INTERNATIONAL STATISTICAL INSTITUTE

INTERNATIONAL ASSOCIATION OF SURVEY STATISTICIANS

SURVEY STATISTICIAN N°2

September 1979

CONTENTS

1. EDITOR'S LETTER

The Editor and the Editing Board of the Survey Statistician thank the members of the Association for their comments on the issue n°1 of the Survey Statistician published in February 1979. The observations and suggestions made were quite relevant. They mainly deal on the presentation, translation of the papers and the contents of the sections. The propositions will be submitted to the Board at the Manila Session ; this is the procedure that will be followed as much as possible for the next issues of the publication.

2. GENERAL INFORMATION

2.1 Reports on sessions and workshops

2.1.1. Meeting of the I.A.S.S. Council, Voorburg, Netherlands, June 22, 1979

Were present : Dr. P.V. Sukhatme, President of I.A.S.S., Pr. T. Dalenius, President elect, Dr. M.N. Murthy; Mr S.L. Diop, Dr. E. Lunenberg, Director of the ISI Permanent Office, Dr. S.S. Zarkovich, Dr. Sanchez-Crespo, President of the 1981 Program Committee, Mr G. Théodore.

A) IASS schemes in the field of survey methodology

Profiting by the presence of the Director of the ISI Permanent Office, the participants have had a large discussion about the preparation of the "U.N. Household Survey Capability Program" and the "Workshop on survey methodology for developing countries" held in London on June 21-23, 1978.

The IASS members present have reiterated their regrets that theAssociation had not been invited as such, but it has seemed more useful to see how the IASS could contribute to the various projects being prepared or considered (like the one described by Mr Duncan). The members of the meeting have also been informed by Dr Lunenberg that Messrs B. Davies and Rao's tour was meant to sensitize the authorities of visited countries and potential institutions giving financial funds. A meeting of the latter has been held besides in New York City on June 22-23, on this subject.

On the ground of these informations and various observations made, it has been decided that a brief would be prepared by the IASS, containing a couple of schemes liable to interest financing bodies. These schemes should be in the line of decentralization, since such seems to be the philosophy of the program of promotion of household surveys.

This brief will then be submitted to the ISI Executive Committee and transmitted to potential financing agencies. The ISI is the body in charge of organizing these consultations.

On the practical side, the IASS President will send a letter to Pr. Dalenius entrusting him officially with preparing a draft of this brief of schemes, in cooperation with Messrs Hansen, Duncan, Fellegi and Murthy (especially on the occasion of his journey to the United States in August), on the basis of the report on the meeting in London turned in by Dr. Lunenberg.

This brief will then be submitted to Dr. Kish and to any person close to the U.N. and whose access would facilitate the feasibility of these schemes.

The brief, thus completed, would then be sent to the IASS President for official transmission to the ISI Executive Committee. This transmission could take place after the Manila Session so that a final touch could be put on the occasion of the IASS Council meeting.

B) Sankhya series C

Various difficulties seem to hinder the edition and circulation of this review. It will be proposed, in Manila, that the Survey Statistician be the support of the articles previously intended for Sankhya series C. It will be suggested that articles be published in extenso in their original languages with a summary in English and French.

C) Cost sharing in the printing of the ISI Newsletter

Dr. Lunenberg explains that the financing asked from the various sections is proportional to the number of copies printed. The question will be raised in Manila (and should be put on the agenda of the ISI Permanent Office) that this cost be equal to the marginal cost of supplementary copies. This implies that if all the sections present the same request, the financial share of the ISI be increased.

D) Manila program

A few misprints in the Bulletin n°2 (for instance the omission of the date of the General Assembly of the IASS) make it necessary to provide for the insertion of a reminder of the IASS special events in the Survey Statistician n°2. The General Assembly should, if possible, be scheduled on Friday December 7 in the afternoon.

There will be no workshop organized before the session, but only informal meetings.

Twelve papers only have been received as yet. The deadline is postponed till September 30 and telegrams will be sent to late contributors.

E) Mexico Session

The officialization of the appointment of Dr. Sànchez-Crespo as President of the Program Committee has been reminded.

The composition of the Committee appears in Appendix I.

The Council Members present have approved this list which will be presented at the Manila Session.

It has been provided that the Scientific Secretary be ex-officio member of the Program Committee.

Dr Sanchez-Crespo will draft a document showing the classification of topics proposed for Manila. It will be transmitted to Pr. T. Dalenius as well as to the Council members in order to receive all complementary suggestions, so that after its final draft in November 1979, it can be submitted for approval in Manila.

F) The Dubrovnik plan

After discussion, the meeting has come to the following conclusions::

The topic retained must deal with errors in sampling surveys.

The meeting will have the format of a workshop or a symposium, not a seminar. This procedure will make the discussion easier.

The plan of this meeting will be set up by Dr. Zarkovich and Dalenius.

Duration : one week (probably in October 1980).

It would be useful to invite a statistician of the European Economic Commission in Geneva.

An official letter will be sent by the IASS President to the Statistical Institute of Beograd, informing them of the interest that the Association takes in the accomplishment of this plan.

G) Content and presentation of the Survey Statistician

A synthesis of the observations made about the Survey Statistician n°1 will be presented in Manila.

During the same session, the President of the IASS will officially appeal to members so that they contribute more papers and information.

In this view, the IASS Secretariat will prepare a form to be distributed systematically to all participants in the session. A second form will also be distributed asking IASS members to state their exact addresses. The Secretariat will provide the statistics of envelopes returned for change of address.

H) Board of Editors ISI Review

The name of Mr Lemel is approved as participant to this Board. An official proposition will be made to Mr Lunenberg.

I) Vote
The result of the vote appears in Appendix 2. It will be notified officially to due members.

J) Miscellaneous questions
The problem has been raised of the changing of our usual interlocutors in the Statistical Office of Mexico.
A topic should be provided, among those of Mexico, able to interest the local authorities (the problem of migrations might be raised).
The ISI Permanent Office will send to the IASS an invitation to take part in the meeting of the WFS group of experts which is to be held in London, on the two topics of mortality and migrations.

Appendix I

Composition of the 1981 Program Committee

J.L. Sanchez-Crespo, Chairman, Spain
K.R.W. Brewer, Australia
S.L. Diop, Senegal
M.E. Gonzalez ,USA
C. Hayashi, Japan
I.B. Thomsen , Norway

Appendix II

Results of votes on the election of IASS Officers and Council members

The ballots of the vote on election of Officers and members of the IASS Council that took place in April 1979, were counted on June 22, 1979 by the IASS Secretariat.
272 ballots were received ; 272 votes were counted.
The following people have got the position of :

President elect (1980-81)	S.S. Zarkovich (Yugoslavia)
Vice-presidents (1980-81)	J. Desabie (France)
	L.E. Rowebottom (Canada)
Scientific secretary (1980-81)	A.L. Finkner (USA)
Members (1980-83)	E.K. Foreman (Australia)
	K.T. de Graft-Johnson (Ghana)
	G. Kalton (United Kingdom)
	M.N. Murthy (India)
	H. Nisselson (USA)
	J.N.K. Rao (Canada)

2.1.2. IASS Workshop, Honolulu, January 24-February 16, 1979

An IASS Workshop in Survey Sampling was held February 5 to 7 with participants from Afghanistan, Bangladesh, Hong Kong, India, Indonesia, Malaysia, Mexico, Nepal, Philippines, Republic of Korea, Singapore, Sri Lanka, Taiwan, Thailand and the United States. We took the advantage of the presence of these statisticians, most of them from their national statistical offices, in a Census Sampling Workshop held on January 24-February 16, 1979 at the East-West Center, Population Institute, (L.J. Cho director), Honolulu, Hawaii, coordinated by L. Kish. Ten new members joined the IASS, and plans were made to form a regional coordinating committee. We urge that other regional workshops be similarly organized when the opportunity offers itself of sampling statisticians from several countries of a region meeting together. This may be more often feasible than finding funds specifically for regional workshops.
If you plan to organize a workshop, please inform the Committee on Workshops, Leslie Kish, chair.

2.1.3. European seminar on information systems for regional planning, Madrid, June 11-15, 1979

The European Ministers responsible for Regional Planning decided to develop technical cooperation in the field of regional planning. An Expert Committee on "Cartography, Statistics, Terminology" was charged with the implementation of the Ministerial resolutions. This Committee has in the past organized the following two European interdisciplinary seminars :
a) In 1975 in Enschede, Netherlands, the seminar on cartography and regional planning
b) In 1977 in Toulouse, France, the seminar on remote sensing and regional planning
The third seminar in this series dealt with "Information systems for regional planning" and was held in Madrid, Spain, during the days 11 to 15th of June.
The seminar is organized by the Instituto Geografico Nacional (Presidencia del Gobierno) and the Direccion General de Ordenacion y Accion Territorial (Ministerio de Obras Publicas y Urbanismo).
The theme of the seminar was : information systems for regional planning. The program included the seven following themes :
a. Regional planning : its variety of information needs and the factors influencing this
b. Regional planning : different types of information systems
c. Methods of analyzing the information needs
d. Editing and integration of the data needed into the different systems
e. Sources and storage of data
f. Analysis and output of data
g. Operational information systems

2.1.4. 8th meeting of the U.N. Statistics Advisory Committee for Food and Agriculture, Rome, April 23-30, 1979.

The 8th meeting of the U.N. Statistics Advisory Committee for Food and Agriculture took place in Rome on April 23-30 1979.
The following topics were talked about :
- food consumption survey
- social indicator program
- statistical program linked with the rural development and agricultural reform.
- methodology on the collection and reliability of post-crop cereal loss estimation

2.2 Manila Session

On December 6, 1979, the IASS General Assembly will be held from 12 to 2 p.m.

The IASS meetings are as follows :

1. Rotation and other resampling schemes — Organizer : I.P. David (Philippines)
   Invited papers :
   - Successive sampling - a review — B.D. Tikkiwal (India)
   - Rotation sample biases and their effects on estimates of change — B.A. Bailar (USA)
   - Some practical problems in successive sampling — G. Kulldorff (Sweden)
   - A paper (title missing)

2. Incomplete data — Organizer : W.G. Madow (USA)
   Invited papers :
   - Handling nonresponse in sample surveys — D. Rubin (USA)
   - Methodology and application of adjustments for nonresponse — R.A. Platek, G.B. Gray (Canada)

   Discussants : C. Hayashi (Japan), D. Singh (India)

3. Market intelligence surveys — Organizer : S. Tulya-Muhika (Uganda)
   Invited papers :
   - The treatment of nonresponse and incomplete response in market intelligence surveys — L. Frankel (USA)
   - Some statistical methods in market surveys with special reference to exploratory categorical data analysis — C. Hayashi (Japan)
   - Market research in a developing country - Indian case — P. Tandon (India)

   Discussants : K. Saito (Japan), Sir Maurice Kendall (UK)

4. Surveys in developing countries — Organizer : M.N. Murthy (UN, India)
   Invited papers :
   - Some methodological problems in large-scale sample surveys — F. Azorin-Poch , J.L. Sanchez-Crespo (Spain)
   - Problems in designing an integrated sample survey system for Thailand — N. Purakam (Thailand)
   - A paper (title missing) — P. Singh (Kenya)
   - Sample survey design in developing countries - Four illustrations of methodology — A.G. Turner, H. Woltman, R. Fay, C.D. Jones (USA)

   Discussants : B.T. Onate (Philippines), K.T. de Graft-Johnson (Ghana), C.S.O. Scott (UK)

5. Technical information exchange on statistical software (TIESS) with emphasis on survey and census processing — Chairman : R. Gnanadesikan (India); Organizer : I. Francis (New Zealand)
   Invited papers :
   - Software for processing — K. Foreman (Australia)
   - A comparison of software for processing and analyzing surveys — I. Francis (New Zealand)
   - A software system for analyzing surveys — B.V. Shah (India)

6. Sampling for consumer price indexes — Organizer : B.A. Boyes (USA)
   Invited papers :
   - Experiences in conducting and using family budget survey in the construction of a new consumer price index for major urban areas of Venezuela — B.A. de Khan, C. Leon (Venezuela)
   - Brazilian national price index program — F. de Melo (Brazil)

   Discussants : M. Hansen (USA), D. Levine (USA)

7. Enterprise statistics — Organizer : G.K.G. Forbrig (GDR)
   Invited papers :
   - Analysis and medium term forecast of the development of industrial branches in the People's Republic of Bulgaria — D. Balewski (Bulgaria)
   - Analysis of efficiency productivity and profitability of industrial enterprises — J. Balint (Hungary)
   - Actual problems of statistics in industrial enterprises and craftman's establishments — Hoang-Trinh (Vietman)
   - Statistical sampling in auditing — J. Kriens (Netherlands)

   Discussants : A. Donda (GDR), L. Biggeri (Italy)

8. Techniques for the processing or analysis of large data sets — Organizer : I. Francis (New Zealand)
   Invited papers :
   - Numerical techniques for large problems — W.M. Gentlemen (Canada)
   - Numerical edit and imputation — G. Sande (Canada)
   - Software for data management and analysis of large-scale biomedical studies — F. Stitt (Australia)

9. Privacy and confidentiality issues in surveys — Organizer : T.B. Jabine (USA)
   Invited papers :
   - Inferable individual information in released statistics — O. Frank (Sweden)

. Techniques for preserving statistical confidentiality — L.H. Cox (USA), G. Sande (Canada)

Discussant : — T. Dalenius (Sweden)

2.3 Other information

Sampling course with application to household surveys

An intensive course on probability sampling with applications to household surveys took place in Madrid on May 7-18, 1979.
The course was held in the headquarters of the the Spanish Central Statistical Office (Instituto Nacional de Estadistica) within the programme of technical cooperation with the Latin-American and Caribbean countries.
Both theoretical and practical points were developed. Methodological aspects related to sampling frames, use of supplementary information, and measurement and control of sampling and non-sampling errors were stressed.

3. PAST AND CURRENT SURVEYS

1975 Occupational Employment Survey (Canada)

The major impetus to the Canadian Occupational Employment Survey (OES) came from the Departments of Employment and Immigration and Statistics Canada in the early seventies. The general objective of this survey was to provide occupational labour distribution amongst the paid wage-earners in Canada.

Traditionally, this type of data had always been provided by the decennial Population Censuses. However, between 1961 and 1971, a new occupational dictionary was developed. The effect was to make the 1961 and 1971 Censuses Occupational distributions non comparable. Several sources for Occupational data were investigated. The resulting choice for cost reasons narrowed down to an Occupational Employment Survey using established Labour Surveys as a frame.

Employment

Occupational employment data represent all employees on the payroll of the reporting unit in the reference period. The employee concept used in the OES is consistent with that used in the Monthly Employment and Payrolls Survey ; hence, working owners or partners in unincorporated businesses and non-working directors are not included in the OES.Counted as employed, therefore, are all wage earners and salaried employees, regular and part time or casual, who are paid for their services or are on paid absence during the reference period.

Sample Design

Primarily for the sake of economy, the design of the Occupational Employment Survey (OES) was conditioned by the already existing structure of operations in two other major Labour Division surveys in the employment area : the monthly surveys of employment and payrolls (ES-1 and ZS-2) and the semi-monthly Job Vacancy Survey (JVS).
The OES is an interview survey of employers in which the reporting unit, which is contacted only once during the year, is usually the establishment; that is, the smallest industrial unit which is a separate operating entity capable of reporting all elements of basic industrial statistics including information on inputs and outputs needed to calculate value added as well as data on employment and payrolls. In order to provide provincial statistics, the survey often requires that the reporting unit be a sub-establishment unit which, although not capable of reporting all basic statistics, can report occupation, employment and payroll data for that particular location. The OES reporting unit is referred to as the Occupational Employment Reporting Unit (OERU).

For the purpose of drawing the sample, the population of OERU's is divided into six sectors. These are : ES-1, consisting of Business Establishments Employing 20 or More Persons ; ES-2 consisting of Business Establishments Employing Less Than 20 Persons ; Educational and Other Institutions ; Federal Government ; Provincial Governments and Municipal Governments.

The Occupational Employment Survey is designed as a stratified two-stage sample involving an interview phase of the reporting units ; no mailout is used in the OES. Within each sector, the population of OERU's is stratified by province, industry group or type (1) and size. For all sectors except the ES-2 sector, the first stage of the sample design takes the following form : each stratum, based on provincial location and industry type, is divided into two substrata which are referred to as the take-all and partial substrata. The take-all substratum may be briefly described as comprising the largest OERU's which contain 40 % to 50 % of the total employees in that stratum. The partial substratum is composed of the remaining OERU's. The take-all substratum OERU's are sampled at 100 % while the partial substratum OERU's are sampled at rates roughly proportional to the stratum's size. Once the substrata elements have been sampled, they are assigned systematically to one of the four calendar quarters (each quarter corresponding to three consecutive months) and then further subdivided to provide sample replication for subsequent variance estimation. Each selected unit is assigned to a cluster within a Statistics Canada regional office area.

Clustering is a subdivision of this area corresponding to the semi-monthly travel capabilities of the interviewers; hence, the workload within each quarter is divided into six groups, or clusters, one for each of the six collection periods or occasions in that quarter.

The ES-2 sample used for the OES is basically the Job Vacancy Survey interview sample. The ES-2 strata are sampled at prescribed rates and the sampled portion is allocated systematically to a multiple of at least 24 ES-2 panels. The number of panels in the sample is the same for all occasions and strata, namely,12. On each survey occasion, one panel rotates into the sample, stays for 12 occasions and then rotates out.

For those OERU' where total collection of the occupational data is not economical, a second stage of selection is used. This second stage, referred to as the within-firm sampling stage, involves a systematic random selection of occupational data with a prescribed sampling rate depending upon the size of the OERU.

The weights used in the estimation procedure for the ES-2 portion are suitably modified weights from the JVS's weighting system, while the weights for all other sectors are simple inverses of the appropriate sampling fractions.

The distribution of employees by occupation found by the OES is applied to the total number of employees obtained from the ES-1 and ES-2 surveys. This prorating is performed at either the industry or stratum level by quarter or averaged over all quarters depending on the tabulation required.

Sample size

Across all industries and sectors at the Canada level, some 80,000 reporting units were included in the 1975 sample; of these, approximately 64,000 were small firms of less than 20 employees. The survey resulted in the acquisition of job titles covering approximately 2,800,000 employees. When this value is related to the total paid worker employment reported by the Monthly Employment and Payrolls Survey, the size of the OES sample approximates 36 % at the Canada level.

Response Rate

The response rate is calculated by subtracting non-response from the sample and converting the value to a percentage of the sample. At the Canada level, across all sectors and industries, the response rate was approximately 87 %. The lowest response rate occurred in Sector 2, Small Firms, where the response rate approximated 84 %. In general, the effect of non-response in this survey is negligible due to the design requirement that distributions obtained from the responses are converted to absolute values based on the Monthly Employment and Payrolls Survey.

Calculation of Estimates and Variance

The estimates for any occupational class can be termed as being a combined domain ratio estimator. The associated variance estimator is made up of two parts : a variance due to inter-variation between reporting units and a variance due to sub-sampling within reporting units. The between variance is obtained via a two-replicate method. Further information concerning this survey may be obtained by writing to M.A. Hidiroglou or K.P. Srinath, 124 Coats Tower, Business Survey Methods Division, Statistics Canada, Tunney's Pasture, Ottawa, Ontario, K1A 0T6.

Transmitted by Dr M.P. SINGH
Statistics Canada

4. PAPERS

4.1 An error profile for the current population survey employment statistics

Introduction

The Subcommittee on Nonsampling Errors of the Federal Committee on Statistical Methodology decided to illustrate the ways in which non-sampling errors can affect survey statistics by constructing an error profile. The error profile constructed relates to the Current Population Survey (CPS) and, specifically, to the statistics on the level of employment. The error profile has been published as "Statistical Policy Working Paper 3 : An error profile : employment as measured by the current population survey" prepared by C.A. Brooks and B.A. Bailar of the Subcommittee with the advice of the other Subcommittee members and considerable help from other employees of the Bureau of the Census.

Because of the importance of the possible effects of nonsampling or measurement errors on survey statistics, it seemed important to the Subcommittee to identify all survey operations and then to look at any differences between the specification of an operation and its execution. Tore Dalenius, an advisor to the Subcommittee, suggested that an error profile give a systematic and comprehensive account of the survey operations that lead to a survey statistic and then include an assessment of the presence or absence of a deviation between the design and execution, the size of any deviation and a measure of the impact, if any.

The CPS was selected as the object of a profile because it is a survey with a long history, much has been written about it, a considerable amount of research has been done on the survey methods used, and it is a survey that produces data on important statistics. Though the focus is on employment statistics, much of the discussion relates to other statistics produced from the survey.

The purposes of the error profile are as follows :

- to illustrate how an error profile is created in an effort to encourage other statisticians to provide error profiles for the major recurrent survey statistics
- to compile in a single document the sources of error and the information that is available about these sources of error and their impact
- to illustrate the need for controlled experiments to measure nonsampling errors because of the lack of knowledge of the impact of these errors
- to stimulate development of a mathematical model that will reflect the ways in which the errors from different sources interact.

This article gives a brief description of the organization of the report, the highlights of the findings for the major survey operations, and some conclusions about the usefulness of the error profile.

Organization of the report

The report is organized into chapters, most of which discuss a major survey operation. A description is given of the process and then a discussion of what is known about possible errors or biases arising from the process follows. The major operations studied are :

A. Sampling design
  1. Frames
  2. Sample selection
  3. Quality control of sampling process
B. Observational design
  1. Data collection procedure
  2. Questionnaire design
  3. Data collection staff
  4. Interviewer training
  5. Quality control of field work
C. Data preparation design
  1. Data input operations
  2. Cleaning, editing and imputation

3. Quality control of data processing

D. Production of estimates

1. Weighting procedure
2. Estimation procedure
3. Quality control of estimation procedure

E. Analysis and publication

Findings for major survey operations

Certain areas of the survey operations have received a great deal of attention and much is known about the limitations inherent in these operations. Other survey operations have not been studied extensively and little is known about the impact of errors arising from these operations. A brief summary of the highlights of the findings for each of the operations listed above is given.

Sampling design . - The sampling design of the CPS and its implementation has been the focus of a considerable amount of work at the Bureau of the Census. Efforts to improve the sampling frame are ongoing activities. The coverage bias resulting from the sampling frame not completely covering the entire universe of households is estimated to be on the order of three percent. However, there are missed households for reasons other than sampling frame deficiencies, and missed persons within households, as well. The impact of the undercoverage in the 1970 census and the additional undercoverage in the CPS is discussed and some data are presented about the possible impact on employment statistics.

Observational design . - The potential sources of error associated with the elements discussed in this chapter are many, but little is known about the existence or the impact of any errors. This is an area in which the need for methodological studies is clearly illustrated.

The interviews are conducted in the CPS by means of a mixture of personal visits and telephone calls. No definitive data are available comparing the effect on labor force statistics of these two types of interviewing methods, though evidence is available from other studies that for some statistics, the mode of interviewing has an effect on the data. A controlled experiment is currently underway to compare telephone with personal visits for collecting labor force data.

Proxy respondents are permitted in the CPS. In fact, a household respondent answers for all persons in the household. Experiments done several years ago show that the use of proxy respondents may result in slightly fewer persons being classified in the labor force. No recent data are available, but a controlled experiment is now in process.

A fair amount of work has been done to develop and test the questionnaire used in the CPS. However, not enough is known about the response quality when supplemental questions are added to the questionnaire or the effect of placement of questions. For example, data are shown that a particular set of additional questions on the interview schedule may result in more persons being classified as employed. A quality control program on interviewer performance yields a series of useful measures. However these measures cannot be translated into a summary measure to describe quality of published data. A study of interview variability showed that there was considerable variability among interviewers on employment items.

The method of training interviewers may be related to the quality of survey statistics. The results of a study following a training session designed to find out about the extent of errors made by interviewers is described.

This chapter is long and encompasses many survey procedures. The findings are quite variable, depending upon the complexity of the procedure. The findings are fragmentary and illustrate those areas where knowledge is lacking. More importantly, there is no model by which the various pieces of information can be combined to describe their joint effects on data.

Data preparation design . - The microfilming and FOSDIC operations are under tight controls, and error rates are presented. There is extensive documentation on this phase of the survey operations. The effects on the survey data are minimal.

Less is known about the effects of the editing and imputation procedure. Very little quantifiable data are available. A study on this is now in process.

Production of estimates . - The estimation process is extremely complex. A great deal of effort is made in the estimation procedure to reduce the variance of the estimates. However, some aspects of the estimation procedure may themselves increase the level of nonsampling error, while other aspects may reduce the level. For example, the "rotation group bias" or the bias associated with the length of time a respondent is in the sample; is larger for estimates of level with the estimation scheme used than with an intermediate estimation procedure.

The effect of the weighting procedure is discussed, some problems in the application are shown, but the impact on employment statistics is unknown.

The data are seasonally adjusted. A discussion of the seasonal adjustment process is given, but little is known about the impact of the process on the estimates.

Analysis and publication . - Analysis and publication of the labor force estimates take place at the Bureau of Labor Statistics. Potential sources of error are discussed, but no data are available on the possible impact.

Conclusions

The Subcommittee found the project useful. It was possible to list the various potential sources of error and hypothesize about the possible effect. It was very useful to compile in a single document all the information available about these sources of error and their impact. The information can be used as an aid for :

1. a total survey design
2. establishing a foundation for a comprehensive methodological research program ; and
3. the improvement of some problem procedures.

It was extremely useful to see where the gaps in knowledge about the impact of nonsampling errors occurred. Also, the project has spurred new interest in developing a mathematical model that will reflect the interaction of errors from various sources.

B.A. BAILAR, C.A. BROOKS
Bureau of the Census, US Department of Commerce

4.2 Correlation analysis from multi-stage sample surveys

The estimation of correlations or covariances between survey responses to different questions is an important input to many multi-variate analyses of the survey data. For example, factor analysis is often performed on the responses to a series of attitudinal questions on a particular topic.

Such data is often collected by a survey using a multi-stage sample design, in the calculation of population correlations, the "clustered" multi-stage nature of the design is ignored, and analysis assuming a simple random sample has been taken. What effect has this assumption on the accuracy of the analysis ?

Holt (1977) gives a model that reflects the multi-stage design in the calculation of population correlation coefficients. This note empirically compares these estimates of correlation with the estimates based on the assumption of a simple random sample.

The study is based on a data set of 385 observations from a pilot test for a Survey of Housing. Many attitudinal questions were asked and it was desired to undertake factor analysis on these questions to assist with their interpretation. There were two stages of selection with at the first stage 39 blocks being selected probability proportional to size and at the second stage households selected within blocks using within PSU sample intervals such that the overall sample was self-weighting. In total, 222 households were selected. Generally all individuals within selected households were interviewed but there was some non-response from individuals within selected households.

Following Holt (1977) we have used the model :

$$y_{hijk} = a_h + u_{hi} + w_{hij} + \varepsilon_{hijk} \qquad (h = 1,2\,;\ i = 1,\dots,a\,;\ j = 1,\dots,b_i\,;\ k = 1,\dots,n_{ij}) \qquad (1)$$

where the subscript h refers to variable number, i to block, j to household, and k to person.

The random variables u_{hi}, w_{hij} and ε_{hijk} are assumed to have zero means and variances σ_{h1}^2, σ_{h2}^2 and σ_{h3}^2 and for fixed h, to be uncorrelated within each set and between each set. But for each individual, correlated block, household and individual effects permit the introduction of correlation between two variables. Therefore, the assumed variance structure is :

$$\begin{aligned}
\operatorname{cov}(\varepsilon_{hijk}, \varepsilon_{gpqr}) &= 0 \qquad (i,j,k) \neq (p,q,r) && (2)\\
&= \sigma_{h3}^2 \qquad (h,i,j,k) = (g,p,q,r) && (3)\\
&= \rho_3\, \sigma_{13}\, \sigma_{23} = \gamma_3 \qquad g \neq h \quad (i,j,k) = (p,q,r) && (4)
\end{aligned}$$

$$\begin{aligned}
\operatorname{cov}(w_{hij}, w_{gpq}) &= 0 \qquad (i,j) \neq (p,q) && (5)\\
&= \sigma_{h2}^2 \qquad (h,i,j) = (g,p,q) && (6)\\
&= \rho_2\, \sigma_{12}\, \sigma_{22} = \gamma_2 \qquad g \neq h \quad (i,j) = (p,q) && (7)
\end{aligned}$$

$$\begin{aligned}
\operatorname{cov}(u_{hi}, u_{gp}) &= 0 \qquad (i \neq p) && (8)\\
&= \sigma_{h1}^2 \qquad (h,i) = (g,p) && (9)\\
&= \rho_1\, \sigma_{11}\, \sigma_{21} = \gamma_1 \qquad g \neq h \quad i = p && (10)
\end{aligned}$$

With these assumptions, it is possible to show by analysis of variance techniques that estimates of $\sigma_{h1}^2, \sigma_{h2}^2, \sigma_{h3}^2$, γ_1, γ_2 and γ_3 are given by :

$$\hat{\sigma}_{h3}^2 = \Big[\sum_{i,j,k} y_{hijk}^2 - \sum_{i,j} \big(y_{hij.}^2 / n_{ij}\big) \Big] \Big/ (N - b) \qquad h = 1,2 \qquad (11)$$

$$\hat{\gamma}_3^2 = \Big[\sum_{i,j,k} y_{1ijk}\, y_{2ijk} - \sum_{i,j} \big(y_{1ij.}\, y_{2ij.} / n_{ij}\big) \Big] \Big/ (N - b) \qquad (12)$$

$$\hat{\sigma}_{h2}^2 = \Big[\sum_{i,j} \big(y_{hij.}^2 / n_{ij}\big) - \sum_{i} \big(y_{hi..}^2 / n_{i.}\big) - (b-a)\, \hat{\sigma}_{h3}^2 \Big] \Big/ \Big[N - \sum_{i,j} \big(n_{ij}^2 / n_{i.}\big) \Big] \qquad h = 1,2 \qquad (13)$$

$$\hat{\gamma}_2 = \Big[\sum_{i,j} \Big(\frac{y_{1ij.}\, y_{2ij.}}{n_{ij}} \Big) - \sum_{i} \Big(\frac{y_{1i..}\, y_{2i..}}{n_{i.}} \Big) - (b-a)\, \hat{\gamma}_3 \Big] \Big/ \Big[N - \sum_{i,j} \big(n_{ij}^2 / n_{i.}\big) \Big] \qquad (14)$$

$$\hat{\sigma}_{h1}^2 = \Big[\sum_{i} \big(y_{hi..}^2 / n_{i.}\big) - \frac{y_{h...}^2}{N} - \Big(\sum_{i,j} \big(n_{ij}^2 / n_{i.}\big) - \sum_{i,j} n_{ij}^2 / N \Big) \hat{\sigma}_{h2}^2 - (a-1)\, \hat{\sigma}_{h3}^2 \Big] \Big/ \Big[N - \sum_{i} \frac{n_{i.}^2}{N} \Big] \qquad (15)$$

$$\hat{\gamma}_1 = \Big[\sum_{i} \Big(\frac{y_{1i..}\, y_{2i..}}{n_{i.}} \Big) - \frac{y_{1...}\, y_{2...}}{N} - \Big(\sum_{i,j} \big(n_{ij}^2 / n_{i.}\big) - \sum_{i,j} n_{ij}^2 / N \Big) \hat{\gamma}_2 - (a-1)\, \hat{\gamma}_3 \Big] \Big/ \Big[N - \sum_{i} \frac{n_{i.}^2}{N} \Big] \qquad (16)$$

$$N = \sum_{i,j} n_{ij} \quad , \quad b = \sum_{i} b_i$$

Once estimates of the variance and covariance components are calculated, estimates of various correlations can be made if we make the assumptions implicit in the model. Firstly, it is of interest to examine whether there is any clustering in the variables being considered. Estimates of the variance components are shown in Table 1. It can be seen from the size of the first two variance components that there is significant clustering in the variables being studied.

Table 1 : Estimates of variance components

Variable	Variance component		
	$\sigma^2_{R_3}$	$\sigma^2_{R_2}$	$\sigma^2_{R_1}$
Rating or locality	0.434	0.311	0.290
Air pollution	0.301	0.325	0.130
Traffic noise	0.337	0.414	0.284
Shopping facilities	0.434	0.238	0.173
Doctors & dentists	0.895	0.562	0.101
Feelings towards dwg	0.265	0.246	0.037
Layout of dwg	0.251	0.289	0.045
External design	0.380	0.230	0.045

The following estimated correlations are tabulated in Tables 2 and 3.

a. The correlation between answers to two questions given by the same individual $(\hat{\rho}_A)$

b. The correlation between an individual's answer to one question and that of a second individual, resident in the same household, answering a second question $(\hat{\rho}_B)$

c. Same as b except for individuals in different households within the same block $(\hat{\rho}_C)$

d. The correlation between individuals' (within the same household) answers to the same question $(\hat{\rho}_D)$

e. Same as d, only for individuals in different households within the same blocks $(\hat{\rho}_E)$

Also tabulated is $\hat{\hat{\rho}}_A$, the estimated correlation between answers to two questions given by the same individual, assuming the data comes from a simple random sample. The figure in brackets is the estimated standard error on $\hat{\hat{\rho}}_A$ assuming a simple random sample and using the well known result that $\frac{1}{2}\log_e\frac{1+r}{1-r}$ is approximately normally distributed with mean $\frac{1}{2}\log_e\frac{1+\rho}{1-\rho}$ and variance $1/(n-3)$ where r is the estimate of the correlation coefficient, ρ its' true value and n the sample size.

Table 2 : Correlations with "rating of locality"

Variable	Correlations					
	$\hat{\rho}_A$	$\hat{\hat{\rho}}_A$	$\hat{\rho}_B$	$\hat{\rho}_C$	$\hat{\rho}_D$	$\hat{\rho}_E$
Air pollution	-0.215	-0.219 (0.048)	-0.163	-0.156	0.602	0.172
Traffic noise	-0.208	-0.208 (0.048)	-0.180	-0.137	0.674	0.274
Shopping facilities	0.305	0.307 (0.046)	0.189	0.057	0.487	0.205
Doctors & dentists	0.208	0.207 (0.048)	0.174	0.053	0.425	0.065

Table 3 : Correlations with "feelings towards dwelling"

Variable	Correlations					
	$\hat{\rho}_A$	$\hat{\hat{\rho}}_A$	$\hat{\rho}_B$	$\hat{\rho}_C$	$\hat{\rho}_D$	$\hat{\rho}_E$
Layout house	0.609	0.607 (0.033)	0.506	0.075	0.572	0.078
External design	0.497	0.496 (0.039)	0.291	0.046	0.420	0.069

As can be seen from Tables 2 and 3, there is close agreement between the two estimates of correlation to two questions answered by the same individual. The effect of the multi-stage clustered design is not significant despite the fact that the first two variance components are quite high, indicative of clustering on the variables.

of interest. This clustering is also reflected by the correlations $\hat{\rho}_a, \hat{\rho}_c, \hat{\rho}_b$ and $\hat{\rho}_e$ which are generally significantly different from zero.

This is only one set of data, but these is no reason to suspect it is atypical. As is well known from Kish and Frankel (1974) and others, design effects on the estimates of variances of correlations are significant and the sample design should be taken into account when estimating variances, making significance tests etc. However, results of this study suggest that ignoring the sample design when using factor analysis and other multivariate techniques that depend on the population correlations, will not seriously jeopardise the accuracy of the analysis.

Bibliography

Holt D. (1977), Correlation analysis using survey data, Paper presented at the 41st session of the International Statistical Institute, New Delhi, and to be published in B.I.S.I.

Kish and Frankel M.R. (1974), Inference from complex samples, JRSS B, 86, 1-37

S.B. GERHARDY, D.J. TREWIN
Australian Bureau of Statistics

4.3 An appraisal of the sample survey of small-holding perennial crops, in Papua New Guinea, 1975/1976

Introduction

In order to meet the demands within the Planning, Economics and Marketing Branch of the Department of Primary Industry as well as to help the Provincial Rural Development officers (P.R.D.O.) in developing their future area programmes, a sample survey on number of growers and area under perennial crops cultivated by smallholders in Papua New Guinea during 1975/76 was embarked by the Statistics Section of the Department. The scope of the survey was restricted only to main perennial crops, viz, coffee (Arabica and Robusta), cocoa, coconut and rubber, which covered more than 80 % of all agricultural exports during 1975/76. As it became apparent that the existing resources within the Statistics Section would not permit a sample large enough to yield meaningful estimates at census divisional levels, estimates only at district and provincial levels were attempted.

Survey Methodology

The type of sample employed at district levels was a stratified two-stage one, the stratification made according to crop zones (up to a maximum of five zones in a district). 5 villages were chosen by simple random sampling (without replacement) as first stage sampling units, the number of villages in each zone being proportional to the total number in that zone within the district. The list of villages were made available from recent electoral rolls. After the first stage units, i.e., villages were selected, the electoral rolls were consulted again and the perennial crops by the electors within each selected village were recorded with the help of village elders and other knowledgeable persons. The second stage units, i.e., 5 electors growing a particular crop were selected again by simple random sampling (without replacement) from the list of crop growers in respect of each crop thus prepared. While the total number of trees in all "gardens" (i.e. plots) of each selected respondent (i.e. crop growers) along with his partners, if any, were counted, another subsample was drawn from each garden, a random triangular "cut" (i.e. a specified area enclosed by pegs and ropes to form a right-angled triangle), to estimate the total area of the garden, as well as the proportion of trees according to maturity, i.e., bearing and non-bearing (the latter including new plantings and senile trees). The area of the "cut" varied, depending on crops and/or mixtures. The area considered suitable for coffee was 0.01 ha, that for cocoa (sole planted) was 0.02 ha whereas for other crops and mixtures an area of 0.05 ha was considered.

Organisation

The actual enumeration was carried out by the district rural officers under the guidance of the P.R.D.O's and overall supervision by the Statistics Section of the Department of Primary Industry. For every crop, five booklets were used, one for each selected village. On each booklet, relevant information was required to collect on crop zones, villages, electors, growers and their partners as well as gardens, numbers of trees - total as well as numbers within specified triangles with breakdown according to maturity for each grower. The training programme was conducted in the provinces by the staff of the Department between July and September 1976. Major part of the field survey was over in early 1977. The Statistics Section was responsible for the design, execution and analysis of the survey. A desk computer "CANOLA SX100" was used for analysis. The summary of information is provided in the table below.

Table 1 : Area under small-holding perennial crop in Papua New Guinea (1975-76) (area in hectares)

Crop	Bearing	Immature	Senile	New plantings	Total	Coefficient of variation (%)
Coffee-Arabica	24,782	3,745	1,257	2,390	32,174	8.4
Coffee-Robusta	6,939	1,051	91	350	8,431	18.3
Cocoa	40,987	15,631	3,026	13,701	73,345	17.1
Coconuts	105,629	41,578	5,102	15,037	167,346	13.9
Rubber	930	2,113	42.	500	3,585	38.2

The results in more details are available at a provincial and district levels in D.P.I. Office. For variance estimates Cochran (1) is referred.

Comments and critical appraisal

As it has been mentioned earlier, the scope and coverage of the present survey were restricted mainly due to limited resources available within the Department of Primary Industry. Apart from administrative problems, however, some of the methodological and operational problems that became apparent during fieldwork and subsequent analysis are briefly mentioned in the following paragraphs :

a) Sample size. Limitations in sample size was the main contributor towards a large variance in the survey. Coefficient of variation of the order of 50 per cent or more have been observed in several important crop-growing provinces (e.g., Arabica coffee : Southern Highlands-0.55 ; cocoa : Morobe-0.63 ; rubber : Western-0.54, etc.). It was apparent that when the between village variations are substantial , a sample of 5 out of a total of more than

100 villages in majority of districts is not expected to yield much precise results. This was also demonstrated by the results of two-earlier sample surveys in the country by Wilson-Evans (1975)[2] -coffee and H. Coulter (1976)[3] - cocoa. One way to circumvent this problem could be to "cluster" the first-stage units which, though likely to sacrifice some efficiency would still be quite practical from operational point of view and has been used with success in a number of cases.

b) Sampling scheme. The sample considered at the present survey was a stratified two-stage, one, stratification being made according to crop zones within a district. The scheme encountered a lot of operational problems due to difficulty in crop-zoning. Some alternative forms of stratification could be considered like stratification of villages, according to number of electors. This was tried by Coulter (1976)[3] in his cocoa survey. Secondly, when the first-stage units vary largely in size (i.e. crop area), a sample with probability proportional to size (or estimated size) briefly, p.p.s. or p.p.e.s.) may lead to a more efficient estimate [1]. Lastly, it also appears that enumeration of five respondents individually for each crop makes the whole exercise unnecessarily burdened at least in some provinces. It would probably be operationably more convenient to select, say 10 respondents from a selected village and then enumerate all the specified crop grown by them.

c) Area measurements. In the present survey, area of a "garden" belonging to a respondent has been measured through indirect means; in other words, estimation of area was based on the count of a small triangular area. A direct measurement of area, say, with the help of compass and measuring tape [4] is definitely preferable which could help in reducing the error of estimation. Secondly, more research seems to be necessary to determine an optimum shape and size of the "cut" (i.e. right angled triangle of specified size) though similar experiments have been carried out in other countries, particularly for yield estimation through "crop-cutting" trials.

References

1) Cochran, W.G.-Sampling techniques, 2nd Edition : Chapter 11 ; J.W. Wiley & Sons, New York, U.S.A. (1963).
2) Wilson, L.D. and Evans, G.B.A.-Sample Survey of Small-holder Coffee Producers; Rural Economics and Commodity Marketing Branch, D.A.S.F. (Presently P.E.M. Branch, D.P.I.)(1975).
3) Coulter, H.-Sample Survey of Small-holder Cocoa, North Solomons ; D.P.I. Area Office, North Solomons (1976).Mimeographed.
4) Dalenius, T.-Estimation of Areas in Agricultural Statistics; F.A.O., Rome (1965).

Appendix : Estimation Procedure

The formulae for estimation and variances are based on Cochran's Sampling Techniques (1963) : Chapter 11. (pp.304-305). Let us consider estimation for a stratum (i.e. zone) which has a two-stage sample, in which village is the first stage and grower the second stage unit.

Let t_{ij} = Number of trees within the specified triangles from all gardens under a particular crop belonging to j^{th} grower in i^{th} selected village.

T_{ij} = Corresponding total number of trees from the whole of gardens.

n_{ij} = Total number of partners operating for the crop.

and A = Area of the specified triangle. (A = 0.01,0.02 or 0.05 ha. depending on crop or crop-mixture).

Then, $y_{ij} = \frac{T_{ij}}{n_{ij}\, t_{ij}} A$ the estimated area under the crop for j^{th} grower within i^{th} selected village. (1)

$(i = 1, 2 \text{ ---------- } n \; ; \quad j = 1, 2 \text{ ---------- } m_i)$

Also, let M_i = Total number of crop-growers within i^{th} selected village. ($i = 1, 2$ ---------- n).

Then, estimated area under the crop of the i^{th} village is :

$$z_i = \frac{M_i}{m_i} \sum_{j=1}^{m_i} y_{ij} = M_i \bar{y}_i$$

where : $$\bar{y}_i = \frac{1}{m_i} \sum_j y_{ij} \qquad (2)$$

therefore, estimated total area of the stratum is :

$$\hat{Y} = \frac{N}{n} \sum_{i=1}^{n} z_i = \frac{N}{n} \sum_i \frac{M_i}{m_i} \sum_j y_{ij} \qquad (3)$$

The estimated area can then be totalled over strata to provide estimates for the districts which in turn could add to provincial and finally to country totals.

Variance calculation

For reasons of simplicity, y_{ij} is considered to be a random variable, ignoring the fact that it is here essentially a ratio estimate. This may lend a small negative bias for variance estimates.

Let : $$s_{wi}^2 = \frac{1}{m_i - 1} \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2$$

and , $$s_b^2 = \frac{1}{n-1} \sum_{i=1}^{n} (z_i - \bar{z})^2 \qquad (4)$$

Where , $$\bar{z} = \frac{1}{n} \sum_i z_i$$

Then, the variance estimate of $\hat{Y}$ is :

$$\hat{Var}(\hat{Y}) = \frac{N}{n} \left[(N-n) s_b^2 + \sum_{i=1}^{n} \frac{M_i (M_i - m_i)}{m_i} s_{wi}^2 \right] \qquad (5)$$

Therefore the estimated standard error (S.E.) of $\hat{Y}$ would be :

$$\hat{S} = \sqrt{\hat{Var}(\hat{Y})} \qquad (6)$$

Since the stratum variances are additive, estimated variance for the district may be obtained as a total and those for the province and the country similarly.

Prepared by C. KOLEY
Department of Primary Industry
Papua New Guinea

4.4 Note on random sampling with unequal probabilities without duplication

When making n random selections with p_i probabilities in a population whose units are refferred to as i=1,2,... N, it sometimes happens that some units (generally with high p_i) be drawn at least twice. The probabilities of inclusion are P_i ; np_i represents the number of times (i) is expected to be drawn. Let us deal with the ordinary case when n = 2. The unbiased linear estimator of $\sum_i x_i$ (where x is an additive variable on the units i) usually is :
$S\ x_i/2p_i$, i.e. $x_j/2p_j + x_k/2p_k$ (j and k referring to the sample). If j = k, there are many different methods to obtain two distinct units (rejective, successive, Midzuno's, Sampford's, etc...). All of them lie under the disadvantage of altering the probabilities and therefore the estimator. Very often, we will ignore this and be inclined to under-represent the units with a large p_i (in favour of the others).

Here is one more method to overcome the difficulty elegantly. If (i) is drawn twice, it is agreed that in the second random selection, (i) is replaced by (i - 1) or (i + 1) (alternatively), so that the probability p_i^2 of the cell (ii) of the table disappears, distributed between $p_{i-1}\ p_i/2$ (cell i,i-1) and $p_{i+1}p_i/2$ (cell i,i+1). If the population is infinite (i varying from $-\infty$ to $+\infty$), the p_i are infinitely small at both ends and should not be taken into account. If the population is finite, it is closed on itself, i = 1 being the successor of i = N.
Thus $P_i = p_i + p_i\ (1 - p_i + \frac{1}{2}\ p_{i-1} + \frac{1}{2}\ p_{i+1}) = 2\ p_i + D_i$ the deviation D_i being, by the way, equal to $p_i\ D^2 p_{i+1}/2$.

In order to disturb the P_i (near $2p_i$) the least possible, as we have no linear function of i for p_i, it is suitable (or so it seems) to order the p_i according to their decreasing values. [This looks like WILK's technique (ISI Session, Brussels, 1958) where we make pairs of units having almost equal p_i, replaceable if one is drawn twice] . In order to get rid of the terminal difficulty (i,N pair) we should rather set up two sub-populations : i=1,2,... M with increasing p_i, i = M+1, ... N with decreasing p_i (with p_1 and p_N close).

In case of duplication, WILKS assigns the probability 1 and 0 to other units, SAMPFORD assigns them probabilities $(N-1)^{-1}$; we suggest 1/2, 1/2 and 0. That is all.

P. THIONET
Université PARIS IX, France

4.5 An Assessment of the Reliability of the Indonesia Fertility Survey Data

1. THE INDONESIA RELIABILITY STUDY

1.1 Introduction

Indonesia participated in the World Fertility Survey (WFS) as part of the Intercensal Survey 1975-1976. Little information is available on the reliability of data obtained through surveys, and in Indonesia such information is lacking altogether. A special study was therefore conceived as a follow-up of the Indonesian Fertility Survey/WFS, to obtain an indication of the reliability of the data generated by that survey.

In consultation with WFS headquarters, it was decided to carry out a reliability study in all six provinces covered by the IFS/WFS, using a sub-sample of the successfully interviewed respondents.

1.2 Objectives

The primary objective of the Indonesia Reliability Study (IRS) was to measure the reliability of data on fertility, dates and ages. Another objective was to measure the reliability of some additional selected variables : these were factual data (background variables and data on marital status and situation) and opinion and attitudinal data. Finally, it was intended that the study should obtain indications of the possible sources of response discrepancies.

It was decided that the study would measure the reliability of the data in their crude form, that is as data collected in the field. The data were therefore not office edited, and hence differences reported could be attributed to the field operation. However, this report is based on coded and punched data, and it is thus possible that there are some coding and punching errors. Procedures were used to minimize these errors. The study is thus basically concerned with response variability and to a lesser degree with the sources of discrepancies.

1.3 The design

1.3.1. The design chosen.

To measure the reliability of data through response variability, one needs to obtain an independent replication of the field procedures.

To obtain an indication of the source of the discrepancies, reconciliation of discrepant results is needed. The design chosen for this Study was therefore a re-interview of a subsample of the IFS with reconciliation of discrepant results.

To reduce the probability that interviewers might remember the answers a respondent had given in the IFS, it was decided that interviewers should not interview the respondents they had interviewed in the original study, and in addition, they should not have previous knowledge of the results of the original interview. To enhance the validity of the reconciliation interview, it would be carried out by field supervisors who were considered to be more suited to this difficult task.

1.3.2 The questionnaire and other field documents

Since it had been decided that the interviewers for the re-interview should not have previous knowledge of the results of the original interview, it was sufficient that only the coded data of the original interview should be available in the field for reconciliation.

1.3.3. Criteria for the reconciliation interview

The reconciliation interview involved only selected questions. In line with the objectives of the IRS, most important were the group of questions on fertility, age and dates. For the fertility questions, whenever a discrepancy was found reconciliation was attempted. For the age and date questions, reconciliation was attempted only if the difference between the results of the two interviews was more than two years. Background variables are generally considered to be simple data, not subject to change, and it was thought that they might therefore serve as a standard of comparison for the reliability of the fertility questions. Before the field work it was decided that any discrepancy in the background variables and in opinion and attitude questions should be reconciled.

During the field work this policy was modified. Owing to the large number of reconciliation interviews which would have had to be held, it was decided that if the only inconsistency found was of a background variable no reconciliation interview would be held. If other inconsistencies were found which required reconciliation, then inconsistencies in the background variables, if any, would also be reconciled. This restriction had to be introduced in view of the shortage of time available for the field work. This is regrettable because the results on background variables from the reconciliation interviews were interesting.

1.3.4. The sample

The Indonesia Fertility Survey was based on a probability sample of noninstitutional households resulting in individual interviews with ever-married women between the ages of 10 and 49 years. In the 10,504 households selected, 9,136 eligible respondents were found and interviewed in detail.

The sample for the IRS was a sub-sample of the main survey. It had been decided that the number of interviews should be evenly spread over the six provinces covered in the IFS. The total sample size was determined at approximately 500 respondents, on the basis of the available time and man-power. Except for Jakarta, which is wholly urban, the sample was evenly distributed over urban and rural areas in each province. Because the size of the urban and rural clusters differ, different numbers of clusters were selected for rural and urban areas.

In each province the required number of clusters was selected purposively from those selected for the IFS. Special care was taken to select clusters that were "typical" for the urban and rural population of the province. In each selected cluster all of the respondents successfully interviewed in the IFS were included for the IRS interview. Therefore, the sample of the IRS is not a probability sample but is well spread among the provinces and the urban and rural areas.

1.4 The field work

Sample outcome. Of the 532 respondents selected for the re-interview, 498 gave a complete re-interview.

2. MEASURING RELIABILITY

2.1 The concept of reliability

In general terms, reliability can be defined as "the extent to which a measurement remains constant as it is repeated under conditions taken to be constant". From this definition it becomes clear that when analysing reliability the individual measurements must be considered ; that is, if the measurement is reliable the distribution over the different categories of the variable studied will also be constant. However, constancy does not guarantee reliability. In social surveys, measuring reliability adequately is further complicated by the fact that the respondent may recall her previous response, and may simply repeat this. This effect cannot be eliminated in social surveys, but is partly dealt with in the design of the reliability study.

The above description of reliability is based on two crucial assumptions : constancy of the phenomena and constancy of conditions. In general for surveys constancy of the phenomena is assumed, or taken into consideration, when designing an instrument (in this case, a questionnaire). Constancy of conditions is difficult to achieve because many conditions affecting the execution of a survey cannot be controlled by the researcher. Nevertheless, some conditions can be influenced by the researcher. These are the proper organization of the field work, adequate training of the field staff, and appropriate field and survey documents and procedures.

2.2 Measuring reliability

To measure the reliability of an instrument at least two independent readings must be obtained. In survey practice, one replication is considered sufficient. When the same respondent is interviewed twice, and consistent answers are obtained, then the answers are considered, for lack of other criteria, to be reliable. Consistency is taken as an indication of reliability.

The consistency of the data can be studied best from a table in which the results of the two interviews are cross-tabulated. By examination it can be determined how many units are classified differently in each category, and the patterns of shifts, if any, can be detected. To obtain an indication of the reliability of the instrument, in social surveys reliability of each question is considered. Summary measures have been developed, some of which take into consideration the level of measurement of the variable involved.

Generally, if the same respondent is re-interviewed after some time and different answers are recorded, then this can be explained by one of the following :

1) Change in situation.

The respondent gives a different answer because the objective situation has changed. The variable measured may be time-dependent : for example, if there is a question on age in completed years, and if the re-interview takes place two months after the original interview, then approximately one sixth of the respondents, should declare their age as one year older than in the original interview. Also, a variable may be subject to other legitimate changes ; for example, if there is a question on number of children at home, and if a child has been away at the time of the original interview, but has returned home before the re-interview, then the number of children at home will be different by one. These different answers are not real discrepancies, but reflect the changing situation. For the analysis of the reliability of the data they have to be taken into consideration to allow proper interpretation of the resutls.

2) Change of attitude or opinion.

The respondent gives a different answer because she has changed her attitude or opinion about a certain topic ; for example, the respondent may for some reason change her opinion about the ideal number of children or she may change her attitude towards a family planning programme. Also there are cases of different age or date estimates in societies where dates are generally not correctly known. Although the difference in response can be legitimate, the fact that a number of questions refer to topics that are unstable should be of concern to the survey designer and the the survey data user. If data are unstable - that is, subject to unpredictable or frequent changes - the results of a survey are of limited value as they reflect a chance situation at a specific time. Thus, if the data are too time-specific, they are of limited value for long-term use. The survey designer and the survey analyst will therefore, other things being equal, prefer questions that are not susceptible to (easy) change of mind on the part of the respondent.

3) Errors

A different answer is recorded because an error is made. The error may be committed by the respondent or the interviewer. It may or may not be caused by a faulty design in the questionnaire. Even if the questionnaire is properly designed it is possible that the interviewer will commit errors. This can be remedied partly by better training and it is, therefore, important for the survey designer to know whether some errors are attributable to interviewers. Even when the interviewer is properly trained, errors may be made by the respondent, either because the respondent does not know the correct answer and guesses, or because the respondent deliberately gives a wrong answer. Errors caused by the respondent are of concern to both the survey designer and the data user. Improvements in questioning techniques, better probing, and the like, can partly reduce this form of error. It is, therefore, important to know not only the nature of the error but also why it occurred.

2.3 Use of the reconciliation interview

To obtain an indication about the nature and source of the discrepancies, reconciliation of discrepant answers is needed. If discrepant results are found when a respondent is interviewed twice, and legitimate changes in the situation have been accounted for, it is assumed thateither the respondent has changed her opinion or an error has been committed. To clarify this, a reconciliation interview is carried out in which the respondent is confronted with the two discrepant results and an attempt is made to establish whether the respondent has changed her opinion or whether a mistake was made. Because reconciliation interviews are often conducted by better qualified interviewers, it is sometimes assumed that a correct answer is obtained in the reconciliation interview in the sense that it is (closer to) the true value of the characteristic measured. If this assumption is accepted then by comparing the results of the original interview with those of the reconciliation interview, an indication of the response bias can be obtained. In the present study, the results of the reconciliation interview are used mainly to establish the nature and source of the discrepancy, and not to measure response bias.

2.4 Indices of reliability

Different summary measures have been considered for indicating the consistency of the results of the two interviews. Some of these measures or indices take into consideration the level of measurement of the variable involved ; thus, there are indices for attributes and unordered categorized data, and for ordered categorized and metric variables.

For the following discussion the notation, defined below, is used :

n_{ij}, $i, j = 1, \ldots, L$ is the data matrix,

$n_{i.} = \sum_{j=1}^{L} n_{ij}$ is the ith row sum,

$n_{.j} = \sum_{i=1}^{L} n_{ij}$ is the j th column sum,

$n = \sum_{i=1}^{L} \sum_{j=1}^{L} n_{ij}$ is the total number of elements.

It is desirable to have a single index of reliability which would allow the assessment and comparison of reliability of questions. However, this is not possible for reasons which will be clarified below.

Inconsistencies between answers in the two interviews are reflected by non-zero counts n_{ij} off the diagonal of the data matrix, that is, with $i \neq j$. This in turn may lead to differences in the marginal distribution of counts for the two questions, $\{n_{i.}, i=1, \ldots, L\}$ and $\{n_{.j}, j=1, \ldots, L\}$.

However, the attainment of identical marginal distributions of counts does not imply consistency, since this can be achieved, for example, by one or more respondents switching from category i to category j and an equal number switching from category j to category i.

The indices of reliability considered here do not measure changes in the marginal distribution. That is, they assume that conditions at the interview and re-interview are constant, and there is no shift in the distribution of responses between interviewers. This assumption is sometimes called marginal homogeneity, and can be verified by comparing the observed margins of the table.

A simple and natural measure of reliability is the index of crude agreement :

$$P_c = \sum_{i=1}^{L} n_{ii}/n \qquad (1)$$

which represents the proportion of correctly classified units. Although this index has considerable descriptive value, there are two reasons why it is not suitable as a single analytic measure of reliability. The first is that no allowance is made for the fact that some units will be classified correctly by chance even if there is no association between the two responses. This proportion of agreement "expected by chance" depends on the marginal distribution of the variable under study. For example, the following two tables give the expected number classified correctly out of 100 units, for two marginal distributions, (90,10) and (50,50) :

81	9	90
9	1	10
90	10	

25	25	50
25	25	50
50	50	

The first table gives P_c = 0.82 and the second table gives P_c =0.50 although in both cases there is no association between the two measures, and so the reliability is effectively zero.

One way of overcoming this deficiency is to define an index of the form :

$$1 - \frac{\text{observed disagreement}}{\text{expected disagreement}} = 1 - \frac{1 - P_o}{1 - P_e} = \frac{P_o - P_e}{1 - P_e} \qquad (2)$$

where P_o = sum of observed proportions reflecting agreement, P_e = sum of expected proportions reflecting agreement. One such measure is Kappa defined as :

$$\hat{K} = 1 - \frac{\sum_{i \neq j} n_{ij}}{\frac{1}{n} \sum_{i \neq j} n_{i.} n_{.j}} = \frac{\sum_{i=1}^{L} \frac{n_{ii}}{n} - \sum_{i=1}^{L} \frac{n_{i.}}{n} \frac{n_{.i}}{n}}{1 - \sum_{i=1}^{L} \frac{n_{i.}}{n} \frac{n_{.i}}{n}} \qquad (3)$$

Both the tables above yield $\hat{K} = 0$, as required.

This adjustment does not answer the second weakness of the index of crude agreement as a measure of reliability, namely, that for ordinal or metric variables the index does not take into account the distance between categories, that is, the magnitude of disagreement. A related problem is that for variables, such as age at marriage the value of P_c or $\hat{K}$ is highly sensitive to the choice of grouping of th variable. For comparative purposes it is useful to have an index which is not sensitive to the choice of grouping.

This leads to a generalization of Kappa to a weighted form $\hat{K}_w$, which is given by :

$$\hat{K}_w = \frac{P_o^* - P_e^*}{1 - P_e^*} \qquad (4)$$

where :

$$P_o^* = \frac{1}{n} \sum_{i=1}^{L} \sum_{j=1}^{L} w_{ij} n_{ij}$$

and :

$$P_e^* = \frac{1}{n} \sum_{i=1}^{L} \sum_{j=1}^{L} w_{ij} n_{i.} n_{.j} \qquad (5)$$

Landis and Koch discuss a number of choices of weights. The unweighted Kappa (3) clearly corresponds to the weights :

$$w_{ij} = \begin{cases} 1, & \text{for } i = j \\ 0, & \text{otherwise} \end{cases}$$

For ordered categorized data, Cichetti has suggested using weights of the form :

$$w_{ij} = 1 - \frac{|i-j|}{L-1} \qquad (6)$$

For metric variables a suitable form of the weights is :

$$w_{ij} = 1 - (i-j)^2 \qquad (7)$$

To facilitate the analysis an interactive program was written to calculate consistency indices. In the program, six options are available. The first three options are for $\hat{K}_w$ using the basic form as given in (2) and the weighted form (4) with weights as defined in (6) and (7) for ordered categorized data and variables respectively. The fourth option was an adaptation of the unweighted form (3) and developed especially for dates and age data. In age and date reporting an error margin of two years was considered acceptable, "agreement" was defined if the difference between the two interviews was two years or less. The weights were defined as follows :

$$w_{ij} \begin{cases} = 1, & \text{if } |i-j| \leqslant 2 \\ = 0, & \text{otherwise.} \end{cases} \qquad (8)$$

For completeness the variances of $\hat{K}_w$ for the four options were also calculated. Option 5 and 6 are the calculations of the product-moment correlation and the intraclass correlation coefficient respectively.

This program is written in FORTRAN and is easily adaptable to all interactive systems. There are built-in checks on the consistency of input data. Nevertheless, when large tables are to be analysed the input process is tedious and prone to error. Thus, a non-interactive program was also written which uses data from two interviews, composes a table, claculates and prints out all six options.

2.5 Consistency index used

Weighted Kappa with weights given by (7), $w_{ij} = 1-(i-j)^2$ has the following properties :

a) For a 2x2 table is is identical to the unweighted Kappa.

b) For metric variables with identical observed margins, it is exactly equal to the product-moment correlation coefficient calculated on the integer-valued categories.

c) If it is assumed that the observer effect is random, weighted Kappa is asymptotically equal to the intraclass correlation coefficient.

The weighted Kappa is suitable for metric variables. In this report only binary or metric variables are concerned. Hence, there are no tables with more than two categories with no ordering among the categories and so weighted Kappa - using weights as given by (7) - is used for all variables considered.

In view of the relation with the intraclass correlation coefficient, for metric variables, weighted Kappa has the following important interpretation : it reflects the proportion of the total variance which is due to the inherent variance attributable to differences between subjects ; the remaining proportion being attributable to within subject response variance.

To summarize, it is proposed that the following two measures be used :

1) The crude index of agreement, P_c which is a simple descriptive measure.

2) The weighted Kappa (4) with weights given by (7), which has the analytic property just mentioned.

Both measures are insensitive to the sample size n , although clearly their sampling error decreases as n increases. For metric variables the product-moment correlation and the intraclass correlation coefficient were also calculated ; however, in all cases these were both very similar to the Kappa, as predicted by the properties given above.

4.6 Imputation methodology for missing household survey data

The reliability of survey estimates is governed by many factors, one of which is the effect of missing data due to non response. In this paper, estimates of totals in a stratum or over several strata by the Horvitz-Thompson estimation formula are dealt with. It is assumed that each unit responds with some probability which may or may not be known and that the response probabilities may vary from unit to unit. If the unit responds, the observed value of some characteristic of the unit may be subject to response error, resulting in a possible response bias and response variance for each observed value. If the unit fails to respond partially or completely, some adjustment for non-response may be carried out by means of an imputation procedure.

In this paper, four imputation procedures have been considered for study and they include (i) zero substitution method (where missing data are simply ignored), (ii) weighting method in a balancing area (stratum, group of strata, or a cluster) or in a weighting class (one of several post-strata according to classes of non-respondents, based on partial information), (iii) duplication in lieu of weighting in (ii), and (iv) historical or external source data substitution method (when available), followed by the weighting method to enlarge a deficient sample when historical data are not available for missing information.

The estimation formula pertaining to each of the four methods of imputation is developed and expressions for the bias and the variance of the estimate according to each imputation procedure are derived. In the case of the weighting method, the inverse selection probability is inflated by the inverse response rate in each balancing area.

The bias of an estimate of a total in an area will be subject to response bias, imputation bias due to the use of historical or external source data (similar in concept to the response bias of the data, relative to current values) and an imputation bias due to the use of the weighting method. There will be no bias due to the use of weighting method apart from the usual ratio estimate bias, unless the response probabilities are correlated with the values of the characteristics of the units. Thus, if the response probabilities are constant, the responding subsample is identical to a simple random selection of sampled units in another stage of selection. An overview of the biases of the estimates pertaining to the different imputation procedures is provided.

Whatever the imputation procedure among the four methods under study, the variance of the estimate at a balancing area or weighting class level, will be subject to the following components : (i) sampling variance (unless a Census is taken), (ii) simple response variance, (iii) correlated response variance, (iv) variance resulting from the variation in the events of responding or not responding among the units and (v) covariance resulting from the covariance between the events of responding or not responding, pertaining to pairs of units. In most cases there will be covariance between the estimates of different balancing areas or weighting classes with the presence of the sampling covariance (unless a Census is taken or unless the balancing area corresponds to a stratum or group of strata) and covariance terms similar to (iii) and (v) above.

The methodology of the estimation formula and variances applies to any sample design where Horvitz-Thompson estimates are used and has been specially adapted to a two-stage design with clusters selected with ppswor and households selected as srswor. The purpose of the adaptation is to apply the methodology to a hypothetical numerical example used to compare the four imputation methods.

Finally, a hypothetical example of a stratum with 50 households, grouped into 8 clusters of three to ten households in each, was set up. The household size was the characteristic under study and the sizes of the households were distributed among the 50 households with about the same distribution as in the Census, including vacants, Response probabilities of about 7 to almost one were assigned with a tendency toward smaller response probabilities in small households. Response errors in the observed counts of persons were assigned with associated probabilities of observing each count and errors from -2 to +2 as well as an erroneous observation of a household being vacant were assumed in the hypothetical example. The bias and the variance of each estimate from different sample sizes were calculated, using the formulas adapted to two stage sample and comparisons were made between the four methods.

Further details concerning this article may be obtained by writing to R. Platek or G.B. Gray, Methodology Division, Statistics Canada.

R. PLATEK, G.B. GRAY
Methodology Division, Statistics Canada

5. TERMINOLOGY

Terminology for repeated surveys

Prepared by Leslie Kish, October 1978, for the Panel on Data Collection of the Committee on Population and Demography (ABASS), National Academy of Sciences (USA). Also for the Committee on Terminology of the International Association of Survey Statisticians.

The aim of the author is practical rather than pedantic : to make the diverse terms commonly used in this field serve us better, by having clearer distinctions between them. To bring closer one -to -one correspondence between terms and uses instead of multiple uses for the same term and multiple terms for the same use. The author was guided partly by current usage (which is not uniform), partly by connotations of words, and partly by the diverse needs of the field. The Committe of Terminology of IASS and the author would be grateful for friendly suggestions.

Repeated surveys denote "similar" observations on the "same" population, but without specifying designs for overlapping coverage of the same set of units. Periodic surveys refer to surveys repeated at specified regular intervals of time. The "same" population needs definition because populations change over time both in extent and in content, e.g., cities and countries change shape and size, and units (persons, adults) are born, die, and migrate. "Similar" observation must also be defined.

Overlapping designs refer to covering the same sampling units in repeated surveys. The overlapping units need to be defined ; they may be the elements of analysis (individuals, persons), or they may be larger units, such as area segments. Units such as families, households, composed of distinct elements, present special problems of change. Designs may require either complete or partial overlapping ; the latter permits gradual changes of the sampling units.

Panels (or longitudinal) surveys refer to studies with repeated observations on the same individuals. These face problems of identification, also of fatigue, learning, and moving by individual respondents, and often higher costs; but they are needed for detecting the gross changes. (And these get confounded with errors of measurement). On the other hand, for measuring net changes it may be easier and clearer to use simpler units of sampling (such as area segments) and still retain much of the gains in the variance from correlations. These gains are also retained partially in partial overlaps. Some studies have done both : retain segments for clear net changes, and follow moving individuals for gross changes.

A third change of panel studies is for obtaining incidence of new events between two (or more) dates (periods). These are called multi-round surveys by some (e.g. demographers) and prospective* studies by others (e.g. health scientists). The collection of data on new events is sometimes aided with records (diaries, budgets) kept by respondents, or by others, or by machines, etc.

---

* in contrast to retrospective studies

L. KISH
Institute for Social Research, USA

6. NEWS FROM THE ASSOCIATION

Financial and Membership Situation of I.A.S.S., on December 31, 1978

In 1978, incomes arose to 20 971 FF (US $ 4 876) and expenses amounted to 20 590 FF (US $ 4 788). On December 31, 1978, the credit balance arose to 47 667 FF (US $ 11 085).

In 1978, there were 121 new members and 71 members were struck off the list for not paying dues (68) or for death (3). On December 31, 1978, there were 1223 members.

7. BIBLIOGRAPHY

- SHISKIN J., PLEWES T.J.

"Seasonal adjustment of the US unemployment rate" - The Statistician. Journal of the Institute of Statisticians - Vol. 27, Nos 3/4, Sept./Dec. 1978, 181-202.

- STRECKER H.

"The efficiency of sample models with equal and with varying probability of selection" - Jahrbücher für Nationalökonomie und Statistik.- Vol. 194, N°2, 113-148.

- RAPAPORT E.

" The Swedish party preference survey, May 1979" - National Central Bureau of Statistics, Sweden. Survey Research Institute for Statistics on Living Conditions - N° 1979 : 182 1, June 13, 1979.

- PACCOU Y.

" Le recensement des nomades mauritaniens" - Groupe de démographie africaine - Paris 1979.

- KROTKI K.P., HILL C.J.

" A comparison of correlated response variance estimates obtained in the 1961, 1971 and 1976 Censuses" - Survey Methodology , Statistics Canada - Vol. 4, N°1.

- BLACKER J.G.C.

" A critique of the international definitions of economic activity and employment status and their applicability in population censuses in Africa and the Middle East" - Population Bulletin of the United Nations Economic Commission for Western Asia - N° 14, June 1978.

- HEDGES B.M.

" Question wording effects : presenting one or both sides of a case" - The Statistician. Journal of the Institute of Statisticians - Vol. 28, N°2, June 1979, 83-99.

- HOLT D., SMITH T.M.F.

"Post-stratification" - Journal of the Royal Statistical Society - Series A, Vol. 142, Part 1, 1979.

- SURVEY METHODOLOGY - Statistics Canada - Vol. 4, N°2, Dec. 1978

Contents :
The evolution of a national statistical agency - J. SPEAR
Non-response and imputation - R. PLATEK, G.B. GRAY
The application of a systematic method of automatic edit and imputation to the 1976 Canadian census of population and housing - C.J. HILL
Large scale imputation of survey data - M.J. COLLEDGE, J.H. JOHNSON, R.PARE, I.G. SANDE
Some methods for updating sample survey frames and their effects on estimation - J.D. DREW, G.H. CHOUDHRY, G.B. GRAY
Alternative estimators in PPS Sampling - M.P. SINGH

- CATALOGUE C.E.A. - Bibliographie de documents choisis de la C.E.A. Nations Unies, Commission Economique pour l'Afrique - Dec. 1976

- SAMPLE SURVEYS OF CURRENT INTEREST - 12th, 13th reports, UN Statistical papers, Series C, Nos 12, 13

- STATISTICAL POLICY, WORKING PAPER 4- Glossary of nonsampling error terms : an illustration of a semantic problem in statistics - US Department of Commerce - Dec. 1978

- MANUEL DE TRAITEMENT INFORMATIQUE D'ENQUETES STATISTIQUES - par un collectif d'auteurs sous la direction de L. BREAS, J.L. BODIN - INSEE - Paris 1978

The purpose of this handbook is to make easier the exchange between the statistician and the computer expert by giving the situation on the survey processing problems and methods. Three chapters deal with the classical computer processing : file, data-cleaning chain, i.e. set of operations that contribute to set up an exhaustive and error-cleaned file, setting up of statistical tables. Other chapters deal with circulation problems, the organization of the manual processing workshop. In annex is presented the INSEE survey processing software LEDA.

- PLAN DE SONDAGE DES ENQUETES INSEE AUPRES DES MENAGES DEPUIS 1977 - INSEE.-Paris, 1979.

It is described in details the way INSEE determined the master sample it realizes household surveys since 1977 with, from 1975 Population Census data, except for the labour surveys and how the master sample is used to determine the dwellings visited by the interviewers. A chapter is devoted to the study of the representativeness of the master sample for various variables known at the 1975 Population Census and for the building of dwellings.

INTERNATIONAL ASSOCIATION OF SURVEY STATISTICIANS

The International Association of Survey Statisticians (IASS) was created in 1973. The IASS is a section of the International Statistical Institute.

The objectives of the Association shall be to promote the study and development of the theory and practice of statistical censuses and surveys and associated subjects and to foster interest in theses subjects among statisticians, organizations, institutions, governments, and the general public in different countries of the world.

OFFICERS AND MEMBERS OF THE COUNCIL

SECRETARIAT

Mr Paul DAMIANI, IASS, c/o INSEE, 18 boulevard Adolphe Pinard, 75675 PARIS CEDEX 14, France

THE "SURVEY STATISTICIAN"

The Survey Statistician has replaced the IASS Newsletter. It is published periodically by INSEE, in English and French

The Survey Statistician and Sankhya C represent the IASS official publications. While Sankhya C contains theoretical papers, the Survey Statistician includes practical information in the field of sampling surveys (surveys in progress, summaries of technical and methodological papers, answers to questions, bibliography).

<u>Editor</u> : J. Desabie, Inspecteur général de l'INSEE, INSEE, 18 boulevard A. Pinard, 75675 PARIS CEDEX 14, France

<u>Editing Board</u> : Messrs Dalenius, Desabie, Kish, Théodore

INFORMATION FOR PAPERS' AUTHORS

The papers sent to the Journal should not exceed 5 pages. They have to be typed with a double space. Two copies of these papers (in French or English) have to be sent to Mr Damiani, at the IASS Secretariat (see address above).

IASS REPRESENTATIVES

M. Adehossi (Niger)
O.O. Ajayi (Nigeria)
H. Al Shabarti (Iraq)
M. Athanassopoulos (Greece)
V. Balaban (Yugoslavia)
F.E. Betham (Western Samoa)
L. Biggeri (Italy)
L. Bodart (Belgium)
B. Boyes (USA)
S.L. Diop (Senegal)
H. Duong (Madagascar)
H. El Bashir (Sudan)
F. El Boustani (Syria)
I.P. Fellegi (Canada)
E.K. Foreman (Australia)
M. Fracalossi (Argentina)
K.T. de Graft-Johnson (Ethiopia)
J. Harewood (Trinidad-Tobago)
C. Koley (Papua-New Guinea)
A.A. Le Roux (Great-Britain)
R.N. Lodhia (Fiji Islands)
M. Maaroufi (Tunisia)
F. Madsen (Denmark)
H. Midzuno (Japan)
G. Nathan (Israel)
S. Nour Eldin (Egypt)
M. Nze Faustin (Gabon)
H.K Oh (Thailand, Korea)
E. Rapaport (Sweden)
J.L. Sanchez-Crespo (Spain)
M. Simonpietri (Chad)
D. Singh (India)
H. Strecker (Fed. Rep. of Ger.)
G. Théodore (France)
I.B. Thomsen (Norway)
S. Tulya-Muhika (Uganda)
M.A. Wahab (Pakistan)

IASS CORRESPONDENTS

B.A. Bailar (USA)
E. Cansado (Chile)
E.K. Foreman (Australia)
M.R. Frankel (USA)
M. Gonzalez (USA)
K.T. de Graft-Johnson (Ethiopia)
C. Hayashi (Japan)
T.B. Jabine (USA)
Ms Kantorowitz (Israel)
M.N. Murthy (Japan)
Ranjan Kumar Som (USA)
J.L. Sanchez-Crespo (Spain)
M.P. Singh (Canada)
F. Madsen (Denmark, Finland, Norway)