Journal of Official Statistics
Vol. 1, No. 1, 1985, pp. 87–96

Book Reviews

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

Kessler, R.C. and Greenberg, D.F., Linear Panel Analysis: Models of Quantitative Change. Quantitative Studies in Social Relations, Series. Academic Press, New York, 1981, cloth, ISBN –12–405750–0, ix + 203 pp., $ 22.50.

This book is a monograph on the use of linear models for the analysis of panel, or longitudinal, data. The presentation is intended for social scientists who are familiar with the basic linear model. Unfortunately, "familiarity with the basic linear model" means different things to different social scientists. Most readers will find the level varying from trivial to quite advanced. Some of the difficulty is due to the different jargons which different social scientists have developed to describe the same models. This language barrier causes problems for anybody trying to reach all social scientists. At least with respect to the economics literature, the authors did a good job covering literature outside their field.

Anyone working with panel data should learn something from this book. The authors cover most of the material completely enough so that it is not necessary to constantly refer to the numerous references. There is a lot of discussion about the assumptions necessary to make valid causal inferences from linear panel models. Such emphasis is appropriate in view of the large number of data analysts who appear blissfully ignorant of these required assumptions. Unfortunately, there are many situations where the assumptions required for causal inferences and identification are clearly false. The authors offer little guidance about how to proceed in these situations. The only logically consistent approach is to look for better data or a new problem to study. A more popular choice is to proceed with a "standard" model while hoping that referees don't notice the bad assumptions. It would have been better if the authors had more strongly advocated the logically consistent approach.

The beginning chapter provides an introduction and historical overview. It alludes to, but does not discuss, the recent debate in the economics literature on inferring causality from correlations. Interested readers should see the articles by Granger (1969), Sims (1972), and Chamberlain (1982). The second

chapter gives definitions of stability and change based on autocorrelation coefficients. The first appendix to this chapter gives a number of useful special cases, and the second appendix gives an introduction to the LISREL program. Throughout the rest of the book the authors mention how LISREL can be used to estimate the various models. This is a very useful feature for those with real data to analyze.

The chapter entitled "The Causal Analysis of Change", begins with the standard difficulties of inferring causality using cross-section data from uncontrolled experiments. The superiority of panel data is then shown, although the authors take great care to show that strong identifying assumptions are needed in this case as well. There is no discussion of causal inferences from pure time series data. This omission can be justified by the authors' correct claim that pure time series data with reasonable sample sizes are rare in the social sciences. This chapter concludes with an appropriate example of a panel analysis of crime and arrest rates. The fourth chapter concerns various decompositions of partial correlation coefficients. Although the discussion is quite clear, Vector Autoregressive models (see Fomby et al. (1984) pp. 576–590) are not mentioned here. These models appear to generalize all of the techniques presented in this chapter.

The fifth chapter is a long attack on the Cross-Lagged Panel Correlations (CPLC) technique. This is longer than required to convey the message that CPLC is an inferior technique. The next chapters discuss identifying level and change effects, and problems caused by serial correlations in the error terms. These are followed by a more data-analysis oriented discussion on choosing lag specifications from correlograms. Systematic use of these techniques leads to data mining, but that is perhaps not as bad as completing an entire study with a grossly misspecified lag structure.

Two useful chapters on sensitivity analysis and measurement error are followed by a chapter devoted to alternatives to the methods described in this book. The authors correctly take a dim view of pooling data, but their views on continuous time models seem overly optimistic given recent work by Heckman and Singer (1984). The final chapter gives a concise overview of panel design issues. Various common designs are described and critiqued. There are also short discussions on choice of interval length and panel attrition.

Overall, the authors did a good job of producing a useful and accessible monograph on a difficult subject. They made a strong case for an increase in the use of panel data in the social sciences, while being careful to point out the assumptions required to use panel data correctly. The large majority of the errors in this book are errors of omission, which are inevitable given the scope and pace of innovation in this field. A related problem is that the material in this book is at least four years old. Readers should therefore check for recent additions to the literature before beginning any major research project based on this book. A good reference for this purpose is Chamberlain (1984). In spite of its age, this book provides a good introduction and overview of linear panel analysis for a broad range of researchers in the social sciences.

References

Chamberlain, G. (1982): The General Equivalence of Granger and Sims Causality. Econometrica, 50, pp. 569–582.

Chamberlain, G. (1984): Panel Data. In Z. Griliches and M.D. Intriligator (eds.), The Handbook of Econometrics, Vol. II, North Holland, Amsterdam.

Fomby, R.B., Hill, R.C., and Johnson, S.R. (1984): Advanced Econometric Methods. Springer-Verlag, New York.

Granger, C.W.J. (1969): Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. Econometrica, 37, pp. 424–438.

Heckman, J. and Singer, B. (1984): A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. Econometrica, 52, pp. 271–320.

Sims, C.A. (1972): Money, Income, and Causality. American Economic Review, 62, pp. 540–552.

David Brownstone
University of California, Irvine
and
Industrial Institute for Economic and Social Research, Stockholm

Patel, J.K. and Read, C.B., Handbook of the Normal Distribution. Statistics: Textbooks and Monographs, Vol. 40, Marcel Dekker, Inc., New York, 1982, ISBN 0-8247-1541-1, ix + 337 pp.

According to the authors, the intention of Handbook of the Normal Distribution is to present a compilation of results related to the normal distribution. They have indeed succeeded: for the first time the many properties of the important normal distribution and its sampling statistics are now collected in one volume. As its title implies, this handbook will be one of the most valuable reference books for both theoretical and applied statisticians as well as for students.

The survey is concise and systematic. References are given to all results – not always to the original sources but instead to modern and easily obtainable ones. The book, therefore, can easily be used in individual studies.

As the flow of information increases, the need for books of this kind also increases. One merit of such books is their reference lists. In this book the authors have given a separate reference list for each chapter including between 50 to 90 references (one chapter has 15 references) in each. Because tables of distributions are available in other books, the authors have listed various sources indicating their accuracy and coverage. They discuss the univariate and bivariate distributions only while omitting the general linear model and regression models.

The book starts with the early historical development of the normal law: Chapter 1 includes a historical list of early tables of normal functions including sources such as Bernoulli, 1770–71, and Kramp, 1779. Forty-one of the 85 references are from the eighteenth and nineteenth centuries. In addition to this historical chapter, the book has the following nine chapters; Ch. 2: Some Basic and Miscellaneous Results; Ch. 3: The Normal Distribution: Tables, Expansions, and Algorithms; Ch. 4: Characterizations; Ch. 5: Sampling Distributions; Ch. 6: Limit Theorems and Expansions; Ch. 7: Normal Approximations to Distributions; Ch. 8: Order Statistics from Normal Samples; Ch. 9: The Wiener and Gaussian Processes; Ch. 10: The Bivariate Normal Distribution.

Chapter 2 presents the basic definitions and properties of the normal distribution (abbreviated N henceforth) such as moments, cumulants and distributions related to it (lognormal and uniform). The chapter also lists the general properties of some well-known families of distributions (e.g., the exponential, Pearson, stable, and infinitely divisible) which also hold for N, a member of these families. Compound, mixed, folded, and truncated N's are also discussed, as are the von Mises distribution and the wrapped N.

Statisticians who need tables of the standard N, its density, quantiles, derivatives, and Mill's ratio, will find a list of seven standard sources in Chapter 3. In addition to references, indicated in the table are functions, coverage, decimal places, and significant figures. There is also a discussion of various algorithms for computing N and related quantities such as quantiles. The many techniques for evaluating N and its density function are discussed both for expressions not directly involving N (power series expansions and rational approximates) as well as for those which are in terms of Mill's ratio. The approximation error for each approximation is given and compared. Approximations of N by other distributions, such as the Burr and Weibulls, are also given.

Characterizations of the normal law involve to a large extent distributional or independence properties of linear and quadratic forms. These are some of the topics of Chapter 4. Other topics are characterizations based on conditional distributions, characteristic functions, polar coordinate transformations, sufficiency, information, and decision theory.

Distributions (and their properties) generated by random samples from N are discussed in Chapter 5. Among the results are sampling distributions related to chi-square, t, and F. The chapter also includes a discussion of the sample mean, mean deviation and moment ratios.

Various central limit theorems for both independent and non-independent random variables are reviewed and compared in Chapter 6. Results for L-statistics and U-statistics are included. Unfortunately, the authors have not included any discussion of von Mises statistical functionals, which are useful in robust estimation and in obtaining asymptotic results for statistics. I believe that a second edition of this volume should include this topic. Convergence rates (Berry-Essen

theorems) and formal Gram-Charlier and Edgeworth expansions are also dealt with here.

The subject of Chapter 7 is normal approximations to distributions. The distributions under consideration are the binomial, Poisson, hypergeometric, Neyman's Type A, beta, von Mises, gamma, and central and non-central chi-square, t, and F. As an example, the approximation of the binomial distribution starts with the classical approach with continuity corrections and then follows various refinements such as the recommendation given by Molenaar. Examples of approximations other than classical ones with refinements are e.g. the angular, Camp-Paulson, Borges, and Peizer and Pratt. Refinements and recommendations are also given for these approximations. This chapter also contains various normalizing transformations such as the square root, inverse, and power transformations.

Chapter 8 is devoted to the topic of order statistics from normal samples. It includes a six-page list of tables of normal order statistics, which will be helpful for anyone seeking a table of functions of order statistics such as expectation, range, Studentized range, etc. Another table gives the exact values of $EX^{a}_{(r:n)}$, $a=1(1)4$, $n=2(1)5$, where the parent population is standard normal. The Wiener and Ornstein-Uhlenbeck processes are used as models to describe Brownian motion, as is discussed in Chapter 9.

Chapter 10 discusses, in less detail than the univariate N, the bivariate N. The chapter is more or less a condensed version of the book and so it serves as a very good survey of the topic.

I have attempted here to give a brief description of the contents of this excellent book. You may well find in it many useful results that you had not known before, and the book may thus serve as a reference book, and as a stimulus to research and statistical applications.

Arne Sandström
Statistics Sweden

Raj, B. and Ullah, A., Econometrics, A Varying Coefficients Approach. Croom Helm, London, 1981, ISBN 0–7099–0313–8, 372 pp., Approx. 12 £.

Quantitative techniques for structural analysis, prediction, and control of economic and social systems are generally based on the assumption that the parameters of the system are stable over time, in space, etc. It is obvious, however, that such assumptions are often inappropriate, and even wrong. Numerous examples of structurally unstable social and economic systems are given in, e.g. Johnson (1977, 1980). One fundamental reason as to why stability assumptions are still formulated is that essential knowledge about how parameter instability should be identified and characterized is unavailable. Raj's and Ullah's book Econometrics, A Varying Coefficients Approach is therefore an interesting alternative to the abundant supply of econometric textbooks where the problem of structural instability is usually only casually treated, if at all.

This book discusses econometric analysis of linear models (uni-relational as well as systems models) with unstable parameters. The authors give three objectives: (i) to treat different alternative models for parameter instability in a simple but systematic way, (ii) to illustrate econometric methodology for analysis of such models by presenting a number of applications, and (iii) to produce a comprehensive bibliography of current econometric research regarding structural instability.

The short introductory chapter includes a brief discussion of different reasons for parameter instability in economic and social systems. The following eight chapters deal with the problem with respect to uni-relational models. The discussion concentrates on parameter estimation in simple and in multiple regression models. In particular, various alternatives of the generalized least-squares estimation are considered. Briefly studied are alternatives such as likelihood estimation, Bayesian estimation, etc. Like most textbooks in econometrics, this book extensively treats some different problems that often characterize econometric analysis, e.g. serial correlation and multicollinearity. One consistent basic assumption here, however, is that the parameters of the model are allowed to be unstable. This allows for an interesting opportunity to analyse several specification problems

simultaneously. The evaluation of the current estimators is mainly based on asymptotic judgements, but the more realistic small-sample evaluations are given a certain amount of space in Monte Carlo studies as well as in approximate analytical studies.

The last four chapters deal with econometric analysis of systems models, such as vectorial regression systems, and recursive and interdependent systems. The estimation problems are emphasized, even if the possibility of identifying interdependent systems is also discussed. This discussion is, however, incomplete basically because the difficulty of finding conditions for identification becomes accentuated for systems with unstable parameters. The present identification problems are not thoroughly investigated in the existing econometric research. Finally, one chapter discusses the empirically very important systems models which are based on time-series as well as cross-sectional data.

Although the book cannot be said to be theoretically advanced, it is somewhat dominated by algebraic exercises. The vital discussion of realism in assumptions and the implications of obtained results, is often missing. This is unfortunate, not least because the authors themselves state that the book is written in particular for economists and empirical econometricians. The book certainly includes some applications, which might have been discussed more thoroughly. Above all, the implications of the econometric analysis for the economic theory are missing. What is the meaning of the observed parameter instability? How should it be interpreted and explained? The main shortcoming of the book, however, is the almost total lack of discussion on statistical tests regarding the presence of parameter instabilities. To economists, this part of econometric analysis is perhaps even more important than estimation. The presentation of the estimation methods is also unnecessarily narrow. Several approaches, which are often empirically applied, are not discussed at all by the authors. Finally, there are quite a few misprints in the book.

I fully agree with the authors that there is a need for econometric literature which systematically treats parameter instability. Their attempt to put together such a textbook, thus, is commendable. But their presentation is not complete. To take one example: the authors hold out the prospect of a comprehensive bibliography of the research area, but such is almost entirely missing. The present book therefore satisfies only a fraction of the urge for econometric literature on parameter instability among economists and other social scientists.

References

Johnson, L.W. (1977): Stochastic Parameter Regression: An Annotated Bibliography. International Statistical Review, 45, pp. 257–272.

Johnson, L.W. (1980): Stochastic Parameter Regression: An Additional Annotated Bibliography. International Statistical Review, 48, pp. 95–102.

Anders H. Westlund
University of Umeå

Salsburg, D.S., Understanding Randomness, Exercises for Statisticians. Lecture Notes in Statistics, Vol. 6, Marcel Dekker, Inc., New York, 1983, ISBN 0–8247–7057–9, 120 pp., SFr 48.

In the analysis of statistical data, one usually tries to estimate the parameters of some underlying model that describes the central tendency of the data, as well as a cloud of random aberrations that disturbs the nature of that model. The statistician's role then is to decide what components of the data can be used to fit the model and what components can be considered random noise. When working with real life data, this decision is never simple and straightforward, but rather depends to a great extent on the statistician's "feeling" for what is random noise and what is not. The primary purpose of this book of exercises, which is designed to be used together with a standard textbook of statistical analysis, such as Dixon and Massey (1969), is to provide the student with such a feeling.

The book is divided into 6 chapters. In each of the first five chapters, a data material is introduced, consisting of 2 to 16 pages of tables. These data are computer-generated in order to provide examples of homogeneous random structures, each with a single, well-

defined aberration. This is followed by exercises of various degrees of difficulty, concerning testing hypotheses, the estimation of variance or confidence limits, for example, or simply a visual inspection of the tables in order to find patterns in the data. It is presumed that the testing and estimation techniques used are to some degree known by the reader, but many of them are also described briefly in a section following the examples, where references to standard textbooks are also given. A concluding section in each chapter contains selected solved examples.

Chapter 1 deals with uniform distributions, Chapter 2 with continuously distributed random variables with positive skewness, Chapter 3 with paired data, Chapter 4 with contingency tables, and Chapter 5 with models versus data.

Chapter 6 concerns real data that is not as well-behaved as the data in the previous chapters. Following a practice from late 19th century algebra text books, the author presents exercises "designed to tease the student. Some of them are obvious to a good student. Others require ingenuity and skill a little beyond the merely good student. And, some may never have been solved even by the authors." Thus, the author gives no hints or solutions to the 9 exercises, each of which are performed on with separate data.

The author makes different suggestions about the prior knowledge necessary to assimilate the text: one semester course in statistics is claimed in the foreword, and two semester courses on page 2 and page 4 (random noise?). The reviewer believes that two semester courses will be necessary for understanding the entire book; some of the exercises deal with rather sophisticated techniques which, although they are described in the text, can be hard to understand even on the higher level.

The book is concise and well-written (some misprints do not hamper the understanding of the text), and the exercises are well suited for providing the reader with a feeling for underlying patterns and random noise in data materials. When performing the exercises, the student will also learn a great deal about how various analysing techniques work in different situations. Thus, the book can be recommended as a supplement to statistical courses. It would also be a convenient aid for self-instruction by anyone needing to improve his or her ability to analyse statistical data.

Finally, one suggestion for future editions: although the book is designed to be studied in combination with a standard textbook, it could be self-contained for a well-educated student if statistical tables necessary for solving the exercises were included. The book ends with 11 empty pages – why not use them for these tables?

Reference

Dixon, W. and Massey, F. (1969): Introduction to Statistical Analysis. 3rd Ed., McGraw-Hill Book Company, New York.

Gösta Forsman
Statistics Sweden

Tufte, E.R., The Visual Display of Quantitative Information. Graphics Press, Box 430, Cheshire, Connecticut 06410, USA, 1983, 197 pp., $ 34.

Over the past decade, many papers and books have been published showing how quantitative information should or should not be presented graphically. Among these publications, Tufte's magnificent contribution The Visual Display of Quantitative Information is a landmark. It deals with, in a readable and illustrative way, the history of displaying data, lying with data, and graphical practice in general. It also presents a base theory of data graphics.

The book itself is beautifully designed and perfectly printed and illustrated. It opens your eyes to data graphics, and once you have read it, you will have a new and more critical view of graphic illustrations.

The book is divided into two parts: Part 1 is devoted to "Graphical Practice" (Chapters 1–3), and Part 2 to "Theory of Data Graphics" (Chapters 4–9 and an Epilogue).

Chapter 1 includes a historical survey of data maps from the Chinese cartographic work of the twelfth century, up to our century. Besides the survey of data maps, it also includes the development of time-series, narrative graphics of space and time, and rela-

tional graphics. Among the graphics shown is the classical work of C.J. Minard depicting Napoleon's fate in Russia 1812–1813.

One of the inventors of modern graphical design was William Playfair (1759–1823). In his book The Commercial and Political Atlas (London, 1786) time-series using economic data were published for the first time. Playfair contrasted his graphical method to a printed table: "Information, that is imperfectly acquired, is generally as imperfectly retained; a man who has carefully investigated a printed table, finds, when done, that he has only a very faint and partial idea of what he has read; and that like a figure imprinted on sand, is soon totally erased and defaced. ... On inspecting any one of these Charts attentively, a sufficiently distinct impression will be made, to remain unimpaired for a considerable time, and the idea which does remain will be simple and complete, at once including the duration and the amount."

The book by Playfair included 43 time-series plots and one bar chart. Playfair was skeptical to the bar chart, his second invention: "This Chart is different from the others in principle, as it does not comprehend any portion of time, and it is much inferior in utility to those that do: for though it gives the extent of the different branches of trade, it does not compare the same branch of commerce with itself at different periods; nor does it imprint upon the mind that distinct idea, in doing which, the chief advantage of Charts consists; for as it wants the dimension that is formed by duration, there is no shape given to the quantities."

His ideas such as graphically illustrating time-series by *simple* plots and in viewing the plotting of small and noncomparative data sets in a bar chart skeptically, are highly valid today. As is illustrated in Tufte's book, the modern computer graphics are usually overhelmed with "chartjunks", such as vibrating graphics.

Another inventor of graphical design was J.H. Lambert who advanced general relational graphics in 1765: He plotted the evaporation rate of water as a function of temperature. Thanks to the work of Lambert and Playfair, any quantitative variable can be put in relation to any other quantitative variable (measured for the same unit).

Until the recent work of John Tukey (EDA), the development of graphics during our century has mainly followed two paths: "First, that graphics had to be 'alive', 'communicatively dynamic', overdecorated and exaggerated (otherwise all the dullards in the audience would fall asleep in the face of those boring statistics). Second, that the main task of graphical analysis was to detect and denounce deception (the dullards could not protect themselves)." (Tufte p. 53).

In the late 1960s, John Tukey invented new designs and used them for exploring complex data and analysing quantitative information. Since Tukey's work has become available, graphical work, including analysis, has gained a better reputation than earlier in the twentieth century. The book by Tufte is a good example of this.

In Chapter 2, Tufte illustrates and discusses "lies" in statistical charts and he also introduces a measure of their extent. This measure, or *Lie Factor* (LF), is defined as

$$\text{LF} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}},$$

and he states that an LF greater than 1.05 or less than .95 indicates substantial distortion. One example, illustrating the "lie", has an LF equal to 15.1. In another example, where a surface area is used as a measure, the LF is 9.4; *but* when the volume is used as a measure, the LF is 59.4!

In Chapter 3, Tufte discusses such questions as "Why do artists draw graphics that lie?" and "Why do the world's major newspapers and magazines publish them?". He gives three answers all indicated in the titles of subsections, viz. i) Lack of Quantitative Skills of Professional Artists, ii) The Doctrine That Statistical Data Are Boring, and iii) The Doctrine That Graphics Are Only for the Unsophisticated Reader.

The theory of data graphics is introduced in Chapter 4. The fundamental principle of good statistical graphics and a theory of data graphics is stated as: *Above all else show the data.* The *Data-ink ratio* (DIR) is defined by the author as a measure of this principle:

$$\text{DIR} = \frac{\text{data-ink}}{\text{total ink used to print the graphic}} =$$

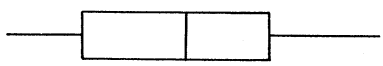
$$= \text{proportion of a graphic's ink devoted to the non-redundant display of data information.}$$

Tufte's rule of thumb for the DIR is to maximize it, within reason. To illustrate this, an original graphic is shown with three reprints. The original has a DIR of .8 to .9. The first reprint has a DIR of nearly zero, the second .7 and the third exactly zero! The author applies this rule to various graphics and shows how the DIR can be increased simultaneously with the graphic becoming more simple and readable.

The fifth chapter is devoted to *Chartjunks*. A low DIR is often due to chartjunks as optical art (vibrations and movements) and grids. As is illustrated in the chapter, much graphics are excruciating to look at and do not give any real information.

To illustrate this the author has examined newspapers, magazines, scientific journals, and books. The percentage of graphics with vibrations in scientific journals was between 2–21 percent, and in textbooks and handbooks of statistical graphics, between 12–68 percent!

In Chapter 6 the author applies his principles to many graphical designs, such as box plots, bar charts, histograms, and scatterplots. This results in new designs! One example is the box plot which step-by-step is simplified as below:

box plot: 

Tufte's preferred form:

The preferred form includes the same information as the box plot and it uses the ink effectively. This 'preferred form' is also used to redesign scatter plots with range-frames where the axes are proposed to have this form.

In Chapter 7, the idea of data-ink maximization from the sixth chapter is extended to complex and multivariate data, and in Chapter 8 data density is discussed. As a measure of graphical performance, Tufte proposes the *data density of a graphic* (DDG), defined by

$$\text{DDG} = \frac{\text{numbers of entries in data matrix}}{\text{area of data graphic}}$$

The rule of thumb is to maximize data density and the size of the data matrix, again within reason. In the extreme, one illustration has a DDG of .02 numbers per square centimeter and the current record, a map of the galaxies, has a DDG of 17 000 numbers per square centimeter! In some selected publications from 1979 – 1980 the extremes were *Pravda* with a median DDG equal to .03 and *Nature* with a median DDG equal to 7.4.

The last chapter discusses aesthetics and various techniques in graphical designs. The two main elements of a good design are stated as "Graphical elegance is often found in simplicity of design and complexity of data" (Tufte, p. 177). In this sense, Minard's 'Carte Figurative' of Napoleon's fate in Russia is perhaps the best statistical graphic design ever made. In this chapter, the author also discusses the choice of design, the combining of words, numbers and pictures, proportions and scales.

After reading this book, you will indeed feel like trying to design new graphics and charts on your own. The book shows how to display data in a good way, by using the measures proposed. One spin-off of this excellent work is a recent paper by Wainer (1984), where he discusses rules on how to display data *badly*. Twelve rules are given, some based on Tufte's measures. It is worth reading as a complement to Tufte's book.

I highly recommend Tufte's book, to all statisticians, but also to every scientist working with quantitative information.

Reference

Wainer, H. (1984): How to Display Data Badly. The American Statistician, Vol. 38, No. 2, May, pp. 137–147.

Arne Sandström
Statistics Sweden

Wilburn, A.J., Practical Statistical Sampling for Auditors. Statistics: Textbooks and Monographs, Vol. 52, Marcel Dekker, Inc., New York and Basel, 1984, ISBN 0–8247–7124–9, x + 410 pp., SFr 113.

This book is intended for financial managers, auditors, accountants, and all others who design, supervise, apply, and evaluate statistical sampling applications of financial and accounting data. The book contains 15 chapters, which cover standard methods such as simple random sampling, systematic sampling, stratified and cluster sampling, but also more special methods such as detection or exploratory sampling, work or random-time sampling, monetary-unit sampling, and flexible sampling. The two last chapters describe basic statistical concepts, estimation and evaluation and, therefore, include statistical formulas (with practical examples).

All sampling methods are presented in a non-mathematical manner. Still, the author does not avoid theoretical or practical considerations. He discusses special problems in audit sampling such as the use of prior information in the decision on stratification and levels of accuracy, and problems in sampling for rare events such as irregularities in the records (the "needle in the haystack" problem). Since the book is primarily intended for persons who work with audit sampling, there are many applications from the audit area. For the reader with at least some knowledge of statistical sampling, the book could be used as a handbook.

For readers not as familiar with sampling, it may be difficult to select the right method as the author does not discuss how to choose the best sampling method clearly enough. He could also have included a more thorough discussion on the value of sampling information vs. other sources of information, and one on the choice of level of accuracy when the auditor has other information.

As a whole, the book is easy to read and follow and should be of great practical value not only to auditors, but to everyone who investigates financial (economic) data.

Per-Olov Edlund
Stockholm School of Economics

The Sixth Scandinavian Demographic Symposium, 16–19 June 1982 in Kungälv, Sweden. Scandinavian Population Studies 6. Stockholm 1984. – 1. Studies in Migration, 195 pp. 2. Studies in Fertility, 215 pp. 3. Studies in Mortality, 107 pp.

As the title indicates these three volumes contain the papers, which were given on the occasion of a Scandinavian Symposium on demography. Like the previous similar ones this conference was organized by the Scandinavian Demographic Society. The volumes in all contain 33 papers.

The first volume deals with *migration;* by the way this is the first time a Scandinavian Demographic Symposium takes up migration as a special topic. Although the Scandinavian countries are quite small, attention is primarily paid to internal migration. Six of the thirteen papers consider problems related to making satisfactory projections, either for municipalities as a whole or for residential areas within a municipality. Four additional papers consider life-cycle theories, i.e. the starting-point is the decision of the individual to change place of residence. The three remaining papers can be said to refer to general and theoretical aspects of migration, both internal and international.

Of the thirteen papers in the second volume, which is wholly devoted to *fertility,* six deal with the great fertility decline, which in Scandinavia took place during the latter part of the 19th and the early part of the 20th century. The papers are a mixture of descriptive and explanatory exercises, thus having the ambition to tell when and why and among what groups of the population that family limitation started.

Of the remaining papers on fertility three deal with present day fertility behaviour and three with the consequences of a fertility level below replacement level. As regards the former topic it is found from research in both Norway and Sweden that the recent decline in fertility primarily has implied that births with higher parity than two are strongly reduced, whereas the number of childless and one-parity mothers so far has only declined slightly. The fact that fertility has been considerably below the level of replacement for a good number of years and in the case of Denmark has resulted in a net natural decrease of the population has given rise to some discussion in all four

dominating Scandinavian countries, but on the whole this discussion has been without much political strength and has so far not resulted in any strong measures in order to stimulate fertility. As regards Norway a voluminous Government report on population is now on the way.

The seven papers devoted to *mortality* can in brief be characterized as descriptive. In one of the papers it is made clear that even small deviations with regard to the mortality assumptions in a population projection may have considerable effect on the future size of special population groups, notably for the population over 75 years of age, a group of great importance because of its great demand for care. Because of the efficient system for registration of the population and its changes it has in all the Scandinavian countries been possible to carry out studies, in which linkages between causes of death data and occupational data in the censuses have been established.

It is often maintained that the rich variety of statistical data both for contemporary studies and for historical analysis imply a considerable advantage to Scandinavian demographers (including Non-Scandinavians studying the Scandinavian countries). This is an opinion which is not shared by Sune Åkerman, professor of history with a special interest in population at the University of Umeå and presently chairman of the Scandinavian Demographic Society. In an opening statement at the Symposium, which is now in revised form printed in Vol. 1, Åkerman says that the exaggerated interest for statistical data among demographers has implied a reluctance to pose the most important questions.

To my mind the argument that statistical data are harmful is absurd. How could a study of the demographic transition for a total population or for some subgroup be carried out without statistical data, preferably of a good quality, about the level of mortality and fertility during the period under study? There are, true enough, examples of demographer-statisticians who have become so obsessed by statistics that they limit themselves to descriptive surveys, including a study of errors in the data. Such persons will normally not claim that they have been able to throw any light on the determinants of the development. But their work is of great value to other persons who have the ambition to answer the question why development has been what description has told. In particular, with regard to historical studies demographic data which can shed any direct light on the determinants are not available.

What has just been said should not be taken to imply that statistical data on population could not be misused. As an example, I am not quite sure that the historical demographers whose papers are included in Vol. 2 above, are all aware of the risks when drawing conclusions from geographical correlations based upon small regional units, when many factors are at work simultaneously.

Åkerman also says that demographers work in isolation and that a more interdisciplinary approach is required. In this respect I can follow him. As an example, the demographer taking an interest in the development of residential areas should not limit himself to making projections, e.g. of the future population at school age, but should – in collaboration with town planners – have the ambition to tackle the whole problem of how to plan a town.

The three volumes were published two years after the symposium was held and just before the following symposium. With presently available photo-offset methods for printing, a much shorter time-lag would be possible, provided that the papers are produced in final form on the occasion of the conference. An internal standard for the design of papers (number of lines per page, length of lines etc.) would be of help. Hopefully the new data-producing typewriters will lead to such a standard.

Erland Hofsten

Printed in March 1985.