

Book Reviews

Books for review are to be sent to the Book Review Editor Jan Wretman, Statistical Research Unit, Statistics Sweden, S–115 81 Stockholm, Sweden.

BRORSSON, B. and WALL, S., Assessment of Medical Technology: Problems and Methods <i>Ronald G. Marks</i>	209	LINDER, F.E. and MORIYAMA, I.M., Improving Civil Registration <i>Jack Harewood</i>	214
CONVERSE, J.M. and PRESSER, S., Survey Ques- tions: Handcrafting the Standardized Question- naire <i>Kent H. Marquis</i>	211	Recent Development in the Analysis of Large- Scale Data Sets <i>Steven B. Cohen</i>	217
JAMBU, M., and LEBEAUX, M.-O., Cluster Analysis and Data Analysis <i>Ove Frank</i>	212	YSANDER, B.-C., (Ed.), Two Models of an Open Economy <i>Anders H. Westlund</i>	218

Brorsson, B. and Wall, S., Assessment of Medical Technology: Problems and Methods. Swedish Medical Research Council, Stockholm, 1985, ISBN 91-85546-36-4, 129 pp.

The authors state that the purpose of their book is “to contribute to the advancement of knowledge concerning: the ethical, institutional, and economic conditions under which testing and assessment of medical technology occurs, the available arsenal of methodology, and how the choice of outcome measures and research design determines which conclusions can be drawn and their practical value.” Although this is a broad, sweeping objective, the authors do an excellent job of presenting information in a clear and concise manner that should benefit medical administrators, educators, and consumers of research, as well as provide valuable knowledge for people who will be engaged in research design and analysis for assessing medical technology. Throughout the book, the authors make liberal references to practical statements from published research which support the points they wish to

emphasize. One aspect of this book that I found particularly important is the constant weighing of scientific versus practical considerations in biomedical research, which certainly must occur in all research projects and cannot be overlooked. There are four major areas covered in this book: the criteria to be used in assessing medical technology, the validity and deduction process, research design concepts, and synthesis methodology.

In the discussion of criteria for assessing medical technology, the authors present an overview for assessing the safety and efficacy of medical technology. They very nicely contrast the need for carefully designed and well-documented research with the problems inherent in human research due to the lack of control over the environmental factors that can oftentimes induce additional uncontrolled variability in the resulting data. They also present an eloquent discussion of a major problem in biomedical research – the quest for immediate knowledge and evaluation of technology that, unfortunately, sometimes requires years to adequately follow-up and

evaluate. In this environment, we run a great risk of modifying technology before a proper evaluation is complete. The result can be that a true evaluation may never be obtained for a specific technology.

I was happy to see that in the section for criteria, the authors dedicate eight pages to the formulation of the research question and the determination of appropriate measures of safety and efficacy and nine pages to ethical considerations in evaluating new technology. There is a long presentation of the concerns for conducting quality research while also providing the participant with quality medical care. Numerous references from the appropriate literature are cited. The need for informed consent, especially in the area of risks and benefits, is stated and contrasted with the need for justice.

The second major area of discussion is validity and deduction that involve the analysis and interpretation of research results. The authors discuss the concept of variability and the need to measure it. They discuss various measures of validity and the sensitivity and specificity of diagnostic tests. They point out the different levels of measurement as nominal, ordinal, interval, and ratio and that the choice of statistical analysis and assumptions depends on the level of the measurement. They also provide a clear description of the principles of statistical deduction including the elements of hypothesis testing, the meaning of statistical significance, and the interpretation of statistical results. The points made in this chapter are important for the analysis and interpretation of research results. (More detail can be found on these topics in Haack (1979) and Marks (1982a, b).)

Although the third major section of this book deals with research design, I believe the second section on validity and deduction should also be included in design since it deals with issues that must be resolved during the planning stage of the study and before any data is collected. The section on research design provides an informative comparison of randomized and nonrandomized experiments and nonexperimental studies which are often referred to as observational studies. Issues of cost, the amount of information obtained, and identification and control of potential biases are contrasted for the different types of studies.

The fourth subject dealt with is synthesis methodology which is defined by the authors as "the compilation and evaluation of available knowledge." The purpose of synthesis is "to establish the direction for research" and "to extract clinically applicable information from a wealth of different studies." This is the most philosophical part of the book. The reader is given many points to consider on how one should synthesize the available information on research and experience to arrive at a final decision on the acceptability of a given technology. As in the other areas, the concept of synthesis methodology involves weighing scientific and practical concepts.

Overall, I believe this is an excellent book for both active biomedical researchers and consumers of research who are interested in learning more about the process of logical and critical thinking. To its credit, the book contains virtually no mathematical concepts, which I believe are not necessary for one to understand the research process. The points made in this book closely parallel my own approach to teaching students and researchers; see Marks (1982a, b). Do not, however, think you will get all the answers from this book. The authors do not provide the level of detail needed to conduct the research introduced in this book. Yet, the authors do provide many references where more information about each of the concepts is available. I recommend this book as an excellent introduction to research design methodology. It provides a solid foundation which you can build on as needed.

References

- Haack, D.G. (1979): *Statistical Literacy: A Guide to Interpretation*. Duxbury Press, North Scituate, MA.
- Marks, R.G. (1982a): *Designing a Research Project: The Basics of Biomedical Research Methodology*. Van Nostrand Reinhold, New York.
- Marks, R.G. (1982b): *Analyzing Research Data: The Basics of Biomedical Research Methodology*. Van Nostrand Reinhold, New York.

*Ronald G. Marks
University of Florida
Gainesville, FL
U.S.A.*

Converse, J.M. and Presser, S., *Survey Questions: Handcrafting the Standardized Questionnaire*. Sage University Paper Series on Quantitative Applications in the Social Sciences, 070-63. Sage Publications, Beverly Hills, 1986, ISBN 0-8039-2743-6, 80 pp., \$6.00.

This small publication is easily read and will give the reader some different perspectives on developing an interview questionnaire. Both authors represent what might be called the "Michigan" point-of-view that once dominated academic survey research in the United States and which has survived through several generations of professional "question crafters" at the Survey Research Center.

As a reviewer, I find it difficult to identify an audience for the monograph. The monograph is not a scholarly work about theories. Nor does it deal with the profound and complex issues of a goal-oriented, total survey design. It is not a comprehensive review of the literature on measurement or questionnaire development. Nor is it an instruction manual on how to handcraft a questionnaire. The editor's comment notwithstanding, this is not a book in Stanley Payne's tradition of art guided by experiment. From a methodological perspective, the book presents no new data, no replications of prior work, and no new thinking about the existing body of knowledge.

Rather, this is a description of the concerns and advice that people with day-to-day responsibility for operating surveys might give to those who are trying to design their first questionnaire. The advice is clearly meant to be helpful and reassuring although perhaps not optimal in a statistical sense.

The monograph is in three chapters. Chapter 1 starts rather unexpectedly by portraying the stereotype of the highly educated, academic client whose first efforts produce an unrealistically arcane questionnaire. The chapter is about some simplification strategies based on ordinary common sense.

The chapter tells us that the professional question crafter must simplify the document by shortening the items, avoiding confusing English-language constructions such as double negatives, avoiding technical jargon, and defining ambiguous terms. And, wherever possible, we should avoid questions that require recall over long periods or questions that are of a hypothetical nature. In

other words, some questions are hard to understand or difficult to answer; it is the question crafter's job to eliminate the difficulties.

But what happens when the research evidence runs counter to the conventional wisdom? There is one example early in Chapter 1. The evidence cited clearly indicates that long, redundant questions are better than short, succinct ones! The authors say that they do not know why and advise us to ignore the research lest our questionnaires become "bloated with too few, too fat questions." Although the logic seems faulty, I am more concerned that they failed to mention a critical, oft-replicated finding of the question length research, that the beneficial effects are achieved only among the more highly educated segments of the U.S. population.

In Chapter 2, the authors derive general principles of question writing from the published research literature. The subheadings convey the main messages (e.g., "Specific Questions Are Better Than General Ones," "Offer a No Opinion Option," and "Use Forced-Choice Questions, Not Agree-Disagree Statements"). There is a discussion of the unpredictable effects of question wording and order. The authors offer no simple prescriptive writing strategies here, instead they suggest design strategies (split-panel experiments and random probes) that can make inferences more trustworthy.

I could not help feeling that this section was too superficial and consequently sometimes misleading. For example, the discussion of question specificity rests almost solely on a 1962 article by Belson and Duncan. While the article offers support for the conclusion, one of the main strengths of that article is that it provides concrete evidence about two negative consequences of being specific: (1) that items (magazines) not on a list of cues have an even lower probability of being mentioned than in response to a general question and (2) that people endorse "trap" items only when they are specifically listed on a questionnaire.

At the end of the section, the authors make a major conceptual and strategic shift. Instead of prescribing how to write questions, they recommend protective strategies for becoming aware of interpretation problems that respondents may have. But this set of recommendations was a real disappointment to this reviewer. Obviously, the authors were

approaching the important concept that questionnaire designers can obtain protection against the idiosyncratic problems of single items by using robust designs.

The essence of robust design strategies in measurement is not to ask questions about questions as the authors recommend. It is to ask several questions, in several different ways, about the same thing. Then, using well-known statistical techniques, extract what is common in these answers and use it as a basis for assigning a numeric value for the respondent on the dimension (e.g., attitude or opinion) of interest. The authors explicitly reject this approach mainly on the unsupported generalization that the item intercorrelations might change over time.

The third chapter is about pretesting questionnaires, a practice which the authors recommend highly. They develop ten criteria on which the pretest should be evaluated and sketch the rules for each of two required pretests (the developmental pretest and the dress rehearsal pretest). They point out six sources of information (presumably information about the ten evaluation dimensions although the link is not explicit).

But this somehow misses the client's central concerns about measurement such as the quality of the data (reliability and validity/problems with nonresponse), the costs of acquiring data of adequate quality, and perhaps the timeliness of the estimates based on processed questionnaire data. The recommended pretest strategy strikes me as a ritual whose rationale disappeared long ago. Today's questionnaire design concerns are addressed in much different ways than the 50-case pretest that was appropriate when the cross-section survey was a new and untried method.

The practical nature of the authors' concerns leads them to warn that the client is going to get new insights from the pretest and will want to add some new questions to the final questionnaire. They recommend erecting high barriers to such frivolous requests (e.g., extra pretesting requirements to be carried by whomever proposes a new item without help from anyone else).

The book provides a useful list of nine compilations of survey questions for those interested either in borrowing questions or in learning how well selected measures have fared in the past.

If you already have some experience designing questionnaires or in running household interview surveys, read the book and let it challenge your own assumptions, the folklore you use, and the rituals that you might be following.

Kent H. Marquis
U.S. Bureau of the Census
Washington, DC
U.S.A.

Jambu, M. and Lebeaux, M.-O., Cluster Analysis and Data Analysis. North-Holland Publishing Company, Amsterdam, 1983, ISBN 0-444-86634-5, xxiv + 898 pp., \$85.00.

This book contains over 900 pages and consists of two parts. Part 1 (by M. Jambu) is a comprehensive account of concepts, methods, and algorithms used in classification. Part 2 (by M. Jambu and M.-O. Lebeaux) lists 40 Fortran programs based on 356 subroutines for data editing, analysis, and interpretation. Part 2 occupies more than 70 per cent of the book. The following abridged table of contents provides an idea of the wealth of material that is covered and of the special approach taken here.

<i>Part I</i>	<i>Methods and Algorithms</i>
Chapter I	Automatic Classification and Data Analysis
Chapter II	Practice of Automatic Classification and Factor Analysis of Correspondence Tables
Chapter III	Definitions and Properties Concerning Hierarchical Classifications
Chapter IV	Distances and Proximities in Automatic Classification
Chapter V	Classical and Fast Algorithms for Hierarchical Classifications
Chapter VI	Completion Algorithms for Hierarchical Classifications
Chapter VII	Algorithms for Determining Partitions and Overlapping Clusters
Chapter VIII	Interpretational Aids and Validity in Classification

Part 2 *Computer Programs and User's Guide*

Chapter 1 Programs for Preparation and Description of Data Tables

Doubling a table of scores. Putting questionnaire replies in complete disjunctive form. Recoding a set of heterogeneous data. Usual description of a correspondence table. Multiple correspondence tables and adjacent tables. Usual description of a set of continuous variables. Histograms and elementary statistics. Rearranging rows and columns of a table according to their marginal distributions. Transformation of a cloud of points provided with masses. Choice of weightings for a heterogeneous data set composed of homogeneous subsets for subsequent correspondence analysis. Transposition of a data table. Calculating distances from a table of logical data. Calculating distances from a data table.

Chapter II Programs for Data Analysis and Classification

Correspondence factor analysis. Principal components analysis. Factor analysis of a finite set provided with masses and distances. Correspondence factor analysis using a division into subtables. Ascending hierarchical classification of a finite set provided with masses and distances. Ascending hierarchical classification of a set of variables using a distance between subtables. Accelerated ascending hierarchical classification of a finite set provided with masses and Euclidean coordinates. Accelerated ascending hierarchical classification using a criterion based on information theory. Ascending hierarchical classification based on a criterion of preorder. Tree representation associated with a hierarchical classification. Method of partitioning by the 'NUEES DYNAMIQUES' algorithm. Method of partitioning by clustering around moving centres with radius restrictions. Regression methods using factor analysis and neighborhood definitions.

Chapter III Programs of Interpretational Aids

Complementary analysis of results obtained by correspondence factor analysis. Complementary analysis of results obtained by principal

components analysis. Representation of the classes of a hierarchical classification and their ellipses of inertia in a factorial space. Representation of the classes of a partition and their ellipses of inertia in a factorial space. Complementary graphical representations of a hierarchical classification. Rearranging rows and columns of a table as a function of the ranks associated with the factors. Histograms and correlation matrices for each class of a hierarchical classification. Calculation of mutual contributions between classes of a hierarchy and factors of a factor analysis. Calculation of mutual contributions between classes of a hierarchy and a correspondence table. Calculation of mutual contributions between variables and classes of a hierarchy. Calculation of mutual contributions between classes of a partition and factors of a factor analysis. Completion of a hierarchical classification according to the metric of inertia of classes. Simulation study of the level measures of hierarchical classifications.

Most English texts do not take this approach to cluster analysis and classification. The French school of data analysts put emphasis on the need to distinguish between different kinds of data by using good data type systematics and efficient coding procedures. Data-editing is given appropriate attention, and it is made clear how important this is to the subsequent choice of data analysis methods.

However, the technical quality of the book is not good. Photocopying from typewritten pages is particularly hard to appreciate in a book of this size. The lack of a good typography makes it almost impossible to get a clear overall picture of how the material is structured. It is not easy to identify which sections are of general importance and to distinguish them from sections with specific and illustrative material. This makes it difficult to use the volume as a quick reference. The organization into sections is complicated. For instance, 3.1.3.9 labels the last Subsubsubsection of Subsubsection 3.1.3 of Subsection 3.1 of Section 3 of Chapter V of Part 1 of the book. Some of these sections are given in the list of contents, but not all. If a selection principle has been applied, it is not clear what it was. For instance, Chapter V of Part 1 has Subsubsection 4.4.2 but not Subsubsection 3.1.3 listed in the Contents. In Part 1, chapters are labeled by roman numerals and their sections by arabic numerals, whereas in Part 2,

chapters are divided into subchapters labeled by arabic numerals and their sections are labeled by roman numerals. A book on classification should not have such a confusing classification of its topics. The interesting approach to classification presented in this book deserves a layout and typography of higher quality so that its systematics would appear clearly and beneficially.

There are numerous misprints, omitted words, wrong paragraph separations, inconsistent notation, and so forth. References appear in the wrong order. The translation from French to English has left "et" in some of the formulae. To some extent such mistakes are inevitable in a long text, but here the proof-reading has been extremely poor.

There is a strong need for a revised and more readable edition of this important and useful text.

*Ove Frank
University of Stockholm
Stockholm
Sweden*

Linder, F.E. and Moriyama, I.M. (Eds.),
Improving Civil Registration. International
Institute for Vital Registration and Statistics,
Bethesda, MD, iv + 242 pp.

As is evident from its title, this book examines and recommends steps to improve the quality and completeness of civil (or vital) registration in developing countries whose registration is deficient. It is concerned with the major purposes of registration: (a) to establish the civil status of individuals, and (b) to provide the base data for the national vital statistics.

The book is published by the International Institute for Vital Registration and Statistics (IIVRS), a nongovernmental, international organization, established in 1974. Indeed, apart from a short introduction, the book consists of 10 of the 27 papers previously published by IIVRS. They have been republished in book form because it is believed that they "are a useful force for improving these (civil registration and vital statistics) systems, and deserve a wider distribution."

The introduction begins with a short history of both civil registration and the actions of international agencies. There is also information on IIVRS, including its objectives and an enumeration of its publications. The last section is a brief description of the present status of civil registration and vital statistics with recommendations for their improvement (with references to where each topic can be found in the book).

It would have been useful to have had a more specific statement in the introduction on the purpose of the book and its intended audience. Nor are the various papers effectively tied together, and the book does not function as an integrated study.

It would also have been useful to group the papers, excluding the first, under a few major topics. For example, without much change in their current order, the following groups could have been used. (A) The Importance of Civil Registration and Vital Statistics (Papers 2, 3, and 9); (B) Obstacles to Improvement in Civil Registration and Vital Statistics (Papers 7 and 8); and (C) Recommendations for Improving Civil Registration and Vital Statistics (Papers 4, 5, 6, and 10).

The first and most recently published paper is on "Actions for the Improvement of Civil Registration and Vital Statistics." It covers five broad topics: (A) A General Statement of the Problem; (B) Analysis of the Problem; (C) International Activities; (D) National Programs; and (E) Plan of Action. It is thus a well-rounded paper on its own, but since it is not written as a review paper, it duplicates large sections taken from other papers in the volume. This emphasizes that the book is a collection of separate papers with little attempt to integrate the material. However, the paper does, particularly in the last two sections, offer useful guidance on the steps that developing countries can take to attain an efficient system of civil registration and vital statistics.

The paper by the Committee on Legal and Organizational Requirements for a Civil Registration System in Latin America (Paper 2) was specifically prepared for high government officials, particularly those in charge of national development policy in Latin America. To this end it needed to be simple and concise, and judging from the English translation it has achieved this objective. The

main topics covered are: (i) The Doctrine, (ii) The Utility and (iii) The Organization of Civil Registration, plus Inter-Agency Coordination and How to Approach the Process of Change.

The paper excellently highlights the importance of civil registration and vital statistics for many administrative purposes, including social security and military draft. But in stressing the importance of vital statistics, the paper gives the impression that population censuses are of limited value because they "may only furnish a static statistical picture" and therefore do not meet the "demands of sociologists and demographers and are of practically no use to public health specialists." This is unfortunate and misguided. Censuses and vital statistics are complementary. Surely the most complete and up-to-date vital statistics would be largely useless without the base population data obtained from censuses.

The paper goes on to discuss fully the shortcomings in the existing legislation and organization in Latin America and makes an effective argument for serious and high-level action. But the proposal to involve top officials in an ad hoc commission at the presidential level seems unlikely to succeed. High-level staff have many commitments and this approach often fails, as discussed in this paper.

At the other extreme, the importance given to local registrars should be commended, since the most elaborate plans can falter if attention is not given to the people who perform the civil registration.

The first paper by N. Powell (Paper 3) gives an excellent and comprehensive review of the social indicators derived from vital statistics in accordance with the United Nations recommendations. The author points out that as many as 36 of the 149 recommended social indicators are derived from vital statistics.

The final section of this paper discusses the unavailability of reliable vital statistics in developing countries and the various UN recommendations for improvement. There is a useful annex of the social indicators from vital statistics.

In another paper (Paper 9) N. Powell presents an impressive argument for the importance of vital registration with reference to

human rights. This is, of course, a relatively new "purpose" of vital registration, but one which will increase in importance. The author and IIVRS are to be commended for the attention paid to this aspect of vital registration.

Several papers deal with the obstacles to an improved civil registration and vital statistics. Paper 8 by N. Powell is a major contribution. It is a report on an enquiry by IIVRS, conducted in October 1979. Officials from developing countries discuss civil registration and/or vital statistics in their own countries. The reported obstacles fall into seven categories: (a) *funding* – all but one country reported inadequate funding as a major obstacle; (b) *the legal basis of the system* – the obstacles here range from the complete absence of a registration law to the need for amending the existing law. The absence of a central registration authority is considered perhaps the most important legal deficiency, another is the need to extend the statutory period for registration, given the difficulties encountered in reaching the registrar; (c) *organization of the system* – lack of understanding on the part of high-level government officials, lack of coordination among the various agencies responsible for different registration functions and the lack of local registration services in remote rural areas; (d) *registration functions* – the difficulty of getting to registration offices especially in rural areas and the illiteracy of the population are examples; (e) *statistical functions* – the delay, on the part of local registration offices, in submitting the appropriate returns to the statistical authority, and the lack of adequate data processing facilities; (f) *staff and training* – shortage of qualified staff and inadequate training are serious obstacles. An important point not mentioned here, but noted in a later paper is the lack of supervision for local registration staff; (g) *the public* – ignorance of registration requirements, resistance to registration (e.g., of illegitimate births), and an unmotivated public are also obstacles.

Despite the above serious obstacles, the responses did give some indications of improvement in the completeness and quality of data over time.

The other paper in this group, (Paper 7) is the article by T. Makannah of the Economic

Commission for Africa, which deals with obstacles in the African countries. Most of the problems named above occur in many African countries where civil registration has been introduced only after attaining independence in the 1960s. This paper includes eleven useful tables which describe the present status of registration in African countries, including the date of institution of civil registration, registration coverage, and the contents of birth registration forms. These can prove useful to anyone who wants to improve the system in his or her own country.

Among the papers dealing with improving civil registration and vital statistics, that by F. Linder (Paper 4) on developing countries is short, but discusses a number of important issues. Linder raises the important point that it is not necessary for a civil registration system to be 100 per cent error proof for it to be an adequate source of information on civil status or vital statistics. He maintains that this criterion is not required for other statistical systems.

Linder proposes that the obstacles to improvement in civil registration be subdivided in the following way: (a) relatively intractable, e.g., problems relating to the geography of the country, and that cannot be quickly solved: in particular, the necessary changes are outside the power of the registration system; (b) solvable but requiring technical assistance; and (c) solvable by the country with only small financial outlays. For an international agency this sub-division appears meaningful. For a national agency, however, it may be useful to subdivide the third group. The first subgroup would be problems that can be quickly corrected by the agency (or agencies) responsible for vital registration and vital statistics, e.g., by improving the administrative system. The second subgroup would be those problems that require action by other agencies, often at a higher level in government, e.g., a change in legislation, and therefore must take more time.

Attention is drawn to the heterogeneity of patterns of national organization, and some of the resulting problems. Because this heterogeneity is seen as a major obstacle to successful international assistance, and because some patterns of organization must be more

efficient than others, it would have been useful if the author had recommended the ideal type of organization, as another paper does.

I. Moriyama (Paper 5) investigates the vital registration system as a source of data on mortality, and especially for differential mortality analysis. He begins with a brief history of vital registration, starting with the ecclesiastic registers. Next he discusses the "death registration process and the uses of both individual records and statistics derived from the death certificates." This includes some illuminating examples of the uses of these data for understanding and improving the health of a nation. Thus, it is not only an argument to administrators on the importance of registration, but also an indication to health technicians of what they can do with mortality data. He concludes with a discussion of the advantages and disadvantages of civil registration for the production of national mortality statistics, and states that despite its few disadvantages it "offers data possibilities that cannot be achieved in any other way" such as single-round and multi-round surveys or the relatively recent indirect estimation methods based on characteristics of the living population obtained in a census or a survey.

In his second paper (Paper 6) Moriyama has collected recommendations on civil registration and vital statistics from a number of regional seminars over the past 25 years. They are classified under 14 subject headings including: legal provisions, administrative organization, personnel, training, technical assistance, and others. In addition, he gives a brief summary of the actual recommendations. This could be extremely useful for officials in developing countries since these recommendations are unlikely to be easily available, if at all, to many who would benefit from them.

The last paper in the book also deals with recommendations for improving civil registration and vital statistics. It is a review of the recommendations made by the National Committees on Vital and Health Statistics to coordinate the statistical activities within each country. The recommendations were also intended to serve as a link between the national medical-statistical institutions and the Expert Committee on Health Statistics of WHO. The review summarizes the principal

recommendations of the two international conferences of National Committees in 1953 and 1973, and discusses the development of these committees around the world. It stresses the potential of the committees for advancing vital and health statistics where there is a strong commitment.

Although, as indicated earlier, the volume would have been improved by a more effective attempt to mold it into an integrated study, the separate papers nevertheless are, indeed, a "useful force" for improving civil registration and vital statistics systems in developing countries, and is to be highly recommended.

*Jack Harewood
University of the West Indies
St. Augustine
Trinidad, W.I.*

Recent Developments in the Analysis of Large-Scale Data Sets. Proceedings of a Seminar held in Luxembourg, 16-18.11. 1983. Office for Official Publications of the European Communities, Luxembourg, 1985, CA-AB-84-006-EN-C, 323 pp., \$9.00.

This volume consists of a collection of 14 papers which were presented at an international symposium, sponsored by the Statistical Office of the European Communities at Luxembourg in 1983. In general, the book takes an expansive view of the analytical role of the national statistical organizations responsible for large scale data collection and data dissemination. Considerable coverage is given to log-linear analytical techniques and to adjustment strategies for survey design complexities, which often characterize large scale data sets. Although the papers focus primarily on the analytical approaches which are specific to distinct economic and social surveys, the generalizability of study findings to related large scale data sets is most evident.

A general discussion of the role of modelling in official statistics is provided by J. Nelder. The desirability of parsimonious models is emphasized, to eliminate unnecessary complexities in the summarization of relationships which characterize a population. A set of related papers concentrate on mathematical modelling techniques to summarize data presented in multidimensional contingency tables. An application of log-linear modelling is provided by M. Aitkin and R. Healey to summarize information on unemployment in the EEC Labor Force Survey. The goodness of fit of an additive logit model specification, that considers the log odds of unemployment as a function of age group, industry, and region, is scrutinized. In addition, the model's stability across European countries is examined. A comprehensive series of related log-linear analyses are presented by H. Stronkhorst and J. Pannekoek, using data from two large scale school surveys conducted by the Netherlands Central Bureau of Statistics in 1964 and 1977. Their investigation identified relationships between student's school choice, teacher's recommended choice, scholastic achievement, socio-economic status and gender, in addition to detecting changes in associations over time.

Large scale data collection efforts are often characterized by survey design complexities which include unequal selection probabilities and related sampling weights, in addition to multi-stage clustered sample designs. Due to the departure from simple random sampling assumptions, special procedures are required for variance estimation and the analysis of complex survey data. D. Binder et al. provide an excellent summary of adjustment strategies which control for survey design effects in goodness of fit tests, tests of independence, and in log-linear and logistic regression analysis. The authors also focus on limitations in the availability of statistical software for complex survey data analysis and suggest alternative strategies. In a related paper, R. Fay discusses replication approaches to the log-linear analysis of complex survey data. In particular, a jackknife procedure is used for estimating standard errors of estimated model parameters, and a jackknifed chi-square test is considered for the determination of goodness of fit and the

contribution of specific model parameters.

Another set of related papers concentrate on specific economic analyses of income data. Using data from a survey of the structure of earnings in France, D. Depardieu and J. Payen consider a multiplicative model to examine the relationship of sex and professional qualifications on wages. A time series analysis of wage distributions in the United Kingdom is presented by A. Airth, which suggests that observed distributions are distortions from a latent distribution of constant form. The virtue of standard, easily interpretable statistical summarization techniques is emphasized in J. Bibby's presentation of the results from the British New Earnings Survey. In addition, a comprehensive series of exploratory data analyses are conducted by A.Z. Israëls et al. on earnings data, providing detailed examples of multivariate techniques appropriate for the analysis of discrete data.

The role of correspondence analysis as a technique for exploratory data analysis is discussed by L. Lebart. The method has great potential for revealing unanticipated relationships and generating new research hypotheses. J. Daudin et al. provide an illustration of the complementary nature of correspondence analysis and log-linear modelling in relation to the summarization and interpretation of survey data. The text also provides coverage of new developments in the field of information theory (G. Longo et al.) and includes a presentation of the adjustment of input-output tables as a function of partial information (A. Bachem et al.), and an examination of alternative measures of concentration within an industry (A. Jacquemin).

In summary, the book should serve as a good reference for statisticians and researchers responsible for the analysis of large scale data sets.

*Steven B. Cohen
National Center for
Health Services Research
Rockville, MD
U.S.A.*

Ysander, B.-C. (Ed.), *Two Models of an Open Economy.* Almquist & Wiksell International, Stockholm, 1986, ISBN 91-7204-265-6, 161 pp., SEK 250.00.

This book discusses two large-scale econometric models designed for analyzing various economic policy problems in an open and small economic system, such as the Swedish economy. In particular, the models may serve as instruments for studying the important problem of how the world market economy influences the structure of the Swedish economy.

The two models stem from very different theoretical and methodological modelling concepts. ISAC (Industrial Structure And Capital growth) is a disequilibrium model, basically specified in line with Keynesian economic theory. ELIAS (Energy, Labor, Investment, Allocation, and Substitution), on the other hand, represents an equilibrium macro-model. In addition to presenting the models, and discussing how each of them can be used for structural adjustment analysis, the book also compares and contrasts the two approaches with respect to methodology and empirical results.

The book consists of three parts. The first part, by the editor, B.-C. Ysander, is entitled "Structural Change as an Equilibrium or Disequilibrium Process – An Introduction." A brief history of Swedish economy, reflected by medium-term macroeconomic models, is first outlined. It is stated that during the 50s and 60s medium term policy focussed on the distribution of gains in industrial productivity, but that the focus of medium-term problems changed drastically during the 70s. Increased instability in price and exchange rate, partly caused by the oil price hikes, gave rise to dramatic swings in industrial production and investment activities in Sweden. This meant new challenges for macroeconomic modelling, challenges which are partly taken up by the two modelling exercises. Some general principles, or criteria, for comparing the models are also suggested. These principles are (i) realism, (ii) relevance, (iii) resilience, and (iv) robustness. With resilience is meant the ability of an economic system to absorb temporary shocks from outside and readjust to a path of stable growth. The problem of resilience is important in dis-

equilibrium as well as in equilibrium models. Robustness is another important aspect, i.e., structural stability should be related to the parameters of the models. How sensitive the model performance is to parameter changes is further discussed when evaluating ISAC. Whether parameters change or are constant is, however, not analyzed at all.

In "ELIAS – A Model of Multisectoral Economic Growth in a Small Open Economy" by L. Bergman, a six-sector multiperiod equilibrium model is described. It incorporates price and wage formation, and explicit links with the world market. Furthermore, it determines consumer demand by linear expenditure functions. The interest in ELIAS is focussed on analyzing the mechanisms of long-term structural adjustments, while disregarding the short-term aberrations and stabilization problems on the demand side. The model's name originates from its emphasis on the analysis of substitutability of capital, labor, and energy, and the sectoral allocation of energy, labor, and investments. In general, the model devotes more detail to the supply side than to the demand side. ELIAS is not derived from econometric analysis, but from combining production and preference functions and assumptions about optimizing behavior. Thus, the model does not represent a framework for an analysis of stochastic properties and structural uncertainties within the actual economic system.

Before describing the complete multisectoral model, Bergman first introduces an aggregated one-sector version. In that way, the reader gets an overview of the model more easily. The three markets, namely, the commodity, labor, and capital markets, are briefly described. The disaggregation into several sectors requires the consideration of interindustry transactions, allocation of consumer expenditures, and allocation of investment funds. The model is briefly, but still clearly, presented. It is, however, claimed that the model is designed particularly for the analysis of structural adjustments in the Swedish economy, due to changes in the world market prices (e.g., energy prices). The presentation should have been more elaborate in discussing how such structural adjustments are to be studied within the actual model.

After a careful description of the algorithm for solving the multisectoral model, Bergman concludes by discussing various aspects of implementation and model behavior. As the model includes, due to economic theoretical requirements, a number of variables which are difficult or impossible to observe, most parameters will not be estimated by econometric procedures. Instead, parameters are "estimated" according to functions between parameters and exogenous and endogenous variables. Thus, of course, the statistical significance of parameter estimates cannot be determined, which in turn renders econometric model evaluation impossible. This is a shortcoming, as in many empirical economic studies.

The third part of the volume is "ISAC – A Model of Multisectoral Economic Growth in a Small Open Economy" by B.-C. Ysander, T. Nordström, and L. Jansson. Here, ISAC, a 36 sector disequilibrium model is presented. It allows for disequilibria in foreign trade and in factor markets, and provides quantitative answers to short-term questions and investigates how the management of short-run problems can effect the process of long-term structural change. The model is specified in such a way that market disequilibria will often characterize projections based on it. Thus, the analysis emphasizes the interdependence between short-term instability and long-term growth within the Swedish economy.

The main structure of the model is illustrated by diagrams. That helps the reader identify the main variables and submodels, as well as the most important relationships between them. As in the case of ELIAS, a one-sector version is first presented. Basic elements of the model are related to the product market, production technology, production capacity growth and depreciation, foreign trade pricing, consumption demand, and labor market and wage determination. The model is disaggregated into 23 producing branches within the business sector. Contrary to the case of ELIAS, most of the parameters in this model are statistically estimated from observed historical data.

Since ISAC is specified for the analysis of stability and stabilization problems, the book includes a thorough, important, and interesting discussion of stability problems. Structur-

al changes in manufacturing and shifting government shares in consumption are two important aspects that are analyzed by the model. It is shown that ISAC is structurally stable in the sense that small isolated changes in some coefficients will imply just small displacements of the simulated growth paths. Thus, elasticity of prices with respect to unit costs is assumed to increase. Second, the short-run elasticity of export with respect to the relative prices is also increased. Another

test assumes that the degree of compensation for inflation in the wage function is reduced. Structural stability is now assumed to be present if these parameter changes do not cause dramatic changes in simulated trajectories of endogenous variables. Another important question which is not dealt with in the present study is whether structural parameters actually have changed, or will change in the future. To answer this question, other econometric test techniques must be used.

*Anders H. Westlund
Stockholm School of Economics
Stockholm
Sweden*