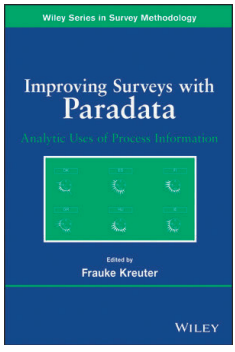

Frauke Kreuter (Editor) (2013):
Improving Surveys with Paradata. Analytic Uses of Process Information
Wiley Series in Survey Methodology
ISBN: 978-0-470-90541-8
416 pages
€65.70

Improving Surveys with Paradata. Analytic Uses of Process Information by Frauke Kreuter is an edited volume containing timely contributions on the use of paradata in survey research, survey management and data analysis. Due to its wide scope and high quality, the book provides helpful information to survey methodologists, survey managers and analysts on how to make valuable use of this young data source.

Kreuter defines "paradata as additional data that can be captured during the process of producing a survey statistic. Those data can be captured at all stages of the survey process and with very different granularities." (Kreuter, 2013, p. 3). In other words, paradata are data collected during the survey data collection process either automatically through computerized systems or individually by interviewers. The book covers a great number of different types of paradata, the most central examples of which are keystrokes, time stamps, call record data and interviewer observations.

Different types of chapters make up this edited volume: reviews of existing studies using paradata to investigate survey errors, applications of paradata in the context of survey production and contributions on the measurement properties of paradata. Finally, the book draws attention to possible applications that still lack empirical evidence, thus formulating a future research agenda.

While some chapters are pitched at an introductory level looking into data structures of paradata and addressing data preparation concerns, other chapters target a statistically advanced readership and cover more complex models for the analysis of paradata. Although these complex models are illustrated by means of mathematical formulas and therefore demand corresponding reading abilities, these formulas are derived and explained step-by-step, thus enabling even less experienced readers to comprehend them.

The book comprises a total of 15 chapters, the content of which is briefly reviewed in the following. In Chapter 1 the editor provides an introduction by defining para-

data and relating them to metadata and auxiliary data as well as an outline on the structure of the book. The remaining chapters are divided into three parts each focusing on the use of paradata in a different setting.

The first part *Paradata and Survey Errors* includes four chapters that address the use of paradata to investigate and adjust for errors occurring at different stages of the survey process. More precisely, chapters 2, 3, 4 and 5 focus on three error components taken from the total survey error framework (see Kreuter, 2013): nonresponse, measurement and coverage error. For each of these error components the aforementioned chapters name relevant paradata and inform about previous research including these data.

In Chapter 2, Kreuter and Olson give a review of the kinds of paradata that have been used to analyze nonresponse error to date. Furthermore, this chapter discusses the possibilities of including paradata in nonresponse adjustments. Chapters 3 and 4 both deal with paradata and measurement error. In Chapter 3, Olson and Parkhurst describe which paradata are important for assessing measurement error and how they differ by mode of data collection. In Chapter 4, Yan and Olson report studies using these paradata to investigate measurement error, therefore supporting the findings of the previous chapter with empirical evidence. Finally, in chapter 5 Eckman focuses on coverage error. For this purpose, she gives an overview of existing research with paradata on frame errors and coverage bias, and points to additional possibilities of using paradata to study coverage error.

The second part *Paradata in Survey Production* consists of five chapters demonstrating how paradata may be used either after data collection to correct for survey errors or directly during survey data collection to guide changes in survey design to increase survey quality. Although these five chapters each consist of individual case studies, the findings can easily be transferred to other survey settings.

In chapter 6, Kirgis and Lepkowski give insight into design changes in the 2006-2010 National Survey of Family Growth based on paradata from a previous data collection. Paradata from the ongoing survey were then used to monitor these design changes and informed responsive design elements. In chapter 7, Wagner reports a number of studies carried out to optimize calling strategies. For this purpose, he includes paradata into models estimating contact probabilities. In chapter 8, Sakshaug summarizes the use of paradata to study nonresponse regarding additional requests within surveys, such as consent to record linkage. In addition, he points out how paradata may be used to intervene on respondents probably not giving consent to those requests. In chapter 9, Jans, Sirkis and Morgan show that paradata can play a decisive role in monitoring the quality of ongoing data collections; more specifically, they demonstrate how they are implemented in graphical control displays. In chapter 10, Schouten and Calinescu present original research

into paradata as predictors of nonresponse and measurement error using data from the Dutch Labour Force Survey.

The third part *Special Challenges* contains five chapters dealing with the measurement properties of paradata. The chapters in this part cover mode-specific paradata, demonstrate methods for analyzing paradata with particular data structures and focus on quality issues.

In Chapter 11, Callegaro provides a typology of paradata that can be collected in web surveys. Furthermore, he gives examples of how these paradata have been used thus far. Chapters 12 and 13 both deal with analytical challenges resulting from the data structure. In Chapter 12, Durrant, D'Arrigo and Müller describe the use of multilevel modeling to analyze hierarchically structured call record data. In Chapter 13, Schafer shows how penalized spline models can be employed to monitor quality indicators with means shifting over time. The two remaining chapters 14 and 15 both focus on the quality of paradata. In Chapter 14, West and Sinibaldi provide a literature review on the quality of various paradata and discuss potential error sources. In Chapter 15, West demonstrates by means of several simulation studies how erroneous paradata may affect nonresponse adjustments.

Each chapter of the book is self-contained. Thus, the reader may choose to read the book from front to back or just individual chapters without missing context. To the reader's convenience, there are many cross-references linking the individual chapters.

In conclusion, *Improving Surveys with Paradata. Analytic Uses of Process Information* provides a comprehensive overview of the possibilities for using paradata in survey methodological research and field management. The book illustrates a variety of areas paradata may contribute to, including advances in academic research on and with paradata as well as more survey practical settings, such as during fieldwork monitoring. As research into paradata is relatively new there is still a need for further analyses. In this way, the book valuably summarizes the current state of research. Whether researchers in academia or survey practitioners, whether experienced or just starting out, we can highly recommend this book to anyone interested in integrating paradata in their daily work.

Literature

Kreuter, F. (2013). Improving Surveys with Paradata: Introduction. In F. Kreuter (Eds.), *Improving Surveys with Paradata. Analytic Uses of Process Information* (pp. 1-9). Hoboken, New Jersey: Wiley.

Franziska Gebhard & Annelies G. Blom
DOI: 10.12758/mda.2015.006