Introduction

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

1. In the previous sections, sampling and non-sampling errors that arise in household surveys were examined in order to gain a better understanding of the quality of survey estimates. In almost all types of such errors, there are methods that can be used to reduce the size of the error. The implementation of those methods, however, often entail an additional cost. Since surveys have fixed budgets to cover expenses, devoting additional resources to reduce one source of error means shifting resources from one area to another procedure. Survey design involves constantly trading off costs and survey error.

2. For example, suppose that in a particular household survey, there is a subgroup of the population speaking a language for which there is no translation of the survey questionnaire. The survey designers may decide initially to exclude this group from the survey, creating a coverage problem. Alternatively, they may decide to decrease the sample size to reduce survey costs, and then use the saved costs to translate the questionnaire into a new language, hire interviewers who speak that language, and bring those households back into the survey.

3. Given that survey design is often a series of such trade-offs, in order to make sound decisions, good information must be available about the nature and size of errors arising from different sources (such as sampling variance and non-coverage bias, in the previous example) and about the costs associated with different survey procedures. The previous sections examined error sources and sizes of errors. In the present section, the nature of survey costs will be examined.

4. Cost considerations in a survey arise at three levels. The first is in the planning phase of a survey when costs must be estimated in advance. Cost estimates in the planning or "budgeting" phase are difficult to obtain, unless one has prior experience to build on. Continuing survey operations can provide relevant cost data for planning new rounds of a survey, although cost considerations at the next level - the monitoring of survey costs - often interferes.

5. Survey organizations, or even others that conduct surveys occasionally, seldom have well-developed systems for tracking costs in such a way as to enable the cost data to be used for planning. Costs are assembled in an accounting system, but those systems do not categorize costs into the kind of categories that a survey designer needs for planning purposes. In instances where such cost monitoring is attempted, it may add to the cost of the survey itself if new systems must be added to the operations.

6. If costs are being monitored in an ongoing operation, it is possible to consider, more systematically, changes in survey design during data collection. Cost information can be used to

project how large both the savings in one operation, and the impact of the reallocation of resources to another area, might be.

7. Reallocation of resources in survey planning is determined by considering trade-offs between cost level and error across multiple sources of error. Sample design development is one area where these trade-offs can be and are made formally to find an optimal solution to the resource allocation problem.

8. For example, as discussed in chapter II, surveys that are based on clusters drawn in an area probability sample from a widely spread population must consider limiting the number of clusters in order to reduce data-collection costs. Limiting the number of clusters however means that the number of observations made in each sample cluster must go up in order to maintain overall sample size. However, this increase in the size of the subsample in each cluster increases the variability of sample estimates. In other words, as costs go down, by taking fewer clusters, sampling variance goes up. What is needed is guidance on how many clusters to select so that the costs can be minimized, given that a specified level of precision is to be achieved, or that the sampling variance is to be kept as small as possible for a given cost. In sample design, there is a mathematical solution to this problem.

9. The cost-error trade-off arises in other aspects of survey design as well. For example, one method for reducing household non-response in a household survey is to visit more than once households for which no response is obtained on a single visit. An interviewer can be instructed to visit households during the survey data collection period as many as four or five times in order to obtain a response. Making repeated visits to some sample households reduces the number of sample households that can be included in the sample. The cost of repeated visits to reduce household non-response limits sample size. The cost of greater non-response reduction efforts to reduce non-response bias thus increases sampling variance. Again, the cost-reduction efforts in one area requires that resources be reallocated, and introduces the potential for an increase in error in another area of the survey design.

10. The chapters in this section consider a number of issues centred around planning, monitoring and reallocation of costs in survey design. They use data from household surveys in developing and transition countries to illustrate the types of costs incurred in survey data collection and, to some extent, the size of the costs. Since survey operations vary so widely from country to country, and even more so across continents, the specific cost information provided may not be useful for planning a survey in a given country. It is hoped, however, that the cost sources and cost levels presented in the following chapters will help survey designers across diverse settings understand survey costs and cost-error trade-offs more fully in their own surveys.