Introduction

James Lepkowski
University of Michigan
Ann Arbor, Michigan
United States of America

1. The previous sections and chapters of the present publication have examined, for the most part, sampling errors that arise when a representative probability sample is taken from a population. A number of other errors that arise in household surveys are considered in the present section. Some of these errors are, like sampling error, variable across possible samples, or across possible repetitions of the measurement process. Others are fixed, or systematic, and do not vary from one sample to the next.

2. In the sample design framework, variable errors are usually referred to as sampling variance. There are fixed sampling errors, some of which have already been mentioned, which are referred to as bias. For example, the deliberate exclusion of a subgroup of the population introduces non-coverage of the population subgroup, and an error that will be present, and of the same size, no matter which possible sample is selected.

3. Non-sampling errors involve non-observation errors when there is a failure to obtain data from a sampling unit or a variable, or measurement errors that arise when the values for survey variables are collected. Non-observation errors are usually fixed in nature, and lead to considerations about bias in survey estimates. Measurement errors are sometimes fixed, but they may also be variable.

4. Among non-observation errors, two sources of error are most important: non-coverage and non-response. In probability sampling, there must be a well-defined population of elements, each of which has a non-zero chance of selection. Non-coverage arises when an element in the population actually has no chance of selection; the element has no way to enter into the selected sample. Non-response refers to the situation where no data are collected for an element response that has been chosen into the sample. This may occur because a household or person refuses to cooperate at all, or because of a language barrier, a health limitation, or the fact that no one is at home during the survey period.

5. Measurement errors arise from more diverse sources -- from respondents, interviewers, supervisors and even data-processing systems. Respondent measurement errors may occur when a respondent forgets information needed and gives an incorrect response, or distorts information in response to a sensitive question. These respondent errors are likely to constitute a bias, because the respondent consistently forgets, or distorts an answer, in the same way, no matter when he or she is asked a question. These errors can also be variable. Some respondents may forget an answer at one moment, and remember it another.

6. There are four dimensions that survey designers consider in respect of these kinds of errors. One entails a careful definition of the error and an examination of the sources of the error in the survey process, encompassing what part of the survey process appears to be responsible

for generating this kind of an error. The second entails how to measure the size of the error, a particularly difficult problem. Third, there are procedures to be developed to reduce the size of the error, although their implementation often requires additional survey resources. Last, non-sampling errors occur in every survey, and survey designers attempt to compensate for those errors in survey results.

7. Chapters VIII and IX in this section examine from a conceptual viewpoint non-observation and measurement error, respectively, providing some illustration of many different types of these errors. Chapters X and XI offer more detailed treatments of these errors, the former considering the overall impact on the quality of survey results, and the latter providing a case study of these kinds of errors in one country, Brazil.