Journal of Official Statistics
Vol. 9. No. 4, 1993, pp. 827–835

Book Reviews

Books for review are to be sent to the Book Review Editor Gösta Forsman, Department of Mathematics, University of Linköping, S-581 83 Linköping, Sweden.

de Leeuw, E.D., Data Quality in Mail, Telephone, and Face to Face Surveys. TT-Publikaties, Amsterdam, 1992. ISBN 90–801073, xii+168pp., NUGI 659.

As suggested by the title, this book's raison d'être is to explore the effects of mode differences in sample surveys. After a general introduction, there is a chapter on why one might expect mode differences. This is followed by a chapter discussing the literature in terms of a meta-analysis. The core of the book is the description of a field experiment, employing mail, telephone, and face to face methodologies, and the results of that experiment in terms of basic data quality indicators, reliability and scalability, and the effect on multivariate models.

This is a well-designed and analyzed experiment which maximizes the chances of isolating and understanding mode effects. Analyses go beyond the standard indicators of data quality to look at reliability and scalability. Most importantly de Leeuw explores the assumption of "form-resistant" correlations by looking at the implications of mode effects on actual substantive models. It is this last step, which is often missed, which is critical in connecting the methodologist to the practitioner.

The mail survey comes out very well in the comparison experiment. Reading this study has certainly led me to rethink the value of mail surveys. The book is well worth reading whether you are a student of survey research or a practitioner. It is "a must" for anyone faced with designing a survey and considering a choice of mode. Knowledge of psychometric principles and covariance structure models (cf Long 1983) would be helpful for the reader but is not a necessity. The book also represents de Leeuw's dissertation. This adds a somewhat pedantic flavour in places but does not distract from the quality of the work.

The few concerns and reservations I have about the book centre on a desire for more information. There were several places where I would have liked to have seen greater detail on a topic, issue, or decision. I found myself wondering whether the

book as it currently stands represents an abridged or concise version. Similarly, there could have been more discussion throughout, especially on the implications of the findings.

The book tends to combine results from the U.S.A. and Western Europe. Countries vary in the type of sampling available, the feasibility of conducting Random Digit Dialling telephone surveys, cultural views to the telephone and the mail, etc. Although it is up to individual readers to apply their own cultural filters a discussion of what is known about between-country differences with respect to mode would have been useful.

Chapter 2 offers some interesting ideas on why data collected from the various modes might diverge. For example, de Leeuw describes how face to face, telephone and mail surveys differ on a number of factors that are inherent to the social convention associated with the medium of communication. These include such elements as how familiar people are with the medium (i.e., how well accepted the telephone is as a medium of communication), where the locus of control is during the data collection (i.e., mainly with the interviewer, with the respondent, or somewhere in between), social conventions regarding the acceptability of silences in a conversation, and how well the mode allows the sincerity of the research purpose to be demonstrated. Also discussed are how the modes differ with respect to communication channels, presentation of stimuli, and regulation of the communication flow. Several avenues for further research on the effect of communication style on survey results could be built from these points, say through other disciplines such as ethnography. For example, the survey assumption that interviewer neutrality is necessary to ensure unbiased answers may not be congruent with some cultural norms of conversation. In cultures in which communication and social interaction largely rely on shared understandings and a high degree of reciprocity, respondents may balk at the unnatural interaction style (Salo, Campanelli, and Martin 1990).

This chapter is a bit weak in its discussion of the literature on the effect of interviewers on data quality. There is a long tradition of research on interviewers, spanning the areas of interviewer expectations and attitudes (cf Sudman, Bradburn, Blair, and Stocking 1977), interviewers' socio-demographic characteristics (cf Schuman and Converse 1971), and interviewer training and behaviour (cf Cannell, Miller, and Oksenberg 1981; Durbin and Stuart 1951; Fowler and Mangione 1984). De Leeuw also does not mention the fact that interviewers can be conceptualized as contributing to response bias as well as response variance.

Chapter 3 summarizes the literature on survey mode effects. This was done via a meta-analysis, a refreshing alternative compared to the standard literature review. The meta-analysis suggested that there are currently only small effect differences between well-conducted face to face and telephone surveys. The differences between mail surveys and interview surveys, although small, were not negligible, e.g., both the overall nonresponse and item nonresponse rates are higher in mail surveys.

Chapter 4 contains details about the empirical study. It is clearly a fine piece of work with careful planning throughout. Great care was taken to ensure comparability with respect to questionnaires, samples, and procedures. Possible sources of confounding were also considered. For example, the same interviewers conducted both the face to face and telephone interviews and the order in which these

took place was switched for a random half of the sample.

The chapter also contains the startling results that the face to face survey had a significantly lower response rate than either the mail or the telephone survey. This point is discussed only briefly and the discussion appears four chapters later in the conclusions chapter.

The results of the different modes on various data quality indicators are given in Chapter 5. For example, the mail survey resulted in more item nonresponse, but interestingly, also in more self-disclosure on sensitive topics. Thus it appears that the physical presence of the interviewer, which is considered a great asset to an interview, may also be a drawback. As de Leeuw suggests, perhaps the social pressure to answer an interviewer produces less missing data while at the same time inhibits self-disclosure.

Also of interest is the finding that respondents in the mail survey had a tendency to report income more precisely (i.e., in guilders and cents) while in the face to face interview respondents were more likely to give qualified answers.

The author found some differences between modes with respect to the number of statements given by respondents to open questions. This occurred in the hypothesized direction with interview surveys providing longer answers than mail surveys. The author goes on to note that the lack of stronger findings could in part be due to the "blandness" of the open questions which were used in the study.

There were no clear mode differences detected for acquiescence, but a small recency effect was found. Also, respondents in the telephone condition had a tendency to choose the extreme positive answer more often than respondents in the mail condition. Importantly, de Leeuw found no consistent differences between the telephone and the face to face surveys.

Chapter 6 shows the surprising finding that both the internal consistency and scalability of items tend to be higher in mail surveys. De Leeuw had hypothesized that this finding should be expected for the mail survey, because it is in mail surveys that respondents experience the least time pressure and have the greatest opportunity to relate their answers from one question to their answers to other questions. (This second hypothesis is not completely convincing. The opportunity to relate questions to each other might lead to greater consistency, but not necessarily to a greater proportion of true score variance.)

In Chapter 7, de Leeuw investigates the effect of the differences between modes in the multivariate context with both a path model and a factor analytic model. The results suggest that all models show the same pattern and the same dimensionality (i.e., the same structure) despite the mode of data collection. In addition, similar conclusions are reached under each mode with respect to whether or not a particular variable influences another one. The author goes on to warn, however, that the relative importance of some estimated parameter values varied considerably across data collection methods, which could lead to different conclusions concerning the importance and strength of the influence of one variable on another. This warning should perhaps be tempered somewhat as I suspect that one might see a similar magnitude of differences with different samples under the same mode.

All in all, this experiment makes a fine contribution to the literature and this book is definitely worth acquiring. The topic of this book is highly relevant as choice of data collection mode is strongly linked with cost and therefore the sample

size which can be afforded, with telephone surveys being generally more expensive per case than mail surveys and face to face surveys being generally more expensive per case than telephone surveys.

References

Cannell, C.F., Miller, P.V., and Oksenberg, L. (1981). Research on Interviewing Techniques. In Sociological Methodology, ed. S. Leinhardt, San Francisco: Jossey-Bass.

Durbin, J. and Stuart, A. (1951). Difference in Response Rates of Experienced and Inexperienced Interviewers. Journal of the American Statistical Association, 114, 163–195.

Fowler, F.J. Jr. and Mangione, T.W. (1984). Standardized Survey Interviewing. Paper presented at the Joint Statistical Meetings of the American Statistical Association, Philadelphia, PA.

Long, J.S. (1983). Covariance Structure Models: An Introduction to LISREL. Beverly Hills, CA: SAGE Publications, Inc.

Salo, M.T., Campanelli, P.C., and Martin, E.A. (1990). Survey vs. Ethnographic Interviewing of Hard-to-Reach Populations. Paper presented at the International Conference on Measurement Errors in Surveys, Tucson, Arizona.

Schuman, H. and Converse, J.M. (1971). The Effects of Black and White Interviewers on Black Responses in 1968. Public Opinion Quarterly, 35, 44–68.

Sudman, S., Bradburn, N.M., Blair, E., and Stocking, C. (1977). Modest Expectations: The Effects of Interviewers' Prior Expectations on Responses. Sociological Methodology and Research, 6, 171–182.

Pamela C. Campanelli
Joint Centre for Survey Methods
SCPR, London, U.K.

Foreman, E.K., Survey Sampling Principles. Marcel Dekker Inc., New York, 1991. ISBN 0–8–247–8407–3, viii+497pp., $126.00.

Having devoted a number of years to the practice of both business and social survey methods, I read this book with great interest. It is a book that is rich in the language of the why's, the what's and the how's of survey sampling. This knowledge can only be gained over many years, and Foreman's considerable experience shines through.

There are fifteen chapters as outlined below:

1. Introduction
2. Some basic concepts
3. Simple random sampling and unbiased estimates
4. Ratio and regression estimates and estimates of ratios
5. Stratified sampling
6. Selection with probability proportional to size
7. Multistage sampling: Concept
8. Multistage sampling: Application
9. Further sampling techniques
10. Control of nonsample error
11. Sample design
12. Sample frames, selection systems, and master samples
13. Household survey master samples
14. Survey design
15. Design of series of surveys

The book can be regrouped into three main parts that will be labelled:

I Introduction to Survey Sampling (Chapters 1 to 8)
II Specialized Topics (Chapters 9 and 10)
III Sample Design Practices (Chapters 11 to 15)

Part I covers the standard introductory material found in many textbooks on sampling. The reader is taken from simple

random sampling to multistage sampling. The examples provided in these chapters are informative and easy to follow. Chapter 2 displays well the meaning of sampling and expectation. Concepts such as bias, variance, confidence intervals and mean squared error are carefully explained via the uses of tables and figures. Chapter 3 provides an excellent overview of simple random sampling and unbiased estimation. The main differences between sampling with and without replacement are clearly explained, with respect to the resulting effect on variance estimation. The notion of subgroup (or domain) is introduced in Chapter 3. The author confines himself to the estimation of proportions only. However, I would have liked to have seen more elaboration on subgroup estimation. Understanding subgroup estimation is very important for practising survey statisticians. It is not well understood by beginners, and a chapter dedicated to this topic would have been good.

Chapter 4 discusses two estimators that use auxiliary information: the ratio and the regression estimators. Note that the use of auxiliary information can be cast in a general context via regression models. The discussion on regression estimation will undoubtedly become more important in survey sampling as it is easily amenable to computer programming. Most of the known and widely used estimators in survey sampling, such as the ratio, regression, post-stratified and raking ratio estimators, can easily be derived in this context. The measurement of change is straightforward. But, the computation of its associated reliability is more difficult because one has to deal with population and samples changing in time (births, deaths, rotations). Readers should be cautioned in directly applying the variances of ratios given in this chapter to measure change reliability. Tam (1984) and Laniel (1987) provide a detailed explanation of how to deal with this problem. Benchmarked estimates, for ratio estimation, should be used cautiously for measuring both level and change; significant birth effects cannot be incorporated into such estimation procedures. The term "benchmark" is also introduced in this chapter. More generally, it should be noted that the term benchmarking is used when sub-annual surveys estimates are calibrated to a sequence of known annual estimates. It is assumed that the latter is the more precise measure of level.

Chapter 5 confines itself to ratio estimation, different types of allocations and the design of a stratified sample. Standard allocation procedures, such as N-proportional, X-proportional and Neyman allocations, are reviewed. Their advantages and disadvantages are discussed. I found the section on special-purpose allocations to be interesting. These include the square root allocation, as introduced by Carrol (1970). This type of allocation plays a role in deciding the relative importance of overall and subgroup estimates. Bankier (1988) provides a good detailed discussion of this topic under the heading of power allocations. The tool chest of stratification is not complete without partitioning populations on size measures. The inclusion of this topic would be greatly appreciated in a future edition of the book. For mildly skewed populations, Dalenius and Hodges (1959) is the standard reference. For very skewed populations, the existence of a take-all stratum is paramount to the reduction of the variance of the estimates. Techniques for delineating populations into certainty and non-certainty strata have been proposed by Glasser (1962), and Lavallée and Hidiroglou (1988).

Chapter 6 provides a good description of

the mechanics involved in drawing an unequal probability sample. There are hints of some of the problems encountered in drawing proportional to size and without replacement. These include the necessity to have "working probabilities" to take into account the realization of previous draws. The reader interested in finding out more about this difficult subject is referred to Brewer and Hanif (1983). Inclusion probabilities (π_i) are described, but joint inclusion probabilities (π_{ij}) are not even mentioned. The π_{ij}'s enter the computation of the variances in without replacement sampling.

Chapters 7 and 8 describe well the principles on which multistage sampling is based. The step by step description of how to design a two-stage sample introduces the many problems encountered in multistage sample design. These include cost considerations, balance in terms of cost and variance between the different stages, and the choice of the ultimate sampling unit. I would have liked to have seen a little more on the general theory for the computation of components of variance for the multistage sampling designs. This topic is well treated in Raj (1966) and Rao (1975).

Part II (the specialized topics) covers further sampling techniques and control of nonsample error. Chapter 9 includes a kaleidoscope of modern sampling topics. These include: two-phase sampling, repeated surveys and sample rotation, estimating medians and percentiles, raking ratio estimation, estimation using multiple and curvilinear regression, and synthetic estimation. There is a wealth of topics in this chapter. Essentials are highlighted without going into detail. Again, I wish that more modern and important papers would have been cited to provide the reader with a more up-to-date list of references. This chapter could easily be expanded into a book. Chapter 10 itemizes well the different types of bias and nonsample errors encountered in survey practice. Various techniques for handling nonresponse in one time as well as repeated surveys are described. The techniques provided in this chapter form a compact package of what nonsample error is and how the survey sampler should handle these errors.

Part III (sample design practices) unlocks the secrets of the survey sampler trade. Chapter 11 provides a good overview of sample design processes. These processes result as a balance of requirements, precision and costs. This chapter offers the bread and butter techniques for actual survey practices as applied in large and small organizations. The important use of design effects for designing multistage surveys is well explained. Chapter 12 deals with frame construction and sample selection. It contains some good advice on how to build and maintain frames for both household and business surveys. The important topic of survey coverage rules used to identify and delineate target population units is introduced. This topic is not discussed in any other sampling book to this depth. The difficulties encountered in designing a sample selection system for repeated surveys are also introduced in this chapter. These topics, however, are still the tip of the iceberg for carrying out such surveys. They have the potential to be elaborated in greater depth.

The design and maintenance of a household survey master sample is provided in Chapter 13. Steps for selecting a multistage sample are realistically illustrated with the aid of actual survey control forms. The handling of frame problems, such as births, deaths and size changes, is outlined. Chapter 14 discusses the principal procedures to be used in survey design. These include a good understanding of the

client's needs, translation of these needs into a plan of action so that the resulting survey is cost effective, efficient and timely. The importance of pilot testing to help finalize the survey design strategy is also stressed. Chapter 15 offers a good description of the required machinery for sustaining a series of surveys. The importance of well trained staff, good methodology and supporting systems is underlined. The chapter concludes with a design strategy for series of surveys. The big hint to accomplish this task is to develop generalized systems based on common statistical standards. This approach is currently being put in practice at Statistics Canada.

The book contains few typographical errors and is well written. It will probably be most useful to readers with some background in survey practice. Many principles mentioned in the book can only be learnt through practice. These principles will re-enforce good survey practice, which is gained by trial and error. If the book is to be used as a classroom text, many more recent references should be added in a future edition. Also, since it is not theoretical, a student of survey sampling should consult other books to develop the statistical backbone behind this subject.

Michael A. Hidiroglou
Business Survey Methods Division
Statistics Canada
Ottawa, Ontario, Canada

References

Bankier, M.D. (1988). Power Allocations: Determining Sample Sizes for Sub National Areas. The American Statistician, 42, 174–177.

Brewer, K.R.W. and Hanif, M. (1983). Sampling with Unequal Probabilities. New York: Springer-Verlag.

Carrol, J. (1970). Allocation of a Sample Between States. Technical Report, Australian Bureau of Census and Statistics.

Dalenius, T. and Hodges, J.L., Jr. (1959). Minimum Variance Stratification. Journal of the American Statistical Association, 54, 88–101.

Glasser, G.J. (1962). On the Complete Coverage of Large Units in a Statistical Study. Review of the International Statistical Institute, 30, 28–32.

Laniel, N. (1987). Variances for a Rotating Sample from a Changing Population. Proceedings of the Business and Economic Statistics Section, American Statistical Association, 246–250.

Lavallée, P. and Hidiroglou, M.A. (1988). On the Stratification of Skewed Populations. Survey Methodology, 14, 33–43.

Raj, D. (1966). Some Remarks on a Simple Procedure of Sampling Without Replacement. Journal of the American Statistical Association, 61, 393–397.

Rao, J.N.K. (1975). Unbiased Variance Estimation in Multi-stage Design. Sankya, Ser. C, 133–139.

Tam, S.M. (1984). On Covariances from Overlapping Samples. The American Statistician, 38, 288–292.

Lessler, J.T. and Kalsbeek, W.D., Nonsampling Error in Surveys. John Wiley & Sons, New York, 1992. ISBN 1–86908–2, 432pp., $69.95.

A characteristic of survey methods as a field of study is that it has been developed and used by practitioners and researchers with many backgrounds. The literature is vast and appears in a great variety of journals, conference proceedings, and government and technical report series, many of which are difficult to access. The authors of this book have done an excellent job of pulling together this literature and organizing it in

a way that allows the reader to compare similar methods and recognize alternative solutions to design problems.

The authors classify every nonsampling survey error as arising from one of the following sources: frame problems, nonresponse, and measurement. Most of the book is organized around these classifications. Chapter 1 gives a brief history of surveys. Chapter 2 describes each survey task in roughly the order it is carried out, followed by a list of its potential problems and possible effects on data quality. The tasks are quite detailed (e.g., budget consideration, preparation of coding instructions, computation of sampling weights), but the authors state that their intent is to "describe what and where things can go wrong in a survey" and not in "describing in detail how one goes about doing a survey."

Chapters 3 through 5 address errors from frame problems. Chapter 3 discusses various frame structures and classifies frame problems into one of six types: missing population elements, inclusion of nonpopulation elements, multiplicity problems, failure to account for clustering, incorrect auxiliary information, and incorrect accessing information. The effect (i.e., bias or MSE) of these problems on estimates of means and totals is discussed in Chapter 4. Methods for identifying the presence of the problem, or estimating its seriousness, are given. Chapter 5 describes methods for conducting surveys using imperfect frames. It includes a clear discussion of the theory of multiple frames and of methods for eliminating or adjusting for frame multiplicity.

Chapters 6 through 8 cover nonresponse. Chapter 6 includes a categorization of reasons for nonresponse, along with their frequency of occurrence in some example surveys (telephone, personal visit, mail). Chapter 7 discusses the effect of nonresponse on estimators of means, totals, variances, covariances, and regression coefficients. The difference in modeling nonresponse as stochastic (in which $0 < p_i < 1$ where $p_i =$ the probability that unit i responds) or deterministic (p_i is either 0 or 1) is discussed at length. Chapter 8 covers methods for dealing with nonresponse, including both prevention and adjustment strategies for unit and item nonresponse. This book is the only published source I know for details on use of weighting class and post-stratification adjustment. Numerical examples would have been helpful in explaining the differences in many of the similar-appearing weighting methods. Bayesian and other model-based methods and many imputation methods for item nonresponse (e.g., hot-deck, distance function matching, regression methods) are covered, but mostly qualitatively.

Chapter 9 focuses on the nature of measurement. Chapter 10 examines models for respondent bias in both numerical and categorical data, and the effect of bias on estimators of means and totals. It describes methods that have been used to study specific types of respondent bias, such as telescoping. It thoroughly covers methods for measuring bias using double sampling. Chapter 11 compares various models for variability of measurement errors, uses them to show the effect of errors on means and totals, and discusses experimental methods (including reinterview and interpenetration) for estimating the models.

Finally, two very general models for total survey error are given in Chapter 12. This is followed by a compendium of nonsampling error terminology, whose purpose is "to indicate how the terminology linked to various sources of survey error are defined by individuals and groups that have contributed to the research literature."

There is a focus on terminology throughout the book.

The authors suggest that the book is designed both for new and advanced students of survey methods, and as a reference guide for practitioners. The book contains wide variation in the technical level of its presentation. The last two chapters would probably be difficult reading for any but advanced graduate students and researchers in survey methods. However, the sections on frame error and nonresponse are quite readable and contain much material that would be an excellent supplement to a text such as Cochran for sampling students. These sections also contain much practical advice, as well as excellent references, for the practitioner who is interested in finding a solution to a design problem in a mail, telephone, or personal interview survey. The book would also be useful for helping the researchers in survey methods to recognize gaps in the literature.

My major objection to the book is that developments in the literature beyond about 1985 are not covered well. In fact, in a (systematic) sample of 100 references from the 22 pages of references, only 19% were from the years 1982–87 and none were from 1988 or beyond. There have been many contributions to the literature in that time. For example, one area that I am familiar with are models for interviewer effects for categorical variables. The omission of this topic seems odd since models for categorical and numerical variables are covered separately in the measurement error section. A reason for this is probably that one of authors' goals was to historically trace the major contributions to the nonsampling error literature. The historical material may have crowded out newer methods. Perhaps a sequel is needed already.

S. Lynne Stokes
Management Science and
Information Systems
University of Texas
Austin, TX, U.S.A.