Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
International Journal of Forecasting 23 (2007) 717–720
international journal of forecasting
www.elsevier.com/locate/ijforecast

Book reviews

Nonparametric Econometrics: Theory and Practice, Q. Li, Jeffrey S. Racine, 2006, 746 pp., ISBN: 0-691-12161-3, Hardcover, $85.00

Nonparametric and semiparametric methods have attracted a great deal of attention in recent years in the econometrics discipline. However, these approaches have been largely confined to continuous data, with no nonparametric or semiparametric methods applicable to ordinal or nominal levels of measurement. The latter is a major feature of *Nonparametric Econometrics: Theory and Practice*, though one which has been largely ignored by theoretical and applied econometricians. Additionally, with the advances in computing power, the implementation of these methods allows for their use in empirical research. Readers should thoroughly understand the concepts of parametric econometrics at the level of Greene (2003) and Wooldridge (2006) to read this book with comprehension. The appendix does provide a comprehensive refresher on mathematical statistics which will illuminate the reader's understanding of the underlying derivations of the methods used. Each of the five sections has its own theme related to nonparametric and semiparametric methods.

Part 1 covers the basics of nonparametric estimation methods, and which should be required reading for those with a limited knowledge of nonparametric estimation methods. Chapter 1 delves into density estimation (both univariate and multivariate approaches) and includes a good discussion of cross-validation methods for selecting the appropriate bandwidth (smoothing factor). Chapter 2 covers the concept of nonparametric regression, since regression methods are widely used empirical methods in social science research. With parametric methods, the researcher must specify the functional form in a regression model to describe the underlying data. The first part of this chapter reviews the concept of parametric regression, then presents the approach of nonparametric regression. The discussion is primarily about local regression methods, which is almost the same as an OLS estimation procedure, but with one major difference: OLS estimates a regression line through the entire data set, whereas the local linear regression estimates a weighted regression through each datum. The methods applied in this chapter are for continuous data.

In chapters 3 and 4, the discussion shifts to the implementation of nonparametric methods for mixed data, i.e., continuous and discrete data using the Li-Racine Generalized estimator. Nonparametric methods from chapter 2 can be applied to models using mixed data; the primary disadvantage of this is that earlier nonparametric methods do not smooth the data very well. Consequently, the methods in chapters 3 and 4 remedy these deficiencies in smoothing mixed data. Chapter 5 deals with conditional density estimation (CDF), the backbone of modern statistical analysis; however, CDF is often not modeled directly in parametric empirical models and occurs even less frequently in nonparametric modeling. The CDF provides for the conditioning of regressors, even though researchers are estimating a CDF rather than a regression model. Sometimes when estimating a conditional mean regression, these models do not capture the full effects of the variations in the data such as censored data. Consequently, a quantile regression would be a more appropriate regression for estimation. This chapter examines these models for both continuous and mixed data.

Part 2 deals with a variety of semiparametric models such as partially linear, single index, additive, varying coefficient, censored, and sample selection

models. Chapter 7 covers the semiparametric partially linear models, but generally speaking, some of the specification is parametric while the remaining parts of the model are not specified. Put in another way, it is a combination of kernel methods with ordinary least squares (OLS). Because of the simplicity of these models, this is given its own chapter. Chapter 8 presents the single index models applied to binary choice models. This chapter also provides discussion of some recently developed methods, notably multinomial discrete choice models. Chapter 9 presents the popular semiparametric regression; however, these models are subject to criticism, as are the parametric models, since part of the modeling process presupposes a functional form in the empirical analysis. The popularity of these models also allows for more interpretable results than a complete nonparametric model. This chapter covers the additive, additive partially linear, and semiparametric varying (smooth) coefficients models.

The selectivity models, which are largely attributed to Heckman, are introduced in chapter 10. This chapter also discusses the semiparametric and nonparametric methods of addressing this issue, as well as variations of the Tobit models. Das, Newey, and Vella's nonparametric approach is briefly presented. Chapter 11 delves into censored regression models, which are commonly used in empirical analysis. The first part of the chapter briefly reviews the censored regression models, as applied to parametric models; the methods are then developed using semiparametric procedures. The concluding part of the chapter develops nonparametric methods.

Part 3 moves onto model misspecification tests. One of the drawbacks of parametric regression is that a model can be misspecified because of selecting an incorrect functional form that may not fit the data as well as was supposed. Chapter 12 provides kernel-based misspecification tests. Additionally, misspecification tests are developed using the smoothing tests, and these are covered in chapter 13. After a brief review of the misspecification tests under the parametric assumption, the methods are developed for nonparametric methods for both continuous and mixed data.

Part 4 provides methods related to nonparametric nearest-neighbor and series methods. Chapter 14 focuses on the *k* nearest-neighbor methods, widely used in nonparametric data analysis since these methods adapt to the local information available at each datum and surrounding data points. That is, the more local information there is available, the smaller the range of smoothing that occurs. Once the kernel functions have been specified, the leave-one-out estimator is applied to determine the window width. The selection of the appropriate window width is an important component of nonparametric estimation. If a small value is selected, not enough points will be covered, and the result will be undersmoothing (low bias, high variance). On the other hand, if a large value is selected, too many points will be covered, resulting in oversmoothing (high bias, low variance). Chapter 15 provides an alternative that uses techniques such as splines and power series. The series method requires the researcher to select the number of "knots" or terms in the power series, which is analogous to the selection of the bandwidth (smoothing factor) in kernel methods. As the number of knots increases, the fit becomes more flexible. The main advantage of these methods is that they allow greater flexibility to incorporate any number of constraints into the model.

Like parametric models, the regressors can be endogenous, which makes the OLS estimator look rather less appealing. Chapter 16 applies semiparametric methods, using instrumental variables (IV) to handle the problem of endogeneity. Chapter 17 deals with modeling simultaneous equations using nonparametric methods. Recall that a set of equations is said to be simultaneous when the dependent variables appear in other equations as a regressor. Chapter 18 discusses weakly dependent data, and chapter 19 then presents the semiparametric and nonparametric methods for panel data. The methods used are the partially linear semiparametric and fully nonparametric panel data models and the panel discrete choice and censored regression models with unknown error distributions. The coverage of panel data methods includes the standard topics of fixed- and random-effects models, as encountered under parametric methods. Chapter 20 covers a smorgasbord of topics relating to nonparametric and semiparametric methods and applications that do not fit into the other chapters. It discusses nonparametric time series methods, average treatment effects, and the development of nonparametric methods for auction models. For the average treatment

effects, these models are commonly used to examine such wide-ranging topics as the relationship between dose–response relationships in medicine, human capital losses from war, and the effectiveness of job training.

This book is a major contribution to econometrics and social science research because of its thorough coverage of nonparametric and semiparametric methods as they apply to economic models and other social science models. More importantly, this book makes clear much new material relating to discrete data. Even econometricians with little or no background in nonparametric methods could apply these methods to their research. It contains specifics that clearly spell out the steps to implement these methods, and many of these methods were recently implemented in an R package. Consequently, researchers can now easily apply these nonparametric and semiparametric methods in their empirical research. Each of the chapters includes exercises which can easily be integrated into graduate courses in econometrics, but it will be most useful to researchers in econometrics who need a thorough survey of the nonparametric and semiparametric literature.

As the title implies, this book does contain theory, and many of the difficult proofs are relegated to the appendix of each chapter. Each of the chapters provides examples to illustrate the methods used in the chapter, and often these applications are from other areas of the social sciences. Providing a broad range of applications shows the potential applications of these methods to areas other than economics.

The authors have done a remarkable job in providing a thorough presentation of the myriad of nonparametric and semiparametric methods useful in econometrics and the social sciences.

References

Greene, W. (2003). *Econometric analysis* (5th Ed.). Prentice Hall.
Wooldridge, J. (2006). *Introductory econometrics: A modern approach* (3rd Ed.). Thomson and South-Western.

Brian Sloboda
University of Phoenix, United States
E-mail address: bsloboda@email.phoenix.edu.

doi:10.1016/j.ijforecast.2007.07.006

Sales Forecasting: A New Approach, Thomas F. Wallace, Robert A. Stahl. T.F. Wallace & Co., (2006). ISBN: 0-9674884-1-9 (paper), $44.95, 166 pages

This is a book aimed squarely at company-based practitioners, with no glances from the authors as to possible academic readers. To justify its subtitle of 'a new approach', the authors emphasize the importance of teamwork and process improvement rather than formulas and forecast accuracy. After an introduction stressing why companies need to be concerned with forecasting and how difficult it is to avoid, the second chapter briefly discusses accountability, locating the organisational ownership of the sales forecasts firmly with the sales and marketing departments. Chapter 3 discusses the process of producing a final forecast based on various customer categories and the economic and market influences that affect sales. The conversion of these inputs to a forecast the authors see as arising from people, software (including statistical models), and a structured process. Accuracy is then discussed, with illustrations as to why it is difficult to judge forecasting performance using MAD as a measure. "Stamp out bias", the authors add (a discussion point on their Amazon site, I note). Practical issues such as the level of detail required when forecasting, and the organisational linkages between forecasting, master scheduling and demand management occupy much of the rest of the book. The final substantive chapter is on implementation. The appendix includes a limited list of software companies and an introduction to simple exponential smoothing.

The one thing the book does quite well is focus on some of the practical issues faced by organisational forecasters. The principles the authors offer are mostly sensible. Many company forecasters seem to have the forecasting baton thrust upon them with no hint of training. A reader with just such a task would recognize many of the issues s/he faces in producing a forecast. A naïve student carrying out a project would gain some insight into the messy world of demand forecasting. However, neither would learn very much. Some of what they read would be misleading and inadequate. The book's failure to offer an elementary analysis of trends and seasonal factors, helpful graphs, or the role of judgement are examples of its serious limitations.