DE GRUYTER
OPEN

Journal of Official Statistics, Vol. 34, No. 3, 2018, pp. 787–789, http://dx.doi.org/10.2478/JOS-2018-0037

Book Review

Stephanie Coffey[1]

Barry Schouten, James Wagner, and Andy Peytchev. *Adaptive Survey Design*. 2017 Boca Raton: CRC Press, ISBN 978-1-4987-6787-3, 252 pp, USD 89.95.

Adaptive survey design (ASD) is earning increased attention as a framework for maintaining or improving survey quality. Noncontact, nonresponse, and the cost of carrying out data collection are all increasing at varying rates. At the same time, the increased computerization of survey operations, along with the increased processing power of computers, means it is possible to generate, aggregate, and process more paradata and survey data, helping survey methodologists, managers, and statisticians understand characteristics of data collection at a more detailed level. *Adaptive Survey Design*, the timely new book by Barry Schouten, James Wagner, and Andy Peytchev, places itself squarely in this environment, providing both motivation for and detailed guidance on implementing ASDs to improve survey outcomes.

This is the first published book addressing the developing field of ASD, and as a result, the authors cover a significant amount of material across five major sections. Section I, *Introduction to Adaptive Survey Design*, lays the foundation for the rest of the book through the introduction of several concepts that return throughout the text. Standard survey methodology topics including survey costs, survey errors, and the variability of survey implementations are tied together to motivate the need for the flexibility to adapt data collection protocols in order to improve survey outcomes. The authors also thoughtfully discuss the nomenclature of ASDs, responsive designs, and their interaction. Their definitions and context provide clear boundaries for what the authors will discuss throughout the book. This is useful for survey practitioners, whether new or familiar with the material, as what qualifies as an ASD is not always consistent in the working literature. This section ends with the introduction of several case studies that return throughout the book to illustrate concepts.

Section II, *Preparing an Adaptive Survey Design*, discusses three components required before implementing an ASD: strata, design features, and models for nonresponse. First, this section covers the process of stratification – stressing that strata should be based on covariates related to survey variables and likelihood of response to different data collection protocols. Data collection features, such as incentives, mode of contact, or case prioritization, are tailored to specific strata and are discussed second. The hope is that the application of particular data collection features to specific strata will result in superior survey outcomes to those obtained through a traditional, non-adaptive designs. Last, this

[1] U.S. Census Bureau, Center for Adaptive Design, 4600 Silver Hill Road, Washington, DC, 20233. U.S.A. Email: stephanie.coffey@census.gov

section addresses statistical models that may be used during an ASD. Two main categories of models are discussed – those that can be used to support changes in data collection procedures, and those that are used to monitor the potential for nonresponse bias. The authors acknowledge that the quality of the strata and models for nonresponse are directly tied to the predictive power of the available auxiliary data, which can vary considerably. Short discussions of potential sources of additional data (such as commercial data or paradata) and methods for stratification (such as response propensity variation, influence on estimates, or machine learning algorithms) point the reader in helpful directions, while making it clear that preparing for an ASD will have some unique elements from survey to survey.

Section III, *Implementing an Adaptive Survey Design*, discusses other aspects of implementing ASDs, and opens with a discussion of costs and logistics. The authors note that, in order to assess cost-quality tradeoffs, survey practitioners need cost models, but they are often very difficult to estimate. The authors help the reader conceptualize a cost model designed for an ASD, and then illustrate the estimation of those model parameters through regression. They also mention the use of expert opinion in constructing cost parameters. This section also discusses the optimization of adaptive survey design, and distinguishes between trial-and-error and numerical optimization, either through mathematical optimization or simulation. Each of these methods have strengths and weaknesses, and the authors suggest using them to validate one another, to the extent possible. Lastly, this section addresses the robustness of ASDs. Both Section II and Section III discuss the estimation of various design parameters, including response to various data collection features, and the related costs of those features. Here, the authors address how inaccuracy in those design parameters can impact the success of ASDs.

Section IV, *Advanced Features of Adaptive Survey Design*, includes the most statistical content of the book, and addresses two main topics. The first chapter reviews the more common indicators of nonresponse bias used in the literature, and classifies them into two types – those that rely only on covariates and a survey response indicator, and those that additionally rely on response data. Again, this requires consideration of the quality of the available covariates or auxiliary data. The second chapter addresses the "during or after" argument – that is, is it worth it to undertake the statistical and operational complexities to design and execute an ASD, or can the same reductions in nonresponse bias be attained through nonresponse adjustments using available covariates? The authors discuss some theoretical evidence of the potential for ASDs to reduce nonresponse bias, even after nonresponse adjustment, and the conditions required for bias reduction. Illustrative examples are provided using the introduced case studies and are particularly helpful here.

In Section V, *The Future of Adaptive Survey Design*, the authors propose a research agenda to further ASD, and itemize nine areas in three categories requiring further experience or research. The first category focuses on proving the utility of ASDs through accumulating evidence of success of ASDs across a range of designs. The second category includes topics related to the implementation of ASDs, including the need to explore statistical methods to inform adaptation (such as Bayesian models for incorporating information) and the need for flexible survey software that can accommodate more complex adaptation. The last category focuses on methodological advances for furthering ASD, including designing paradata to meet the needs of ASD, the optimization of decision

making through numerical methods, and ASD's ability to address sources of error beyond nonresponse and cost. This section also offers a more expanded discussion of ASD for reducing measurement error, in particular. The authors make it clear through this agenda that there are open questions for exploration throughout the survey lifecycle.

The book is a success due to its accessibility and applicability. *Adaptive Survey Design* is written to appeal to a broad range of survey practitioners: methodologists, managers, and statisticians. The only real prerequisite is an understanding of the survey lifecycle process – how surveys are designed, conducted, processed, and analyzed. This is by design – the authors are clear that implementing an ASD *should* involve individuals throughout the survey process, as their knowledge and involvement is key to implementing adaptation successfully. At the same time, the authors provide mathematical detail for those interested, and clearly identify gaps in the existing research and unanswered questions for survey practitioners to consider.

Beyond clear organization and communication, what sets this book apart is the inclusion of case studies from the authors' own experiences with adaptive design. The examples include a random-digit-dial telephone survey, a multimode survey that can be linked to administrative data, and an in-person interview made up of a screener and a personal interview. By including such varied examples, the content and examples in the book are applicable to a wide range of data collection designs. Many aspects of surveys will evolve in the future, from cost and nonresponse patterns to available auxiliary data and design features. However, the underlying concepts of adapting data collection that are detailed in this book will continue to be valuable.