them. Chapter 2 is an article by medical ethicist Kenneth F. Schaffner on general ethics and alternative designs for clinical trials. We note that some of his sections on rudimentary hypothesis testing and Bayes's rule are really not appropriate for a book targeted to statisticians. Chapters 3 and 4 go into statistical considerations for the KSS design, including the concepts of admissibility of treatments and likelihood-based inference, the latter being the Bayesian's justification for ignoring the design of an experiment in its analysis. We avoid commenting on Bayesian/frequentist views of inference except to quote from p. 116: "The Bayesian-likelihood approach can also be a valid basis for interference!" (Overall, the book contains few misprints.) Part 2, "Test Case: Verapamil/Nitroprusside," is the heart of the book and describes the actual application of the KSS method to a clinical trial. Chapter 5 provides a brief introduction by Kadane. Chapter 6, written by two anesthesiologists, covers medical issues and internal review board approval. Chapter 7 is a comment on Chapter 6 by John L. Coulehan, a clinician. Chapter 8 describes the development of computer software to implement the KSS methodology. Chapter 9, "Being an Expert," was written by the five medical experts whose "opinions" were elicited in the trial. We were amused that the five experts could muster among them only three pages. (Is that sufficient degrees of freedom?) Chapters 10 and 11, written by statisticians, discuss the verapamil/nitroprusside trial. Chapters 12 and 13 report results of this trial. Of the 42 pages in these two chapters, 28 are devoted to graphics of not particularly good resolution. Part 3, "Other Issues," goes mostly into (American) legal aspects of clinical trials. Chapter 14, an article by lawyer David Kairys (complete with 104 footnotes!), is followed by two commentaries on the article and the author's response (Chaps. 15–17). Finally, before the brief Epilogue, Chapter 18 claims to be an article from the patient's view on whether to participate in a clinical trial, but it appears to be more concerned with statistical theory.

We strongly believe that ethically motivated alternative designs for clinical trials are not considered often enough (Palmer and Rosenberger 1996). In that regard, we were very interested in this book and applaud the intent behind the KSS method. It has other positive features, notably the incorporation of covariates important to prognosis (other than treatment) into the design. Although important covariates are not always known beforehand and incorrect selection may bias results, we feel that the methodology is a significant step in the right direction. Clearly, this design would be appropriate only in the scenario where there is sufficient experience with all experimental therapies as well as experts who have that experience. It would seem that relatively few clinical trials meet these criteria. One possible scenario suggested is a clinical trial of segmental versus total mastectomy in treating breast cancer. The elicitation of expert opinion is inherently controversial. Who should choose the experts? What constitutes an expert? Are the opinions of the experts independent of one another? Often, clinical investigators are chosen for their facilities and resources, for their proposal writing skills, or for political reasons, and sometimes the most experienced investigators may not be involved in the trial at all. Furthermore, because one goal of clinical trials is to gain conclusive evidence to influence the medical community and regulators rather than rely on each individual physician's feelings and opinions, do we want to incorporate such feelings and opinions into a design of a clinical trial? (Just considering the variety of responses of the experts to the elicitation procedure—"boring," "fun," "I don't see where it was going," "interesting" [p. 161]—may lead to some doubts.) These are important matters that should be digested thoroughly. Kadane does attempt to answer some of these questions in Chapter 1, though not always to our satisfaction. It should be noted that the "expert opinions" are just priors updated by the actual data, and they will be downweighted if they do not correspond well with the accruing data.

We particularly enjoyed Schaffner's (Chap. 2) discussion synthesizing ethical theory and clinical trials. However, we disagree with his assessment of adaptive designs, which are dismissed on the grounds that as the trial progresses, a point may be reached at which the allocation proportions are 9:1 in favor of one treatment. Hence there is a 10% chance that a patient will be assigned to what by then is recognized as the inferior treatment. We respond with two comments: First, whether a standard clinical trial design or an adaptive design was used, external monitoring would most likely have terminated such a trial under either scenario; second, 9:1 allocation at that stage is considerably more ethical than 1:1 allocation. In addition, Schaffner's comments on adaptive designs need updating, as he ignores some recent adaptive clinical trials (see Rout, Rocke, Levin, Gouws, and Reddy 1993, another anesthesia trial, and Tamura, Faries, Andersen, and Heiligenstein 1994, a trial on depression).

It took a chance meeting on an airplane for the verapamil/nitroprusside to get underway (p. 129). We look forward to the day when more ethical designs of clinical trials do not have to rely on such a "happy circumstance" to move from theory into practice. Despite the limited focus of this book and its stylistic unevenness (not surprising, being a compilation), we applaud this book's diverse authorship and its ethical motivation. We welcome this book into our libraries.

William F. ROSENBERGER
University of Maryland

Christopher R. PALMER
University of Cambridge

REFERENCES

Kadane, J. B. (1986), "Progress Toward a More Ethical Method for Clinical Trials," *Journal of Medicine and Philosophy*, 11, 385–404.

Palmer, C. R., and Rosenberger, W. F. (1996), "Ethics and Practice: Alternative Designs for Phase III Clinical Trials," unpublished manuscript.

Rout, C. C., Rocke, D. A., Levin, J., Gouws, E., and Reddy, D. (1993), "A Reevaluation of the Role of Crystalloid Preload in the Prevention of Hypotension Associated With Spinal Anesthesia for Elective Cesarean Section," *Anesthesiology*, 79, 262–269.

Spiegelhalter, D. J., Freedman, L. S., and Parmar, M. K. B. (1994), "Bayesian Approaches to Randomized Trials" (with discussion), *Journal of the Royal Statistical Society*, Ser. A, 157, 357–416.

Tamura, R. N., Faries, D. E., Andersen, J. S., and Heiligenstein, J. H. (1994), "A Case Study of an Adaptive Clinical Trial in the Treatment of Out-Patients With Depressive Disorder," *Journal of the American Statistical Association*, 89, 768–776.

Measurement Error in Nonlinear Models

R. J. CARROLL, D. RUPPERT, and L. A. STEFANSKI. London: Chapman and Hall, 1995. xxi + 305 pp. $54.95.

Over the last 15 years, substantial effort has gone into understanding and compensating for the effects of covariate measurement error in nonlinear regression models. The number of papers in the area is quite large, as evidenced by the bibliography of this book, and continues to grow. This is the first book on this subject, and the authors have done an excellent job organizing and presenting many of the results of this research.

The monograph is organized so that it is quite accessible and useful to statisticians interested in applications or research. The writing style is clear and easy to understand. Chapters that present methodology typically are organized so that the basic ideas and methodology are first described and then illustrated by example. Technical details are often relegated to a chapter appendix. The overall mathematical and statistical level is not as high as that in the text by Fuller (1987).

The book contains 14 chapters and an appendix. The first five chapters form the heart of the book. Chapters 1 and 2 are introductory. The methods presented in Chapters 3–5 tend to be more widely applicable and less technical than the methods in subsequent chapters. Also, SAS procedures and S-Plus routines that implement these methods can be obtained from the authors.

Issues regarding modeling measurement error are often not clear from the literature. Chapter 1 clarifies the situation by describing seven examples, most from applications in epidemiology. The examples illustrate a variety of measurement error models and the assumptions and types of data available to assess measurement error.

Chapter 2 uses the linear model to review the effect of measurement error on estimation of regression parameters and to review basic strategies to correct for these effects. The first two chapters provide an excellent introduction to measurement error modeling. Anyone interested in measurement error, in either linear or nonlinear regression models, will benefit from reading these chapters.

Chapters 3 and 4 describe the default methods for estimation. Regression calibration, probably the easiest method to understand and implement, is described in Chapter 3. SIMEX estimation, a relatively recent and inno-

vative addition to the measurement error literature, is described in Chapter 4. Chapter 5 describes approaches to using instrumental variables based on regression calibration. Chapter 6 describes corrected score and conditional score estimators and gives a comparison of the methods.

Chapter 7 describes estimation using the likelihood of the observable data, and Chapter 8 describes Bayesian methods. Chapter 9 focuses on techniques that require validation data. Chapter 10 describes methods for generalized linear models with unknown link functions and Chapter 11 gives some results for hypothesis testing. Chapter 12 details results for density estimation and nonparametric regression, and Chapter 13 describes methods used when the response variable has measurement error. The final chapter, 14, provides a brief description of a variety of topics, including nondifferential measurement error and estimation in proportional hazards models.

The estimators described in the book are solutions to estimating equations, and most formulas for standard errors are obtained from estimating equation theory. Alternatively, it is often straightforward to compute bootstrap standard errors. The appendix provides a clear synopsis of estimating equation theory and bootstrap methods. It also provides an introduction to likelihood methods and quasi-likelihood and variance function models.

There are some typos in the book, but the ones I found are fairly obvious and rare enough not to cause aggravation. The monograph could be used as a text for a special topics course at the graduate level but there are no exercises.

My only quibble with the book is that the authors do not always state their preferences among estimators that are applicable to the same data. For example, a least squares and a sliced inverse regression estimator are described in Chapter 10, but no recommendations are provided concerning situations where one estimator may be more appropriate than the other. For some methods, recommendations are made, or it is acknowledged that it is not clear which estimator is preferable.

I think the authors have written a wonderful book. It provides a thorough and concise description of several useful strategies and an overview of many others. It should prove useful to those interested in accounting for measurement error in regression analyses as well as researchers in this field. To those interested in measurement error, I strongly recommend *Measurement Error in Nonlinear Models.*

Jeffrey BUZAS
University of Vermont

REFERENCE

Fuller, W. A. (1987), *Measurement Error Models*, New York: Wiley.

Models for Uncertainty in Educational Testing

Nicholas T. LONGFORD. New York: Springer-Verlag, 1995. xiv + 285 pp. $44.95.

This book is an account of a collection of statistical problems encountered by the author as a statistician at the Educational Testing Service, and of the solutions that he supplied for those problems. Some of these have been published previously as journal articles; others have not. The problems generally involve large and complex datasets, and the solutions generally use linear models giving wide-sense analyses of variance, in a context of assessing uncertainty of educational measurements. In this book "uncertainty" is the preferred term for what is referred to in test theory, perhaps loosely, as errors of measurement. Specific topics of current interest in psychometric research include estimating abilities from ratings, differential item functioning between populations, and equating scores on alternative test forms, as needed, for major testing instruments such as the Graduate Record Examination.

The book is offered as a statistical text, with deliberate minimization of educational measurement terminology and of the need for background in test theory. It is not very demanding technically, and if used in a specialist seminar would not present difficulties to students with a regular undergraduate statistics training. How it could be placed as a textbook remains unclear. For this collection of problems and solutions to be considered for an advanced seminar in educational statistics, it would need to be preceded by a comprehensive graduate course on classical (linear) and modern (nonlinear) models for item responses. In such a case, the instructor would have to supply a great deal of linkage to conventional test theory and would regret the author's minimization of such background and his tendency to write known test theory models rather idiosyncratically and without reference to previous work. (For example, the standard three-parameter logistic model is rewritten in a form that disguises the intent of the parameter that models guessing.) The analyses are of very large datasets, with no database and no exercises or teaching supplements or account of the S-Plus programming that, we are told, underlies all the analyses.

In the absence of a unified topic of general significance, it is not obvious why these contributions have been put together as a text. In designing a seminar, whether for students of educational statistics or for general statisticians, instructors probably would prefer to make a list of readings from journal articles and book chapters. Quite possibly, some chapters from this volume might find a place on such a reading list.

The best thing to be said about this book is that it represents the work of a competent statistician meeting the challenges offered by large and complex data sets, and exemplifies the statistician's task. The treatment is not beyond criticism. I object to the use of linear models for differential item functioning and for equating tests, and would like to have seen at least an acknowledgment of alternatives from item response theory and their possible advantages. Even without an ultra-careful working through of the derivations, one can be startled, for example, by the casual statement that a certain matrix (p. 71) is symmetric, which, on the face of it, cannot be true. And some remarks fall in the unbridged gap between accepted test theory and these contributions that seem problematic; for example, the superficial account of reliability (p. 134) and the assertion that "the only available alternative to unidimensionality [nowhere defined] is the presence of two dimensions" (p. 137).

Models for Uncertainty in Educational Testing should be of interest to researchers actively engaged with test theory, especially those who have a concern with the problems of large-scale ability testing. Longford is to be respected for these contributions, and the volume that brings them together may prove a useful resource. I suspect that this book will not be in demand as a textbook, but with supplemental materials it might succeed as such in a suitable specialization.

Roderick P. MCDONALD
University of Illinois

Mathematical Statistics for Economics and Business

Ron C. MITTELHAMMER. New York: Springer-Verlag, 1996. xviii + 723 pp. $54.95.

Most economists—econometricians, in particular—have a love–hate relationship with statistics textbooks. We understand division of labor and anticipate that statisticians will do a better job of writing statistics textbooks than economists do. Furthermore, we appreciate the time, thought, and experience that go into these texts. However, there is a seemingly endless and fruitless search for the perfect text. Our lament is like the cry of the jilted spouse, "They don't understand us." No one text seems to satisfy all our needs, and the available options (multiple textbooks, substantial supplemental material, etc.) are less than ideal for both instructor and student.

The ideal text comprises a crib sheet or second voice for students who are not keeping up with lectures and a handy set of examples and exercises to reinforce learning. If I can be permitted one sweeping generalization, two types of textbooks are generally considered by economists teaching statistics to their graduate majors. The first is the mathematical statistics text along the lines of the venerable Hogg and Craig. The second has the generic title "statistics and other tools useful for doing econometrics." This can appear as a separate text or comprise the first handful of chapters in an econometrics text. The typical material covered is a large battery of matrix operations, properties, and so on, followed by an extension of univariate point estimation and hypothesis testing to the multivariate case.

Neither class of textbook is truly satisfactory for graduate students in economics. I have typically assigned the basic math statistics text supplemented by a thin text on real analysis to flesh out the weak probability theory and preordering the subsequent econometrics text to get the full flavor of econometrics-minded multivariate statistics.