sample of the New Zealand population. The information, which complements existing data from administrative sources, fills a critical gap in the health information base.

The use of the Household Labour Force Survey, as a vehicle for conducting the health survey, has demonstrated that this provides an effective means of producing high quality statistical data to satisfy important public policy demands.

******

International Blaise Users Conference[1]

Tony Manners
OPCS
Social Survey Division

1. Introduction

In October, the British Office of Population Censuses and Surveys, Social Survey Division

(SSD) hosted a 3-day international conference for users of Blaise, the computer assisted interviewing (CAI) package developed by the Netherlands Central Bureau of Statistics. Outside NCBS, OPCS is the largest user of Blaise. It used Blaise to pioneer CAI for face-to-face (CAPI) interviewing in British households, starting three years ago with the Employment Department's Quarterly Labour Force Survey; it also uses Blaise for telephone interviewing (CATI).

Since its development in the second half of the 1980s, Blaise has been adopted by virtually all European government statistical agencies, as well as by others around the world. It has increasing numbers of users in the universities and market research. The International Blaise Users' Conference (IBUC/93) in London was attended by 80 delegates from 22 countries including, besides those in Europe: Australia, Mexico, New Zealand, South Africa and the United States. The purpose of the conference was to bring users together for an intensive period of discussion with each other and with the software developers; to report the results of explorations in this new technological area; to generalise useful experience; and to highlight priorities for development. A major new version, Blaise III, was demonstrated by NCBS.

Bill McLennan, Director of the Central Statistical Office and Head of the Government Statistical Service, UK, opened the conference by stressing the importance of international cooperation in the development of statistical methods and the better use of technology. He noted the leading role of government agencies in CAPI survey methodology and use. Particular tribute was due to NCBS for developing Blaise and for its interest in the experience and wishes of its international users. CAI has now been used on major surveys for some years, and it is vital to think through the organisational implications of re-engineering the data collection processes. It was impressive to see that Blaise is easily used in a variety of organisations.

The keynote speaker for the conference was Wouter Keller, Director of Statistical Methods and Development at NCBS, who was the original architect of Blaise. He reminded the conference that one of the driving ideas of Blaise was to transfer control of the computing process in data collection, analysis and dissemination from computing specialists to statisticians and others who were conducting substantive research. The aim was to improve efficiency by enabling researchers to work more directly with their large-scale, complex data through intuitive interfaces of the kind they were familiar with in word processing and spreadsheets. He described how the current challenge was to extend the scope of this idea to cover and make use of the significant development of practical client-server systems. Blaise itself, which had hitherto been principally a data entry and interviewing tool, had become in Blaise III a "control panel" for defining and manipulating data through all the stages from design of the data collection to analysis.

The contributed papers for the conference covered a range of topic areas; agencies' experiences in converting to CAI with Blaise; technical innovations in survey practice which CAI made possible; using Blaise for very complex surveys; Blaise tools for output and dissemination of data; the need for CAI standards; and organisational issues in agencies where CAI was becoming the predominant mode of data collection and processing. The papers tended to follow the same pattern. They described the significant progress made with Blaise compared with traditional methods; some limitations imposed by current versions (2.4 and 2.5) of the package; ways that these had been overcome by integrating Blaise modules with other software such as MSDOS database packages and Pascal, in which Blaise is written; and suggestions for improvements. Many of the authors had already discussed their wishes with NCBS and contributed indirectly to the specification of Blaise III.

2. Conversion of existing surveys

A common issue for statistical agencies is how to convert existing continuous surveys from paper and pencil interviewing (PAPI) to CAI. Papers from the French Institut National de la Statistique et des Etudes Economiques (INSEE) and the United States National Agricultural Statistics Service (NASS) described their experiences with their initial projects using Blaise. These provided some intriguing similarities with and differences from the early UK experience at OPCS and the Central Survey Unit, Northern Ireland. The projects described were the French Labour Force Survey and the US quarterly Agricultural Survey. In both cases, the need to adapt to an existing regional office organisation and interface with a final stage of central mainframe processing were important considerations, although both envisaged phasing-out the mainframe stage as soon as the reliability of the CAI system was established.

Both agencies had adopted a strategy of phasing in the conversion of each survey to CAI, using it in part of the sample, rather than converting a whole survey at one time. OPCS has always converted whole surveys, and the CSU in Northern Ireland converted most of its existing surveys at the same time. The paper from INSEE noted that in France where one-ninth of the sample had been carried out in CAI in the first year, and one-half of the remainder in the second year, the need to maintain two parallel systems "turned out to be a severe constraint".

Statistics New Zealand has taken a different approach from INSEE and NASS. It has recently undertaken wholesale reorganisation, and adopted Blaise as a component of its overall strategy for Generalised Automated Survey Processing. By March 1993, nearly 10% of all questionnaires at the agency were processed in Blaise, and a range of major surveys under development with Blaise were described in the paper, including the Migration Survey, the Household Expenditure and Income Survey, and the Quarterly Business Survey.

3. Technical advances

A paper from NCBS described how interviewing time could be shortened and data quality improved in the Netherlands Family Expenditure Survey, which has a panel design, by checking for changes in the situation recorded at the previous interview rather than asking questions afresh at each wave. Blaise facilitated this process with tools which made it easy to provide the information on screen in a suitable form. A different use of prior information, from population registers, was explained in a paper from Statistics Finland. Call scheduling in the CATI unit could be based on respondents' characteristics.

John Polak from the Transport Studies Unit, University of Oxford, described the advantages brought to conjoint measurement (stated preference) research by the ability to display options on laptop computers. Blaise had provided a more cost effective solution than questionnaire development in either a conventional programming language (too expensive) or a specialist conjoint analysis package (too flexible).

Hans Stol of NCBS set out his view of a design for automated transmission of business statistics from firms' administrative systems to NCBS through "EDI-Blaise". A start had been made with a computer assisted self interviewing system which is used to transfer the foreign trade data of some 14,000 Dutch firms to NCBS each month.

SIA Ltd demonstrated the use of Blaise tools to provide a system suitable for market research applications.

Two papers, from NCBS and Statistics New Zealand, examined the improved facilities for computer assisted coding in Blaise 2.5. A new searching algorithm, using dictionary matching based on "trigrams" (all 3 character combinations) in the text to be coded, complements the existing alphabetic and hierarchical methods.

4. Complex surveys

Blaise is being stretched to its limits in versions 2.4 and 2.5 by some very complex surveys. OPCS described how it had overcome the problems presented by detailed financial surveys, even when the editing requirements were as stringent as for the UK Family Expenditure Survey. Jim O'Reilly of Battelle, USA described a health survey with a 3 hour interview, and the consequent need to link 11 Blaise questionnaires to form a single interview instrument. Both organisations had used database packages to provide interviewer interfaces to multiple Blaise questionnaires.

5. Standards

Papers from NASS and OPCS looked at different aspects of the need for standard procedures within agencies if they were to make efficient use of CAI as its use spread. The main problem facing NASS was to produce 45 regional versions of the same questionnaire, and they had developed highly automated procedures for doing this and for accepting many different kinds and combinations of inputs. OPCS discussed, in particular, standards for questionnaires writing in Blaise and screen designs which would help its interviewers across the range of its CAI surveys.

6. Outputs

NCBS described the latest version of Statview, a Blaise-related system for electronic dissemination of statistical data. Very large amounts of aggregate data can be stored on floppy disk and accessed in the form required through a simple table specifying interface with drop down menus. James Gray of OPCS explored the potential conflict between data

structure requirements of interviewing and analysis, and the drawbacks of traditional solutions to this problem. He made some suggestions for the use of Blaise tools, including Statview, for the later stages of processing and output to customers.

7. Organisational Issues

The last two papers of the conference discussed the impact of CAI in organisations where many, if not all, surveys would use Blaise rather than paper questionnaires. One of the main issues for the Central Survey Unit in Northern Ireland was the need for a new kind of IT support, which would focus less on carrying out computing for statisticians and more on providing them with means to do it themselves. At OPCS, the rapid spread of CAI across surveys had led to organisational changes to recognise its implications. For example, much of the traditional edit is now carried out in the interview, so skills in interviewing, coding and editing need to be combined. It is logical and cost-effective to combine the previously distinct field and coding sections. In general, it is essential to have a vision of how the organisation should change over the next few years to maximise the opportunities for better, cheaper and faster data collection offered by CAI.

8. Blaise III

Blaise III is an almost completely rewritten version of the package, as indicated in its new designation with Roman numerals. It focuses on ensuring that data anywhere in the survey system from questionnaire design to analysis are accompanied by adequate descriptions ("meta-data") which will allow them to be used flexibly. This idea is integrated survey processing: "a concept that attempts to avoid the problems generated by repeated specification of data and meta-data. The idea is to offer you a system that asks for just one specification of the meta-data. As soon as you have specified the meta-data, the integrated system manages and controls the collection of data, and the subsequent steps in the survey process".[2] Specification of the meta-data is achieved essentially by writing Blaise statements of a similar type to earlier versions. A Blaise questionnaire description can be used in all further stages of processing.

For many users the key difference is the facility to deal with hierarchical data without any constraints except the limits of the computers used. This would remove the chief problems mentioned by users at the conference. There are numerous other improvements, such as (to take two examples that users had requested) increased speed and the ability to switch languages to suit respondents. NCBS distributed test versions of the package to the conference delegates, inviting comments. The aim is for the production version to be available in January 1995.

9. International Blaise User Group

It was agreed that Social Survey Division, OPCS, would continue to take the lead in coordinating the group and its activities such as the newsletter, for the next year. Statistics Finland announced its sponsorship of the next users' conference, in Helsinki in spring 1995.

References

1 The conference papers are published in *Essays on Blaise, 1993. OPCS, 1993.*

This article is also published in *Statistical News, Issue 103, Winter 1993.*

2 Netherlands Central Bureau of Statistics, 1993. *Blaise III, An overview.*

******

Measurement Errors and Survey Data Analysis

Chris Skinner
University of Southampton, UK

The importance of measurement errors in surveys has long been recognized within the IASS. Papers from the part-IASS-sponsored International Conference on Measurement Errors in Surveys, published in Biemer et al (1991) and the 1992 (No.3) issue of the *Journal of Official Statistics*, provide a recent survey of the 'state of the art'. The regular appearance of papers on measurement errors at ISI meetings, such as the influential paper of Hansen, Hurwitz and Bershad (1961), bears witness to long-standing concerns. Much of this literature has as its ultimate purpose the design of data collection procedures to reduce measurement error. In this note I shall look instead at approaches to data analysis which accommodate the presence of measurement errors. My aim is just to survey what I see as a few key distinctions and issues and to point readers towards some relevant literature and software.

First, it seems necessary to emphasise that an attempt to accommodate measurement error in data analysis does not in any way belittle the importance of reducing the error in the first place, nor imply poor practice on the part of the survey agency. It just recognises that measurement error can never be avoided entirely. Survey agencies can collect data of high quality without claiming perfection. For instance, work on the World Fertility Survey acknowledged that some infant and child deaths were likely to go unreported, while presenting evidence that the quality of coverage in the surveys was better than in vital statistics sources (Goldman et al, 1985).

Before proceeding to look at approaches to handling measurement error, it also seems necessary to acknowledge the conceptual and indeed practical difficulty which often arises when we are faced with the question: what do we mean by measurement error? Or, if we define measurement error as the difference between the observed value and the true value, what do we mean by true value? Even for apparently straightforward 'factual' variables like employment status, it is often easy to find examples where the application of the variable definition is ambiguous. Much has been written on this issue - see, eg Groves (1989) and Biemer et al (1991). From the point of view of the data analyst, however, even if we have a satisfactory notion of a true value there remains a further key empirical question: what aspects and effects of measurement error can the analyst reasonably expect to be able to allow for, given the information available (together with plausible assumptions)? Such information might be external to the main survey, such as from pilot work, reinterviews or checks against administrative records, or may be internal, derived from multiple indicators or repeated questions or perhaps from imaginative use of the survey data, eg from comparisons of estimates from different rotation groups in a rotating panel survey. With minimal information, there may be little more than can be done than estimate characteristics of the observed variables. With some information it might, for example, be possible to allow for measurement error effects on standard errors. With yet more information it might be possible to make useful bias adjustments and so forth.

Let us turn then to summarising methods for accommodating measurement error. Most methods can be represented as dependent on the fitting of certain models. The classical error model assumes that the measurement error has zero mean and is uncorrelated with the true value of the variable. In descriptive surveys, where the aim is to estimate parameters like means and totals, such classical error typically affects the variances of conventional estimators but not their biases. As discussed by Wolter (1985, App D), random group (or other) variance estimators can, in certain circumstances, be constructed to handle the variance inflation arising both from complex sampling designs and from measurement errors which are either independent between responding units ('simple response variance') or which apply to more than one unit, as eg for interviewer effects ('correlated response variance'). The use of replicated and interpenetrating designs assists the assessment of such effects (Hansen et al, 1961).

Measurement error with non-zero mean will generally lead to bias in descriptive surveys. Such bias effects are less commonly addressed than the variance effects described above, but when more accurate measurements are available for a random subsample (double sampling), conventional ratio or regression estimation methods may be used for bias reduction (Lessler and Kalsbeek, 1992, Sect 10.6).

For analytical surveys, where the aim is typically to estimate parameters of models such as regression coefficients, the presence of classical zero-mean measurement error will generally lead to bias in conventional estimators. The most well-known example is the simple attenuation of a correlation coefficient. But bias effects can be much more perverse than this. For example, measurement error in a confounding variable can lead to failure to control properly for selection effects and hence to quite misleading conclusions about causal effects of 'treatment' variables in multiple regression modelling. In many other applications, eg longitudinal analysis, the bias effects can also be severe. Historically, two broad approaches to bias adjustment may be distinguished.

In the first approach, external information about the characteristics of the measurement error are combined with data on the measured variables from the main survey. A well-known example is 'errors in variables' regression in econometrics, where one or more of the explanatory variables is measured with error and estimate(s) or parameter(s) of the error distribution are assumed available and usable for bias correction. Another application area, where there has been much recent interest in related methods for non-linear (eg logistic)

regression models, is medical statistics (see, eg the special issue of *Statistics in Medicine*, 1989, Vol 8).

In the second approach, no external information is assumed. Instead certain assumptions are made about the relationships between certain measured variables (or indicators) and certain latent (unobserved) variables, which may include true variables. The combined set of assumed relationships is sometimes termed a 'structural equation model'. Examples of such models include the factor analysis model in psychometrics and the model which underlies instrumental variables estimation in econometrics.

From a formal statistical point of view, there is no fundamental distinction between the two approaches. External information must come from somewhere. Thus, a reinterview study does not magically produce a measurement error variance. Rather, it is necessary to specify a model relating the interview value, the reinterview value and the hypothesised true value, just as in the second approach. Hence both approaches come down to fitting models to one or more sources of data under more or less plausible assumptions.

Methods based on the classical measurement error model are most suited to continuous variables. There is also a tradition of methods for categorical variables which are subject to misclassification. An example of the first approach above would involve estimating misclassification rates by double sampling. An example of the second approach would be to assume that observed categorical indicators are related by a latent class model.

There are a number of books which those interested in applying such methods might like to refer to. Fuller (1987) provides a thorough discussion, starting with the first approach above but extending this to include, for example, factor analysis models. This account is particularly useful background to the program EVCARP (for reference to this and other software see the end of this article). Bartholomew (1987) provides a general account, following more the second approach above and incorporating both continuous and categorical variables, the latter discussion being useful background for the program PANMARK. Bollen's (1989) account might also be viewed as following the second approach and his book is useful background for the programs LISREL and EQS.

It is very important to note, however, that learning how to handle measurement error in survey analysis does <u>not</u> simply involve learning how to use software such as LISREL. What methods can sensibly be used depend fundamentally on what information is available about the measurement error. In the absence of much information one may only be able to make very unrealistic assumptions, for which it would be quite inappropriate to claim with any confidence that measurement error has been satisfactorily dealt with.

Unfortunately, for many surveys very little validating information is collected about the nature and extent of measurement error. Even the quality control exercises that are carried out to supervise fieldwork may provide little useful information. To paraphrase Forsman and Schreiner's (1991) conclusions, reinterview (and other validation) studies may be unpopular, not only because of their resource implications but also because they may be the bearers of bad news about the degree of error. To avoid conducting such studies for these reasons is, however, short-sighted. Validation studies designed to evaluate and monitor measurement errors can have significant benefits in the long run in terms of improvements in data quality as well as, for our purpose, by enabling appropriate adjustments to be made for measurement errors in data analysis.

Finally, for those readers whose interests are more methodological, I should like to suggest two topics, where I think there have been interesting recent developments and where there still seems room for survey statisticians to make useful contributions. First, continuing the theme of the previous paragraph, there seems a need for more work on the design of reinterview and other validation studies to estimate components of measurement error. Secondly, there seems more to be done on developing more realistic measurement error models and incorporating these into data analysis methods. Most validation studies reveal the inadequacy of the classical measurement error model. There is much evidence in the survey methodology literature on other types of measurement error for specific types of data, for example telescoping or heaping of the timing of events or the under- or over-reporting of counts. Data analysis often seems to proceed in apparent ignorance of such evidence. It would be nice to see further development of data analysis methods which more realistically reflect the nature of survey data.

Computer Software

LISREL

This program fits a wide range of linear structural equation models containing both observed and latent variables. The user specifies relationships between the variables. Examples of models include factor analysis and simultaneous equations models. The program is distributed by SPSS and has the advantage that it can be used in conjunction with SPSS/PC+ data and file management facilities. A Windows version will soon be available. For a review see the *American Statistician 45*, 1991, pp 68-73. For an illustrative price, a single user license in the UK was recently quoted as £495. For further information, contact your local SPSS distributor, eg in the UK:

SPSS House
5 London Street
Chertsey
Surrey KT16 8AP, UK

EQS

This program fits a similar broad class of models to LISREL and has many common facilities. The review of LISREL referred to above also assesses EQS and concludes that 'we feel comfortable recommending either program'. The latest version of EQS runs under Windows as well as on several other systems. It is distributed by BMDP. For an illustrative price, EQS/Windows was recently offered at £450 in the UK. For further information contact

BMDP Statistical Software
Cork Technology Park
Model Farm Road
Cork, Ireland or

BMDP Statistical Software Inc
1440 Sepulveda Blvd, Suite 316
Los Angeles, CA 90025, USA

EVCARP

This program conducts regression analysis, allowing for measurement error in the explanatory variables. Either estimates of characteristics of the measurement error distribution are required or instrumental variable estimation is used. The program can be used in conjunction with the program PCCARP which enables complex sampling designs to be handled. For a review see *Applied Statistics 38*, 1989, pp 531-3. EVCARP runs on PCs. The cost is $300 (or $500 with PCCARP) but it will be provided to institutions in developing countries for only the shipping cost. Further information from:

PCCARP
Department of Statistics
Iowa State University
219 Snedecor Hall
Ames, IA 50011, USA

PANMARK

This program is suitable for categorical panel data. It fits mixed Markov models (see eg Van de Pol and Langeheine, 1990) where the observed states at each occasion may be subject to misclassification. The misclassification probabilities are estimated internally from the data using model restrictions rather than from external information. A review will appear in *Applied Statistics 43*, 1994. PANMARK runs on PCs. The price is dfl 98 inside the Netherlands, about dfl 135 outside and free for developing countries. Further information from:

Netherlands Central Bureau of Statistics
Automation Department
Statistical Informatics Unit
PO Box 959
2270 AZ Voorburg
The Netherlands

References

Bartholomew, D J (1987) *Latent Variable Models and Factor Analysis.* Oxford: Oxford University Press

Biemer, P P, Groves, R M, Lyberg, L E, Mathiowetz, N A and Sudman, S. Eds (1991) *Measurement Errors in Surveys.* New York: Wiley.

Bollen, K A (1989) *Structural Equations with Latent Variables.* New York: Wiley

Forsman, G and Schreiner, I (1991) The design and analysis of reinterview: an overview. In Biemer at al, Ch 15.

Fuller, W A (1987) *Measurement Error Models.* New York: Wiley

Goldman, N, Rutstein, S O and Singh, S (1985) Assessment of the quality of data in 41 WFS Surveys: a comparative approach. *WFS Comparative Studies No 44*, Voorburg, Netherlands.

Groves, R M (1989) *Survey Errors and Survey Costs.* New York: Wiley

Hansen, M H, Hurwitz, W N and Bershad, M A

(1961) Measurement errors in censuses and surveys. *Bull Int Stat Inst 38*, No 2, pp 359-374.

Lessler, J T and Kalsbeek, W D (1992) *Nonsampling Errors in Surveys*. New York: Wiley.

Van de Pol, F J R and Langeheine, R (1990) Mixed Markov Latent Class Models. In C C Clogg, Ed *Sociological Methodology*. Oxford: Blackwell, pp 213-247.

Wolter, K M (1985) *Introduction to Variance Estimation*. New York: Springer Verlag.

******

Using Survey Research Skills in Program Evaluation: A Study of the Quality of Student Financial Aid Programs

Benjamin F King, Florida Atlantic University and Ronald S Fecso, National Agricultural Statistics Service, USDA[1]

The U.S. Department of Education requested that a panel of experts be convened at the National Academy of Sciences to study the quality control of federal student financial aid programs. The panel had representation from the fields of survey research, quality management and financial aid. The Department was motivated by concerns about the accuracy of payments to students in the need-based award programs and the system of accountability under the various programs. Payment inaccuracies and procedures that do not effectively provide for accountability raise questions about the equity with which federal student aid is distributed and could foster public misconceptions about waste, fraud, and abuse in federal government. The panel reviewed the methods used to collect data and the availability of information used by program managers. The panel's report (*Quality in Student Financial Aid Programs: A New Approach*, Ronald S Fecso, ed., National Academy Press, 1993) goes beyond a review of the use of such information in the current quality inspection approach used by the Department of Education. It focuses on the use of statistically-based performance monitoring systems to create a management system focused on taking appropriate corrective actions to prevent errors.

The panel reviewed previous research reports related to those issues prepared for the Department of Education and other federal bodies. Many of the studies relied on survey research techniques. Although problems in the specification, design and estimation in the surveys were pointed out, the panel found useful suggestions in many of the research reports. Yet, for the most part, the reports assume that major aspects of the delivery and control systems (e.g. school-based verification of student supplied information) would remain essentially unchanged. The panel, influenced by the recurring nature of several problems found in previous studies, proposes several changes to these systems. The major recurrent problems are the following:

- Making the educational institution responsible for identifying and correcting errors in information supplied by students creates inefficiencies and adversarial relationships. Further, clear evidence is lacking that, by increasing inspections in current accountability processes, schools can be regulated into achieving error removal that would significantly reduce the national student error rate.

- The Department of Education's use of regulations that require actions by participants to solve problems found in only a few cases is inefficient and unfair.

- Research and development of quantitative tools to guide management of the programs lacks adequate technical guidance. Especially troublesome was the lack of organization and use of information gathered in the costly auditing and program review activities.

The panel reviewed related practices used in a number of public and private organizations. Many of these organizations have recognized that a system based on prevention of errors rather than inspection is necessary to improve quality and operate efficiently. Key elements in a philosophy of continuous improvement include the following:

- A focus on improving quality rather than only measuring it.

- Identifying and focusing on the customers' needs for a product or service now and in the future.

- Emphasizing thinking about the entire system rather than individual operations in the system.

- Making decisions based on data.

- Anticipating and accepting change.

- Emphasizing nonhierarchical teamwork for decision making and implementation.

- Understanding variation in measurements of the process and making a commitment to reduce it.

- Identifying and clearly communicating the aims and purpose of the enterprise.

- Fostering the above activities through commitment and leadership of top management.

The report provides a study of how these principles are as valid in the Department of Education and other governmental agencies as they are in private industry. The report points out the opportunities for "reinventing government" that are possible in an approach to administering student financial aid that seeks continuous improvement in important outcomes of the process--outcomes such as the accuracy of the award, customer satisfaction, and accessibility of aid for eligible students. New procedures for more effectively and efficiently determining need-based financial aid award levels are suggested. They may eliminate mistakes before they are made as well as lowering costs, providing faster processing, and reducing red tape, thus bring improved services and greater value to all the customers and stakeholders --the Department of Education, Congress, postsecondary institutions, students, parents, and taxpayers.

The report can be used as a case study for many organizations embarking on quality improvement efforts. It also provides examples of how organizations need the kind of knowledge and skills provided in survey research training to collect, organize, and use information necessary in quality improvement efforts.

[1] B King was chair of the Panel on Quality Improvement in Student Financial Aid Programs and R Fecso was study director.

******

Question/Answer Column

Leslie Kish has requested that we reprint the answers to questions 29.1 and 29.2 because the versions in the July 1993 edition of the Survey Statistician contained errors for which I take full responsibility and apologise.

D Lievesley

29.1 Question

Are ratio estimators in any way related to the topic of post-stratification? If yes, how? Sent by Sv Terp, Assoc Prof, Institute of Economics & Statistics, Aarhus, Denmark.

Answer

You are right to note the relation as some others have. Alert readers of textbooks notice the similarity, then wonder why the two topics are so widely separated, with post-stratification early in several sampling books and ratio estimation much later, sometimes without cross references.

Post-stratification may be viewed as a special case of ratio means. These may be written as $\bar{y}_r = \Sigma(X_h/\Sigma X_h)(y_h/x_h)$; here X_h is an auxiliary (ancillary) variable for adjusting the random variables y_h and x_h, which are sample totals, within control classes h. Often the X_h and x_h are counts of cases in the subclasses, and $X_{hi} = 1 = x_{hi}$ when unweighted. This is often called a "separate ratio estimator," to distinguish it from the combined ratio estimator $r = \Sigma y_h/\Sigma x_h$.

This is equivalent to a post-stratified mean $\bar{y} = \Sigma W_h y_h/n_h$, where the weights $W_h = N_h/\Sigma N_h$ are applied to the sample results y_h/n_h obtained by a simple random sample of $n = \Sigma n_h$, a ratio estimator because n_h is a random variable. This alternative to proportionate stratified element sampling is often introduced in an early chapter on stratified element sampling. It serves as a simple heuristic device to show how estimation can serve as a substitute for selection procedures. This is a valuable lesson, because stratifying variables may become available only after the selection of sample cases, and often after completion of the survey. Thus post-stratification is also called "stratification after selection" sometimes.

In actual practice, simple random samples are not often (ever) selected, but ratio estimation is often applied to data from complex multistage samples. The weights from "known" population counts are applied to ratios of sample statistics. Usually more important than adjustments for sampling variation are corrections for nonresponses and

especially for noncoverage, though the technical presentation may obscure this blemish, rather than clarify this important usage.

Ratio estimators also take many other forms in survey sampling and they are also related to weighted estimates of diverse types. "Synthetic estimation" for small domain estimates is one important type. "Standardization" to target population is another. Then with iteration to fit marginal distributions we get to "raking" or "rim weighted" or SPREE estimates.

29.2 Question

"Random routes" appear as a reasonable procedure for probability selection of sample dwellings from sample blocks, because they are a cheaper alternative than complete listings of dwellings. How valid and how efficient is this procedure for multistage area sampling?

Answer

This question arises sometimes when area sampling is used to identify sample blocks (or other areas) with clear boundaries. Assume further that these blocks are selected with probabilities proportional to measures of size (MOS) in numbers of dwellings that are reasonably up to date and accurate, but not perfect. If mappings or lists of the dwellings were available, they would be used for selection. However, such lists often are not available and it is felt to be too expensive to prepare those lists. The procedures typically designate a starting spot, usually a corner, and instruct the interviewer to proceed in some designated direction (e.g. clockwise) and select into the sample the next fixed number n_b of dwellings. In some modifications the interviewers are instructed to obtain some fixed number n_b of dwellings by some other procedures; e.g. every tenth dwelling to reduce homogeneity, but this may be even more difficult to operationalize in the field. Furthermore, in some cases only first calls are made, leading to more biases.

Most of us would not care to use this procedure, because it poses practical obstacles to the interviewer that prevent its conversion into a feasible probability selection. The chief problems:

a) The designated starting point is often a corner dwelling, which must be found

b) The dwellings often do not lie in a simple, straight line; and this needs complex instructions. We cannot go into the diverse complexities used, but I have not seen satisfactory procedures.

c) The actual number of dwellings ($N_a \neq MOS_a$) is not known and this results in unequal, and often unknown probabilities for the dwellings.

On the other hand prelisting all the N_a dwellings may be a costly procedure, which the "random route" tries to avoid. The cost of listings is reduced:

a) When the ratio of sample dwellings n_a/N_a is large, and even more by repeated uses of the listings by several surveys; and such repeated use is common.

b) When the variation in cluster sizes N_a is reduced by combining small blocks, and by splitting the very large blocks (a small percentage of all blocks), especially for repeated uses of the sample of blocks (clusters).

Thus random routes, if used at all, should be confined to single use surveys, within single cities where interviewers may be closely supervised and controlled, and where good measures of size are available. I advise against them, but there are a number of organizations that use random routes, some of them prominent. Perhaps a reply from one of them would be welcome and informative.

30.1 Question

You said: "The many sources of error that affect surveys are commonly divided into two main kinds: first, the survey biases mainly due to definitions, measurements and responses; and second, the variable errors of sampling." (Kish 1993.) Is that not crude, because there are also sampling biases; and also variable errors of measurement, which may sometimes be randomized?

Answer

Two people raised this question at the ISI/IASS session in Firenze and they are right: sampling biases also exist, and some variable errors of measurement have been measured at times. However: "We may begin by merely contrasting the variable errors of sampling with the nonsampling biases of measurement. That familiar distinction delineates roughly the most influential types of survey errors... A more complete and complex model in the next section also includes sampling biases and variable nonsampling errors" (Kish 1965, Section 13.1).

The many sources of errors of survey data are commonly divided into biases and variable errors; then, mostly conceptually, they are combined into the total error in the root-mean square error, RMSE = $\sqrt{(Bias^2 + V.E^2)}$. I present here an informal list of five pairs of terms used for that division, based on conversations with experts and perusal of the literature; and I invite comments. I used measurements/sampling errors (A) as a first approximation above. In the sampling literature most commonly the "essential survey conditions" are separated from other sources of errors, some of which are randomized (B). For example, the essential survey conditions include the definition of the population, also the frame for selecting the sample, the major domains of analysis, as well as definitions for strata and sampling units in the design. But randomized selections are used to select actual units, from PSUs down to elements, for the sample. The measurement biases of individual interviewers may be included within the sampling errors, as computed from the PSUs.

What are the sources of essential survey conditions? They may originate in external constraints (natural conditions, or laws, or traditions), or with "authorized personnel" (a ministry or a committee), or they may be influenced by the survey statistician. These sources should be examined more critically, especially because statisticians should have more input in these decisions.

For example, the timing of surveys and censuses is being widely debated and reexamined. "Rolling censuses" are being proposed to collect (at least some of) the data of decennial censuses. The variation between weeks in periodic surveys is a statistical problem.

Thus, instead of passive acceptance of essential survey conditions, we may consider all potential survey variables with a view of controlling some and randomizing others. This view (C) has been expounded in detail elsewhere in connection with experimental designs (both "true" and "quasi") (Kish 1987). But to be controlled we must have knowledge in available forms, and a division of "known" versus "unknown" effects has also been proposed (D). Finally in experimental designs the concepts of "fixed" versus "random" effects have been used to deal with components of the variance (E).

BIASES	VARIABLE ERRORS
A) Sources	
Measurements	Sample Selection
B) Survey sampling literature	
Essential survey conditions	Randomized errors
C) Disturbing variables in experimental design	
Controlled variables	Randomized variables
D) Available knowledge	
Known effects	Unknown effects
E) Effects on experimental design	
Fixed effects	Random effects

Perhaps it is less difficult to divide the source of errors in specific practical situations than as abstract generalizations. However some guiding principles may be useful, and the five types of pairings above may help. Though related, the five are not entirely similar, and that shows that more work is needed for separating biases from variable and randomized errors. Comments are welcome.

******

COUNTRY REPORTS

EDITOR - Gordon Brackstone

AUSTRALIA (from Susan Linacre)

In 1992 the ABS was commissioned by the Government to develop a new survey to

improve the type and quality of forward looking data available for economic forecasting and policy evaluation purposes.

In the private sector, surveys of business expectations have been conducted in Australia for many years, mainly by employer organizations either alone or in collaboration with banks. However, the reliability of these surveys has been limited by their small sample size, limited coverage, the absence of any weighting of responses by size of business and industry and hence the inability to measure the magnitude of expected change. This makes the data difficult to use in economic forecasting.

The new ABS survey collects the expected percentage change for the next calendar quarter and for the current quarter of the following year for the following business performance indicators: Sales of Goods and Services; Employment; Investment; Selling Prices; Wage Costs; Non-wage Costs; All Other Expenses; Stock Levels; Profits; Imports and Exports. As the survey does not collect actual levels of reported activity, benchmarks had to be obtained from other sources for use in weighting individual business responses to obtain aggregates.

The solution has been to use a statistical design in which the survey sample is drawn as a sub-sample of a recently developed Economic Activity Survey (EAS) which obtains annual actual data on most of the business performance indicators included in the Business Expectations Survey.

A synthetic estimate for profit/loss needed to be devised as all the estimators tested performed poorly due to the volatility of reported values and frequent sign changes. Profit/loss is derived as the difference between the new benchmark for the item sales of goods and services and the new benchmark for the sum of the items wage costs, non-wage labour costs and all other expenses.

The new benchmark is calculated by applying the expected change percentage for each component to the EAS benchmark for those components. An additional question is included on the survey form to provide a more up-to-date indication of whether the business is in a surplus or deficit situation. This information will be used to enhance the estimates derived.

Unlike current practice in some expectations surveys, the ABS survey seeks information from respondents without any adjustment made by them for any seasonal effects, even though the short term expectations will be at times greatly affected by seasonal activity. The series will be seasonally adjusted once a suitable time series has been collected.

The Business Expectations Survey is considered a good candidate for use of techniques such as smartfax (receipt and electronic lodgement of survey forms via facsimile), use of a full OCR survey form, and/or Computer Assisted Telephone Interviewing (CATI).

For more information, please contact Susan Linacre, Australian Bureau of Statistics, FAX (06) 2525172.

CANADA (from Gordon J Brackstone)

Statistics Canada is preparing the **National Population Health Survey** (NPHS). The NPHS will be a comprehensive longitudinal household survey covering a variety of aspects related to health.

Data will be collected for the following general areas: a) Measures of health status including a health index, identification of health problems, measures of mental health and of functional limitations; b) Determinants of health which include behavioural, lifestyle, economic and social factors; c) Use of health services and prevention practices; and d) Socio-demographic information including educational attainment and labour force characteristics of household respondents. Most of the information in categories a) to c) will be collected for a randomly selected household member. Selected members will form a longitudinal panel to be surveyed every two years for a maximum of twenty years. Data collection for the first wave will take place from May 1994 to February 1995 and use computer-assisted personal interviewing. The sample of about 22,000 households will allow national, provincial and territorial estimates. An institutional survey component has been proposed to cover long-term residents of hospitals and residential care facilities. The target sample size of the institutional component is of 3,100 residents in 1994. For more information, contact Jean-Louis Tambay, Social Survey Methods Division, Statistics Canada, Tunney's Pasture, Ottawa, Ontario K1A 0T6 (tel.: (613) 951-6959).

Fast Track Annual Surveys: Traditionally, at Statistics Canada, annual business surveys have been conducted as Censuses by

surveying the largest units and obtaining the data for the smaller units using administrative data sources. This traditional approach has drawbacks. The annual surveys are perceived as burdensome, expensive and of limited use because of tardiness. The results of annuals are published some eighteen months after the reference period.

Consequently, a new approach is being implemented for the 1993 Annual Retail Trade Survey. This approach is known as the Period 13 survey. The objective of the Period 13 survey is to estimate census value added on a calendar year basis at a reasonable level of detail using the monthly sample. Estimates are to be published within six months of the calendar year end. The advantage of a quick annual is its timeliness and cost effectiveness. Period 13 will be based on the Monthly retail Trade Survey sample. Period 13 will use some existing functions of the Monthly Retail Trade Survey for sample selection and data collection facilities. Also, some preliminary edits will use data from the Monthly Retail Trade Survey to detect and follow-up significant differences between the Period 13 survey and the Monthly Retail Trade Survey.

Data for the Period 13 survey will be collected early in 1994. It is recognized that calendar year statistics collected this way may be less reliable than fiscal year end statistics. Other annual business surveys conducted at Statistics Canada may follow the Period 13 route. It is also expected that every three to five years, a census-like "annual" business survey will be carried out to update existing business structures. For more information, contact M.A. Hidiroglou, Business Survey Methods Division, Statistics Canada, Tunney's Pasture, Ottawa, Ontario, K1A 0T6 (tel.: (613) 951-4767).

HUNGARY (from Adam Marton)

The Hungarian Central Statistical Office (HCSO) started its social statistical surveys in 1963. In 1976, the Unified System of Household Surveys (USHS) was formed. Based on the 1990 Census of Population and Housing (CPH), a new master sample was designed and selected in 1992 for the USHS.

Besides traditional, continuous or periodic surveys such as the Family Budget Survey, the transition into a market economy has raised demand for studying new subject matters, in particular, unemployment (unknown in Hungary up to the late 1980's), small enterprises (their number is increasing rapidly), health problems of the population, etc. To ensure reliable estimates for the parameters of interest played an important role in the sampling design.

The target population is all private households. The sampling frame is the register of dwellings from the 1990 CPH, which contains nearly 3.8 million addresses. The number of enumeration districts (ED's) in the CPH was about 90,000. The sample is a three-stage cluster sample. Primary sampling units are the localities. In the first stage, about 550 of them were selected with probability proportional to size (the number of dwellings) without replacement. Secondary sampling units are the ED's. After having been stratified according to area type, about 10,000 of them were selected, also with probability proportional to size. Ultimate sampling units (housing units) are selected with equal probability from the addresses in each of these ED's for the actual surveys during the decade 1992-2001.

At the end of 1992 the HCSO had to make a 17% reduction of the sample size for budget reasons. Representativeness of the reduced sample was an important requirement, and it was assessed using certain census data (sex, age group, family status, educational attainment and economic activity). The results showed that the representativeness of the sample was satisfactory.

The most important survey, the Labour Force Survey, takes place continuously on monthly subsamples of 8300 housing units. Each household (together with persons living there) entering the sample is asked to provide information for six consecutive quarters through personal interviews. In each quarter one sixth of the sample is rotated.

This was the survey in which laptop computers were first introduced. The questionnaires were simulated by the aid of BLAISE on the computers. in the majority of cases the experiences have been favourable, though in certain situations the use of the computers is not possible or advisable. In such cases regular questionnaires are being used. At present nearly one third, and in the near future more than half, of the interviewers will be provided with computers.

For more information on the data collection system and the various surveys, please contact Adam Marton, Survey Sampling Section, Central Statistical Office, P.O. Box 51, 1525 Budapest, Hungary (tel.: 361-115-9689).

ISRAEL (from Malka Kantorowitz)

Starting in 1989, the number of immigrants reaching Israel, mostly from the former USSR, has risen sharply, constituting at present, about 10% of the population. Thus, considerable interest focuses on the absorption of these new immigrants, especially the degree to which they have been able to integrate into the labour force. In January 1991, a longitudinal **Immigrants' Employment Survey** started, in which samples of immigrants from selected periods of arrival are investigated at various intervals. Besides former and present labour force characteristics, housing conditions, household composition and command of the language have also been investigated.

The investigating unit is a dwelling, comprising all households and their members aged 15 and over, occupying the dwelling. However, the available frame for sampling was of individual immigrants. Since a sample of individuals would not be efficient for the survey purposes, clusters of persons were formed, utilizing the information available, so that the sampling unit should, as far as possible, match the investigation unit, in order to balance between sampling errors and cost. Due to the dynamic nature of household structure, where it is not uncommon for families to reunite or break up after arrival, and the consequent changes in housing density, selection probabilities have to be adjusted in each wave.

Since this survey, like many other longitudinal surveys, is aimed to provide gross as well as net changes over time, the issue of how to use the survey data for both purposes, is now being examined.

Israel's first **Time Use Survey** was conducted recently. Two types of personal diaries were used for either a 24-hour period of the weekdays, or for a 48-hour period of the weekend: a retrospective "Yesterday-diary" filled in by the interviewer, and a "Tomorrow-diary" left for the interviewee to complete by himself and to mail back. This two-phase mode of data collection was chosen, within the limits of cost, as a compromise to balance between the different non-sampling errors of each of these two types of diaries. The survey data will be used to evaluate the relative merits of the two diary types, although the "Tomorrow-diary" was obtained from only a quarter of the respondents. An individual questionnaire was also used for obtaining retrospective information on various infrequent activities related to reference periods of different lengths. All persons aged 14 and over in a sample of households were to be interviewed. One reason for this was to provide estimates on collective household activities, e.g., on the way household duties are shared between its members.

Since not all respondents provided both a questionnaire and at least one diary, two separate series of estimates were provided from the two data sets independently, based on separate weighting procedures. Also, the number of households where at least one person did not respond was not at all small, so to obtain estimates for households, a third weighting procedure is required. The alternative of one set of weights for all estimates will be examined at the expense of increasing the overall non-response rate.

MALAWI (from S.C. Babu)

The National Sample Survey of Agriculture (NSSA) is conducted once in 10 years in Malawi to provide information for agricultural planning and policy making. The latest one, for the year 1992-1993 (crop season), is in the process of completion. The previous NSSA was conducted in 1980-82. This was followed by smaller annual surveys of agriculture conducted with the objective of updating the once in ten year surveys. The NSSA involves six different modules: household composition; crop production and garden measurements; livestock production; extension use and technology adoption; social dimensions of adjustment; and household food security and nutrition.

A two stage sampling procedure was followed to select the sample households. At the first stage, the country was divided into 7,000 enumeration areas, 600 of which were randomly selected for the survey. In the second stage, 20 households were chosen from each selected enumeration area using the total household listing, yielding a sample size of 12,000 households for the survey.

The NSSA integrated the annual survey of agriculture which collected farm management data every year from 1982 to 1989, and food security and nutrition monitoring surveys conducted twice a year starting May 1990. Under the food security and nutrition monitoring system the enumerators, the supervisors, the data entry clerks and the data analysts have been trained at the district level. The NSSA involved close collaboration between the eight agricultural development districts (ADD) under the Ministry of Agriculture and the National

Statistical Office. The analysis of the data will be done in the ADDs where the evaluation economists will use the data at the decentralized level. About 40 evaluation officers have been trained in policy analysis using the SPSS package. This approach ensures speedy use of data for development planning and avoids time delays between data collection and actual use of the information.

A major addition to the current survey is the module on the Structural Dimension of Adjustment (SDA) which collected information on the variables that affect welfare among rural households that are influenced by structural adjustment and stabilization policies implemented by several countries in Sub-Saharan Africa over the past decade. A characteristic unique to Malawi in the southern African region is the human and institutional capacity that has been developed recently to collect, process and analyze data from sample surveys through the food security and nutrition monitoring system implemented by the Ministry of Agriculture. It is expected that the complete set of data and preliminary reports of analysis from NSSA will be available to the public by the end of 1994. For more information on the NSSA, please contact: Mr. M. Golozi, Senior Statistician (NSSA), Agricultural Section, National Statistical Office, P.O. Box 333, Zomba, Malawi.

PAPUA NEW GUINEA (from Mizanur Rahman)

The Rural Statistics Branch of the Department of Agriculture and Livestock has conducted a Nation-wide Food Crop Market Survey in 1988 and 1989.

The main objective of the survey was to determine the volume (both quantity and value) of locally produced agricultural commodities traded through the major urban and rural markets in the country. The survey also investigated the origin of produce, transporting costs, frequency of seller's visits to the market, reasons of sellers for not bringing more produce at the market, sellers' decisions with regard to unsold produce at the end of the day, etc.

The data were collected through a sample survey using a systematic sample design. Sellers who used to bring produce at the market for sale on survey days, theoretically constituted the survey population. The number of sellers in the sample varied from one market to another, depending on the total counts. Enumerators, upon arrival at the markets at selected hours of each survey day, made a thorough count of sellers in order to fix the sample size for interview and weighing of their produce.

The survey dates were pre-determined on a 3 days per month basis. The survey continued for the whole of 1988 and 1989.

The report presents estimates of the total quantity and value of agricultural commodities offered, sold and unsold at the market, including the weighted average prices.

To obtain more detailed information, please contact Mizanur Rahman, Chief Rural Statistician, The Rural Statistics Branch, Department of Agriculture and Livestock, P.O. Box 417, Konedobu, NCD, Papua New Guinea.

PHILIPPINES (from Gervacio G. Selda, Jr)

The National Statistics Office (NSO), the country's general purpose data gathering agency, has proposed recently the updating of the master sample for its household surveys. The proposal was made in line with the objective of making the master sample more efficient given budgetary constraints. Among the general recommendations forwarded by the NSO to the National Statistical Coordination Board (NSCB), the country's policy-making body on statistical matters, were: (a) the trimming down of the overall sample size for the regular household surveys; (b) the institution of more quality control measures to reduce non-sampling errors in household surveys, and (c) the revision of sample allocation. Currently, the Technical Committee on Survey Design (TCSD), a body composed of experts on survey design, is reviewing the proposal of the NSO, especially the recommendations to: (a) reduce the PSUs from 2,100 to 1,800; (b) use the same number of SSUs for each urban and rural PSU; (c) allocate 1,100 PSUs and 700 PSUs to urban and rural stratum, respectively; (d) keep a minimum of 80 PSUs for each region; (e) maintain a 20 per cent sample for annual rotation; and (f) reduce the allocated PSUs to the principal region to only 150. Also, the committee is looking into the proposal of the NSO to expand the master sample in case resources are made available so as to meet the need for more disaggregated data (i.e., provincial-level statistics). On this aspect, the NSO has recommended the use of a larger number of PSUs in one quarter of the year to generate provincial-level statistics and a smaller sample to provide regional estimates for the

rest of the year. The TCSD will soon come out with its recommendations on the NSO proposals for final consideration by the NSCB.

POLAND (from J. Bielecki)

In Poland political and economic changes have created demands for statistics for decision making in the market economy. Data users require more information based on household surveys.

There is an increasing need for data on the labour market, since the unemployment problem has become crucial. To satisfy this demand the Statistical Office started labour force surveying for the first time in May 1992. During the survey preparation the experience of some western countries in this field was taken into account. The standard concepts of EUROSTAT were introduced into the questionnaire design, to enable international comparisons. The survey is nation-wide, and is based on a household sample with panels that last at least one year. The survey is repeated every three months. The observation time is the week that includes the 15th day of the middle month of every quarter of the year.

The sample design is a two stage stratified sample. The PSUs are "Statistical Districts" (in Polish: Rejony Statystyczne) composed of approximately 350 dwellings. The sampling frame is the list of Statistical Districts (SD) and Census Enumeration Areas taken from the 1988 Population Census, updated every year.

The population of SDs is stratified regionally by voivodships (administrative regions) and by the size of towns and communities. For the whole country 194 strata have been formed. From each stratum the number of SDs chosen is proportional to the total number of dwellings in the stratum.

In total, 1238 SDs are selected with probability proportional to the number of dwellings in the SD. The secondary sampling units are the dwellings belonging to the selected SDs. On average about 15 dwellings are to be chosen from each selected SD. This number is modified with the purpose of obtaining a self-weighted sample with constant probability p=1/606. In practice, from 10 to 18 dwellings are selected from each SD in the sample. The total sample size does not exceed 18,500 dwellings distributed across the country.

It should be noted that before the survey started some preparatory work lasting six months, and ending with a pilot survey, was undertaken. Furthermore, an experimental survey was conducted in six selected voivodships (among 49) to obtain more information for the survey design.

For more information on the Polish Labour Force survey consult "Central and Eastern Europe Employment : Trends and Developments", Edited by Commission of the European Countries, in the series "Employment in Europe", No. 4, 1993.

REPUBLIC OF KOREA (from Seung-Kon Lim)

Data stored on an Information Management System Data Base amounts to 800,000 series spanning 30 years and 20 subject areas, such as population and social indicators. The database is updated the same day as the availability of the data is announced. Among these data, about 350,000 series in 20 subject areas are open to the public; this is the Korean Statistical Information System (KOSIS). The Central Government is developing the Administrative Information System (AIS) to synthesize various government computer systems. This system will be most useful in this computer age. Six major systems were selected for inclusion, one of them being KOSIS. To allow the public to use KOSIS, the National Statistical Office's mainframe is linked to the AIS's main computer using TCP/IP. Users need the following equipment: a PC supporting the Korean alphabet; MS-DOS as operating system; a VT 220 as communication emulator, and a modem with a capacity of 2400 BPS or more.

The NSO uses KOSIS to provide the following:

(i) Major domestic statistical information on population, social indicators, trade and industry, etc.

(ii) Statistical information from the U.N., IMF, and individual countries for international comparisons.

(iii) A database system to allow systematic loading of statistical data.

With KOSIS, users can retrieve detailed lists of statistical data, recent statistical data by classification, and figures for long term trends. Further, KOSIS can compute balances, rates of increase or decrease, component ratios and shares of contributions. It can reclassify analytical data according to value size and retrieve information about which agency

released the data, release date and time series. For this service NSO has five telephone lines and three DACOM-NET service lines. At present, 250 terminals in 140 institutions, including government agencies, research institutes, universities and public and private institutes, are linked to KOSIS. Oracle is a relational database system used to complement the hierarchical database. It uses a fourth generation language such as SQL, thereby providing users more friendliness. IMS/DB stores mainly time series data while raw data and summary data are stored in the Oracle relational database. Since 1989, about 70 application programs for data entry, tabulation, report and system management have been developed. The Oracle system is used for the "Economically Active Population Survey" and the "Mining and Manufacturing Survey".

For more information, contact Seung-Hee Han, Director, Information Management Division, Data Processing Bureau, National Statistical Office, Hanta Building, 647-15, Yoksam-dong,

Kangnam-gu, Seoul 135-080, Korea. (Tel.: (822) 222-1966, Fax: (822) 538-3874).

SWEDEN (from Anders Christianson)

Decentralization of the Swedish Statistical System

In the Survey Statistician July 1993, it was reported that major changes to the Swedish statistical system were underway. A committee chaired by the new General Director for Statistics Sweden, Mr Jan Carling, was appointed to implement the changes due to new laws concerning funding and coordination of national statistics. This committee has now published its proposals for fiscal year 1994/95. The proposal implies that Statistics Sweden maintains the responsibility for major branches of statistics, like labor market, population, national accounts, prices and consumption, transport and communications, and also for basic survey methods research (including publishing of the Journal of Official Statistics).

However, in other fields it is proposed that agencies that are identified as main users will become authorized to provide official statistics. This authorization implies that the agency is made responsible for production of statistics and for the control of their quality. The fundings are consequently directed towards the authorized agencies. The largest of these authorized agencies (in order of proposed magnitude of redirected fundings) are:

National Board of Housing, Building, and Planning, National Board of Health and Welfare, National Board for Industrial and Technical Development, National Agency for Higher Education, National Council for Crime Prevention, National Institute of Occupational Health, National Environment Protection Agency, National Institute of Economic Research.

The proposed size of the fundings redirected from Statistics Sweden to these and a few other agencies is 90 million Swedish kronor.

UNITED KINGDOM (from Roger Thomas)

Thirty six methodologists and statisticians from 14 countries gathered in Bath (U.K.) on September 6-9 for the 4th International Workshop on Household Survey Non-response. Research papers and country reports were presented and a series of group discussions was held. Among the main topics discussed were: research on doorstep introductions; the psychology of survey participation; the role of the interviewer in response/non-response; adjustment of results for non-response; and the results of a questionnaire survey on non-response addressed to survey organisations in a number of countries. Jean Morton Williams presented material from her recently published book "Interviewer Approaches".[1] (Shaun Cove, Office of Population Censuses and Surveys, tel: 071-396-2309.)

Demographic and post-enumeration checks carried out by the Office of Population Censuses and Surveys to estimate the coverage of the 1991 Census indicate that the enumeration may have missed around 3% of the resident population. It can be deduced that those missed were predominantly aged under 40 and predominantly male. It is believed that problems of under-enumeration have become appreciably worse since the previous census was held in 1981. An attempt was made to enumerate "roofless" and other homeless persons, but more generally there may have been some increase in deliberate evasion. In 1993 OPCS published first results of a post-enumeration check on the accuracy of answers given on census forms. (John Craig, OPCS, tel: 071-396-2136, Patrick Heady, OPCS, tel: 071-396-2240).

The census post-enumeration survey was also used to check the completeness of the Register of Electors in Great Britain. It is estimated that about 7% of those qualified to vote are missing

from the Registers, but in London the proportion rises to around 20%. It is also high for ethnic and other minorities. These findings have important implications for the use of the Electoral Registers as a survey sampling frame. (S Smith, OPCS, tel: 071-396-2328.)

A report on a survey of health and health-related behaviour amongst adults in the general population, carried out in 1991 by OPCS on behalf of the Department of Health, was published in Spring 1993. The survey comprised anthropometric, blood pressure and blood chemistry measurements, as well as an extensive interview. From 1993 the survey is continuous, covering a sample of 17,000 adults annually. Following application of government policy on competitive tendering the 1994 survey will be conducted by Social and Community Planning Research, an independent survey research institute. The main focus of the 1991 survey was on risk factors related to cardio-vascular disease, but other aspects of health and nutrition are being covered in subsequent rounds of the survey. (Mrs Joy Dobbs, OPCS, tel: 071-396-2010; Ms Patricia Prescott Clarke, SCPR, tel: 071-250-1866.)

(1) "Interviewer Approaches" Jean Morton Williams, Dartmouth 1993.

UNITED STATES (From D Kasprzyk)

The National Center for Health Statistics (NCHS) and the National Immunization Program of the Centers for Disease Control and Prevention (CDC) will conduct a telephone survey to obtain data pertaining to the immunization coverage of preschool children across the United States. Since 1992, CDC has awarded grants to 78 state and local areas to establish Immunization Action Plans (IAPs) to improve vaccination levels of children by their second birthday. The State and Local Area Immunization Coverage and Health Survey (SLICHS) is designed to monitor progress on these initiatives. It involves 1) the conduct of four annual surveys to be fielded between April 1994 and December 1997 in each of the 78 IAP areas; 2) descriptive analysis of the data; and 3) preparation and documentation of public use data files. Each of the four annual surveys comprises a survey of households requiring approximately 50,000 screening interviews each month using random digit dialing and computer assisted telephone interviewing methods. The survey questionnaire will be based on the 1993 National Health Interview Survey's Immunization Supplement. The target population for the initial 1994 immunization survey is children 19 through 35 months of age in each of 78 geographic IAP areas. The surveys will produce state and local area estimates combined for the last three quarters of 1994 and, thereafter, yearly moving averages produced each quarter for the proportion of children receiving measles, mumps, and rubella vaccine and three or more doses of diphtheria-tetanuspertussis vaccine; polio vaccine; and Haemophilus influenzae B vaccine (Hib). Additional questions regarding other health care initiatives may be added to the surveys during the 54-month period. For more information, write to Peter Hurley, National Center for Health Statistics, 6525 Belcrest Road, Room 1120, Hyattsville, MD 20782.

In 1991, the U.S. Congress established a Bureau of Transportation Statistics (BTS) in the U.S. Department of Transportation. The Bureau was created in response to long standing recommendations to develop a more effective Department of Transportation wide statistical program. The Bureau is responsible for compiling, analyzing, and publishing a comprehensive set of transportation statistics, including information on productivity in the various parts of the transportation sector; travel times, traffic flows, vehicle weights, travel costs of commuting; variables influencing travel behavior; availability of mass transit and the number of passengers served by mass transit; accidents and, in general, the condition of the nation's transportation system. The Bureau's first Transportation Statistics Annual Report is due to the U.S. Congress in January 1994. The report will include information on the items mentioned above. For further information, contact Dr Phil Fulton, U.S. Department of Transportation, 400 Seventh Street, S.W., Room 2104, Washington, DC 20590.

Statistical Policy Working Paper 21, Indirect Estimators in Federal Programs, a working paper of the Office of Management and Budget's Federal Committee on Statistical Methodology, was released in the summer of 1993. This paper, developed by the Small Area Estimation Subcommittee chaired by Wesley Schaible (Bureau of Labor Statistics), provides summary reports on eight U.S. programs that publish indirect estimators, that is, estimators that use values of the variable of interest from a domain and/or time period other than the domain and time period of the estimate being produced. Individual copies of the report are available from Maria Gonzalez, Office of Management and Budget, NEOB, Room 3228, Washington, DC 20503.