DESIGN EFFECT AND COST ISSUES FOR SURVEYS IN DEVELOPING COUNTRIES

Ibrahim S. Yansaneh, U.N. Statistics Division and John L. Eltinge, Bureau of Labor Statistics
I.S. Yansaneh, U.N. Statistics Division, DC2-1526, Two U.N. Plaza, New York, NY 10017; yansaneh@un.org

Key Words: Demographic and Health Surveys (DHS); Incremental cost per interview; Information capacity; Living Standards Measurement Surveys (LSMS); Survey infrastructure.

1. Introduction

1.1 Efficiency in sample design

In survey work, one generally seeks to use a sample design that has two properties:

(1) A satisfactory level of information capacity. Information capacity is generally measured by the variances of the estimators of selected population quantities that are considered to be of principal interest. In some cases, information capacity may also be quantified through related measures, e.g., mean squared error or inferential power.

(2) Costs that are consistent with available budgets and that make reasonably efficient use of resources.

These properties have been considered in depth for surveys carried out in developed countries. See, e.g., Andersen et al. (1979), Cochran (1977), Groves (1989), Kish (1965, 1976, 1995), Linacre and Trewin (1993) and references cited therein. In addition, for a broader discussion of cost and precision criteria as two of many criteria for evaluation of national statistical systems, see deVries (1999, p. 70) and references cited therein.

In keeping with remarks in Groves (1989, Chapter 2), two notable features of this literature are the following. First, formal mathematical development in this area frequently centers on optimization of well-behaved information capacity or cost functions subject to relatively simple constraints. However, due to limitations in available cost and variance information, this optimization approach often should be viewed as providing only rough approximations for the preferred design, or for the information capacity and cost values that will be achieved in implementation. In addition, within-stratum sample sizes at one or more stages of sampling often are small or moderate. Consequently, optimization work must account for the discrete properties of the resulting cost and variance functions

Second, despite the above-mentioned limitations, one often encounters a substantial amount of common structure (in e.g., costs and variance components) across surveys that can be useful in addressing criteria (1) and (2). In some cases, this common structure is limited to qualitative indications of the relative magnitudes of several variance components, or of several cost sources. Other cases support a more detailed quantitative evaluation, including (approximate) optimization work.

1.2 Components of design effect and cost structures for surveys in developing countries

Over the past several decades, a large number of surveys have been conducted in developing countries. Most of these surveys are commissioned by international financial and development agencies that need the data for decision-making on developmental assistance projects. In design work for these surveys, criteria (1) and (2) are important. Thus, in keeping with the literature reviewed in Section 1.1, it is of interest to study the extent to which one may identify common variance and cost structures within groups of developing-country surveys. For some previous work with variances and design effects from surveys in developing countries, see, e.g., Verma et al. (1980), Verma and Lê (1996) and references cited therein.

In the interest of space, the remainder of the present paper will focus on cost factors. Section 2 provides some general background on two programs of surveys currently carried out in developing countries. Section 3 gives a qualitative description of some factors that influence the overall costs of these surveys. Section 4 outlines a general framework for a quantitative analysis of these factors.

2. Surveys in Developing Countries

2.1 Two examples: The Demographic and Health Surveys and the Living Standards Measurement Study Surveys

Two prominent examples of developing-country survey series are the Demographic and Health Surveys (DHS) and the Living Standards Measurement Study (LSMS) surveys conducted in various countries in Africa, Asia, Latin America, the Caribbean and the Middle East.

The DHS is an offshoot of the World Fertility Survey (WFS). It is a program of surveys conducted by Macro International, Inc. for the United States Agency for International Development (USAID). DHS surveys are intended to generate nationally representative samples of households in many developing countries. DHS surveys provide data for a wide range of monitoring and impact evaluation indicators in the areas of population, health, and nutrition. See, e.g., Verma and Lê (1996) and references cited therein.

The LSMS is a multi-topic integrated survey conducted in several developing countries with technical assistance from the World Bank. For some general background, see Grosh and Glewwe (1995). The LSMS surveys provide comprehensive data on most aspects of household welfare, including consumption, expenditure, income from activities in the labor market, household enterprises or agriculture, asset ownership, migration, health, education, nutrition, fertility, and anthropometrics. There are also additional modules on disability and time use for some, but not all, countries. LSMS data are collected at the individual, household, and community levels. Other programs of surveys conducted by the World Bank in developing countries include the integrated surveys and the priority surveys, both implemented under the Social Dimensions of Adjustment program. The integrated surveys are similar to the LSMS and the priority surveys are shorter versions of the integrated surveys, usually without the consumption module, but not always. In addition, there is a new survey, the Core Welfare Indicator Questionnaires (CWIQ) survey, which is designed to complement the other surveys. The CWIQ survey collects information through a short four-page questionnaire on simple indicators such as access to services and basic needs. See, e.g., World Bank (2000) for a detailed description of these surveys.

2.2 Similarities and Differences Between the DHS and LSMS

Sample designs for surveys in developing countries have many common features. This makes information on design effects or survey costs more easily transportable across different surveys and different countries. In other words, information on design effects and costs for one survey in a developing country can provide a useful potential indication of these characteristics for a subsequent survey in the same country or for similar surveys in other developing countries. However, the use of such information in the planning for other surveys must recognize the caveats discussed in section 3 below.

We illustrate the similarities and differences among surveys in developing countries by comparing the DHS and the LSMS. The designs of these two programs of surveys have the following common features:

(i) They are both household surveys.

(ii) Due to the lack of good-quality list frames, they are both based on a multi-stage stratified area probability design.

(iii) Clusters are constructed from geographic areas identified and used in a preceding national population census.

(iv) Explicit stratification is based on administrative regions by urbanicity, with a separate stratum reserved for the capital city, which may or may not have a rural component.

Despite these similarities, there are important differences between the LSMS and the DHS. First, the focus of the DHS is the evaluation of population, reproductive health, and nutrition programs, whereas the LSMS is concerned mainly with poverty and human development issues. The target of the DHS data collection efforts is women aged 15-49. The LSMS collects information at the household level and from all household members individually. However, to reduce respondent burden in instances when the questionnaire is very long, the decision is sometimes made to randomly select one female respondent of child bearing age in the household to respond to the fertility module. In such instances, the random selection of the respondent is done in the field through carefully defined procedures which are taught to the interviewers during the training period for the field work (Grosh and Muñoz, 1996).

2.3 Use of Design or Cost Information from Previous Surveys

It is sometimes the case in developing countries that several different surveys with different sponsors and possibly different objectives are conducted in the same country either simultaneously or in rapid succession. For instance, in Turkmenistan in Central Asia, an LSMS survey was carried out in 1998. The survey was based on a one-stage area cluster sample, with sets of contiguous households serving as the clusters, and with complete enumeration of all households in a selected cluster. In June 2000, a DHS

survey was conducted in Turkmenistan, based on a two-stage stratified cluster design. In this DHS survey, the primary sampling units are again sets of contiguous households, but a given primary unit could be coarser or finer than the clusters used in the 1998 survey. In the current DHS survey, individual households are the secondary sampling units. In December 2000, a multiple-purpose household mini-census survey will be conducted and there is a question as to whether to adopt a one-stage design such as the one used in the LSMS or a two-stage design such as the one used in the DHS. Despite the differences in cluster features and subsampling plans, all three surveys use some common design information based on data from the 1995 census.

These surveys provide an example of a situation in which information from a previous survey potentially can be used to improve the design of a subsequent survey. A one-stage cluster sample design may have been appropriate for the 1998 survey because of the currency of the sampling frame information and the relatively small number of households included in the survey. However, a two-stage stratified cluster sample design would be preferable for the 2000 mini-census survey. This design offers several advantages over a single-stage alternative. First, due to the additional distance in time from the 1995 census and the considerable changes that have since taken place, there is an obvious need for updating the sampling frame via a listing operation. Second, it is less expensive to implement because the cost of travel and implementation of post-survey quality-control procedures are much lower. Third, it requires less time to implement, so data are more quickly available for decision-making purposes. Fourth, it requires fewer people for listing and interviewing, so the people who do this work are more selectively chosen and better trained. Finally, it is easier to monitor the quality of field operations, thereby reducing non-sampling errors to the extent possible. See Yansaneh (2000) for details.

3. Cost Issues

3.1 Comparability of Costs Across Surveys and Across Countries

Discussions of survey costs often approximate an overall survey cost C as a linear function of the numbers of selected primary sample units and selected elements. An example of such a function is

$$C = c_0 + \sum_{h=1}^{L} n_h c_h + \sum_{h=1}^{L} \sum_{i=1}^{n_h} n_{hi} c_{hi} \quad (1)$$

where c_0 represents the fixed costs of initiating the survey; c_h equals the incremental cost of collecting information from an additional primary sampling unit (PSU) within stratum h; and c_{hi} equals the incremental cost of interviewing an additional household within PSU i in stratum h. See, e.g., Cochran (1977, Sections 5.5 and 11.13-11.14) and Groves (1989, Chapter 2).

Note that expression (1) is one of many possible cost functions that could be considered. For example, Cochran (1977, p. 313) discusses inclusion of a separate cost component associated with listing of secondary sampling units within selected primary units, where that component depends on the number of secondary units in each primary unit. Also, for a three-stage design, that is, a design in which persons are randomly selected within households and interviewed, there will be an extra term in (1) above, denoting the incremental cost associated with interviewing an additional person within a selected household.

In general, the cost coefficients c_0, c_h and c_{hi} will depend on a large number of factors that may vary across countries and across surveys within countries. Section 3.2 reviews cost factors that that may be important for cases in which a considerable amount of survey infrastructure is already in place. Section 3.3 considers cases in which there is limited or no prior survey infrastructure. Section 3.4 discusses changes in the cost structure that may result from modifications in survey goals. Section 3.5 provides some related cautionary remarks regarding interpretation of reported survey costs.

3.2 Costs for Surveys with Extensive Infractructure Available

3.2.1 Factors Related to Preparatory Activities

Much of the cost of a one-time survey goes to the financing of preparatory activities (see, e.g., Grosh and Muñoz, 1996, p. 199), and so the funds for such activities are disbursed early in the survey process. Preparatory activities with relatively fixed costs include coordination of survey planning by multiple government agencies; frame development; sample design; questionnaire design; and publicity directed toward potential respondents. Preparatory-activity costs that depend on sample size (either at the primary unit or household level) include the hiring and training of field staff (e.g., listers, interviewers, supervisors, and translators).

The costs of preparatory activities depend on local factors like the size of the survey staff, the amount of equipment, the prices of the items, and whether the

survey is a cross-sectional study being done for the first time, or part of a continuing survey.

3.2.2 Factors Related to Data Collection and Processing

The costs of data collection and processing also involve both fixed and variable components. For the most part, the costs of data collection are variable, i.e., dependent on the number of primary sampling units and households selected. These costs include the costs of listing of households within selected primary units or listing of persons within selected households; interviewing; and field supervision. The cost of data collection also includes the cost of travel both between and within PSUs. These data collection costs depend on the organization of the interview operations, the length of the questionnaire, whether or not interpreters are used, and the number of units to be interviewed.

One option for reducing travel costs and improving data quality is to create national survey teams consisting of supervisors and interviewers and move the teams around from region to region. This approach can also be useful in situations where data collection is carried out on a rolling basis, or when survey operations involve the use of expensive equipment. The model of multiple survey teams has been used in many surveys in developing countries, such as the LSMS series (Grosh and Muñoz, 1996, Chapter 5).

In data processing, edit and imputation work may involve a mixture of fixed and variable costs, depending on the degree of automation used in this process. The other principal costs of data processing are arguably fixed, and include the costs of computing equipment and software; and the development of weights, variance estimators and other data analysis work. For instance, weights would be computed regardless of the number of PSUs or households sampled, and after a weighting procedure has been developed and programmed, the incremental cost of computing a weight for an additional household would be negligible.

The cost of data processing depends on how many levels of analysis are included in the budget. For some surveys, only preliminary analysis is carried out on the collected data in the form of tables. For other surveys like the DHS and LSMS, many secondary analyses are conducted as a basis for policy recommendations for beneficiary governments and donor agencies. Such secondary analyses are compiled in a series of analytical and methodological reports for the DHS, and in a series of working papers for the LSMS. Some examples are included in the reference section of this paper. Some costs are also incurred in various services to other analysts.

3.3 Costs for Surveys with Limited or No Prior Survey Infrastructure Available

In a country with relatively little previous survey infrastructure, it is likely that the sponsoring agency will need to devote a substantial amount of resources to capacity building efforts that would not be required in a country with substantial survey infrastructure (Grosh and Muñoz, 1996, Chapter 8). The costs of preparatory activities, field operations and data processing can all be inflated by a lack of infrastructure.

For instance, capacity building generally involves extensive initial training of personnel. In addition, the time of field personnel tends to be used more efficiently as a survey organization gains experience. Also, in countries with substantial previous survey experience, the need for travel is much lower because the statistical agencies in such countries are likely to have experienced regional data collection teams, or to provide the means of transportation for survey field staff. These advantages provide savings in the cost of transportation, training, and other personnel costs. Countries with no history of previous surveys usually include vehicles in the survey budget and this item may become a major part of the overall cost of the survey (Grosh and Muñoz, 1996, Chapter 8). Other examples of budget items where the existence of some survey infrastructure or history of previous surveys has a substantial impact are computer equipment and maps for identification of households.

3.4 Factors Related to Modifications in Survey Goals

As noted above, many cost factors are linked with features of the survey design, including the sample size; the length of the questionnaire; the number of modules; and specific methods employed in sample selection and listing, pilot testing, and questionnaire design and translation. For a given design, some of the resulting costs are approximately constant across countries.

However, survey designs in developing countries often have to be modified to accommodate ad hoc specifications by beneficiary governments. For instance, a government may decide to broaden the objectives of the survey to include other national priorities. This in turn may lead to:

(i) inclusion of additional modules in the questionnaire; or

(ii) an increase the number of reporting domains if estimates of key variables for sub-regional groups are desired at the same precision level as the national-level estimates.

These modifications can affect trade-offs between cost and data quality in several ways. First, they can lead directly to significant increases in the total amount of interviewer time required for data collection due to an increased mean length of an interview (due to (i)) or an increased number of interviews (due to (ii)). Second, if a survey organization has available a relatively fixed number of well-trained interviewers and field supervisors, then the above-mentioned increases may lead to increased costs due to the need to train additional interviewers, and due to an increase in the amount of supervisor time required per minute of interview time. Third, the above-mentioned increases can lead to an increase in the magnitude of nonsampling error relative to sampling error. For example, inclusion of extra modules in a questionnaire may inflate nonsampling error due to inadequate question testing or respondent fatigue. Nonsampling error may also increase due to the use of a larger number of relatively inexperienced interviewers, necessitated by an increase in the number of interviews or in the mean length of an interview.

3.5 Some Cautionary Notes on Reporting of Survey Costs

Several factors need to be considered to ensure that comparison of costs across surveys and countries are carried out on a reasonably common basis.

First, surveys in developing countries are sponsored by several different organizations, which often have different policies and accounting procedures. For instance, for some sponsoring agencies, it may be important to distinguish between the cost to the sponsoring agency and the overall cost of completing the survey.

Second, it may be important to account comparably for survey support that is provided in kind, e.g., vehicles for transportation of field personnel. In some cases, in-kind support may be provided by the national statistical office, e.g., through assignment of their permanent field staff to an internationally sponsored survey. Although such costs may be considered in-kind and excluded from the itemized budget, they nevertheless represent an opportunity cost in so far as the survey exercise is an additional activity that takes time away from other potential work by the national statistical office.

Similar comments apply to provision of external technical assistance. This latter item can be especially important in countries with no survey infrastructure or history of conducting surveys. For many surveys, such technical assistance is provided in kind by international agencies that conduct or sponsor the surveys, and thus is not included directly in the survey budget. However, sometimes, such technical assistance is contracted out, and thus included in the budget. For instance, the 1998 Turkmenistan LSMS-type survey was conducted by the Research Triangle Institute (RTI), under contract to the World Bank.

Third, due to the hierarchical cost structure (1) given in Section 3.1, it is important to distinguish between the total cost for a survey and the cost per completed interview. For instance, due to availability of greater resources and a greater degree of interest in reliable estimates reported at a subnational level, larger developing countries tend to use larger sample sizes in their surveys. (UNICEF, 2000, Chapter 4). Because of high costs associated with transportation and salaries of a larger number of survey staff, surveys in larger countries tend to have higher total costs than surveys in smaller countries. However, larger countries with higher overall costs may sometimes have lower costs per completed interview, because of economies of scale and the distribution of fixed costs over a larger sample.

Fourth, evaluation of overall and per-interview costs may be complicated by special features of the sample design. For example, costs may be inflated by the use of oversampling or use of screening samples to ensure achievement of precision goals for certain subpopulations that are small or difficult to identify from frame information (e.g., households with children under the age of five).

Finally, for surveys of populations with heterogeneous household sizes, it can also be important to distinguish between costs per contacted household and costs per completed interview.

4. Use of Cost Information

Recall from Section 1 the suggestion that one may be able to identify common cost structures in developing country surveys, and then use those common structures in sample design work. In light of the discussion in Sections 2 and 3, one could consider implementing this idea in the following way. First, consider the coefficients c_0, c_h and c_{hi} in cost model (1) to be functions $c_0(x)$, $c_h(x)$ and $c_{hi}(x)$ of a

vector $x = (x_1,...,x_p)$, say, of indicator or continuous variables related to the principal cost factors, e.g., availability of vehicles, maps, computers and other physical infrastructure, availability of trained survey field staff and supervisors, prevailing wage rates, and the length and complexity of the questionnaire.

Second, if the factors x affect the cost coefficients in relatively simple ways, then one could consider approximation of these coefficient functions through additive models, e.g.,

$$c_0(x) = \beta_{00} + \beta_{01}x_1 + ... + \beta_{0p}x_p \qquad (2)$$

In such cases, one could consider estimation of the coefficients of (2) from previous surveys, and then use the estimated coefficients in subsequent design work. On the other hand, if available information indicates that the coefficient functions involve a large number of interaction terms or (as a special case) factors x_i that are unique to a small number of countries, time periods or surveys, then the qualitative descriptions in Section 3 may offer relatively little insight into quantitative results applicable to design of new surveys. In the latter case, descriptive studies of cost structures in developing countries may help to provide useful general guidelines, but may offer less insight into cost-benefit trade-offs in design work.

Acknowledgements

The authors thank Gary Shapiro, Diane Steele, and Tilahun Temesgen for helpful comments on earlier drafts of this paper. The opinions expressed in this paper are those of the authors and do not necessarily represent the policies of either the United Nations or the U.S. Bureau of Labor Statistics.

References

Andersen, R., Kasper, J. and Frankel, M.R. (1979). *Total Survey Error.* San Francisco: Jossey-Bass.

Cochran, W.G. (1977). *Sampling Techniques, Third Edition.* New York: Wiley.

de Vries, W. (1999). Are We Measuring Up ...? Questions on the Performance of National Statistical Systems. *International Statistical Review* **67**, 63-77.

Fitch, D. (1999). Comment by D. Fitch quoted by V. Verma in "Q/A 40.1," *The Survey Statistician* **41**, 14.

Grosh, M.E. and Glewwe, P. (1995). *A Guide to Liming Standards Measurement Surveys and Their Data Sets.* Living Standards Measurement Study Working Paper No. 120. Washington, DC: The International Bank for Reconstruction and Development/The World Bank.

Grosh, M.E. and Muñoz, J. (1996). *A Manual for Planning and Implementing the Living Standards Measurement Study Survey.* Living Standards Measurement Study Working Paper No. 126. Washington, DC: The International Bank for Reconstruction and Development/The World Bank.

Groves, R.M. (1989). *Survey Errors and Survey Costs.* New York: Wiley.

Kish, L. (1965). *Survey Sampling.* New York: Wiley.

Kish, L. (1976). "Optima and Proxima in Linear Sample Designs." *Journal of the Royal Statistical Society, Series A* **139**, 80-95.

Kish, L. (1995). "Methods for Design Effects." *Journal of Official Statistics* **11**, 55-77.

Linacre, S.J. and Trewin, D.J. (1993). "Total Survey Design – Application to a Collection of the Construction Industry. " *Journal of Official Statistics* **9**, 611-621.

Steele, D. (2000). Personal communication from Diane Steele (World Bank) to Ibrahim S. Yansaneh, June 2, 2000.

UNICEF (February, 2000). End-Decade Multiple Indicator Cluster Survey Manual. United Nations Children's Fund.

Verma, V. and Lê, Thanh (1996). "An Analysis of Sampling Errors for the Demographic and Health Surveys." *International Statistical Review* **64**, 265-294.

Verma, V., Scott, C. and O'Muircheartaigh, C. (1980). "Sample Designs and Sampling Errors for the World Fertility Survey" (with discussion). *Journal of the Royal Statistical Society, Series A* **143**, 431-473.

World Bank (2000). Poverty in Africa – Survey Databank. http://www4.worldbank.org/afr/poverty

Yansaneh, I.S. (2000). "Sample Design for the 2000 Turkmenistan Mini-census Survey". *Mission Report, United Nations Statistics Division.*