Case Study No. 5

Determining the effectiveness of a proposed sample sizes

This case study is from a private sector survey-research firm, Mitra & Associates[†], contracted by DFID Bangladesh to carry out a baseline survey for a radio project. As part of their proposal they prepared a report for DFID to justify the proposed sample size of 1500 people in the study.

Pages 2 to 4 of this case study give the material in exactly the form that it was presented to DFID Bangladesh. It provides a valuable illustration of the way we believe that a study of this type should be planned. In particular it shows the use of the design effect (deff) as a component of the formulae used to estimate the appropriate sample size.

The use of design effects may not be familiar to all readers and hence we have added an appendix which gives further information on the calculation of sample sizes.

We are also using this example to demonstrate and emphasise the advantage of using support from a local research firm compared to using external consultants. It is clear that it is the local knowledge and experience of previous surveys of a similar nature that enabled the local firm to justify their use of a sample of 1500 respondents. This was invaluable in this survey.

Acknowledgement

We are very grateful to Mitra and Associates for agreeing to let us use their material in our case-study series.

[†] Mitra and Associates is a leading private sector survey-research firm of Bangladesh. Established in 1983, it has been successfully carrying out research, evaluation and surveys in the fields of Health; Population; Women and Child Development; Agricultural and Rural Development; and Communication. Its clients include donor agencies; UN bodies; international universities, research bodies and NGOs; and national government agencies and NGOs.

Mitra and Associates has so far undertaken about 90 assignments that include four successive national Contraceptive Prevalence Surveys, 1983-1991; three successive Demographic and Health Surveys, 1993-2000; national media access study 1995; Matlab Health and Socio-economic Survey (MHSS) 1996-97; Baseline Survey for Strengthening Rural Broadcast Project 1999; Agricultural Extension Services Study 2000, Monitoring and Evaluation of Bangladesh Urban Primary Health Care Project 1998-2003 (ongoing); Poultry Model Monitoring System Survey 2000-2002 (ongoing); and Baseline Study on Rice Production for Enhancing Income and Employment Opportunities 2001 (ongoing).

Résumé on
Reliability of the Samples for the Baseline Survey
for Strengthening Rural Broadcast (SRB) Project

Reliability of estimates drawn from the proposed male and female samples will greatly depend on how precisely the following key parameters are estimated from them: the proportion of respondents having access to radio, the proportion listening to radio at least once a week, and the proportion listening to agriculture related radio programmes.

Precision of an estimate obtained from a sample is defined as the amount of tolerated errors in the estimate. It depends, among other things, on the size of the sample. For a given sample size, likely precision of an estimate is usually calculated by using the following statistical equation[‡] on conditions that chances are 95 percent that the sample would provide the estimate within the level of calculated precision.

$$(p - P)^2 / (s^2/n)\ (\text{Deff}) = t^2\, 95\% \qquad (1)$$

where P = the proportion to be estimated

p = estimate of the proportion P

s = standard error of the estimate p

t95% = t-value at the 95% level of t-test

n = the size of the sample

Deff = the ratio of the two variance estimates, the estimate drawn from other than the Simple Random Sample (SRS), divided by the variance estimate of an SRS of same size (n).

The above equation can be written as

$$d^2 = \frac{s^2\,(t95\%)^2\,(\text{Deff})}{n} \qquad (2)$$

or

$$d^2 = \frac{P(1-P)\ (\text{Deff})\,(t95\%)^2}{n} \qquad (3)$$

where n = 1500 for a proposed sample

d = p – P = the amount of tolerated error in p

s^2 = P(1–P)

The proposed sample design for the baseline survey is similar to the design used for the 1996/97 Bangladesh Demographic and Health Survey (BDHS). Based on the experiences of the 1996/97 BDHS, Deff is therefore assumed to be at the most 1.5 for the samples for the

‡ Cochran, William G, 1977, *Sampling Techniques*, New York; John Wiley and Sons Inc.

baseline survey. The equation (3) can be further simplified by putting the value of t95% and the assumed value (1.5) of Deff, as

$$d^2 = \frac{P(1-P)(1.5)(1.96)^2}{n}$$

$$\text{or} \quad d = \sqrt{\frac{P(1-P)(5.76)}{n}} \qquad (4)$$

In the 1995 National Media Survey conducted during November 1994 through January 1995, it was found that 55 percent of rural men and 34 percent of rural women had access to radio; 36 percent of rural men and 23 percent of rural women listened to radio at least once a week, and 12.3 percent of all men and 8.3 percent of all women listened to agriculture related radio programmes. Assuming that the current values of the key parameters will not be larger than by more than 5 percentage points from their observed values, likely error margins (precision) of estimates in the proposed samples were calculated from equation (4) for the key parameters corresponding to both their observed values in the National Media Survey and their assumed maximum current values. The calculated likely error margins for the observed values are shown separately for the male and female samples in Table 1 and those for the assumed maximum current values are shown in Table 2.

While for either sample for any parameter the calculated error margins range from ± 1.71 to ± 3.08 percentage points for both the observed and assumed current values, it can therefore be concluded that the proposed male and female samples each with a size of 1500 respondents are 95 percent likely to provide estimates of the key parameters within error margins ± below 4 percentage points. It is thus seen that the proposed samples are large enough to have a reliable assessment of the impact of the SRB (Strengthening Rural Broadcasting) project at the acceptable level of accuracy.

Table 1

Likely error margins in estimates for the key parameters according to their observed values in the 1995 National Media Survey

Parameters	Male sample		Female sample	
	Estimates	Error margins	Estimates	Error margins
Proportion having access to radio	55.0%	±3.08	34.0%	±2.94
Proportion listening to radio at least once a week	36.0%	±2.98	23.0%	±2.61
Proportion listening to agriculture related programs	12.3%	±2.04	8.3%	±1.71

Table 2

Likely error margins in estimates for the key parameters according to their assumed current values

Parameters	Male sample		Female sample	
	Estimates	Error margins	Estimates	Error margins
Proportion having access to radio	61.0%	±3.02	39.0%	±3.02
Proportion listening to radio at least once a week	41.0%	±3.05	28.0%	±2.78
Proportion listening to radio at agriculture related programs	17.3%	±2.34	13.3%	±2.10

Appendix 1

Calculations of error margins for a proportion in complex survey designs

In this appendix, we give the background to the formulae used in this report. Further information can be obtained from the references given at the end of this document.

Suppose P is the true proportion of people in the population having access to a radio.

Let p be the corresponding proportion estimated from survey results.

Suppose we wish to estimate P by p so as to be more than 95% confident that the absolute difference |P-p| is less than d percentage points. Here d is the error margin we wish to determine. For example, if the true P is 0.40, we may wish the estimate p of P to fall within 0.37 and 0.43. Then d = 0.03.

i.e. Require $\text{Probability}\left\{|P - p| < d\right\} \geq 0.95$

Equivalently, if **se** is the standard error of p, we require

$$\text{Probability}\left\{\frac{|P-p|}{se} < \frac{d}{se}\right\} \geq 0.95 \qquad (5)$$

If p is determined on the basis of a large sample size (1500 is the example used here), we may assume $\frac{P-p}{se}$ follows a standard normal distribution.

Hence we must have from (5)

$$\frac{d}{se} = 1.96 \qquad \text{approximately, from tables of the standard normal distribution.} \qquad (6)$$

Prior to conducting the survey, **se** is unknown. We therefore use the design effect (deff) from a previous study. This is defined as follows.

$$\text{deff} = \frac{se^2}{\text{Variance estimate of a simple random sample of the same size}}$$

So far this is general and applies to any quantity that is being estimated. Here we are dealing with a proportion and hence the variance of the corresponding estimate, calculated from a simple random sample, is P(1 – P)N.

So $$\text{deff} = \frac{se^2}{\frac{P(1-P)}{n}}, \qquad \text{where n = proposed sample size.} \qquad (7)$$

Note that the variance estimate given in the denominator of the above expression, assumes that the finite population correction (i.e. n divided by the size of the population) is small enough to be neglected.

Obtaining **se** from (7) and substituting it in equation (6) gives

$$d = \sqrt{\frac{P(1-P)}{n} \times \text{deff} \times 1.96^2} \qquad (8)$$

Since deff is a measure which is portable across different surveys with the same design for the same, or similar types of variables, the value of deff from a previous survey, similar in design, may be used in the above expression.

In the Bangladesh survey, a previous survey of a similar design indicated that deff = 1.5 was suitable. So for example, if the proportion of men listening to agricultural related programmes was expected to be 17.3% (see last row of Table 2), i.e. P = 0.173, then (8), above, for n = 1500, gives

$$d = \sqrt{\frac{(0.173)(1-0.173)\times 1.5 \times (1.96)^2}{1500}}$$

$$= 0.0234.$$

This is the ±2.34% given in the last row of Table 2 for males.

Some comments:

(a) For a simple random sampling procedure, deff = 1; for a stratified sampling procedure, deff < 1; while with cluster sampling or multi-stage sampling designs, deff >1.

(b) The above expression can also be used to determine the sample size required to achieve a specified error margin d.

(c) The above calculations give absolute errors rather than relative errors. It can be translated into a relative error. For example, an absolute error of 2.34 percentage points for an estimated proportion of 17.3% males listening to agricultural related programmes (see Table 2 in main document) gives a relative percentage error of

$$\frac{2.34}{17.3} \times 100 = 13.3\% .$$

For estimates shown in Tables 1 and 2, the highest relative error is approximately 20% which would generally be considered an acceptable level for the relative error. It is also quite common for surveys to use either absolute or relative errors based on survey requirements.

(d) When sampling from a finite population of size N, the expression for determining sample size becomes

$$n = \frac{1 + \dfrac{V}{P(1-P)\text{deff}}}{\dfrac{1}{N} + \dfrac{V}{P(1-P)\text{deff}}} \qquad \text{where} \quad V = \frac{d^2}{(1.96)^2}$$

References:

(a) "An analysis of sample designs and sampling errors of the Demographic and Health Surveys" by Thanh N. Lê and Vijay K. Verma. Demographic and Health Surveys Analytical Reports No. 3. Macro International Inc. Calverton, Maryland, USA. (May 1997).

(b) "Survey Design and Analysis" by F.R. Jolliffe (1986). Wiley.