INTERNATIONAL ASSOCIATION OF SURVEY STATISTICIANS

THE SURVEY STATISTICIAN

ASSOCIATION INTERNATIONALE DES STATISTICIENS D'ENQUETES

No. 32
July 1995

Table of Contents

Letter From the President

Dear IASS Members:

With this issue, Mike Brick, Westat, Inc., Rockville, MD, USA, assumes the Editorship of *The Survey Statistician*. We are very pleased that Mike has assumed responsibility for the *Survey Statistician*. I would encourage you to send comments and suggestions on the content and format of *The Survey Statistician* to Mike Brick at Westat, Inc., 1650 Research Blvd., Rockville, MD 20850, USA or E-mail to brickm1@Westat.com.

We would also like to express our gratitude to Denise Livesely, University of Essex, UK, for her fine work as the outgoing editor of *The Survey Statistician*. Denise's new responsibilities include the Presidentship of the International Association of Official Statistics, our sister organization.

The Survey Statistician is of central importance to the IASS, and we always strive for continuous improvement of its contents. We have many members in developing countries and for some of them *The Survey Statistician* is an important means of keeping abreast of the international survey world. Many colleagues in developing countries have great difficulty getting information on conferences, current developments, and new literature. We hope that the *Survey Statistician* helps fill this need by providing short articles on timely topics, information on conferences, and other important news.

Some of you might wonder how a council composed of members residing in all corners of the world can work together. Well, such endeavors do entail difficulties not encountered by councils or groups located in one place. For practical reasons, the entire council is never able to meet with every member present, not even at ISI meetings. During recent years, communication has improved as communication technologies grow more varied and sophisticated. Still, the enduring fact is that the bulk of the responsibilities is carried out by a handful of individuals. The IASS Centre is run by the French Secretariat in Paris and Bordeaux, where Alain Charraud, Director, and Anne-Marie Vespa-Leyder, Secretary work. They manage membership, finances, print *The Survey Statistician*, and make many, many other valuable contributions.

Some Council Members are responsible for specific areas, for instance, the President, the Vice Presidents, and the Scientific Secretary. Other Council Members are called upon to oversee special tasks. During my two-year tenure, Kirk Wolter has taken charge of the information pages of the ISI Newsletter; Fred Smith has been instrumental in membership issues; Chris Skinner has organized the short courses to be held in China and overseen the organization of other conferences; and Gösta Forsman conducted the membership survey with the assistance of two of his students.

Other individuals who merit special mention for their contributions are Gordon Brackstone and Leslie Kish. Gordon Brackstone organizes our corps of Country Representatives. These representatives comprise a worldwide net of informants who report on survey developments in their own countries. These country reports appear regularly in *The Survey Statistician*. Country Representatives usually gather at the ISI meeting to discuss matters of common interest. Leslie Kish has compiled a booklet drawing from the questions and answers column he has written for *The Survey Statistician* over the years. This booklet will be distributed to our members.

At the ISI Florence meeting, it was noted that membership was declining. One response to this development was to conduct a membership survey. Our main objective was to tap the ideas and wishes of our membership, thus enabling our organization to more effectively set new goals and steer in new directions. The survey results were presented in the December 1994 issue. What emerged from the survey was that the members thought we had already struck the right course. Members want us to continue publishing *The*

Survey Statistician, organizing conferences, and relaying developments in the survey world. The provision of information was especially important to members in developing countries. Other salient suggestions that emerged from the membership survey are the need for additional committees modeled after the organization of the American Statistical Association and cooperative agreements with publishing companies to offer reduced rates for members in developing countries.

A pleasant turn of events is that membership is on the rise again. Membership coverage is uneven, though. Membership is weak in Eastern Europe, Latin America, and in most of Asia. Members in these countries have special needs and requirements that must be addressed to enable their participation. One such effort is the IASS conference in Budapest in 1996 under the leadership of Ádám Marton. Regular in-between year conferences (i.e., conferences held in the years between the general ISI meetings) is an option which we are currently considering. Further activities have been our participation as one of the sponsors of the Conference on Survey Measurement and Process Quality held in Bristol, April 1995, and of the 1996 CASIC conference to be held in Texas, USA.

All in all, the IASS is alive and well and forward looking. Our new logo is modern and aesthetically appealing; a new brochure is under way, and it should also attract new members. IASS material was presented at the American Statistical Association Meeting in Toronto in 1994 and at the Bristol Conference, 1995. Our financial future looks secure, much due to the generosity of INSEE, who provides a home for our organization.

I wish to express my gratitude for being allowed this opportunity to serve the IASS as President. It has been a valuable experience for me and I can only hope that I have been able to give back to the organization something of what has been given me. As my term draws to a close, I welcome Dennis Trewin as the new President assuming office at the ISI meeting in Beijing this August. It is my desire to assist Dennis with as much support, enthusiasm, and good ideas as Graham Kalton assisted me.

Lars Lyberg
President

A Conceptual Framework for Measurement Error in Surveys (Note on the History of Measurement Error in Surveys)[1]

Colm O'Muircheartaigh

In considering the different contexts in which early users of surveys operated and the different perspectives they brought to their operations, it is difficult to find common criteria against which to measure the success of their endeavors. The history of surveys (in their modern sense) goes back only 100 years, but from the outset there was a great diversity in the settings, topics, philosophies, and executing agencies involved. In the initial stages there was no particular separation between the issues of survey design and execution and the issues of error in surveys.

The concept of quality, and indeed the concept of error, can only be defined satisfactorily in the same context as that in which the work is being done. To the extent that the context varies and the objectives vary, the meaning of error will also vary. I propose that as a definition of error we adopt the following: work purporting to do what it does not do. This redefines the problem in terms of the aims and frame of reference of the researcher rather than an arbitrary (pseudo-objective) criterion. It immediately removes the need to consider true value concepts in any absolute sense and forces a consideration of the needs for which the data are being collected. Broadly speaking, every survey operation has an

[1] This paper is based on the keynote address given at the International Conference on Survey Measurement and Process Quality, held in Bristol UK from 1-4 April 1995, organized by the American Statistical Association in collaboration with a number of associations. A complete version will appear in the monograph of the conference to be published in 1996. That version will contain a set of detailed references.

objective, an outcome, and a description of that outcome. The gaps in which errors or quality failures will be found are the gaps among these three elements.

There are three distinct strands in the development of survey research: governmental or official statistics; academic and social research; and commercial, advertising, or market research. Each of these has its own intellectual baggage, its own disciplinary perspective, and its own criteria for evaluating success and failure.

The International Statistical Institute (ISI) was the locus of debate for official statisticians at the end of the 19th century when A. N. Kiaer, director of the Norwegian Bureau of Statistics, presented a report of his experience with "representative investigations" and advocated further investigation of the field. In this context the evaluation of surveys was largely statistical and the survey was seen as a substitute for complete enumeration of the population. Here the concept of error became synonymous with the variance of the estimator (essentially the variance of the sampling distribution following Neyman's influential paper in 1934). This equivalence of quality and variance and its measurement by repeated sampling, with some acknowledgment of bias, was confirmed by the work of Mahalonobis in India, reported in the mid-1940s, and in particular by his design of interpenetrated samples for the estimation of errors introduced by fieldworkers and others. The influence of statisticians on the conceptualization of error and its measurement has continued in this tradition, and can be found in all the classic texts of survey sampling. It is worth noting, however, that Kiaer's proposal to the ISI in 1895 would almost certainly have been rejected had it not been for the support of the monographers (Cheysson and others) whose work consisted of the detailed and intensive examination of one or a small number of cases (what might today be called the case study approach).

The involvement of the monographers in the debate at the ISI is interesting particularly because it provides a link to the second major strand in the development of surveys. This was the Social Policy/Social Research movement, whose beginnings are perhaps best represented by Booth's study of poverty in London from 1889 to 1903, and the Hull House papers in the USA in 1892. Though not in any way a formal or organized movement, there were certain commonalties of approach and objectives across a wide range of activities. The goal of this movement was social reform, and the mechanism was community description. Here the success or failure of the activity was in the impact of the findings on decision makers and politicians.

The principal influences on this group were the social reform movement and the emerging sociology discipline. Some of the pioneers of sample surveys spanned both official statistics and the social survey; in particular, Bowley (who made a substantial contribution to the development of survey sampling) produced a seminal work on social measurement in 1915 which helped define the parameters of data quality and error for factual or behavioral information. Bogardus, Thurstone, and Likert provided "scientific" approaches to the measurement of attitudes. In this field the disciplinary orientation was that of sociology and social psychology, with some influence from social statistics.

The third strand arose from the expansion of means of communication and the growth in the marketplace. From modest beginnings in the 1890s (Gale and others), there was a steady increase in the extent of advertising and a development and formalization of its companion, market research. The emphasis was on commercial information, originally in the hands of producers of goods and services and collected and evaluated informally; market research, however, developed gradually into a specialized activity in its own right.

Here the impact of psychologists was particularly strong. The work of Link and others in the Psychological Corporation was influential in providing an apparently scientific basis for measurement in the market research area. Though the disciplinary influence was psychology also, experimental psychology took precedence over social psychology here. The terminology and the approach were redolent of science and technology. The term "error" was not used explicitly; rather there was a description of reliability and validity of instruments. This contrasts particularly with the "error" orientation of the statisticians.

Thus the field of survey research as it became established in the 1940s and 1950s involved three different sectors — government, the academic community, and business; it had three different disciplinary bases — statistics, sociology, and experimental psychology; and it had developed different frameworks and terminologies in each of these areas.

In general in describing errors or data quality in surveys, models concentrate on the survey operation itself, in particular on the data collection operation. The models may be either mathematical and schematic (or conceptual). One of the more useful of the latter is that presented by Sudman and Bradburn in their book on response effects in surveys in 1974. This schematic model describes the relationship among the interviewer, the respondents and the task in determining the outcome of the survey interview, and identifies the potential contribution of a number of the key elements in each of these to the overall quality of the survey response. To compare the different approaches to survey research described above, however, it is necessary to provide an overarching framework that encompasses the concerns of all three major sectors.

One possible framework draws on some ideas presented by Kish in his book on statistical design for research. Kish suggests that there are three issues in relation to which a researcher needs to locate a research design. I propose that a similar typology could be used to classify the dimensions that would encompass most sources of error. These are representation, randomization, and realism. It is worth noting that each of these "dimensions" is itself multi-dimensional. Representation involves issues such as the use of probability sampling, stratification, the avoidance of selection bias, and a consideration of nonresponse. Randomization (and its converse in this context, control) covers issues of experimentation and control of confounding variables. Realism arises in two ways: realism in variables concerns the extent to which the measured or manifest variables relate to the constructs they are meant to describe; and realism in environment concerns the degree to which the setting of the data collection or experiment is similar to the real-life context with which the researcher is concerned. These dimensions are related to the ideas of internal validity and external validity used by Campbell and Stanley in 1963 in describing the evaluation of social research.

Four examples may help to illustrate the relevance of these dimensions. In the work on cognitive aspects of survey methods, the emphasis is on understanding the process of response generation and the production of 'stable' predictors of response effects and 'rules' for question wording. The theoretical interest is in the prediction of and explanation of effects rather than in the solution of specific empirical problems. Therefore the key consideration of the researcher is in providing a rigorously controlled environment, and not in either representation or realism. In statistical modeling the impetus is towards incorporation of multiple levels/effects into estimates of relationships; essentially, the modelers attempt to compensate for failure to randomize by achieving control in the analysis. By technical means, they also attempt to address substantively more meaningful questions. Error profiles are a multi-dimensional description of the characteristics of a survey and its outputs. They can present both quantitative and qualitative data, and can over time build up a broad picture of the performance of a survey across a number of dimensions. Computer-assisted methods provide an environment in which there can be great flexibility in the presentation of materials to respondents; and permit the incorporation of fully or partly randomized experiments into the standard fieldwork and processing of surveys. Thus they provide the beginnings of an environment in which both randomization and realism can be achieved simultaneously.

I would argue in conclusion that it is inappropriate to consider a survey without its broader context. Moving towards perfection on a particular dimension is apt to be associated with a movement away on another dimension. Two examples may illuminate the issues.

We may reduce the contribution of clustering to the sampling variance of a survey by designing a more dispersed sample, but pay the price of higher nonresponse due to the higher cost of callbacks for such a sample. Similarly, we may reduce nonresponse by bringing pressure to bear on reluctant respondents, but at the same time increase response error for respondents. A consumer panel recruited by a market research firm may have very high completion rates for its tasks, and the tasks themselves may be competently performed, but the need to select respondents who are willing to perform these tasks may render the sample extremely unrepresentative.

The redesign of the US Current Population Survey (CPS) involved both experiments and appropriately selected representative samples. It was thought undesirable to allow the tests to interfere with the running of the CPS itself and the tests were therefore carried out in parallel to the main survey. Thus, although representation and control were taken into account, there were elements of realism missing. The various tests of the redesigned survey were devised and carried out with great care, and on the basis of these tests it was predicted that the redesigned survey would produce a discontinuity in a particular direction in the estimates provided by the CPS. Despite the high standards of the whole operation, the conclusions of the tests did not carry over to the real-life implementation of the new CPS instrument. I would argue that in this case it may have been the unmeasured differences between the parallel and real CPS that led to this nonportability.

These examples illustrate why it is necessary to step outside our own disciplinary boundaries if we are to appreciate the many possible aspects of quality/error in our data. Survey researchers should become more externalist and less internalist in their criteria for evaluating their work. The early practitioners generally produced work that was grounded in the sector in which they were working and their criteria were related to their objectives. We should now be willing to respect the criteria of others and to recognize that, as there will never be perfect empirical social data, we should be aware of the many environments in which data could be needed or used. There is not a simple hierarchy of errors, only a multi-dimensional array.

For further information, contact:
Colm O'Muircheartaigh, Methodology Institute,
London School of Economics and Political Science,
Houghton Street London WC2A 2AE
E-mail: omuirche@lse.ac.uk

Computer Assisted Data Collection Data Quality and Costs; An Annotated Bibliography

Edith D. de Leeuw and Joop J. Hox[1, 2]

1. Introduction

The question whether Computer Assisted Data Collection methods (CADAC) should be used in social research has by now become purely theoretical: in practice most if not all professional research organizations have adopted them with enthusiasm. This raises the question of what influence computer assisted data collection methods have on the *quality* of the data. To answer this question we compiled an annotated bibliography on the influence of computer assisted data collection on the data quality. Each entry in this bibliography is followed by one or more simple codes indicating the topics investigated in the original articles and papers. The first part of this bibliography contains a short description on how relevant literature was located, the second part contains the bibliography itself.

2. Search for Relevant Literature

We conducted a computer search to retrieve relevant literature. The abstracting services used were Psychological Abstracts, Sociological Abstracts, and SRM. SRM is an international documentation service in the field of social research methodology stationed in the Netherlands. SRM contains abstracts of both American and European studies. In addition an appeal for relevant references was sent to both national and international newsletters, and direct appeals for help were sent to experts in the field. The reference lists of the studies found were then searched to uncover additional studies.

For a successful literature search, the key words used are very important. The emphasis of this bibliography is on data quality and cost, so we started with these key words. We added other important factors such as interviewer acceptance and respondent acceptance. Further operationalization of the search terms led to the following list: logistics, ergonomics, costs, interviewer training, interviewer effect, acceptance by respondent and interviewer, response rate, data quality, item nonresponse, social desirability.

The second task was to explicitly name different forms of computer assisted data collection methods for survey research (e.g., CADAC, CASIC, CAI, CAPI, DBM). This resulted in a taxonomy of different forms of computer assisted data collection. All terms for computer assisted data collection in this taxonomy were added to the key words in the search definition (the taxonomy is available from the authors).

As stated above, we were primarily interested in studies that investigated the influence of computer assisted data collection on data quality and costs. Studies focusing on other aspects of the survey process (e.g., sampling, analyzing, reporting) were omitted. For a thorough bibliography on sampling for household telephone surveys, see Blair et al (1995)[3]. We attempted to be as thorough as possible in searching for relevant studies. However, we do not claim completeness, and we would be grateful if new or absent studies are brought to our attention.

3. Annotated Bibliography

3.1 Description of the Entries

The entries are grouped in four sections: 1) CADAC in general; 2) CATI; 3) CAPI; and 4) CASI. Within each section entries are listed in an alphabetical order according to the first author's surname. When a publication is not in English, a translation of the title is provided and the language is indicated after the original title.

To assist the users of this bibliography in their search for specific topics, codes are used to classify each entry (see Table 3.1). Codes are between square brackets at the end of each entry.

Table 3.1. Codes used to classify entries

Topic:	Code:
Logistics general	L
Software/programming	P
Ergonomics	E
Training of interviewers	IT
Costs	C
Acceptance by respondents	RA
Acceptance by interviewers	IA
Response (unit nonresponse)	RR
Data quality general	DQ
Item-nonresponse	NRI
Social desirability	SD
Sensitive questions	SQ
Open questions	OP
Interviewer effects (variance)	IE
Review	R

4.2 Computer Assisted Data Collection (CADAC) in General

Bethlehem, J.G., & Keller, W.J. (1991). The Blaise system for integrated survey processing, *Survey Methodology*, 17, 43-56.
[L, P]

CBS (1987). Automation in survey processing; CBS select 4. Den Haag: Staatsuitgeverij (CBS: Statistics Netherlands). [P, RA, IA]

Hox, J.J., De Bie, S., & De Leeuw, E.D. (1990). Computer assisted (telephone) interviewing: a review. In: J. Gladitz & K.G. Troitzsch (Eds.), *Computer aided sociological research*. Berlin: Akademie-Verlag. [P, R]

Johnson, R. (1990). Computer interviewing: What's wrong with it?, *Sawtooth News*, 6, 1-2, 6-7. [L]

Nicholls, W.L. II, Baker, R.P., & Martin, J. (1996). The effects of new data collection technologies on survey data quality. In: L. Lyberg et al (Eds.), *Survey measurement and process quality*. New York: Wiley (to appear in spring 1996). [IA, RR, DQ, NRI, R]

Saris, W.E. (1989). A technological revolution in data collection. *Quality & Quantity*, 23: 333-349. [R]

Weeks, M.F. (1992). Computer-Assisted Survey Information Collection: A review of CASIC methods and their implication for survey operations. *Journal of Official Statistics*, 4: 445-466. [L, C, RA, IA, DQ, R]

4.3 Computer Assisted Telephone Interviewing (CATI)

Berry, S.H., & O'Rourke, D. (1988). Administrative designs for centralized telephone survey centers: Implications of the transition to CATI. In: R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls II, & J. Waksberg (Eds.). Telephone survey methodology (pp. 457-474). New York: Wiley. [L, IT]

Bergman, L.R., Kristiansson, K-E., Olofson, A., & Säfström, M. (1994). Decentralized CATI versus Paper and Pencil Interviewing: Effects on the results in the Swedish labor force surveys. Journal of Official Statistics, 2:181-195. [RR, DQ]

Catlin, G., & Ingram, S. (1988). The effects of CATI on costs and data quality: A comparison of CATI and paper methods in centralized interviewing. In: R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls II, & J. Waksberg (Eds.). Telephone survey methodology (pp. 437-456). New York: Wiley. [C, RR, NRI, OP]

Freeman, H.E., & Shanks, J.M. (1983). The emergence of computer assisted survey research; Special issue SMR, Sociological Methods & Research, 12, 2. [L]

Groves, R.M. & Mathiowetz, N.A. (1984). Computer assisted telephone interviewing: Effects on interviewers and respondents. Public Opinion Quarterly, 48: 356-369. [IA, RR, IE]

Groves, R.M. & Nicholls, W.L. II (1986). The status of computer-assisted telephone interviewing: Part II-Data quality issues. Journal of Official Statistics, 2: 117-134. [IA, RR, DQ, NRI, IE, R]

House, C.C. & Nicholls, W.L. II (1988). Questionnaire design for CATI: Design objectives and methods. In: R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls II, & J. Waksberg (Eds.). Telephone survey methodology (pp. 421-436). New York: Wiley. [P]

Nicholls, W.L. II., & Groves, R.M. (1986). The status of Computer Assisted Telephone Interviewing: Part 1- introduction and impact on cost and timeliness of survey data. Journal of Official Statistics, 2: 93-115. [L, C, R]

Thompson, J.M. (1989). CATI's impact on the bottom line, CATI News, 2, 2, 4-6. [C]

4.4 Computer Assisted Personal Interviewing (CAPI)

Baker, R.P. (1992). New technology in survey research: Computer assisted personal interviewing (CAPI). Social Science Computer Review, 10: 145-157. [C, RA, IA, DQ, OP, R]

Baker, R.P. (1990). What we know about CAPI: Its advantages and disadvantages. Paper presented at the annual meeting of the American Association of Public Opinion Research, Lancaster, Pennsylvania. [C, RA, IA, R]

Baker, R.P., & Bradburn, N.M. (1992). CAPI: Impacts on data quality and survey costs. Information Technology in Survey Research. Chicago: NORC. Information Technology in Survey Research Discussion Paper 10 (also presented at the 1991 Public Health Conference on Records and Statistics). [C, NRI, SQ]

Beckenbach, A. (1992). Befragung mit dem Computer, Methode der Zukunft?. Anwendungsmöglichkeiten, Perspektiven und experimentelle Untersuchungen zum Einsatz des Computers bei Selbstbefragung und persönlich-mündlichen Interviews. [In German: Computer Assisted Interviewing, A method of the future? An experimental study of the use of a computer by self-administered questionnaires and face-to-face interviews]. Diplomarbeit [Ph.D. thesis] Universität Mannheim. [E, RA, IA, SQ]

Bennet, D., & Goodger, C. (1993). Interviewer training for CAI at OPCS. Paper presented at the 1993 conference of the Study Group on Computers in Survey Analysis. City University, London. [IT]

Bradburn, N.M., Frankel, M.R., Baker, R.P., & Pergamit, M.R. (1993). A comparison of CAPI with PAPI in the National Longitudinal Survey of Labor Market Behavior/Youth Cohort. Chicago: NORC. Information Technology in Survey Research Discussion Paper 7. [L, RA, C]

Bradburn, N.M., Frankel, M.R., Baker, R.P., & Pergamit, M.R. (1992). A comparison of CAPI with PAPI in the NLS/Y. Chicago: NORC. Information Technology in Survey Research Discussion Paper 9 (also presented at the 1991 AAPOR-conference, Phoenix, Arizona). [RA, RR, NRI, SD, SQ]

Couper, M.P., & Burt, G. (1994). Interviewer attitudes towards computer-assisted personal interviewing (CAPI). Social Science Computer Review, 12, 38-54. [IA]

Couper, M.P., & R.M. Groves (1992). Interviewer reactions to alternative hardware for computer assisted personal interviewing. Journal of Official Statistics, 8: 201-210. [E, IA]

Couper, M,P., Hansen, S.E., Sadosky, S.A. (1996). Evaluating interviewer performance in a CAPI survey. In: L. Lyberg et al (Eds.), Survey measurement and process quality. New York: Wiley (to appear in spring 1996). [IT, DQ, IE]

Edwards, B., Sperry, S., Schaeffer, N.C. (1995). CAPI design techniques for improving data quality. Paper presented at the International Conference on Survey Measurement and Process Quality, Bristol, England. [L, P, E, IA]

Edwards, B., Bittner, D. Edwards, W.S., & Sperry, S. (1993). CAPI effects on interviewers: A report from two major surveys. Paper presented at the U.S. Bureau of the Census Annual Research Conference, Washington D.C. [E, IA]

Hippler, H-J. & Beckenbach, A. (1992). Das persönlich-mündliche Interview am Scheideweg?; Anwendungsmöglichkeiten und perspectiven zum einsatz von Computern in Umfragen [In German: The face-to-face interview at the turning point; possibilities and of the use of a computer in data collection], Planung und Analysen, 5, 44-52. [RA, IA, RR, DQ, R]

Martin, J., O'Muircheartaigh, C., Curtice, J. (1993). The use of CAPI for attitude surveys: An experimental comparison with traditional methods. Journal of Official Statistics, 9, 641-661. [RA, IA, RR, DQ, NRI, SD, SQ]

Olsen, R.J. (1992). The effects of computer assisted interviewing on data quality. Paper presented at the conference, Trento. (see also Bradburn et al, 1992). [NRI, SD, SQ]

Riede, T., & Dorn, V. (1991). Zur Einsetzbarkeit von Laptops in haushaltsbefragungen in der Bundesrepublik Deutschland [In German: Acceptance of laptops for household surveys in Germany]. Wiesbaden: Statistisches Bundesamt. Heft 20 der Schriftenreihe Ausgewählte Arbeitsunterlagen zur Bundesstatistik. [RA, IA, RR]

Sebestik, J., Zelon, H., DeWitt, D., O'Reilly, J.M., & McCowan, K. (1988). Initial experiences with CAPI. Paper presented at the U.S. Bureau of the Census Annual Research Conference, Washington, D.C. [C, NRI]

Sperry, S., Bittner, D., Branden, L. (1991). Computer assisted personal interviewing on the current beneficiary survey. Paper presented at the AAPOR 1991 conference, Phoenix, Arizona. [RA, IA]

Thornberry, O., Rowe, B., & Biggar, R. (1991). Use of CAPI with the U.S. National Health Interview Survey. Bulletin de Méthodologie Sociologique, 30: 27-43. [L, IA, RR]

Van Bastelaer, A.M.L., Kerssemakers, F.A.M., & Sikkel, D. (1987). A test of the Netherlands continuous labor force survey with hand-held computers: Interviewer behavior and data quality. In: CBS-Select 4; Automation in survey processing. Den Haag: Staatsuitgeverij (CBS: Statistics Netherlands). [RA, IA]

Woijcik, M, Yung, A., & Hunt, E. (1993). Using a computerized field management system to control a field-based experiment intended to measure the effects of collecting longitudinal data using a computer-assisted personal interviewing technique. Chicago: NORC. Information Technology in Survey Research Discussion Paper 6). [L, C]

Woijcik, M.S., Bard, S., & Hunt, E. (1992). Training field interviewers to use computers: a successful CAPI training program. Chicago: NORC. Information Technology in Survey Research Discussion Paper 8 (Also presented at the 1991 AAPOR-conference, Phoenix, Arizona). [IT]

4.5 Computer Assisted Self Interviewing (CASI)

CASI General

Beckenbach, A. (1992). Befragung mit dem Computer, Methode der Zukunft? Anwendungsmöglichkeiten, Perspektiven und experimentelle Untersuchungen zum Einsatz des Computers bei Selbstbefragung und persönlich-mündlichen Interviews. [In German: Computer Assisted Interviewing, A method of the future? An experimental study of the use of a computer by self-administered questionnaires and face-to-face interviews]. Diplomarbeit [Master's thesis] Universität Mannheim. [E, RA, SQ]

Evan, W.M., & Miller, J.R. (1969). Differential effect of response bias of computer vs. traditional administration of a social science questionnaire: An exploratory methodological experiment. Behavioral Science, 14: 216-227. [SQ, SD]

Hippler, H-J. & Beckenbach, A. (1992). Das persönlich-mündliche Interview am Scheideweg?; Anwendungsmöglichkeiten und perspectiven zum einsatz von Computern in Umfragen [In German: The face to face interview at the turning point; possibilities and of the use of a computer in data collection], Planung und Analysen, 5, 44-52. [RA, IA, RR, DQ, R]

Martin, C.L., & Nagao, D.H. (1989). Some effects of computerized interviewing on job applicant responses. Journal of Applied Psychology, 74: 72-80. [SD]

O'Reilly, J.M., Hubbard, M.L., Lessler, J.T., Biemer, P.P., & Turner, C.F. (1994). Audio and video computer assisted self-interviewing: preliminary tests of new technologies for data collection. Journal of Official Statistics, 2: 197-214. [L, RA, SQ, DQ]

Waterton, J.J., & Duffy, J.C. (1984). A comparison of computer interviewing techniques and traditional methods in the collection of self-report alcohol consumption data in a field survey. International Statistical Review, 2: 173-182. [SQ, SD]

Disk by Mail/Electronic Mail Survey

Kiesler, S., & Sproull, L.S. (1986). Response effects in electronic surveys. Public Opinion Quarterly, 50: 402-413. [DQ, NRI, SD, OP]

Saltzman, A. (1992). Improving response rates in disk-by-mail surveys. Sawtooth software conference proceedings. Evanston: Sawtooth software. [RR]

Witt, K.J. & Bernstein, S. (1992). Best practices in disk-by-mail surveys. Sawtooth software conference proceedings. Evanston: Sawtooth software. [L, C, RR]

Zandan, P., & Frost, L. (1989). Customer satisfaction research using disk-by-mail. Sawtooth software conference proceedings. Evanston: Sawtooth software. [L, RA, RR]

Telepanel/Teleinterview

Kalfs, N. (1993). Hour by hour; Effects of the data collection mode in time use research. Amsterdam: Nimmo. [RR, DQ]

Saris, W.E. (1991). Computer-assisted interviewing. Sage quantitative applications in the social sciences no 80. Newbury Park: Sage. [L, P]

5. Notes

1 We would like to thank Nancy Mathiowetz, Karl van Meter, Frans Schelbergen, Ger Snijkers, and Mike Brick for their support.

2 This bibliography was compiled in the context of the research program 'Living arrangements and social networks of older adults.' This research program is conducted at the departments of Sociology & Social Gerontology and Social Research Methodology of the Vrije Universiteit in Amsterdam, and the Netherlands Interdisciplinary Demographic Institute in the Hague. The research is supported by a program grant from the Netherlands Program for Research on aging (NESTOR), funded by the Ministry of Education and Sciences and the Ministry of Welfare, Health and Cultural Affairs.

3 Sampling design for household telephone surveys; A bibliography 1949-1995. Compiled by Johny Blair, Shanyang Zhao, Barbara Bickart, & Ralph Kuhn. Maintained by Survey Research Center, University of Maryland at College Park. (E-mail: sampling@cati.umd.edu)

Correspondence to:
Edith D. de Leeuw, Interuniversity Graduate School for Psychometrics and Sociometrics (IOPS), Plantage Doklaan 40, 1018 CN Amsterdam, Holland.
E-mail: v70uedit@hasara11.bitnet

Some Thoughts About Sampling Dwelling Units Within Villages

David J. Fitch

The "correct" way of sampling dwelling units (DUs) within villages is to first map each selected village in sufficient detail so that each DU is identified. A sample of DUs is then selected from such a map and given to an interviewer who returns to the village to collect data. The purpose may be to collect DU data, such as level of iodination of salt used in the DU, or the purpose may be to make estimates for people living in DUs, but to keep things simple here we will think only about making DU estimates. The "correct" method is probably appropriate for a national statistical office with a program of sample surveys; however, there is often a need for a procedure which does not require preinterview visits to the selected villages and production of expensive maps.

In response to this need, the World Health Organization (WHO) Expanded Programme on Immunization developed their EPI method for selecting DUs within villages. The interviewer goes to the center of a selected village, spins a coke bottle and selects randomly a DU in the direction to which it points. The second sampled DU is the one closest to the first, etc. The method does not yield a probability sample as DUs on more populated routes, on the periphery, those located off-road, and in villages that have grown are undersampled, and those DUs in the center, in clusters, and in villages that have shrunk since their measure of size was made, are oversampled (Fitch, Flores & Matute 1994). Also, selecting from only one part of the village typically gives inefficient estimates as compared with sampling from all parts of the village because of homogeneity within a village.

Both the "correct" and the EPI method are inefficient in some sense. The "correct" method is relatively expensive in many developing world situations, and the EPI method with its various problems would likely require a large sample of villages in order to be able to obtain a reasonably precise estimate.

There is a need for a method that can be easily undertaken in villages of the developing world and is also efficient. We present two methods for consideration and invite readers to present additional methods, to comment or challenge the arguments given here, and especially to present experiences obtained in carrying out alternative sampling of DUs in villages. Perhaps we from the developing world can, with help from our colleagues of the affluent world, work out proposals for undertaking research to develop sampling procedures that will help us to better estimate those variables important in our social and economic development.

Before considering alternative methods, let us briefly review how we would estimate the accuracy of the estimates of the means that we make from our survey in a sample of villages. This is necessary as background for considering problems and advantages of alternative methods. We assume villages have been selected with probability proportional to estimated size (*ppes*). To the extent that the size figures are accurate for each village and as an unbiased estimate of the mean for the village is available, then the sample mean for each village when properly weighted by the inverse of probability of selecting is an unbiased estimate of the population mean. Now the estimator for the variance of the population mean is $\upsilon(\bar{y}) = \sum(y_i - \bar{y})^2 / n(n-1)$, under simple random sampling and ignoring the finite population correction. In the case just described, we can use this variance estimator with y_i being the weighted mean of the ith village, $\bar{y}$ the mean of the y_i, and n the number of villages sampled. Unless villages were sampled with replacement, there is a second term but in practice this term can be ignored. It is common that the available village size estimates are not very exact and for this reason and others we do not generally use this estimator. Despite this, it is almost certain that using available size estimates and selecting *ppes* will give more efficient estimates than a *srs* of villages. However, with size measures that depart considerably from the actual values, we should no longer use the simple variance estimator even though it is used by in the EPI method.

In practice we compute a weight, w_{ij}, for DU j in village i and use as our estimate of the mean, the ratio $\sum\sum w_{ij} y_{ij} / \sum\sum w_{ij}$. We estimate the variance of this mean using a program such as PC CARP (Fuller et al, 1986). Understanding how one arrives at these weights can help us better understand potential problems with methods of selecting within villages. If M_i is the available size figure for the ith village in terms of DUs, M the available population total, and if we select with replacement, 30 villages, then the ith village is selected with the probability $30M_i / M$. Let us assume that we have selected, by some method and with equal probability, m_i DUs within the ith village. In order to know the probability by which these m_i were selected, we need to have a count M_i^* of the total number of DUs in the village. M_i, the available size estimate used to select the village *ppes*, will often not be a very good estimate of M_i^*, although if the M_i's are recent and reasonably good, one may decide to use them. Putting this all together, the overall probability of selecting a DU in the ith village is $(30M_i / M)(m_i / M^*)$ and the weight for each DU selected is the inverse of this. These would be the weights used with a program such as PC CARP.

Turning back now to the selection and estimation alternatives, our first candidate method could be called a sector method, thinking of the Spanish word sector. In this method, the sampler-interviewer divides the village into sectors of approximately equal size, randomly selects one sector, and interviews in all DUs of that sector. If the target "take" is *m*, then the village is divided into *M*/m* sectors, with the selected sector containing m_i DUs. In forming the sectors, it will be necessary to draw a simple map that clearly gives the boundaries of each sector.

It is likely that m_i will vary from *m* for one or more of the following reasons: 1) It is not possible to divide the village into sectors all of size *m*; 2) It takes too much time to first count to obtain the exact number of DUs in the village and then to count out sectors of the desired size; and 3) It may be desirable to use the sectoring in a later survey when things have changed in the village, so that sector boundaries must be chosen so that they can be identified at a later time.

Examination of the expression given above for the DU selection probability shows that if we set m_i to m and M_i^* to M_i, and use the resulting probabilities for selection within each village, i.e., m / M_i, then all DUs in the population have the same probability of selection, i.e. $(30M_i / M)(m / M_i) = 30m / M$, and hence the same weight. Please note that the "take" will vary from *m* as the M_i^* vary from M_i, and while the sample mean will be an unbiased estimate of the population mean, it will not be theoretically correct to use the simple variance estimator given. One would use something like PC CARP.

In general, a sample design where all elements have the same weight is the most efficient design so there are inefficiencies where these equalities do not hold. It seems likely that these inefficiencies would be small. Some idea of the problem could be obtained from simulation studies. More important in deciding the desirability of using a sector method is the amount of time needed to construct the sectors as compared with the time required for undertaking other possible methods. Since it is easier to construct larger than smaller sectors, a potential problem would come from the temptation to form sectors that were too large for efficiency. It is likely that maximum efficiency would come from a aimed-for "take" in the range of 10-20, which in a very large village, say of 1,000 DUs, would mean 50-100 sectors. Forming the sectors could take a lot of time. In fact one would likely not divide the whole village into sectors unless the plan was to use such sectorization in a series of surveys. In a one-time sampling operation within a village of say *M** = 450 DUs and an aimed-for "take" of 20, one would likely divide the village in half, select one at random, divide that in half, select one at random, and again select one half of this at random. We would now have about 1/8th of the village which would then be sectored into 3 parts and one would be selected at random which, if everything went according to plan, would contain about 18 or 19 DUs.

We will call a second candidate method a hand held computer method. In this method a computer is programmed to compute intervals which give the target selection probability, m/M_i. The sampler-interviewer walks through the village with someone from the village who will know most of the people, and presses the enter key for each DU. The computer selects one DU at random from each interval. Note that if the aimed-for take was say 20, i.e. $m = 20$, the actual number selected will be less if the village has shrunk in size and more if it has grown.

We consider this method to be good since it gives all DUs in the population the same probability of selection and the same initial weight and avoids the possibilities of bias inherent in a method that selects a fixed number of DUs from each village. It also eliminates the need for a count of the total number of DUs in the villages. The computer can be programmed to retain: 1) the number of DUs in the village; 2) the number of DUs selected; and 3) reasons for nonresponse, and to compute weights which correct for nonresponse of selected DUs. It would also be possible to program in some quality control measures. For example, interviewers may be tempted to substitute more attractive DUs, e.g., no dog, on their side of the creek, for the DU selected by the computer. If they are instructed to enter the initials of the head of household, from the helper person before the selection decision is made, then these initials can be retained for those DUs selected and can be compared with the names on the interview schedule. Undoubtedly other uses of the computer will be found in practice.

As we work with these problems, we have the possibility of making very large contributions. A great deal of money has, and still is being wasted in inefficient sampling. Hopefully we will be able to find ways of improving the survey work on which our development to an important degree rests.

References

Fitch, D.J., Flores, R., & Matute, J. (1994). Proceedings of the American Statistical Association, Section of Survey Research Methods.

Fuller, W.A., Kennedy, W., Schnell, D., Sullivan, G., & Park, H.J., (1986). PC CARP. Statistical Laboratory, Iowa State University, Ames, Iowa.

For further information, contact:
David Fitch Apartado Postal 1188, Guatemala
E-mail: dfitch@incap.org.gt

Adaptive Sampling
Steven K. Thompson

Adaptive sampling refers to designs in which the procedure for selecting units or sites on which to make observations may depend on observations made during the survey. Thus, the sampling plan has the flexibility to change during the course of the survey in response to observed patterns in the population. For example, in a survey of a rare, clustered animal population, whenever unusually high abundance is observed at a sample site, neighboring sites may be added to the sample. In a survey of an environmental pollutant, additional observations may be added in the vicinity of observed hot spots. In an epidemiological survey, whenever a person with the condition is found, close contacts of that person might be added to the sample. Such designs are in contrast to conventional sampling designs in which the entire sample may be selected prior to the survey. References to the rapidly expanding literature on adaptive sampling may be found in Thompson (1992) and Thompson and Seber (1995).

Some types of adaptive designs include adaptive cluster sampling, adaptive allocation, ordinary sequential designs, and optimal Bayes designs. In adaptive cluster sampling, an initial sample is selected by some conventional design such as simple random sampling, stratified sampling, unequal probability sampling, systematic sampling, or conventional cluster sampling. Then, whenever a unit in the sample satisfies a specified condition, such as having a high value, neighboring units are added to the sample. Still more units may then be added because some of the neighboring units also satisfy the condition. With adaptive allocation designs the allocation of sampling effort among strata is determined sequentially during the survey, usually based on observed sample means or variances.

An example of a situation in which adaptive allocation is used is the annual moose survey conducted by the Alaska Department of Fish and Game in interior Alaska. Moose are surveyed from aircraft flying over selected sample plots within a large study region. The study region is stratified based on the quality of habitat. At the end of each day during the survey, sample variances are computed for each stratum and the sampling effort for the next day is allocated to approximate the optimal Neyman allocation. Adaptive allocation designs have also been used and investigated for surveys of commercial fish species including mackerel, orange roughly, anchovy, and shrimp. The inherent problem in such surveys is that the mobility of the populations makes the patterns of abundance unpredictable prior to the survey so that classical allocation formulas are of limited usefulness. In the adaptive version, the pattern is assessed in the initial part of the survey (for example as the research vessel travels westward through the study region) and the sampling effort is reallocated for the second part of the survey (as the vessel returns westward in the example). With adaptive allocation the conventional stratified sampling estimator is not in general unbiased. Unbiased estimation is possible, however, through a variety of methods.

Adaptive cluster sampling designs have been investigated for ecological populations such as birds, sea mammals, fish, and trees, and for household surveys of rare characteristics. Similar designs have been recently investigated for elusive human populations such as drug users. These populations tend to be rare and unevenly distributed in the study region, with aggregation or social linkage tendencies. By adding neighboring units whenever high or interesting values are observed in the sample, the sample total of the variable of interest may be substantially increased — increasing for example the number of animals of the species that are observed — so that unbiased estimation needs to take the adaptive selection procedure into account.

The possibilities with adaptive designs are very wide. Much work is needed to determine effective designs for specific populations to compare the efficiency of adaptive designs with conventional designs and to determine which type of design will be most effective for the population of interest. For rare, clustered populations, adaptive cluster sampling can produce substantial gains in efficiency relative to conventional designs of equivalent sample size or cost. Factors influencing the relative efficiency include within-network variation, rarity within the study region, and cost issues. Adaptive allocation designs have also been shown to give gains in efficiency for unpredictably distributed populations including schooling species of fish and shrimp.

Nonsampling errors such as imperfect detection of animals, variable catchability of nets for fish, and incorrect self reporting of drug use must be taken into account with adaptive as well as with conventional designs. Generally, it is straightforward to adjust estimates for imperfect detectability by dividing observed values by detection probabilities. Imperfect detectability also adds components to the variance which must be estimated.

Motivation for adaptive sampling is provided both by real world situations such as those described above and by results in sampling theory showing that the optimal sampling strategy in many cases will be an adaptive one. The practical motivations for using adaptive sampling procedures for rare, clustered populations, from whales to rare plant species, have led field researchers to suggest adaptive procedures or to use them on an improvised basis. The theoretically optimal strategies tend to be too complex or require too much prior knowledge for practical implementation. Even so the theoretical results are suggestive of practical adaptive procedures with the potential to improve efficiency in surveys of real populations. Consideration of adaptive along with conventional sampling procedures greatly increases the possibilities in survey sampling.

References

Thompson, S.K. (1992). Sampling. New York: Wiley.

Thompson, S.K., and Seber, G.A.F. (1995). Adaptive Sampling. New York: Wiley. In press.

For further information, contact:
Steven Thompson, Department of Statistics, 326 Classroom Building, Pennsylvania State University, University Park, PA 16802
E-mail: SKT@Stat.psu.edu

The Right to Survey?
The Case for IASS Involvement
René Padieu

The IASS has done a great deal for the study and development of survey methodologies (sampling, formulation of questionnaires, nonresponse, analysis techniques, etc.). This continues to be a major part of its mandate. However, to date the association has had little involvement in debates that determine whether a survey can be conducted at all, as in the case of anti-survey campaigns or the drafting of laws and regulations. This paper pleads to the IASS to vigorously develop its involvement in this sphere.

The acceptance of a survey by its respondents depends not only on the ability of those conducting the survey to inspire confidence, also on the reputation of statistics among the general public. In recent years there has been an increase in efforts to denigrate statistics. Efforts which have responded to or even exploited the anxieties of the public concerning electronic data processing. The statistical organizations concerned have combated these problems on an ad hoc basis. But as yet we have seen little in the way of studies to explain the situation or propose an overall strategy.

At the same time, the right to survey has been tightly controlled by national laws, and international texts are being prepared. In many countries, protectors of data are imposing restrictions on surveys, and they do not always confine themselves to ensuring that data will not be used against the persons involved. Sometimes they become involved in deciding on the appropriateness of a survey or the relevance of a question for a proposed study.

Statisticians' reactions to these restrictions vary. Some become indignant, feel that they are misunderstood and are pained to see their professional ethics unrecognized. For their part, legal authorities sometimes suspect statisticians of engaging in unfair practices or at least of being unaware of the dangers that their activities pose. Curiously, support for statisticians is coming from legal authorities. This was the case recently when the population census in Germany was delayed until a judgment of a constitutional court clarified matters. The court recognized that the census was valid and that conversely some of the claims made against it were indefensible.

The issue is taking the form of a challenge to statistics itself. As a result, statistics, and statisticians find themselves engaged in a battle in which they are on the defensive. What is also at issue is the evolving relationship between science and society. This more positive way of looking at the situation calls for a collective effort of reflection. Statistical institutes (both public and private) are to some extent engaged in such a process, but often their energies are taken up with immediate negotiations. Thus it is up to the professional associations to carry this process forward.

The IASS can make a crucial contribution, owing to its international dimension. It, along with the ISI, can lend its prestige to disseminating the findings of this reflection in the statistical community and beyond, to national governments and in the arena of public opinion. However, the problem is posed in different ways in different countries, and it can be extremely enlightening to analyze these differences.

The problem is not posed in the same terms in Anglo-Saxon countries as in the Latin or Germanic countries of Europe. In the former countries, the legitimacy of surveys is not extensively challenged, and the emphasis is mainly on protection of the data as such (physical security, encryption, jamming, etc.). In the latter countries, the controversy focuses on actual data collection, and the debate is either political or ideological. Asia, Latin America and Africa with quite different cultural contexts, the problem takes on still other, quite different forms.

It should be noted that the International Association of Official Statistics could play a similar role, but the IASS has two particular reasons to become directly involved. First, what is at issue is really the initial phase of data collection; and second, private statistics as well as official statistics are equally affected. For all these reasons, I propose that the IASS undertake an international, analytical and comparative reflection. It should also develop a strategy and support positive and resolute efforts by our community to find a satisfactory solution.

For further information, contact:
René Padieu, INSEE, 18bd Adolpha-Pirard - F75675 Paris Cedex 14.

Was 1994 a Good Year for Statistics in France?
Benoît Riandey

Based on three events, French survey statisticians answer "no" without hesitation before qualifying their response. These three events and other activities in France in 1994 are discussed below.

January 1, 1994 saw the disappearance of the Centre d'Etude des Revenus et des Coûts (CERC), a highly independent public research center whose excellent studies and surveys especially its annual report on the evolution of incomes did not go unnoticed. The CERC had increased the number of specialized surveys of social and occupational groups such as the elderly, farmers, physicians and the legal professions in response to the known measurement problems associated with people who do not earn wages or salaries. A number of statistical services have taken over the Center's functions. We wish them the same success and recognition.

The year was also marked by the decision to postpone the next census of population, originally scheduled for 1997. Local communities in France were anxious that a census be held not long after the 1990 census because of the difficulty of making local intercensal population estimates without a national population register. This statistical priority was forgotten in the face of the requirement that public deficits be reduced, one of the famous convergence criteria in the Maastricht Treaty.

However, the European cause and statistics were not hurt by the postponement. The revised census schedule in France should be more in keeping with the European recommendations and the delay should provide an opportunity to invest in methodology to modernize the census and mobilize potential sources. In place of population files, France has excellent files on households (and consequently household migration), that are held by the major utilities such as the power and telephone companies. Some consider using these data as unrealistic because it has never been done and involves major institutional and legal problems. Specifically, legislation that is very protective of individual data, may make it difficult to use these data. This is a subject of much debate among European survey statisticians at a time when a new European directive on the circulation of individual data and the protection of privacy is being discussed.

A third event prompted a lively national debate in which the statisticians' associations took part: in the wake of demonstrations by disaffected young people, the government launched a national consultation of 15 to 25 year olds. A questionnaire with about 60 questions was widely distributed and more than 1.5 million responses were collected, representing 17 percent of the target age group. Leaving the social interpretation of the event to sociologists and political analysts, statisticians strenuously opposed publication of the statistical results of the consultation. They pointed to the irreparable bias of the "sample" and demanded that no results of a consultation be published unless a representative sample survey was conducted simultaneously. The matter would be closed if not for the risk that this type of political communication could be repeated in France or the European Union. Our associations should make their views known because such incidents undermine public acceptance of statistics.

Despite these events, several positive decisions were also reached during the year. In a country with few panel surveys, decisions were made to proceed with a number of major operations, including the French component of the European panel on household incomes and large cohorts for a study of the epidemiology of cancer and cardiovascular diseases. Since 1990, the Institute National de la Santé et de la Recherche Médicale (INSERM) has been conducting a ten-year panel (E3N) involving 100,000 women belonging to a teachers' mutual (MGEN). This is the largest component of the EPIC study coordinated by the World Health Organization in seven countries. Beginning in 1994, SUVIMAX is conducting a monthly followup of 15,000 individuals for ten years using telematics technology. This is an extremely innovative operation that we will have time to monitor.

For further information on these operations and discussions, contact:
Benoît Riandey, Institute National d'Etudes Démographiques, 27, rue du Commandeur, 75675 Paris Cedex 14, France
FAX: 33-1-12-18-21-92.

Country Reports

Australia (from Susan Linacre)

The ABS recently released information on family, culture, health, housing, education, training, employment, income and justice for the indigenous population (less than 2% of the Australian population) from its **Aboriginal and Torres Strait Islander Survey.** Careful consultation with many indigenous organizations and people was supplemented with advice from a committee chaired by an elected representative of the Aboriginal and Torres Strait Islander Commission and a series of technical sub-committees with a majority of indigenous people. A widespread information strategy ensured the value of the survey was respected in indigenous circles. A panel of 90 indigenous interviewers was recruited and trained, and a number of indigenous people joined the survey development and processing team. The success of these initiatives was reflected in the response rate, with completed questionnaires received from over 90 percent of the number of people that the Census suggested should have been enumerated.

The sample design involved stratification of census districts (CD) by location and size of indigenous population at last Census. For small CDs, all households with an indigenous resident were selected, while in the large CDs a sample of households was taken. Two approaches to locating indigenous people within the CDs were tested: 1) interview at each household in the CD to find out if indigenous people lived there; and 2) seek the addresses of indigenous people in the CD from members of the local Aboriginal community. The former approach was clearly better. Testing established this procedure was most effective using the ABS existing panel of interviewers and allowing the size of the newly recruited indigenous panel to be kept relatively small and more thoroughly trained. For more information, contact Glenn Cocking, ABS, Tel.: 61 6 252 7160.

The ABS is undertaking a major redesign of the **Monthly Population Survey (MPS)** (incorporating the Labour Force Survey and supplementary topics). The redesign involves: a review of content; change to a predominantly telephone based CAI collection methodology using BLAISE; a sample redesign to optimize for the revised survey; and development of an integrated set of system facilities for survey development, collection, processing, estimation and data management.

Users have expressed a high level of satisfaction with the main survey outputs, and see undisrupted time series of core data items as a high priority. Thus substantial content change to key series is unlikely. However, areas for review include better information on casual work, hours of work if part time, labor mobility, underemployment, marginal attachment to the labor force, discouraged workers, very long term unemployed, and the impact of labor market assistance programs. CAI also permits enhancements such as better targeting of subpopulations for particular topics and using data from previous interviews to obtain greater consistency and quality of gross flows data.

The sample redesign will look at the relative priority of different data outputs such as national versus regional estimates or the relative priority of original and seasonally adjusted versus smoothed series. Other issues will be optimizing clustering given changed cost structures, stratification for socio-economic as well as geographic variables, rotation rates, and the weighting in estimation.

The redesign will involve a series of small tests during 1995 with a major test in January to June 1996. The need for some form of parallel run at implementation in December 97 will be assessed following this major test. For more information, contact Robin Green, ABS, Tel.: 61 6 252 7166.

Canada (from Gordon J. Brackstone)

The Canadian Consumer Price Index (CPI) basket of goods and services is normally changed every four years to reflect changes in consumer spending patterns. At the same time, concepts and procedures are re-examined and sometimes revised. With the basket updating in January 1995, Statistics Canada made a number of changes to the CPI. The motivation was partly financial since program expenditures are expected to decrease by about 7 percent.

The CPI now targets the total population of Canada. Previously, cities with less than 30,000 people (representing about 30% of Canada's population) had been excluded from the target population. The sampling of retail outlets is modified to better represent provincial price movements with less emphasis on city indexes. The shift in emphasis leads to a reduction in the number of price quotes required during data collection. The commodity classification is reorganized to improve relevance, balance, and international comparability. CPI information is available in a monthly publication, a quarterly publication, and in electronic form. To obtain additional information, please contact Stuart Pursey, Business Survey Methods Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6, Tel.: (613) 951-9432 or FAX: (613) 951-1462.

For the 1996 Canadian Census, a new minimum change hot deck imputation methodology (NIM) will be implemented. It will perform edit and imputation for the variables age, sex, marital status and relationship to person 1 for everyone in a household simultaneously. Missing, invalid and inconsistent responses will be resolved by the imputation actions. The NIM allows, for the first time, minimum change hot deck imputation of a mixture of numeric and qualitative variables simultaneously. The NIM edits are specified through machine readable decision logic tables that are created by the subject matter experts. For the 2001 Census, it is planned to generalize the NIM so that it will process all the Census variables. For further information, please contact Mike Bankier, Social Survey Methods Division, Statistics Canada, Ottawa, Ontario, K1A 0T6. Tel.: (613) 951-6938.

The Canadian Cultural Labour Force Survey was conducted in 1994. This sample survey collected data, through computer-assisted interviews, from 19,000 people working in a large number of occupations and industries related to cultural activities. Results of this survey will be used for identifying work patterns, employment stability, income dynamics, career paths and impediments, and training needs for cultural workers. The survey has a list frame component based on about 250 lists from cultural associations, councils, unions and funding agencies, and another component for surveying workers in selected cultural industries and institutions. For more information, please contact Simon Cheung, Household Survey Methods Division, Statistics Canada, 16-J, Robert H. Coats Building, Ottawa, Ontario, Canada K1A 0T6. Tel.: 613-951-1482, FAX: 613-951-3100.

Central America and Panama (from David J. Fitch)

Preliminary results from the 1994 Guatemalan census give a population of 8.3 million. The projection for 1994 based on the 1981 census was 10.3 million. Population figures are needed soon to determine on the number of deputies for the elections in 1995. The President stated the figures are unofficial. The Vice-President indicated a million and a half people are missing but the figures from INE (Instituto National de Estadistica) have to be accepted. The current president of the Congress (and likely the next president) suggested INE lacks a technical and scientific staff to adequately carry out their work.

Currently underway is a UNICEF Micronutrient Survey that surveys households with a mother and at least one child aged 1-5. Data are collected on salt iodination, hemoglobin status of mothers and children, household food consumption, and vitamin A status of children. For sampling, this survey uses maps prepared by INE in 1992-93 for the 1994 census. These maps are a great improvement over the EPI method of selecting households that are used in many surveys. Additional design features currently being considered on an experimental basis include: 1) Call backs – currently if a mother is out and won't be back until later a substitute household is sought rather than returning at a later time; 2) New construction – There is no provision for including people living in houses constructed since the maps were made, except as substitute households. One possibility is to draw boxes around each house listed so the entire area of the cluster is covered and then to include a newly constructed house if it fells within the box of a sampled house; 3) Probability selection of children – Currently, if there is more than one child aged 1-5, the one most convenient is selected; 4) No substitution – there would be advantages in selecting a sample and sticking to it, and not substituting; and 5) Weighting to reduce bias – if counts of the number of children in a household and the number of households screened could be used to develop weights and reduce the bias of the estimates. For more information, contact David Fitch, Instituto de Nutricion de Centro America y Panama, Apartado Postal 1188, Guatemala, CA, Guatemala.

New Zealand (by Kevin Eddy)

Statistics New Zealand will be using imaging technology to capture, code and edit data from the 1996 Census of Population and Dwelling questionnaires. An estimated 1.2 million dwelling and 3.5 million individual questionnaires will be scanned and processed. A project to assess the feasibility of imaging tick box and numeric responses from 1996 Population Census questionnaires and trials with vendors were conducted. The image trials were critical to the project in providing base line recognition results, and for gaining first hand imaging experience in a census environment. Individual questionnaires from the 1991 Census were used with no modifications to provide a worst case scenario. In the most successful of these trials, an average speed of 3,800 forms per hour was achieved with a maximum rate of 4,300 per hour.

Overall, accurate recognition rates of 99.4 percent for ticks and 74 percent for numerics was achieved. Of the 0.6 percent error rate for ticks, 0.52 percent came from ticks not picked up, and 0.08 percent came from "phantom" ticks picked up. Of the 26 percent fail rate for numerics, 6 percent were recognized incorrectly and 20 percent were not recognized. The key entry error rate for ticks from the 1991 Census was 0.43 percent. The high recognition rate obtained from questionnaires not specifically designed for scanning, and the savings in resources, led to the decision to proceed with imaging technology.

The images and interpreted data will be used for subsequent coding and editing. Instead of referring to responses from the paper questionnaires, images of the responses will be displayed on screen. From this the coder will key enter any unrecognized responses, capture all alphabetic responses, and complete all coding and editing. Operators will be able to view up to 2 pages of the form on 21 inch monitors with the capability of zooming any part of the image.

The development of the capture system is entirely outsourced. The development of the code and edit system is also outsourced but with Statistics New Zealand providing modules for complex coding such as family and households. Scanners and associated hardware are provided as part of the contract with the vendor. This is seen as preferable to purchasing because of the costs and risks associated with maintaining large and quickly outdating hardware.

The capture operation will take two months to complete. Provisional data will be released up to three months earlier than in the 1991 Census. Final coded and edited data will be available by the end of 1996, i.e., within nine months of Census date. For more information, contact Kevin Eddy, Statistics New Zealand, P.O. Box 2922, Wellington, Tel.: 04-495-4690 and FAX: 04-495-4603.

Philippines (by Gervacio G. Selda, Jr.)

A National Convention on Statistics is to be held in Manila, Philippines on December 4-5, 1995 with "Gearing the Philippine Statistical System Towards the 21st Century" as its theme. The convention will include invited talks, contributed papers and commercial displays by organizations offering statistical products and services. The event will provide a forum for the country's theoreticians, practitioners, data producers, and users to exchange ideas and experiences in the field of statistics and discuss recent developments and current issues and problems affecting the statistical system.

Papers to be presented and discussed in the convention will cover: statistics and information systems and statistical computing; statistical application in industries; system of national accounts; statistics on children and gender concerns; environment statistics; statistics for local development planning; statistical theory and methodology; dissemination and utilization of statistics; geographic information systems; population and vital statistics; social statistics; agricultural statistics; labor and employment; and statistical education and training.

Due to representations from local government units and other sectors, a Mid-Decade Population Census involving a complete count of all residents, both Filipinos and foreigners, will be conducted by the National Statistics Office in September 1995. This Census will cover both the household and institutional population including overseas Filipino workers. The count will gather data on the following characteristics of each individual: relation to household head; sex; age; civil status; disability; ethnicity; educational attainment; trade skills; and economic activity. The preliminary results will be made available in December 1995, while the official count will be released in March 1996. The early release of census results will be made possible through the implementation of a decentralized system of data processing using microcomputers.

The Philippine government through its National Statistical Coordination Board has recently adopted a System of Designated Statistics that would facilitate the collection, compilation, processing and dissemination of data by agencies in accordance with an approved statistical calendar. The scheme is expected to address problems with the country's statistical system such as data gaps, delayed release of important data series and duplicating and conflicting statistics. Also, the new system will enable the determination of the most critical and essential statistics required for social and economic planning and analysis. It will establish priorities for data production and provide a means for better allocation of budgets for government statistical activities.

A Redefinition and Classification of Urban and Rural Areas in the Philippines is underway to replace the definitions formulated in 1970. A study is being undertaken jointly by the Statistical Research and Training Center and the National Statistics Office to determine the key factors of urbanization, to recommend an alternative method of defining urban areas, and to update the classification of urban and rural areas. Multivariate analysis is being performed utilizing the 1990 Census of Population and Housing data. It is the intention of the study to come up with an urban redefinition which is simple to apply using regular census results. The United Nations has taken the position that each country is the best judge of what criteria will best define an urban area in its particular circumstances.

For more information, contact: Mr. Gervacio G. Selda, Jr., Executive Director, Statistical Research and Training Center, J&S Bldg., 104 Kalayaan Avenue, Quezon City, Philippines. FAX: (632) 921-7485.

Republic of Korea (from Seung-Kon Lim)

The 15th Population and 7th Housing Census, which have been conducted at five-year intervals since 1925 and 1960 respectively, are scheduled to be jointly carried out during October 1-9, 1995. The National Statistical Office (NSO), the agency in charge of the census-taking, finished preparatory works such as the setting up of enumeration districts (EDs), conducting four pretests and a pilot survey, and the determination of survey items. The total number of EDs delineated was 220,000, comprising about 60 households for each ED.

Concerning the mapping work, NSO will introduce the concept of computerized ED maps for the first time in its census-taking. This means that the boundaries of EDs and outstanding features located within EDs will be produced by the computer using a Geographical Information System.

Three major features characterize the 1995 Census. First, there is a considerable reduction in the number of survey items compared with previous censuses in order to mitigate the burden on respondents. Second, in addition to a short form questionnaire there will be a long form used on a 10 percent basis. Third, some new items such as religion will be included.

An OMR system will be adopted for data capture in the 1995 census. The preliminary report and the advance report will be published in February and June 1996, respectively. The final results will be published by the end of 1996.

For further information on the census, contact: Seung-Kon Lim, the Director of Planning and Management Division (recently changed from the Director of Statistical Standards Division), the National Statistical Office, 647-15 Yoksam-dong, Kangnam-gu, Seoul 135-080, KOREA
FAX: 82 2 538 5343, Tel.: 82 2 222 1822.

Sweden (from Anders Christianson)

Statistics Sweden Launches Annual Research Conference: Statistics Sweden has decided to organize a research conference on survey methodology June 12-13, 1995 in response to the changes made to the Swedish statistical system and the need for coordination of methodology and standards caused by these changes,. The conference is designed to become a forum for discussing methodological issues between the agencies authorized to provide official statistics in the new decentralized system. The conference will also devote particular attention to the issues raised by harmonization of European statistics. One part of the conference will be held in Swedish, another one in English. The English sessions will cover the following topics: Methodological Issues in European Harmonization; Euroharmonization Issues for Particular Survey Areas; Dissemination of Micro and Macro Data; Quality Issues for Official Statistics; Measurement Models and Their Impact on Design and Analysis of Official Statistics; and Survey Quality Evaluation and Improvement.

Syria (from Ibrahim Ali)

The Central Bureau of Statistics has executed the fourth housing and population census as well as simultaneous censuses of establishments and agricultural holdings. These censuses will offer a broad base of statistical data describing the size and development of the population, its geographical distribution, its demographic and socio-economic characteristics, the number and types of dwelling, and their geographical distribution and characteristics. They will provide a picture of changes in agriculture and agricultural holders, as well as a comprehensive frame of economic establishments, including their numbers, their geographical distribution, their activities and their number of employees.

Data on establishments are obtained from an initial enumeration of dwellings that takes place three weeks before the counting period. Data for the housing and population census and for the agricultural census are collected by enumerators in one visit during a six-day counting period. Two questionnaires, the first including housing and population data and the second including data on agricultural holders and holdings are used. The housing and population census questionnaire is divided into three parts. The first part includes data on the housing and its characteristics. The second part includes demographic and basic data, place of residence (usual and current). It also asks if the person has an agricultural holding or not. The third part includes demographic, social and economic data obtained from a sample of 10 percent of families within the district of every enumerator.

Through the design of the two questionnaires, and the ability to link them together, the C.B.S. will be able to describe and analyze the demographic, social and economic characteristics of agricultural holders.

For more information, contact:
Dr. Ibrahim Ali, Director, Central Bureau of Statistics, Damascus, Syria.

United States (from Dan Kasprzyk)

A new survey of the United States non-institutionalized population 70 years of age or older has just entered the data collection phase. The study is being conducted by the Survey Research Center at the University of Michigan and focuses on changes in economics and health status for older persons. The sample for this study, called AHEAD (Asset and Health Dynamics), is derived from the screening activity used to identify the pre-retirement cohort (those age 51-61) currently being studied in the Health and Retirement Survey (HRS). AHEAD aims to achieve a better understanding of the interaction between changing health status and changing asset position. Cognitive functioning, demographic characteristics, housing status, income, net worth, and health insurance coverage will be collected in AHEAD. The sample includes approximately 8,000 persons in roughly 6,500 households. For more information, write to Dr. Thomas Juster, Survey Research Center, the University of Michigan, Ann Arbor, Michigan 48106-1248.

The second edition of *A Guide to Selected National Environmental Statistics in the U.S. Government* is now available from the U.S. Environmental Protection Agency. The Guide is a directory to selected U.S. statistical programs which collect national level, time series environmental statistics. Each entry describes a separate statistical program, giving program purpose, data coverage and collection methods, geographic coverage, agency contacts, pertinent publications, and database access options. Statistical coverage includes data related to the state-of-the environment (e.g., air and water quality), pressures on the environment (e.g., energy use, mining, transportation), human health and welfare issues (e.g., exposure to toxic chemicals), and societal responses to environmental problems (e.g., expenditures on pollution abatement). To order the guide, write to U.S. Environmental Protection Agency, Office of Policy, Planning and Evaluation, Environmental Statistics and Information Division 2163, 401 M Street, S.W., Washington, D.C. 20460.

The Social Security Administration (SSA) conducted the 1991 New Beneficiary Follow-up Survey (NBF) by reinterviewing respondents to the 1982 New Beneficiary Survey (NBS) or their survivors. These data constitute a nationally representative sample of persons who began to receive various types of Social Security benefits in 1980-81. Item nonresponse was cross-sectionally imputed for each survey without information from the other survey. The Survey Research Center at the University of Michigan, under contract to the SSA, will develop longitudinally based imputations for the NBS and NBF using the full range of survey information. For more information concerning this project, write Dr. James Lepkowski, Survey Research Center, University of Michigan, Ann Arbor, Michigan 48106-1248.

The National Health Interview Survey (NHIS) is a cross-sectional household probability survey of the civilian, non-institutionalized population of the U.S. In-person household interviews are conducted throughout the year and each weekly sample is a representative national sample. In 1992, an Immunization Supplement was added to monitor vaccination levels of preschool children across the United States. Vaccination histories for each sample child are obtained from a parent/guardian's vaccination record or from recall.

Due to concerns regarding response errors, the National Immunization Provider Record Check Study was conducted in conjunction with the 1994 NHIS. This study was a collaborative effort of the National Center for Health Statistics (NCHS); the National Immunization Program (NIP) with Westat, Inc. as the contractor. All children 19-35 months of age in the NHIS reported to have received any vaccinations were asked to provide the name, address, and telephone number for up to three vaccination providers and to sign a consent form for the release of the vaccination information from the provider.

Preliminary results using provider records as the "gold standard" and weighing class adjustments, show that national vaccination levels are higher than those previously reported. Work is also underway to create a "best value" for each individual child's vaccinations, where "best values" take into account information from both the household report and the provider record. These "best values" will then be compared to the original NHIS report and measures of net and gross difference rates computed.

For more information on this study, contact:
Elizabeth R. Zell, National Immunization Program, Centers for Disease Control and Prevention, Mail Stop E05, Atlanta, Georgia 30341, USA.
Tel.: (404) 639-8392, FAX: (404) 639-8613.

Question/Answer

Editor's Note: The Question/Answer section of *The Survey Statistician* has published questions raised by members and answers to those questions from 1978 until 1994 due to the untiring efforts of Professor Leslie Kish. Recently, the President and Council of the IASS asked Professor Kish to make all of these sections available so that they could be published in one compendium. The publication of this compendium is now being planned. All of the members of IASS would like to thank Leslie for his great contribution *The Survey Statistician*.

The President and Council have also asked Vijay Verma to support Professor Kish in the Question/Answer section of this and future issues of *The Survey Statistician*. Dr. Verma has agreed to act as the coordinator of the section, calling upon Professor Kish and other experts to respond to the questions. One feature being considered for future issues is a discussion of various aspects of the answers, featuring the comments of Professor Kish and the other experts. The section below is the first with Vijay serving as coordinator.

The following are some of the questions asked at the IASS Workshops on Survey Sampling in Developing Countries (Cairo 1991, Padua 1993). Other questions raised at the workshops will be discussed in this column in future issues. In the meantime, we welcome alternative views on the issues raised, comments on the answers provided, and especially other *good and practical questions*. Please send questions to Vijay Verma, 105 Park Rd., Teddington, Middlesex TW11OAW, UK, FAX: (44 181) 943-4056.

Q31.1 A 2-stage sample was designed with the first-stage units (PSUs) selected PPS with size being the estimated number of second-stage units (SSUs) in each PSU, and a random sample of a fixed sample size of SSUs to be selected within each sampled PSU. This design was chosen to obtain a self-weighting sample and have a fixed workload per PSU as well as a fixed total sample of SSUs. As it inevitably happens, it was found during the survey that the actual number of SSUs within each PSU was different from that used in selecting the sample of PSUs. Is there any method of keeping the design approximately self-weighting without appreciably changing the sample size per PSU and the total sample size?

Answer:
This is common and hence an important problem. With the higher stage units selected with probabilities proportional to estimated sizes, it is important to keep clear the distinction between 1) a *strictly self-weighting design* on the one hand; and 2) a *fixed-take design* on the other. In the former, the ultimate units or elements all receive the same overall probability of selection, and the number of these units selected is allowed to vary to the extent the measures of size are inaccurate.

I recommend that, as a general rule, one should opt *for a self-weighting rather than a fixed-take design*. The same applies to seeking solutions for the particular problem raised in the question. This is because fixed-take designs have a number of practical disadvantages:

1. Arbitrary variations in the ultimate selection probabilities which a fixed-take design involves are generally undesirable and can be inconvenient, especially when the variations are large enough to require weighting of the sample data.

2. It is most common in practice to select the ultimate units systematically from lists. With a self-weighting design, the selection interval can be prespecified for each sample PSU. However, this selection procedure is more complex with a fixed-take design, which can cause problems if the selection has to be done by lower-level staff in the field (see SS Question-Answer 27.3.).

3. As the overall selection probabilities are not fixed in advance, they can be computed only if an account is kept of the actual number of *eligible* units listed in each sample PSU. This may appear straightforward, but *many surveys fail to keep such records in practice.*

4. Comparison of the selected and target sample sizes can identify problems of undercoverage in the frame, such as incompleteness of the list of areas included in the frame, major mistakes in sample selection e.g. leaving out a part of the frame, systematic preferences of fieldworkers to select smaller segments and, perhaps most commonly, systematic omission due to outdated or poor listing of ultimate units in the sample areas. The problem of the selected sample size seriously falling short of the target size has been encountered in many surveys. A major disadvantage of a fixed-take design is that by automatically pre-fixing the sample size, it hides problems of this type.

5. Worse still, a fixed-take design may actually encourage incomplete listing, especially when the same fieldworkers are responsible for both listing and interviewing. I have come across surveys where only a fixed number, at worst only as many as required to get the fixed number of sample cases, were listed in each area, debasing the design and creating an enormous potential for bias against less accessible units.

6. Fixing the sample sizes can encourage uncontrolled substitution for nonresponding

cases and discourage the effort required for persisting with more difficult cases.

7. In any case, fixing the sample sizes is often unattainable because of nonresponse and other problems at the implementation stage. The same applies when the units enumerated are not the same type as the units used for listing and sampling, or when the survey is concerned with a subpopulation not controlled in the design.

Fixed-take designs may nevertheless, be preferred in very 'heavy' surveys (e.g. involving very lengthy interviews, frequent and repeated visits to each household, or elaborate physical measurements), when even minor variations in the number of sample cases is felt to result in unacceptable differences in interviewer workloads. But these are exceptions.

Solutions

1. The advice in general is: stick with the self-weighting design and tolerate variations in sample-takes per PSU and in the overall sample size to the extent possible. We often underestimate the flexibility and tolerance available in practice. Often extra staff can be assigned to areas with exceptionally large workloads. Also there is more flexibility when interviewers work in teams rather than singly within sample areas. After all, even fixing the number of interviews does not necessarily equalize the interviewers' actual workload. As to the overall sample size, a great deal of the variation in individual PSU sizes tend to cancel out.

2. The following steps can be taken to control the overall sample size.

   First, ensure at the outset that the parameters in the selection equation are correctly scaled. This has to take into account:

   - Differences in the nature of units in terms of which the PSU sizes are measured (such as persons or households as defined in the frame), and the units to be enumerated in terms of which the sample size is measured (such as dwelling units or households as defined in the sample survey);

   - Changes in the size of the population since the frame was prepared (which may be several years ago for area frames); and

   - Expected losses due to nonresponse, undercoverage etc. It is not necessary to uniformly adjust all measures of size, as the required overall scaling can be applied in defining b, the target sample-take per PSU. For instance, if it is desired to obtain a sample of size n=ab, with a PSUs selected with PPS at the first stage, followed by a target of b SSUs per sample PSU at the second stage, the selection equation can be written as:

$$f = \left(\frac{aM_i}{\sum M_i}\right)\left(\frac{b'}{M_i}\right) \text{ with } b' = b\left(\frac{\sum M_i}{\sum N_i}\right)$$

where M_i is the measure of size of PSU i, $\sum M_i$ is the sum of measures of size, and $\sum N_i$ is the total of the required ultimate units in the population.

Second, if the listing and the final sample selection are separated in time and intervention is possible between the two operations, the overall expected sample size can be adjusted taking into account the outcome of listing in the sample areas by changing the last stage sampling rate to:

$$\left(\frac{b''}{M_i}\right) \text{ with } b'' = b'\left(\frac{1}{n}\sum_s\left(\frac{L_i}{M_i}\right)\right)$$

where L_i is the number of ultimate units listed in PSU i, and the sum is the separate ratios over the *sample* PSUs. This adjustment can be useful especially if the requirement of proper scaling noted above was not met. However, the adjustment is of limited value, in so far as it affects only the overall sample size, but not the variation of PSU workloads. Also the selection probabilities are now known only in relative rather than absolute terms, though that suffices in sample surveys aimed at estimating ratios or similar statistics rather than population aggregates.

3. More difficult of course is the problem of large variations in sample-takes per PSU. *When the discrepancies are large, it can greatly help to use a combination of approaches*. Here is an example. We can arrange and group the sample PSUs according to the expected sample-take on the assumption of a self-weighting design, i.e. computed in proportion to the ratio for each PSU of its actual size to the estimated measure of size originally used in its selection. *Within each group*, a self-weighting design can be followed letting the sample-takes vary as required. *From one group to the next*, the average sample-take can be increased, but less rapidly than that implied by a self-weighting design throughout, with appropriate increase in the sample weight. Thus, in a group j where the PSU sizes have changed by a factor c_j on the average compared with the measures used in the original selection, the average sample-take according to the self-weighting design would bc_j. We can arrange for this take to differ less from the originally planned b, for instance by using $bc_j^{1/2}$ and compensating with weights $c_j^{1/2}$. For units which have quadrupled in size ($c_j = 4$), the sample-takes as well as the weights would only be doubled. Even the extreme case of c_j=10, the takes and weights are only approximately tripled. This approach utilizes the fact that the effects of weighting (or of variations in PSU sizes) tend to be negligible for small departures from self-weighting, small for moderate departures, but increasingly large for large departures. Prespecification of the groups for the last stage of sample selection with this scheme also permits a prespecification of the sampling intervals to be applied, thus simplifying the sampling operation.

4. *Trim extreme values*. For a self-weighting sample, this means putting limits on the minimum and, more importantly, the maximum number of units to be selected from a PSU irrespective of the magnitude of the discrepancy between its size measure and actual size. As practical limits, perhaps a 10-fold to 15-fold range of (b/3 - 3b) or (b/4 - 4b) will be a reasonable choice in many circumstances. A similar range may be imposed on the weights when weighting is involved such as a fixed-take design. The idea of trimming is to control the increase in variance, even though it may introduce some bias.

5. Special techniques may be used in continuous survey operations, such as the one discussed by Kish (1965, section 12.6c) involving the construction of a 'stratum of surprises'. The basic idea is to accumulate information over survey rounds for the estimation of the total population of the units which have expanded in size by more than a certain factor. That data can then be used to determine the total number of ultimate units to be taken from such units encountered in the present sample, given its required sampling rate.

Q32.1 When good current estimates of total population by age and sex are available, some recommend that the sampling weights for a multi-stage sample of households be ignored and the sample be weighted to the current estimates for poststrata defined by age and sex. What would your recommendation be?

Q32.2 What should be done if post-stratification leads to empty poststrata, i.e. poststrata with no sampled units in them?

Answer:

These two questions can be taken together as both concern certain aspects of weighting the sample data. The first issue is whether or not it is appropriate/desirable to weight the data in particular circumstances. This a major and important practical question, and is not considered here (see SS Q/A 22.2, and fuller discussion in Kish (1992)).

Assuming that the sample data are to be weighted, how should it be done and on the basis of what information?

Information for the computation of weights can come from a number of sources: from the sample design itself (selection probabilities); from sample implementation (nonresponse and characteristics of the nonrespondents); from the frame; and from external sources, of which there may be several. While it may be useful to use some or all of these sources, it is my recommendation that *preference be given to sources closer to the sample over more removed· or external sources* as the basis for computing the sample weights. Furthermore, when the sample data are to be weighted, it is highly desirable to *follow a systematic, step-by-step procedure*. The weights to be applied as a result of each step should be computed separately so that the magnitude and effect of the weights introduced at each step can be identified and evaluated. On examination, some steps may turn out to be unnecessary, while others may require modification such as 'trimming' to avoid extreme values of the weights involved.

This means that, where required, the first candidate as a source for the computation of sample weights is the design itself: *design weights* applied in inverse proportion to the overall selection probabilities of the sample elements. Next would be *nonresponse weights* to compensate for differences in response rates in different parts of the sample. These require comparable information on all units selected into the sample — the responding as well as the nonresponding units. Often some such information is available from the sampling frame; additionally some may also be collected as part of the survey itself.

Only then should we consider weighting the sample data on the basis of information external to the survey. In principle, external adjustments can be applied to any sample, but such adjustments provide no insurance against biases if the sample is basically of poor quality or if the external data are of doubtful accuracy. The correct procedure is to design and implement the sample properly, using strict probability procedures, and introduce correctional weights only if the external source is known to be clearly superior to the sample. It is not a good practice to automatically and routinely apply external weights in every survey. *It is certainly undesirable to ignore the design weights, nonresponse rates and other information internal to the survey, and adjust the survey results purely on the basis of external data.* Furthermore, controlling only a few characteristics does not always suffice. Controlling the age and sex distribution of individuals, for instance, does not ensure that the distribution by other important characteristics such as household size and composition is also improved. 'Correcting' the sample results on just one or a few arbitrarily chosen characteristics may in fact introduce distortions on other variables.

A number of requirements should be examined before the decision is taken to use external information for weighting:

- Establish that such weights are *needed*, as may be the case if the sample is small, or there have been obvious shortcomings in its design and implementation — especially serious departures from probability sampling;

- The external information must clearly be more *reliable* than what is available from the survey itself;

- *Relevant* and *effective* external data are needed to improve the representativeness of the survey results.

- The external information should *cover diverse variables*, to the extent possible.

- The external information should be *consistent* with the survey, i.e. the variables used in the adjustment should be defined and measured in the same way in the sample and external source(s). Consistency is also required between the external sources if more than one are involved.

Often in statistically less developed counties, good and up-to-date external data are not available. Therefore, it is necessary to be cautious in applying external weights to 'correct' the sample results, and avoid application of the correction at too detailed a level. Indiscriminate and elaborate adjustment of sample results on the basis of external data of insufficient quality has, unfortunately been the practice in many surveys.

Weighting on the basis external information basically involves the division of the sample into appropriate weighting classes or poststrata. Within each class, the responding units are weighted by a uniform factor to make the resulting weighted distribution of the sample units on the characteristic(s) involved in the classification the same as the distribution found in the external source. (Similarly, in weighting for nonresponse, the responding units in each weighting class are weighted to make the distribution of the responding cases according to the characteristics involved the same as that for the total sample originally selected.)

The effectiveness of the procedure in reducing biases in the survey estimates depends on the extent to which the characteristics used in defining the weighting classes are correlated with the (often diverse) substantive variables to be estimated from the survey, and the units within any particular class are homogeneous on the substantive variables of interest. This, plus the desirability of controlling many important characteristics simultaneously, would favor the use of many weighting classes or strata, each relatively small and homogeneous. On the other hand, the use of many small weighting classes can result in the application of large and variable weights which can greatly increase the variance of the sample estimates. The choice will depend upon the particular situation, but in any case, *the use of very small classes and resulting large weights must be avoided.* In a normal household survey involving a few thousand interviews for instance, I would hesitate to use any weighting classes or poststrata smaller than 100-200 units. Hence if some of your poststrata are empty or contain only a small number of sample units, the best thing is to re-examine the classification chosen and combine small classes to avoid this problem.

To retain sufficient class sizes when controls by a number of characteristics are involved, one practical solution is to control simultaneously only the marginal distributions, rather than the full cross-classification. This can be done by simple iterative procedures, such as the classical Iterative Proportional Fitting Method of Deming (1943). More elaborate procedures, which bound the weights within specified limits, are also available (e.g. Deville, Särndal, and Sautory (1993)).

References

Deming, E. (1943). Iterative Proportional Fitting Method, Statistical Adjustment of Data, Chapter VII. New York: Wiley. Paperback by Dover Publications, 1964.

Deville, J., Särndal, C., and Sautory, D. (1993). Generalized Raking Procedures in Survey Sampling. *Journal of American Statistical Association*, 88:423, pp. 1013-1020.

Kish, L. (1965). Survey Sampling. John Wiley Sons, Inc., N.Y., N.Y.

Kish, L. (1992). Weighting for Unequal Pi. *Journal of Official Statistics*, 8:2, pp. 183-200.

Budapest Conference

Plans for this regional IASS conference are now underway following positive responses from several transition countries to the preliminary announcement. The conference will focus on problems and experiences of sample surveys in transition countries. It is planned to take place in Budapest on 2-4 September 1996. The conference is to be sponsored by the IASS, the Budapest University of Economic Sciences, the Hungarian Statistical Association and the Hungarian Central Statistical Office. A local Organizing and Program Committee has been set up, composed of the following members:

> László Vita, Budapest University of Economic Sciences (BUES), Chairman of the Committee;
> János Ay, BUES;
> Ádám Marton, Central Statistical Office (CSO);
> Ms. Edina Pál, CSO; and
> György Sebestyén, BUES.

Members will also be invited from other countries.

Tentative topics for invited papers include business register construction, nonresponse and imputation and total quality management in sample surveys, together with the general problems and experiences of sample surveys in transition countries.

It is planned that contributed papers will be organized into the following four sessions:

- I: Problems and experiences with registers;
- II: Coping with nonresponse;
- III: The treatment of sample survey methodology in university curricula; and
- IV: Management of fieldwork.

Inquiries about the conference should be directed to:

> Professor László Vita, Department of Statistics, Budapest University of Economic Sciences, H-1828 Budapest 5, P.O.B.49, Hungary. Tel/FAX 36-1-217-6118. E-mail vita@ursus.bke.hu

---

Reduced Subscription for IASS Members

Members of IASS are reminded that an arrangement between IASS and Statistics Canada allows members to subscribe to *SURVEY METHODOLOGY* at a cost of US $30. This is a saving of over 30 percent off the regular price. Members of IASS in developing countries receive a further reduction of 50 percent off this price through a subsidy funded by IASS. Members wishing to order *SURVEY METHODOLOGY* should do so when they pay their annual IASS membership dues.

The June 1995 issue includes a special section on "Randomized Response" as well as papers on Weighting of Longitudinal Surveys and Undercoverage in Censuses.

If you have any questions regarding your subscription or delivery of your copies, please contact Statistics Canada at (613) 951-9807 or FAX: (613) 951-3100.