DISCUSSION OF SESSION ON SAMPLING IN DEVELOPING COUNTRIES

Gary Shapiro, Westat
Westat, 1650 Research Boulevard, Rockville, Maryland 20850

Key Words: Sample design, Weighting, Cost models

I would very much like to thank David Fitch for organizing this session. There are few papers presented at the Joint Statistical Meetings that deal with sample design and related issues in developing countries, and so it is really good to have a session such as this. Two of the papers in the session provide general advice. The Yansaneh paper deals with optimal survey design with respect to costs and variance components, while the Fitch paper discusses why weights are needed in surveys. A third paper in the session, the one by Proctor, is a mix of advice and a specific case study. Finally, the paper by Abeyasekera is strictly a case study. I will discuss the papers in order of generality, beginning with the Yansaneh paper and concluding with the Abeyasekera paper.

Overall, the Yansaneh paper has provided a very nice framework and layout of issues related to cost and design effects in the design of a survey. There are actually three versions of this paper: A very detailed outline, the oral presentation, and the written paper. The outline is the most ambitious and expansive of the three, but even it is only part 1 of the authors' ultimate goal, which is to provide information on design effects and costs that can be useful for subsequent surveys in the same or another country. Related to this goal is the specification of a cost model, which is presented in the paper as a linear function of principal cost factors. I would very much like to see this model developed. I encourage the authors to continue work in this area, deriving design effects and relative costs that could be easily applied by statisticians in developing countries. The framework presented in the written paper is useful, but for statisticians like Dr. Abeyasekera, there is usually not enough time and no available data to allow application of the framework. If the authors can develop this framework with specific cost models and estimates of design effects and cost factors, this would be of great value.

I have a couple of specific comments on the paper. First, a realistic cost function is frequently a stepwise function rather than a linear function. For example, if 8 interviews can be conducted in a single day, then the addition of a ninth interview requires an extra day of work and thus substantial cost, whereas the addition of a tenth interview may add little cost. Discussion of this idea would be useful. Also, sometimes there are constraints that determine decisions such as the number of sample PSUs. For example, it may be that one would want to spend a full week interviewing in a village. In that case, less than a week's workload would not be feasible, although a double workload equivalent to 2 weeks' work might be possible. In such a situation, the number of sample PSUs would not be directly determined by consideration of costs and design effects.

The Fitch paper contains excellent advice – weights are necessary when units are selected with unequal probabilities. The paper contains very good examples of the bias you can get if weights are not used. However, failure to use weights is not just a problem in developing country surveys: I have seen surveys in the U. S. where weights were not used despite quite unequal selection probabilities. I do agree with the advice of Scott: self-weighting designs should be used unless there is a good reason not to have a self-weighting design. There are, however, times when there is a good reason. Furthermore, it is often desirable to use weights even with a self-weighting design.

Table 1 below throws some further light on the WHO example in the Fitch paper. The table shows fairly modest relative biases. However, the relative biases are mostly larger than the relative standard errors. Bias is important largely because the relative standard errors are small. In most situations, use of weights reduces bias, but at the expense of increases in standard errors. The WHO example is unusual in that the use of weights reduces standard errors as well as bias. This, together with relative biases that are large compared to relative standard errors, results in major benefits from the use of weights.

Table 1. Relative bias compared to relative standard error, for WHO example in Fitch paper

Variable	Relative bias	Relative standard error
Nc	1.9%	2.1%
Edu	4.3%	2.3%
Luz	5.6%	2.4%
Cb	0.3%	2.5%
Cd	4.0%	1.8%
Kc	2.8%	1.8%
Radio	3.5%	1.0%

The Fitch paper discusses a problem in the 1995 Guatemalan DHS when there were a lot of vacant DUs in some PSUs. For example, in a desire to obtain 34 DUs, they selected half this many DUs in a PSU with an expected 50% vacancy rate. This resulted in only an expected 8½ occupied DUs, too small a sample size and too large a weight. The problem could have been avoided by the selection of 34 DUs instead of only 17 DUs.

For the Guatemalan survey, the paper recommends selecting only 1 part of a cluster for interview. Then, if there is sufficient time, the second part of the cluster could be included in sample as well. This introduces a bias problem, in that the second part will only be included when the selected part is small. Thus, the true probabilities of selection are changed and can not be properly accounted for in the weights.

The Proctor paper is very valuable in that in contains really practical advice, like what to take with you on a consulting trip to a developing country. It is very unusual in statistical papers to go beyond advice on technical design issues. Dr. Proctor has provided some sound advice and a good case history.

Unlike the other authors of papers in this session, Abeyasekera had to be responsible for the design of a survey and had to deal with some non-statisticians who could make or influence decisions about the design. The Abeyasekera paper raises the issue as to what extent the excellent advice in the other papers can actually be implemented in real life situations. For example, she did not use cost and design effect information, as discussed by Yansaneh, in making design decisions. To the author's credit, the paper provides a lot of details, making it easy for me to raise questions.

One interesting contrast with the Proctor paper is in the selection of second stage units. Proctor's design is to select the second stage units in the abstract, and only after selection is there an association with a specific geographical area. Abeyasekera, on the other hand, drew grids on a map and selected specific grids.

The objectives and sample design for the described survey are quite complex. Samples are needed for different size estates, so that what is optimal in the first and second stages of selection for sampling large estates is not optimal for sampling small estates. The sampling procedure was to select the first stage units with probability proportional to size and to select second stage units within sample first stage units with the same probability across all first stage units. This does not result in a self-weighting design. Among first stage units, there is a range by a factor of 3.3 in the probability of selection of second stage units. I would have been inclined to select approximately equal numbers of second stage units in each of the first stage units. The sample design used had generally equal sample sizes of estates within each sample second stage unit, which resulted in closer to a self-weighting design for estates than for second stage units. However, a different procedure could have resulted in more equal weights.

Abeyasekera selected second stage units at random, with replacement. I would have been inclined to combine second stage units into strata, and select 1 or 2 units per stratum. Or, alternatively, to have sorted the second stage units on relevant characteristics, and selected a systematic sample. These methods would have resulted in better precision and better control of the sample. Either technique would permit adequate estimates of precision.

In conclusion, I think this was a very worthwhile set of papers and a very interesting session.