Journal of Official Statistics, Vol. 24, No. 2, 2008, pp. 323–338

Book and Software 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

Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey
Safaa Rabie Amer 323
Missing Data: A Gentle Introduction
John Dixon 325
Data Quality and Record Linkage Techniques
Michael D. Larsen 327
The Telephone Interviewer's Handbook: How to Conduct Standardized Conversations
Michael W. Link 330
Margins of Error: A Study of Reliability in Survey Measurement
Stephen M. Miller 333
Register Based Statistics: Administrative Data for Statistical Purpose
Jean Ritzen and Andreas Lindner 334

Roger Jowell, Caroline Roberts, Rory Fitzgerald, and Gillian Eva (eds). *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*. New York: Sage, 2007. ISBN 1-4129-1981-9, 288 pp, £55.

In an era of globalization and as the number of international survey projects keeps growing, there is urgency in addressing the challenges that arise when conducting cross-cultural cross-national studies. Comparative research in this arena requires careful attention to carefully constructed metadata systems. Whether researchers are looking for commonalities or differences, this proves to be a challenging task for many reasons. Most important in that respect is not just achieving high-quality data, but the standardization of access methods and their applicability to make the data user-friendly.

One of the best known books that you will most likely find in the library of survey statisticians who have dealt with international projects is the excellent *Cross-Cultural Survey Methods* (Harkness et al. 2003). A second book that should cross the path of researchers interested in cross-cultural cross-national research is this book on measuring attitudes cross-nationally. However, before you read it, you need to have some background in survey research and sampling because it focuses on hands-on experiences as opposed to theoretical developments. The book provides insights into real-life survey research cross-nationally.

An important characteristic of the book reviewed here and of others on cross-cultural survey methods is the need to unite many authors in order to cover the different aspects of such surveys. This multi-author characteristic entails both pros and cons. One of its major advantages is that it brings up the experience of each author within a specific stage of the survey cycle and allows the authors to highlight their work. A disadvantage is the resulting differences in the level of depth presented by the authors – some providing more details,

formulas, and references than others. For an experienced survey statistician, this would not represent a big challenge, but for a young survey researcher it requires "doing your homework" – seeking more details elsewhere and looking for more references.

I personally found this book to be realistic and educational in many ways. The chapters of the book walk you through the real-life problems faced when dealing with international research comparing nations or trying to carry and adapt methodologies used in one cultural context to a different one. The challenge of standardization and maintaining control over bias and variability arising from these differences among the nations is emphasized.

Chapter 1 provides the reader with some background on the European Social Survey and an overview of the wide range of aspects covered in the rest of the book within the challenging area of measuring attitudes. Chapter 2 carries the reader into inference issues in the European Social Survey based on the sampling methodology, including coverage, frame, sample size, sampling design, design effects, and sample weights. Chapters 3 and 4 cover questionnaire design, content, and translation issues – a nightmare to international survey research if not carefully approached and tested. These two chapters suggest ways to handle problems and reduce bias as much as possible while dealing with a multitude of languages and cultural differences.

Chapter 5 transitions the reader into the longitudinal aspect of the European Social Survey and the fact that it measures social and cultural attitudes across nations which face socio-political swings due to events and happenings at both the national and the international level. Chapter 6 handles differential nonresponse and its effect on the survey quality in terms of data quality and nonresponse bias. The chapter suggests ways to handle field work to deal with noncontacts, refusals, follow-up, and refusal conversion techniques in different cultural settings. Carrying the thought of data quality further, access to the survey data and its metadata are covered in Chapter 7 with an emphasis on the technology and planning needed to reduce the gap between data producers and data users within different regulatory data dissemination environments.

Chapter 8 provides a summary of lessons learned from the European Social Survey as a cross-national effort. It highlights the survey's points of strength such as consistency, transparency, coordination and management, sampling, translation, data access, capacity building, funding, and continuity.

Chapter 9 digs into the content of the European Social Survey as a tool to measure attitudes and represents an effort to explain cultural value orientations. This intended purpose affected not only the design of the study instrument, but would equally affect the analysis of the data for multi-national comparisons. Chapter 10 carries analytical issues further to present some patterns of political and social participation in Europe based on the survey data. Chapter 11, meanwhile, expands beyond comparisons of European countries to compare observed European trends with those observed in the United States General Social Survey.

The complexities lying behind the design of such a large-scale cross-national effort require that analysts and researchers who plan to use the European Social Survey be aware of the challenges and assumptions associated with the survey and its data quality.

This awareness is to be gained not only by looking to the documentation or metadata available through the project website, but also through the chapters of this book. The book is thus recommended to users of the data, even if the users are outside the circle of survey researchers. Users need to be aware of the error and data quality threats and considerations

before attempting to use the data and interpret the results. Furthermore, the book is highly recommended for those who are planning to design other cross-national surveys and aiming to follow good practice.

Stressed throughout this book are standardization, documentation, and quality control – three major driving forces contributing to the success of such a huge cross-national data collection effort. These driving forces cannot be achieved, of course, without good planning, good management, and the availability of adequate funding!

The book is yet another step taken by a group of researchers on cross-cultural survey methods who are planning an upcoming event which is expected to result in another publication in the area of multi-national multi-cultural survey research. This event is the 2008 Conference in Berlin on Multicultural, Multinational, and Multiregional Surveys. I admire their continuing efforts to raise awareness, encourage research, and unite researchers interested in cross-cultural surveys to share experiences and learn from each other.

References

Harkness, J., Van de Vijver, F., and Mohler, P. (eds.) (2003). Cross-Cultural Survey Methods. New Jersey: Wiley.

Van de Vijver, F. and Leung, K. (1997). Methods and Data Analysis for Cross-Cultural Research. Thousand Oaks, CA: Sage.

European Social Survey Documentation http://www.europeansocialsurvey.org/

International Conference on Survey Methods in Multinational, Multiregional, and Multicultural Contexts http://www.3mc2008.de/

Safaa Rabie Amer

National Opinion Research Center (NORC)
At the University of Chicago
Department of Statistics and Methodology
& International Projects Department
4350 East West Highway
Suite 800
Bethesda, MD 20814
U.S.A.
Phone: 202-223-1623
Fax: 202-223-6104
Email: amer-safaa@norc.org

Patrick McKnight, Katherine McKnight, Souraya Sidani, and Aurelio Jose Figueredo. *Missing Data: A Gentle Introduction*. New York: Guilford Publications, 2007. ISBN 1-59385-393-9, 251 pp, $38 USD.

This very readable text is aimed at researchers and research courses in the social sciences. It draws on ideas in several different fields, including psychology, testing, and evaluation. Examples cited throughout the text are from a wide variety of disciplines such as nursing, psychology, and education.

The first chapter provides an overview of missing data issues and describes the organization of the book. The second chapter sets the stage for the remainder of the book with a discussion of the consequences of missing data in terms of different biases. The third chapter describes classification systems for missing data, principally based on Rubin's 1976 Missing completely at random (MCAR), Missing at random (MAR), and Missing not at random (MNAR) distinctions. Missing data mechanisms are also discussed. In Chapter 4 the implications of experimental design with regard to planning for missing data is central to the discussion. Testing of data collection procedures and many aspects of training are discussed. Chapter 5 covers diagnostic procedures, although complete enough for the intentions of the book, proscriptions as to utility and problems with the procedures are deferred to later chapters. Chapter 6 has step-by-step procedures for assessing the missing data problems, and then continues with decision making objectives based on the analysis.

Chapter 7 focuses on the data deletion methods for handling missing data. The authors also discuss specific designs with an emphasis on growth curves. Chapter 8 covers data augmentation methods including Maximum Likelihood (ML), Expectation Maximization (EM), Markov Chain Monte Carlo (MCMC), dummy coded models (shown to be biased), and weighting. Chapter 9 describes and warns about single imputation methods, leading into the next chapter on multiple imputation. The theory of multiple imputation is described, and a worked example is explored. Reporting missing data is covered in Chapter 11, with an emphasis on the American Psychological Association Task Force on Statistical Inference. The authors expand on the task force's recommendations on the basis of their experience. They give examples for each of the stages of a typical study (recruitment, enrollment, treatment assignment, data collection, follow-up, and data analysis). They recommend going beyond just reporting the amount of missing data and emphasize the potential impact on the conclusions of the study.

There are few weaknesses in the text, and most of them are mentioned by the authors as limitations in scope. Guidance on the different methods is often too gentle: e.g., listwise deletion (Chapter 7) could be more consistently warned about, since it is often biased and is the default of so many statistical packages. The authors provide a good explanation of how the methods can be used, but many social scientists would appreciate a "best practices" recommendation for different problems.

While the book covers a great deal, topics outside the scope of the book could have been mentioned. In addition to the data augmentation methods (Chapter 8), propensity scores could have been discussed. One of the desirable aims of the book was to provide practical proscriptions for avoiding the problems associated with missing data (Chapter 4). However, many of the methods not mentioned fall outside the scope of an introductory book, so providing "best practices" based on the material presented was really not possible. A brief mention of those additional techniques for including missing data in the experimental design would have given researchers an idea of some potential solutions to difficult missing data problems. Some of the methods could be follow-up (e.g., sampling the nonresponse in a survey), alternative measures (e.g., indirect measures of weight, where weight may be difficult to measure or often refused), calibration samples (e.g., selecting an additional sample of known characteristics and using it to adjust for the missing data in the experimental sample, such as in dual frame sampling), using measures

in the study which would be useful in developing weights from a benchmarking sample (e.g., asking some of the same questions from a larger survey, such as the Current Population Survey, the American Community Survey, a survey more related to the study's topic).

Although this volume could have contained more, it is a welcome introduction to the topic of missing data. While other texts on research methods mention missing data problems, none combines the readability and scope that this text offers. This would be a useful text in a social science research course or for any researcher who encounters missing data. The authors also have a bibliography of additional reading available as a link from the publisher's web site.

John Dixon

Office of Survey Methods Research
U.S. Bureau of Labor Statistics
Washington, DC
U.S.A.
Phone: 202-691-7516
Email: Dixon.john@bls.gov

T.N. Herzog, F.J. Scheuren, and W.E. Winkler. *Data Quality and Record Linkage Techniques*. New York: Springer, 2007. ISBN 978-0-387-69502-0, 227 pp, 44.95 USD.

Data quality has many dimensions. It could include measurement error, response error, processing error, and issues of nonresponse. In order to improve data quality, one can consider psychological dimensions of asking a question, the larger survey process, or the scientific basis of physical and biological measurements. A statistical perspective would stress the level of bias and uncertainty associated with estimates of well-defined population quantities. Searching the Internet for "data quality" produces links to numerous web sites, including a reference in Wikipedia. Searching an on-line bookseller such as Amazon produces numerous books on the subject. Based on reviewing the book *Data Quality and Record Linkage Techniques* by Herzog, Scheuren, and Winkler in detail and examining in a much more cursory manner the other information on data quality, it seems that the book that is the subject of this review is unique in its focus and coverage of topics. Indeed, the authors are unique in their expertise.

Winkler is the director of the record linkage research staff at the U.S. Census Bureau and a pioneer in probabilistic record linkage and data editing. Herzog is the chief actuary of the U.S. Department of Housing and Urban Development. Scheuren was the 100th president of the American Statistical Association and has wide-ranging experience in government statistics and survey research. All authors have had extensive involvement in very large-scale administrative and survey data collection processes. The book is strong in their areas of expertise and specialties. It is not a comprehensive book about all potentially relevant statistical connections. It could be considered for supplemental reading in a course on data quality and uses of administrative data. The book is a "primer" concerning the more statistical aspects of data quality, but it does investigate some areas, especially record linkage, in more depth.

The book is divided into four parts. Part I concerns data quality: what it is and how to measure it. Part II considers three sets of techniques for improving data quality: editing, imputation, and record linkage. The greatest portion of Part II concerns record linkage and specialized topics related to it. Part III provides case studies of record linkage and data quality. Part IV discusses statistical software and the issue of privacy in the context of the ability to potentially link records to available sources of information. Chapter 20 in the book provides a summary of individual chapters and serves as a useful starting place in addition to the table of contents. Examples from the authors' considerable experience are used throughout the text. Although the examples are primarily from the U.S., the principles and details of methods should be of interest generally to *JOS* readers from industry, government, and academia in many countries.

The authors state that there are three purposes for which the methods and techniques they describe are intended: improving data quality, merging two lists to create a larger list while removing duplicate listings, and merging two databases on the same people so that more variables can be utilized in a single analysis. Regarding the first topic, in order for data to be of high quality, data must "conform to standards" and be "fit for use." The authors discuss five database properties (more are possible) and give practical advice about achieving them. The authors clearly state that preventing poor data quality is strongly preferable to detecting it and then taking steps to "repair" the data. Chapter 3 of the book provides several brief examples from business and government. The examples illustrate advantages of good data as well as disadvantages of bad data and that quality issues can arise in computer storage of data and in analysis and interpretation of data in addition to the basic task of recording data elements. This interesting section of the book is readable by a general audience and provides some useful metrics. It would have been interesting to see detailed comparisons of the metrics in some applications rather than a list of metrics. References to sources on data quality from an organizational and management standpoint are provided in the book's preface.

One who is familiar with imputation for missing data (e.g., Little and Rubin 2002) will find the imputation ideas in this book familiar. The purpose of the imputations in the context of editing and imputation, however, could seem a little bit different than their use to address nonresponse. The focus on data quality in this regard is rather unique. Editing here is akin to outlier detection and logical error prevention. The goal is to have a unified method for checking data, editing data, and imputing alternative values if edit checks are violated (Fellegi and Holt 1976). One naturally wants to change as few values as possible to produce consistent data. Winkler's (2003) procedures have taken this idea as far as anyone has achieved in practice. The authors briefly review standard methods of imputation of missing values and software for editing and imputation. The chapter includes some brief examples and practical advice.

Record linkage is the activity of identifying records that correspond to unique individuals in two or more databases. The challenge arises when data files are very large, there are no unique keys common to all databases on which to exactly match individuals, and information is recorded with possible errors. Fellegi and Sunter (1969) produced a mathematical/statistical model of record linkage that has been used to motivate work on this topic (e.g., Winkler 1995; Alvey and Jamerson 1997; Lahiri and Larsen 2005; and references therein). There is not really a sufficient textbook on record linkage, so Herzog,

Scheuren, and Winkler (2007) is an important contribution on this topic. Newcombe's (1988) book is an important reference in this regard, but it is out of print.

The chapters on record linkage provide details on general methodology, estimation of unknown parameters useful in models for record linkage, and several operational topics (standardizing and parsing fields of information, phonetic coding of names, blocking files into smaller files for computational convenience and better operational performance, and string comparator metrics for comparing character fields). The chapter on editing and imputation and the first two chapters on record linkage utilize ideas found in a mathematical/statistical review chapter and are more technical in that sense than the rest of the book. The review chapter generally is useful and will help a reader with some statistical background connect to some of the topics. If one wants to understand modeling and computational details, then one will need to study material in the cited references. As mentioned, this book serves as a primer for larger areas of study. In the chapters on record linkage, one will note the number of references to work by Winkler and colleagues at the U.S. Census Bureau. Examples are from Census, HUD, and other applications. The chapters on record linkage are complemented by more detailed examples in Part III of the book. The more extensive examples provide a better context in which to judge the contribution of the various statistical procedures.

The book concludes with thoughts on the issue of confidentiality and a brief description of record linkage software. Confidentiality protection (i.e., disclosure limitation, disclosure avoidance, inference control) is a hot topic in statistical and research circles. The goal of providing useful data for research and policy evaluation and simultaneously protecting the privacy and identity of individuals represented in administrative databases and responding to surveys is truly challenging. In this book, one will find a primer on procedures that are either used or that have been proposed and discussion of how record linkage methods should play a role in evaluating the level of confidentiality protection provided by various nondisclosure methods. The discussion of record linkage software can be viewed as providing pointers for further investigation by individuals seriously interested in large-scale record linkage.

In conclusion, the book covers a lot of ground on data quality, editing and imputation, and record linkage, and provides unique insight based on experience. As it is a primer on these subjects, there is not as much detail on some topics as one (at least, this reviewer) would have liked. Understandably it is inherent in the challenge of writing for both a statistical and a general audience on a broad topic. More detail on examples throughout the book and comparison of methods, such as data quality metrics, would have been interesting. Nevertheless, the authors have made a commendable effort and produced a unique contribution. It should motivate study in several areas and help many organizations improve their data quality.

References

Alvey, W. and Jamerson, B. (1997). Record Linkage Techniques – 1997, Proceedings of an International Workshop and Exposition. Washington, DC: Federal Committee on Statistical Methodology, Office of Management of the Budget, http://www.fcsm.gov/working-papers/RLT_1997.html.

Fellegi, I.P. and Holt, D. (1976). A Systematic Approach to Automatic Data Editing. Journal of the American Statistical Association, 71, 17–35.

Fellegi, I.P. and Sunter, A.B. (1969). A Theory for Record Linkage. Journal of the American Statistical Association, 64, 1183–1210.

Lahiri, P. and Larsen, M.D. (2005). Regression Analysis with Linked Data. Journal of the American Statistical Association, 100, 222–230.

Little, R.J.A. and Rubin, D.B. (2002). Statistical Analysis with Missing Data, Second Edition. New York: Wiley.

Newcombe, H.B. (1988). Handbook of Record Linkage: Methods for Health and Statistical Studies, Administration, and Business. Oxford: Oxford University Press.

Winkler, W.E. (1995). Matching and Record Linkage. In Business Survey Methods, B.G. Cox, D.A. Binder, B.N. Chinnappa, A. Christianson, M.J. Colledge, and P.S. Kott (eds). New York: Wiley.

Michael D. Larsen

Iowa State University
Department of Statistics and Center for Survey Statistics & Methodology
Snedecor Hall
Ames, IA 50011
U.S.A.
Phone: 515-294-5189
Fax: 515-294-4433
Email: larsen@iastate.edu

Patricia A. Gwartney. *The Telephone Interviewer's Handbook: How to Conduct Standardized Conversations*. San Francisco, CA: Jossey-Bass, 2007. ISBN 978-0-7879-8638-4 (pbk.), 315 pp, $25.00.

The *Telephone Interviewer's Handbook* is a well-written, succinct volume which stands out against the sparse resource landscape of interviewer training materials. Gwartney's volume provides research interviewers with a solid understanding of the survey process including a useful (but not overblown) amount of theory on survey error and potential bias, some insights on survey workplace considerations, and a good dose of essential interviewer skills. She provides a unique perspective, writing directly to interviewers in the voice of an interviewer trainer. The knowledge and skill building lessons are centered on three primary goals: collecting the highest quality telephone survey data, minimizing interviewer-related error, and avoiding bias. Interviewers are the focus here, serving as the front line of any telephone survey research endeavor, and as such are key drivers of the end quality of the data collected. An underlying theme of the book is that it does not matter how much money, time, and planning go into a survey if the interviewers are poorly trained and unprepared for their task. The volume is punctuated with clear, well-formulated examples. As in the case of many resources of this type, however, there are some areas which are of greater utility than others.

Chapter 1 provides a solid foundation on the survey process itself, emphasizing roles, process, and errors, including cogent explanations of random error versus bias, and the

total error perspective. Chapter 3 covers professional ethics (including a reprint of the American Association for Public Opinion Research's Code of Professional Ethics and Practices) and the important role of professional organizations in the field. Gwartney does a good job providing interviewers with a look at how their role fits within the broader survey context. Calling is covered in Chapter 5, including what to expect on a call, how to initiate a successful call, basic call outcomes or dispositions, and specialty skills such as leaving messages on answering machines and scheduling callbacks.

Chapters 6 and 7 are the "meat and potatoes" chapters, focusing on the "must know" skills for effective interviewing. Chapter 6 provides background on the critical nature of an interviewer's voice – what is said and how an interviewer says it – as a determinant in the decision by respondents to participate in a survey. Interviewer skills such as the importance of memorizing introductory scripts and learning to respond firmly and positively to respondents' questions are examined. The book makes clear that the initial seconds of each call are what will often make or break the entire interview. Learning to engage respondents and deftly respond to questions is what separates successful interviewers from unsuccessful ones. The author also focuses on the importance of "active listening" – that is, paying close attention to the verbal cues a respondent is sending and responding to these appropriately. The chapter also provides a nice set of examples and techniques for responding to reluctant respondents and converting those who may have initially refused to participate. One weakness in these sections is the lack of discussion on differences that may occur with bilingual (primarily English/Spanish) interviewing. Given the growing diversity of society, particularly in the U.S. (which is the primary focus for this handbook), a better understanding of potential differences between English-only and bilingual interviewing would be instructive.

Interestingly the book is also strangely silent on how interviews conducted over cell phones differ in terms of process, procedures and techniques from those conducted over landline telephones. This is a critical skill set which needs to be developed as the use of cell phones continues to proliferate in the U.S.

Chapter 7 deals with techniques for standardized interviewing, reviewing how different types of questions (behaviors, attitudes, beliefs, knowledge, and demographics) each have their own peculiarities in terms of question administration and probing for complete responses. The author provides the reader with some useful background on how respondents' thought processes work for different types of question stimuli. The bulk of the chapter focuses on critical skills in standardized interviewing: verbatim question administration, following specified question order, use of optional language in question administration, remaining neutral, using positive feedback, and keeping respondents on task. There is a nice section on probing techniques with some very useful examples of categories (racial/ethnic demographics and occupational groups) which are notoriously difficult to probe. The glossary is also a very useful reference.

In writing a volume directly to interviewers, however, the book stumbles in some places, offering content which is perhaps better suited for call center supervisors or directors than for the interviewing rank-and-file. This becomes distracting in the sense that some content and approaches advocated are debatable and as such might not be the type of content a supervisor would typically like to provide an interviewer with. Some of the content it would make sense to include if a prospective interviewer were buying this volume off the shelf for themselves, but the reality is that as a handbook for such a specialized profession

the volume is most likely to be purchased by an employer for distribution to trainees. As such the discussion in Chapter 2 concerning what makes a desirable versus less desirable employer seems misplaced. Putting a list of desirable employers by name in the Chapter automatically dates the book as newer, research-based calling centers are established. Additionally, while unwritten, there is a clear implication that if a company is not listed, then the company must not be a desirable employer. Likewise, the section on what employers do and do not need from interviewers also takes a fairly parochial view. For example, the author notes: "A telephone interviewer's appearance makes no difference in his or her ability to be an asset to the organization" (p. 46) – and as such should not be of concern to an employer. While it may be true that there is little relationship between how one dresses and one's interviewing skills on the telephone, the statement misses subtle, yet important points which should also be discussed such as the need to foster an environment of professionalism for client visits and for minimizing potential disruptions in larger call centers (strange as it may seem, minimal dress policies are essential for smooth operation in larger shops). Additionally, many new or smaller survey shops, which have the goal of performing good scientifically valid surveys may be dismayed to find out that some of the attributes of "less desirable" employers include having a small number of individuals run the entire organization (how many university shops were started with just one or two people heading the call center – quite a few!), the owner or senior interviewer supervises the interviewers, and being open for less than one year.

There are similar concerns with Chapter 8, which covers "what to expect in the survey workplace" – again an area where to the degree to which an employer's policies differ from those advocated in the handbook, these sections can be problematic. Much of this material is better suited for a supervisor's handbook (or supplement to this volume), such as setting up a good workplace environment and workplace routines. Sections on "benefits and wages" can quickly date the volume as well, such as the example of interviewer wages drawn from a survey of research organizations conducted in 1998. The chapter does contain sections on quality control and monitoring, but these would have been better as a stand alone chapter, providing interviewers with more detail on why these processes are in place and what to expect in terms of this critical task.

In sum, the volume contains some very valuable chapters which should be readily adopted by those organizations conducting scientifically-based standardized surveys. It also, however, contains several chapters which could prove distracting within a handbook for interviewers, particularly when the topic matter strays from its primary focus on conducting standardized interviewing.

Michael W. Link

Chief Methodologist
The Nielsen Company
1145 Sanctuary Parkway, Suite 100
Alpharetta, GA 30004
U.S.A.
Phone: 770-777-4287
Email: Michael.Link@Nielsen.com

Duane F. Alwin. *Margins of Error: A Study of Reliability in Survey Measurement*. New York: Wiley-Interscience, 2007. ISBN 978-0-470-08148-8, 408 pp, $89.95.

Dr. Alwin presents a comprehensive discussion of measurement error in surveys, which will be of interest to both the experienced researcher, and someone who is new to the field. The overall approach of the book is to present a unified theory of measurement error from essentially a psychometric perspective, and apply it to the survey context – something not unique to this book but which serves as an accessible overview for the interested reader. The book then discusses statistical estimation techniques, which can be used to estimate reliability under a variety of assumptions based on the previously discussed theoretical models. Next, the major part of the book presents a detailed empirical investigation of measurement error, and reliability measures in particular, using 500 survey measures taken from more than 20 years of work at the University of Michigan. A summary of the findings and a review of future directions in research completes the book. I next present some detailed comments on specific sections of the book.

The first two chapters review the nature of reliability, particularly with its inherent link to validity, and present a compelling discussion of why we should be interested in measurement error. A further discussion of the sources of measurement error is framed within the context of six critical elements of the response process: question validity, comprehension, accessibility, retrieval, motivation, and communication. Chapters 3 and 4 then launch into a detailed discussion of Classical True-Score Theory (CTST) and Multitrait-Multimethod (MTMM) approaches, with particular attention to *repeated measures*, *multiple indicators*, and *multiple measures*. The author does a good job of using path diagrams to explain the theoretical models, and presents the ideas of measures and indicators in a very understandable way, especially for those not well versed in psychometric methods. While the author is careful to define his notation, I did find it a little awkward at times, especially for someone who is used to mathematical statistics. Having said that, the notation does not get in the way of understanding the concepts. Chapter 5 presents a detailed discussion of longitudinal estimation techniques. Alwin presents a good discussion of the inherent problems with test–retest methods, and gives a very helpful discussion of Quasi-Markov models. Chapter 6 discusses the empirical analysis which is the main subject of the remainder of the book. A careful categorization of the estimation techniques as well as a careful description of the data sources made this a particularly useful part of the book. Chapters 7 through 10 present reliability estimates, and general conclusions, for the following areas: The Source and Content of Survey Questions, Survey Question Context, Formal Properties of Survey Questions, and Attributes of Respondents. A few of the specific topics covered include: reliability for factual questions, stand-alone questions versus questions in series or battery, open-ended versus closed-ended questions, attributes of scales, and the effects of age and education of respondents. While a survey researcher with a particular area of interest might be tempted to focus on one chapter and skip the others, I found that reading all four chapters together was beneficial since they tend to build upon one another. Chapter 11 changes gears slightly, and discusses reliability estimation for categorical latent variables. This was a very useful chapter, although thematically I might have preferred to see it presented earlier in chapter three, four or five

– but this is purely stylistic. The final chapter does a nice job of summarizing the book and outlining directions for future research.

Overall, I highly recommend the book. I think it is best suited for someone who is acquainted with either psychometric theory or structural equation modeling applied to surveys. It would be well suited as a reference book for graduate students or advanced undergraduates, and will undoubtedly be a well-referenced resource for many years to come.

Stephen M. Miller

U.S. Bureau of Labor Statistics
2 Massachusetts Avenue NE, Room 1950
Washington, DC, 20212
U.S.A.
Phone: 202-691-7379
Email: miller.steve@bls.gov

Anders Wallgren and Britt Wallgren. *Register Based Statistics: Administrative Data for Statistical Purposes*. Hoboken, NJ: Wiley, 2007. ISBN 978-0-470-02778-3, 258 pp, $130.

The use of administrative registers is a very important issue today, on both policy and statistical agendas. The main reasons are the need for a reduction of the administrative burden at enterprises on the one hand and the possibilities of attaining greater efficiencies on the other. During the last decade, considerable research progress has been reported in this field. National Statistical Institutes (NSI's) are moving from traditional single methods into mixed mode methods in which statistical surveys' data collection and register-based surveys are combined. One of the big challenges is to make an optimal and valid translation from administrative concepts into statistical ones for units and variables. Last but not least, IT developments accelerate this process.

Management of very large databases is no longer a barrier. But the change implies the necessity of adapting methodological approaches too. Traditional methodologies of sampling, editing, imputation and grossing up of statistical surveys must be complemented with methodologies related to the use of administrative data of registers in statistics. This is increasingly a core issue of many statistical seminars and conferences. An example is the last International Conference on Establishment Surveys, held in Montréal in June 2007. The Nordic countries (Denmark, Finland, Norway, and Sweden) already have long experience of register-based statistics, although many other countries also have this without being explicitly aware of it.

Anders and Britt Wallgren have done considerable research in this field, resulting in the book under review. This book seems to us to be a long overdue "systematisation" of the need to integrate hitherto separated, individual registers into a consistent set of registers, ultimately – and ideally – leading to an integrated national design of how to use (almost) all available information that can be used for the compilation of statistics. Clearly addressed to national data compilers, it addresses an impressive range of statistical design

issues. We were impressed by the scope and detail of the guidance provided. The book is very broad in scope, broader than the title suggests. After an introduction, many aspects of register-based statistics are dealt with. But some more general aspects are also dealt with, e.g., the protection of privacy and confidentiality. These do not much differ from practices in traditional (census) surveys. In the book, topics are well explained, with a large number of examples and illustrations.

Chapter 1 introduces the aim and content of the book. Four principles regarding how to use administrative data in a statistical register system are introduced. The need for the use of administrative data for statistics is clearly explained. But although other administrative sources are used, NSI's remain fully responsible for the statistical outputs. Advantages and disadvantages of the use of register data are summarised in relation to surveys based on direct data collection. The overview provided is very good and prepares the reader for the following chapters.

Chapter 2 provides highly relevant insight into design issues related to the statistical register system and the underlying administrative sources. Nevertheless, it is – unsurprisingly – based upon the state of Scandinavian registers, which internationally are considered to be good. A more generic discussion of register content would have increased its usability for non-Scandinavian data compilers. The use of the model depends heavily on the national administrative organisation and legislation. The base registers are the core of the system. Therefore, maintenance of these base registers is a concern of all to attain coordination and coherence as explained in Chapter 13. For the creation and maintenance of a statistical register system the existence of unique identification numbers of the basic objects is important for matching data from different registers. Otherwise, other advanced matching methods must be used, which always results in mismatches to be resolved and analysed. These mismatches easily lead to poor quality of the system. Matching of objects and standardising of variables are core activities in the system. The presented business register model (p. 26) is in line with EU regulations on statistical units and statistical business registers. Unfortunately, a reference to the EU regulations is missing.

In Chapter 3 the reader is taken on a journey through the terminology forest – fundamentally unavoidable to prepare the reader to understand common national and international terminology. The authors have gratified the reader by adding a glossary of terms and definitions at the end of the book – a very good idea. The Chart 3.17 addresses well the key issue in register-based statistics: integrate, compile without having to conduct additional surveys, and hence, create new information from different sources. The main technical aspects are importation of data, translation of the data through the filter, and exportation of the results. How traditional sample surveys and register-based surveys can benefit from each other and be complementary to each other is dealt with in Chapter 4. We mention here the role sample surveys can have in register quality assessment, the advantage of the register survey in small area statistics and the strength of the combination in the compilation of virtual censuses (elucidated, for instance, with the Dutch virtual population census 2001).

Chapter 5 can be considered to be the core of the book, well introduced in the previous chapters. It is well-structured and useful for guidance. The steps in the procedure of creating a statistical register are described, but what is missing is the starting point: the statistical program. For economic statistics in the EU this program is mostly determined

by EU regulations like "PRODuction COMmunautaire" (Prodcom), Structural Business Statistics (SBS) and Short Term Statistics (STS), which also influence the research objectives. Quality assurance and documentation of all steps are emphasized.

Chapter 6 continues with the completion of the variables in the statistical register. Editing is also an important activity here. It is mentioned that in the case of invalid values corrections are made in consultation with the administrative authority (p. 103). This seems to us to be a somewhat tricky statement for confidentiality reasons. Many methods for imputation of missing variables are described. The authors rightly stress impediments to longitudinal analysis; there are now ways of tracing and this should perhaps be developed a bit more (but they take a bit too narrow a perspective). They acknowledge the need for longitudinal analysis (e.g., business demography) and this is a strong aspect of this book, responding to a crucial policy need (for example, how to create jobs from the lower end of the spectrum, namely micro, small and medium enterprises). This also makes the introduction of the time dimension in the register so crucial in addition to the necessity of the construction and reconstruction of derivations from the register.

Chapter 7 deals with estimation methods needed because of missing values in register-based statistics. Methods are described and illustrated with clear examples (although an error in the x_{4i} column of Chart 7.10 (p. 123) is confusing for an inexperienced reader). The calibration of weights in register-based surveys is the most complex issue treated in this book.

Chapter 8 focuses on calibration and imputation of missing values, similar to item nonresponse in surveys. It is an important value added to the statistical process that missing values in registers are eliminated by estimation so that "unknown" classes and/or categories will be avoided as much as possible. The techniques are similar to those in traditional statistical surveys. Special attention is given to this phenomenon in time series, where we are confronted with missing values or item nonresponse varying over time. Object (population) undercoverage can only be solved by searching additional sources. Overcoverage can be corrected by estimation using sample surveys or other sources. The strength of the register system is that the objects in different base registers can be combined and linked to each other. This can lead to a huge enrichment or extension of statistical information.

Chapter 9 concentrates on the value of the activity register in the case of many to many relationships of objects. Also the related potential errors are dealt with. A very interesting approach, the so-called functional statistics approach, is presented, although one can question the business statistical unit types used. But this approach also has its limits. We must be careful in generalising this, because many, and especially many economic statistics, have an institutional basis in EU or other regulations. Furthermore, the application of this kind of functional approach requires very high quality of the base registers.

A separate Chapter 10 would have been useful on quality aspects of register-based statistics. Quality needs attention in all stages of the processes. Metadata and documentation are key words. Timeliness and stability of the content (objects and variables) of registers is important because good management of relationships with administrative sources is essential, not only at the operational level. A number of quality characteristics of a register are mentioned, but above these consistency and co-ordination

are important, both nationally and internationally. Fitness for purpose is the basic quality indicator. The treatment of errors could have been moved to another, more technical chapter.

The discussion of metadata that is now part of Chapter 11, combined with IT aspects, could preferably have been part of the quality chapter. The objective of reducing redundant metadata means that metadata on the same item should be documented and stored only once within a clearly structured metadata framework. Central databases of classifications and definitions prevent different local versions. It is important to have and maintain an overall view of the register system. IT-related, this is a kind of Statistical Warehouse which requires good management and handling of very large databases.

The protection of privacy and confidentiality, the topic of Chapter 12, is always associated with statistics. In addition, in a statistical register system a large amount of combined information on individuals is available. This leads to the need of very strict measures to prevent disclosure, not only in the statistical output but also internally in the statistical offices. Some measures are the minimising of the access to the databases and the recoding of the personal identification numbers (PIN's) in the databases in combination with eliminating text information as much as possible. Requests for analysing micro data should be strictly regulated, e.g., by giving permission to do this only onsite and using anonymised databases.

In Chapter 13 co-ordination and coherence are dealt with. Although this is nearly the last chapter, these aspects must be the starting point in the process. Standardisation of populations and variables is a basic requirement, as is consistency within the system.

Strategic design issues require top-down approaches and not bottom-up ones. The topic should be included at the start, not be the closing issue at the end of the process. Coordination and coherence are not restricted to the national level but are growing in importance in the international context, too.

The last chapter, 14, offers the authors' conclusions. The advantages of a statistical register system are briefly summarised. Also the main differences in error approach are mentioned once more. The finishing paragraph, "What more is needed," gives a warning on the use of and dependency on the growing number of private register suppliers. For the sake of data integrity and resource-efficiency, NSI's should be in charge of new key areas of data collection, probably requiring a sustained effort.

In conclusion we can say that this book is very valuable for statistical institutes that are introducing the use of administrative data as a supplement to (or substitute for) existing data collection. In this context we do not need a theory but an internationally applicable set of key recommendations to ensure that this can be translated into national reality – a handbook of good practice. The book must be considered to be a "good practice manual or handbook for national data compilers." Because of the different national administrative systems and structures, the applications must be adapted to national properties. We would suggest more attention to the international dimension – a prerequisite in a globalised world. Statistical systems get more and more embedded in national and internationally integrated information systems. This leads to considerable mutual dependencies and relationships which have to be managed.

We are convinced that this book significantly contributes to further discussions on and developments in the use of administrative data for statistical purposes in a well-founded

methodological way. It is not possible to be fully exhaustive in describing the issues related to the statistical register system. Therefore not all comments on missing aspects must be considered to be negative criticism. But we are convinced that it contributes to the further elaboration of this important subject.

Reference

UNECE (2007). Register-based Statistics in the Nordic Countries; Review of Best Practices with Focus on Population and Social Statistics. (ISBN 978-92-1-116963-8).

Dr. Jean Ritzen

Statistics Netherlands
National statistical services
Kloosterweg 1
6412 CN, Heerlen
The Netherlands
Email: jrtn@cbs.nl

and

Mr. Andreas Lindner

OECD
Statistics Directorate
2, rue André Pascal
75016 Paris
France
Email: andreas.lindner@oecd.org