

interviewers and respondents, the structure of questionnaires etc.) helping to detect the origins of nonresponse or to detect inconsistencies, respectively. The authors justly point out that the selection of contents is a very difficult task. In my opinion, an additional challenge consists in an appropriate adaption of the contents to the different degrees of prior knowledge and when experience of the participants. It is not least a question of financial support to avoid redundancies generating the list of topics. Furthermore, there remains justified doubt as to whether the whole of the courses listed by the authors can be provided in a reasonable time interval.

The JPSM allows its participants to treat specific problems arising from conventional statistical disciplines. The participants have the opportunity to gain a deep insight into the processes and needs of the federal statistical system. It seems very likely that the cooperating universities are responsible for the academic education. In my opinion, the main merit of the JPSM is that they have found a balanced combination of theory and practice. In general, the concepts of life-long learning and dual education on any grade level are very topical (vocational schools, vocational colleges etc.). For instance, in Germany in recent years the lack of statistics in the curriculum of social science studies has been subject to controversy and painfully observed in official statistics. Thus, the idea of supplying what the authors call "short courses" adapted to the specific needs of their participants is worth imitating.

Over the years, the JPSM has periodically re-evaluated its programs. It would be of interest to get further information on the re-evaluation procedure (further remarks on the duration of a period, who has the responsibility for re-evaluation and so on). Does the financial support of the program (and hence its continuation) depend on the results of re-evaluation? Moreover, to bring in internationally recognized experts seems to be connected with some (nonnegligible) financial expenditure.

As a whole, the article is a very interesting contribution to the general context of life-long learning, dual education and the harmonization between federal statistics and empirical sciences. The convincing measures of the Joint Program in Survey Methodology widely exceed ordinary conventional vocational training and other existing approaches like trainee programs, the latter consisting mainly in an agitation of their participants through the departments of the employing institution. Last but not least, in the long run efforts like the JPSM essentially improve the reputation of federal statistics among both the general public and the scientific community.

Rainer Lenz  
Saarbrücken University of Applied Sciences  
Goebenstrasse 40  
66117 Saarbrücken  
Germany  
Email: rainer.lenz@htw-saarland.de

## Book Reviews

Books for review are to be sent to the Book Review Editor Jaki S. McCarthy, USDA/NASS, Research and Development Division, Room 305, 3251 Old Lee Highway, Fairfax, VA 22030, U.S.A.  
Email: jaki\_mccarthy@nass.usda.gov

Propensity Score Analysis: Statistical Methods and Applications	
Sunghye Lee	711
Interpreting Economic and Social Data: A Foundation of Descriptive Statistics	
Brady T. West	713

**Shenyang Guo and Mark W. Fraser.** *Propensity Score Analysis: Statistical Methods and Applications*. Thousand Oaks, CA, Sage, 2010. 370 pp, ISBN 978-1-4129-5356-6, \$64.95.

*Propensity Score Analysis: Statistical Methods and Applications* is a well-organized book that provides a concise yet up-to-date overview of a vast volume of literature around causal inference about treatment effects using observational data. While the book appears to target program evaluation researchers as the main audience, researchers in other disciplines may find the philosophical and historical background of causal inference in early chapters helpful to broaden their understanding of the usage of observational data. In most social sciences, pure randomized experiments are difficult and often unfeasible to implement with human populations, leaving observational studies as the only option. This complicates causal inferences about the effectiveness of social programs or interventions as there may be numerous confounding factors associated with the program or intervention participation, introducing selection bias to treatment effects. Hence, it becomes necessary to balance between treatment and control groups in the observational data for removing selection bias and drawing valid causal inferences. Classical data balancing methods, such as Ordinary Least Squares Regression, matching and stratification, briefly reviewed in Chapter 3, exhibit limitations in correcting for such a bias.

The book introduces two schools of thought for dealing with observational data for estimating treatment effects in Chapter 2. The one labeled as "Neyman-Rubin counterfactual framework" originated from and was developed within the discipline of statistics; while the other, the "Heckman framework," originated from econometrics. As clearly illustrated in Table 2.2, the main difference between the two traditions is that Heckman's approach explicitly models the participation process attempting to develop more generalizable models and requiring strong assumptions, whereas the counterfactual framework does not make such an attempt. The debate between the two traditions continues and is central to advancing causal inferences, as the authors acknowledge.

Instead of contrasting these two camps, the authors do an excellent job of integrating them into a cohesive theme of propensity score analysis and re-categorizing it into four

distinctive yet related methods in Chapters 4–7. These four chapters follow a similar pattern: illustrating the logic, theoretical development and mathematical formulations of each method and providing instructions on how to apply the particular method using popular statistical software (mainly Stata and R), with concrete examples, and concluding with advantages of, and caveats concerning each method shown in the examples. For this reason, this book should benefit practitioners and academics alike. However, these chapters require a certain amount of technical background in order to thoroughly appreciate them. Chapter 4 includes Heckman's two-stage estimation approach that simultaneously considers the sample selection process and the outcome variable in separate models, the treatment effect model, and the instrumental variable model. Propensity score matching in Chapter 5 has seen a large number of developments since its first introduction to the statistical literature in early 1980's by Rosenbaum and Rubin. Therefore, the authors include various extensions in the chapter with different ways of estimating and modeling propensity scores, and of using estimated propensity scores through matching and postmatching analysis. Chapters 6 and 7 include very recent developments in the propensity score literature, namely matching estimators and kernel-based nonparametric propensity score matching.

What I find most useful with this book is how closely it follows the causal inference literature as it reviews highly state-of-the-art techniques, such as generalized boost regression, optimal matching, and fine balance in Chapter 5, matching with replacement in Chapter 6 and local linear regression in Chapter 7. It is imperative for the successful removal of selection bias to select relevant predictors and specify correct models in estimating propensity scores. However, because it involves subjective judgment, this book provides practical guidance to researchers faced with such tasks. Generalized boost regression is presented as a potential solution when the specification of predictors' functional form is an issue. Optimal matching, a new matching method that attempts to minimize the distance of propensity scores at the total sample level rather than the individual level, and fine balance, a technique using a nominal variable (not propensity scores) created with substantively important variables to achieve balance, are introduced as alternatives and complements to greedy matching methods, such as nearest neighbor or Mahalanobis metric matching.

Matching estimators in Chapter 6 are different from propensity score matching as they do not require prediction of propensity scores. This is an advantage over propensity score matching because it reduces the number of analytic procedures in matching as well as the number of subjective decisions in model specification. Instead, this technique uses a vector norm to calculate distance on the observed covariates for matching. Among them, matching with replacement may be a particularly useful tool as it allows individual cases in the control group of the observational data to be used more than once in forming matches, unlike all other matching techniques, which use them at the most once. Ultimately, this leads to higher-quality matches and allows estimators to produce the population level average treatment effect by matching all units in the data.

The nonparametric propensity score models in Chapter 7 developed by Heckman and his colleagues lead to kernel-based matching where each treatment case is compared to all control cases. This approach uses information from all possible control cases, whereby more information is taken from the closer matches by applying differential weights.

Local linear regression is used in determining this weighting function and smoothing unknown and potentially complicated functions.

The authors caution readers about the sensitivity of model misspecification, the difficulty with verifying assumptions and the uncertainty about their utility with small samples. This is summarized in Chapters 8 and 9. Chapter 8 also provides theoretical and hands-on illustrations of how to evaluate sensitivity to elements not considered in models, while Chapter 9 considers common mistakes in observational studies and provides future research directions along with alternatives to propensity score analysis for causal inferences.

Overall, I would recommend this book to statisticians, including those outside program evaluation research, as it presents the concept of propensity score analysis in an easily understandable manner with its application to existing observational data. While the minimal use of mathematics for a complicated issue in the book may appeal to some, those who wish to follow the mathematical works may want to complement the book with some references. While the focus of the book is on program evaluation, it may be worthwhile to attempt to examine the applicability and feasibility of employing new developments in propensity score analysis such as those described above in other areas of statistics where selection bias exists.

Sunghee Lee  
University of Michigan  
Survey Research Center  
426 Thompson St.  
Ann Arbor, MI 48104  
U.S.A.  
Tel: 734-615-5264  
Email: sungheel@isr.umich.edu

**Othmar W. Winkler.** *Interpreting Economic and Social Data: A Foundation of Descriptive Statistics.* Berlin, Springer-Verlag, 2009. ISBN 978-3-540-68720-7, 265 pp, \$119USD.

This insightful text comes from a veteran scholar and targets individuals working with socio-economic (SE) data from business, economics, sociology, and other social sciences. The book provides a refreshing view on how data from these fields should be approached in a realistic manner. The author strongly and consistently advocates the use of straightforward descriptive statistical methods to describe SE realities, rather than forcing SE data to conform to convenient probability models and inferential methods developed by mathematical statisticians for data from the natural sciences (which are much more likely to be governed by natural laws and "true" probability models). *Interpreting Economic and Social Data* will appeal to (and should be read carefully by) students and professional researchers in the social sciences who are responsible for analyzing cross-sectional or longitudinal SE data and generating written reports describing and interpreting the analysis findings. Readers are assumed to have sufficient training in elementary statistical concepts such as simple random sampling, hypothesis testing, descriptive statistics, and simple linear regression.

The Preface and Chapter 1 combine to lay the groundwork for the remainder of the book, discussing the importance of descriptive methods for analyses of SE data and arguing that the probability theory underlying many common statistical methods was developed for the natural sciences and cannot satisfy the needs of SE statistics. The author calls for removing notions of "randomness" (or stochastic properties) from inferential methods for SE data, given that SE processes do not occur at random. The author also notes that modern computing power makes it easy to apply advanced statistical procedures and create an illusion of scientific rigor, even though the procedures are often incorrectly applied and the relevant statistical assumptions are not justified. The Preface quickly introduces the reader to the author's liberal use of end notes, which contain a healthy mix of cited works and very interesting (and well-researched) supplemental information. Some readers may find that the large number of end notes prevents efficient reading of the text, but many of the notes are well worth reviewing.

Chapter 2 clarifies the nature of SE data and emphasizes the need for analysts to have a detailed understanding of the units of analysis producing aggregate statistics, while Chapter 3 details the author's view that aggregate SE statistics should be reported along three important dimensions: history/time, classification (or subject matter), and geographic location. This type of three-dimensional aggregation is emphasized throughout the book for all forms of descriptive analyses discussed, and this emphasis will be extremely helpful for analysts who often aggregate data across these dimensions, potentially clouding important research findings. Chapter 4 focuses on the importance of ratios in analyses of SE data, and connecting statistical aggregates with others to discern the "big pictures" of SE processes. Unfortunately, Chapter 4 does not discuss any methods for stating the uncertainty associated with the estimation of ratios based on sample survey data, which is a notable omission from other chapters as well. Design-based methods for estimating sampling variances for ratios can be found in classical texts such as Cochran (1977) and newer applied texts such as Heeringa et al. (2010).

Chapter 5 begins a three-chapter sequence on longitudinal SE data, with a focus on descriptive methods for various types of time series analyses. The author details components of traditional time series models, and argues (correctly) that the "random fluctuations" in a time series depend entirely on the model posited by an analyst for the time series rather than SE reality. Alternatives to traditional analyses of time series are proposed at the end of the chapter which many analysts of time series will find useful. Chapter 6 discusses the "art" of forecasting based on time series, and focuses on the fact that data obsolescence is a much larger problem for SE data than for data from the natural sciences. The author stresses the need for subject-matter experts to have a solid understanding of the history of a SE process, with more weight being given to more recent SE data in the forecasts. This chapter is not a "how-to" chapter on computing forecasts, but offers several interesting insights on this methodology. Chapter 7 then concludes the three-chapter sequence by focusing on longitudinal price-index statistics. This chapter contains the most economic concepts (including production concepts) and may be tougher reading for those without this background. An interesting insight from Chapter 7 is that a price is actually a feature of a *transaction*, rather than of a product. Accordingly, transactions for a given product and service should be the units of measurement when computing price-indices. As a result, price-index statistics should be

for prices actually paid, for a given type of product, at a given time, at a given location, and those processing SE statistics need to pay careful attention to the definition of a "price" in statistical reports. A discussion of quantifying the uncertainty in price estimates is once again omitted.

Chapter 8 shifts gears and focuses on descriptive analyses of cross-sectional SE data. This chapter begins with a heavy focus on graphical methods for describing frequency distributions, and continues with brief discussions of cumulative distribution functions and measures of central tendency, variability, and asymmetry (the latter being especially important for SE data). An interesting method for quantifying and interpreting the asymmetry in a distribution is described at the end of the chapter, including example calculations. Chapter 9 focuses on the correct interpretation of simple linear regression coefficients from models fitted to cross-sectional SE data, and begins with an interesting story about a sex discrimination lawsuit that changed the author's view about how simple linear regression coefficients for cross-sectional SE data should be interpreted. The chapter concludes with an alternative technique proposed by the author for assessing the bivariate correlations of SE variables, given that simple linear regression is not appropriate for many of the non-standard bivariate distributions found in SE data. Appendix D provides more details on this technique and heuristic tools for comparing different types of correlations. This chapter provides only minimal discussion of multiple linear regression models for SE data, and focuses on the study of bivariate relationships.

Chapter 10 further emphasizes the inadequacy (for analysts of SE data) of stochastic statistical tools developed for the natural sciences. The author lashes out against the notion that observed SE data deviate from some known model based on random errors following a probability distribution, again stressing that SE processes do not occur at random and should be described rather than modeled. This chapter includes brief discussions of probability, random sampling and significance testing, as also of common misuses of probability and statistical significance in analyses of SE data. Chapter 11 is more of a commentary on the many similarities between statistics and accounting, how elucidating statisticians could learn much from the reliance of accountants on objective facts rather than probability models. The author also calls for more emphasis in textbooks on the practice of data entry (the statistical form of bookkeeping), given the importance of this task for correct data analysis.

Chapter 12 discusses the importance of taking geographical location into account when computing SE statistics, which the author claims is often overlooked in statistical textbooks for business and economics. This chapter includes many examples of geographers using SE data in interesting applications, and the author bemoans the fact that statisticians have not used geographical data enough when performing analyses of SE data. Unfortunately, this chapter does not point readers in the direction of software or statistical methods for the analysis of spatial SE data; interested readers might consider Anselin (1992) or Anselin et al. (2010). The book concludes with some interesting "Afterthoughts" from the author, presenting a brief personal background that led to the development of the book, and five appendices providing supplementary material for Chapters 3, 5, 8, 9, and 10.

The major problem with the current edition of this book is the very large number of grammatical mistakes and typographical errors that hinder efficient reading of the

material. These problems become readily apparent no sooner than the first few pages of the Preface and Chapter 1. For instance, a double colon (::) five sentences into the Preface causes hesitation, and sentences like "The powerful influence and the band wagon effect of the developments in statistics in biology, agriculture came to dominate all fields of statistical application." on page *vii* require 3–4 reads to fully understand the intended meaning. These mistakes unfortunately persist in later chapters (e.g., 18 and 18 somehow add to 32 on page 165). Later printings should follow an extensive and detailed review of all chapters, Figures (e.g., the sign of the regression coefficient for females in Figure 9.1 is incorrect), and ancillary end notes (e.g., "Explained Variation" [*sic*] in end note (1) on p. 252) by a professional proofreader or copy editor, with these mistakes fixed. The overall messages in the book are quite clear, but future readers would greatly benefit if this advice were to be followed by the author and the publisher.

Despite the grammatical and typographical issues, this book is an enjoyable read, and it does an excellent job of reinforcing the unique features of SE data and how statistical analyses should be tailored to these features to produce the most meaningful descriptions of SE phenomena.

### References

- Anselin, L. (1992). *Spatial Data Analysis with GIS: An Introduction to Application in the Social Sciences*. National Center for Geographic Information and Analysis, University of California – Santa Barbara, Technical Report 92-10, August.
- Anselin, L., Syabri, I., and Kho, Y. (2010). *GeoDa: An Introduction to Spatial Data Analysis*. *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*: Springer-Verlag, Part 1, 73–89.
- Cochran, W.G. (1977). *Sampling Techniques*: Wiley.
- Heeringa, S.G., West, B.T., and Berglund, P.A. (2010). *Applied Survey Data Analysis*: Chapman and Hall/CRC Press.

Brady T. West  
Michigan Program in Survey Methodology  
Survey Research Center  
Institute for Social Research (ISR)  
University of Michigan  
P.O. Box 1248  
Ann Arbor, MI 48106  
U.S.A.  
Tel: 734-223-9793  
Fax: 734-647-2440  
Email: bwest@umich.edu