Preface

The preface of the 1963 edition of William G. Cochran's *Sampling Techniques* begins: 'As did the previous editions, this textbook presents a comprehensive account of sampling theory as it has been developed for use in sample surveys.' The reader, theoretician or practitioner, who hopes to find some theory and methods for the treatment of nonresponse must patiently await the concluding Chapter 13 of the book, where a few suggestions are given.

Sampling Techniques and other classical sampling texts bear witness to the important scientific advances made in survey theory during the first half of the twentieth century, although without perceiving a need to dwell extensively on estimation in the presence of nonresponse.

Today, nonresponse is a normal (but undesirable) feature of the survey undertaking. Accordingly, the underlying theory should recognize nonresponse from the beginning. Nonresponse can no longer be handled by an adjunct to a theory that would work well if it were not for certain less desirable realities. The present book attempts to build on these premises.

This change in perspective does not entail any criticism of earlier work. It simply reflects the changing times and the changing survey climate.

Writing more than 40 years ago, Professor Cochran had good reason for relegating nonresponse to the end of his book. The book is remarkable, not least because it served as a valuable guide to a generation of survey statisticians. It was about the theory of sampling, about controlling the sampling error, about estimators that work well under controlled randomized selection. Nonsampling errors are summarily mentioned in Chapter 13. They include errors due to nonresponse and missing observations, errors of measurement, errors in editing and tabulation. At that time, nonresponse was low and its effects negligible.

Even today, sampling error is often contrasted with nonsampling error. But it is a dichotomy that feels much less appropriate now than it may have been 40 years ago. For one thing, the distinction is lopsided: it places together in one category the errors attributable to all causes other than sampling and observing only a fraction of the finite population. Some of the nonsampling errors may indeed be greater than the sampling error.

To us, and no doubt to many others, it is unsatisfactory to have a theory that focuses on sampling error and handles nonsampling errors through 'adjustments' that one somehow appends to a theory meant for ideal conditions not met in practice. Today's reality includes high nonresponse and often imperfect frames.

In writing this book, we found inspiration in the extensive literature on nonresponse. Although further development is bound to follow, we believe that our book has something to offer to all categories of professionals engaged in surveys and statistics production, survey managers, subject-matter specialists, as well as specialists in statistical survey methodology.

The background knowledge required to assimilate the contents of this book varies between the chapters. Roughly the first one-third of the book is easy to follow for all categories; the rest requires more of a technical preparation.

Chapters 2 and 3 give a wholly nontechnical overview. They can be read with only a minimal background in statistical science. They emphasize general issues in the treatment of nonresponse and provide a general orientation in the field. Technical arguments and formulae are essentially absent.

Chapters 4–14 form a logical sequence in which one chapter builds on the preceding ones. The theory of estimation in the presence of nonresponse is developed from Chapter 6 onwards. Many of the notions of the classical survey theory are preserved. There is a finite population of identifiable elements with which one can associate individual probabilities of being selected and of responding. The classical randomization theory is obtained as a special case, namely, when the nonresponse is reduced to nil.

The contents of Chapters 4–14 can be described as moderately technical. To fully appreciate these chapters, readers should have been exposed to at least one thorough course in statistical inference, including principles of point estimation and confidence intervals. Familiarity with linear models and regression analysis, including some linear algebra, is important background. It is a considerable advantage also to have followed one or more courses in modern survey sampling theory and techniques.

Calibration is the unifying concept that keeps the chapters together. Calibration is a technique for computing weights to be used in estimation, given an input of auxiliary information.

Persons working actively with survey methodology, in statistical agencies or survey institutes, should be able to easily follow the whole book.

Readers with practical experience with surveys and statistics production are in a favoured position. They can identify with many issues raised in the book. Their practical background will facilitate their reading.

A reader who is a first-year graduate student at university may start from a background derived mainly from a set of undergraduate statistics courses and with little or no exposure to the reality of surveys. Nevertheless, the book should find use in a graduate course within a university statistics curriculum.

Some more technical material – notably derivations and lengthier proofs of certain propositions – is placed in end-of-chapter appendices. These need not be read or understood if the primary objective is to see the practical utility of the methods.

This book evolved from a Current Best Methods (CBM) manual which we wrote in 1999–2001 for Statistics Sweden. Its title is *Estimation in the Presence of Nonresponse and Frame Imperfections*. A far-reaching revision and extension

of the contents of that manual was necessary to meet the objective that we and the reviewers set for this book.

The contents have a definite practical orientation. On the other hand, it would not be appropriate to call the book a handbook. A practitioner should perhaps not hope to find, on a certain page of the book, a tailor-made answer to a particular problem encountered in a survey. He or she should always be prepared to adapt the techniques in this book to fit the context of his or her own survey. This may require familiarity with probabilistic evaluations of expected values and variances, with techniques in regression analysis and with other knowledge mastered by survey methodologists.

We gratefully acknowledge the support of Statistics Sweden in the preparation of this text. Three anonymous reviewers engaged by John Wiley & Sons, Ltd commented on the proposed content. We are indebted to them for their valuable comments, which provided a clear direction for our work.

Ottawa Carl-Erik Särndal*
Örebro Sixten Lundström†
December 2004

*Carl-Erik Särndal, Ph.D., is Consultant to National Statistical Agencies and was formerly Professor, Université de Montréal

†Sixten Lundström, Ph.D., is Senior Methodological Adviser, Department of Population and Welfare Statistics, Statistics Sweden