Survey Research Methods

4 EDITION

Floyd J. Fowler, Jr.

Center for Survey Research
University of Massachusetts Boston

1 APPLIED SOCIAL RESEARCH METHODS SERIES

Edited by Leonard Bickman and Debra J. Rog

SAGE

Los Angeles • London • New Delhi • Singapore • Washington DC

errors equivalent to simple random samples if there is no stratification. Stratified samples can produce sampling errors that are lower than those associated with simple random samples of the same size for variables that differ (on average) by stratum, if rates of selection are constant across strata.

Unequal rates of selection (selecting subgroups in the population at different rates) are designed to increase the precision of estimates for oversampled subgroups, thus

(a) they generally will produce sampling errors for the whole sample that are higher than those associated with simple random samples of the same size, for variables that differ by stratum, except

(b) when oversampling is targeted at strata that have higher than average variances for some variable, the overall sampling errors for those variables will be lower than for a simple random sample of the same size.

Clustering will tend to produce sampling errors that are higher than those associated with simple random samples of the same size for variables that are more homogeneous within clusters than in the population as a whole. Also, the larger the size of the cluster at the last stage, the larger the impact on sampling errors will usually be.

It often is not easy to anticipate the effects of design features on the precision of estimates. Design effects differ from study to study and for different variables in the same survey. To illustrate, suppose every house on various selected blocks was the same with respect to type of construction and whether or not it was occupied by the owner. Once one respondent on a block reports he is a home owner, the additional interviews on that block would yield absolutely no new information about the rate of home ownership in the population as a whole. For that reason, whether the researcher took one interview per block or 20 interviews per block, the reliability of that estimate would be exactly the same, basically proportionate to the number of blocks from which any interviews at all were taken. At the other extreme, the height of adults is likely to vary as much within a block as it does throughout a city. If the respondents on a block are as heterogeneous as the population as a whole, clustering does not decrease the precision of estimates of height from a sample of a given size. Thus, one has to look at the nature of the clusters or strata and what estimates are to be made in order to evaluate the likely effect of clustering on sampling errors.

The effects of the sample design on sampling errors often are unappreciated. It is not uncommon to see reports of confidence intervals that assume simple random sampling when the design was clustered. It also is not a simple matter to anticipate the size of design effects beforehand. As noted, the effects of the

sample design on sampling errors are different for every variable; their calculation is particularly complicated when a sample design has several deviations from simple random sampling, such as both clustering and stratification. Because the ability to calculate sampling errors is one of the principal strengths of the survey method, it is important that a statistician be involved in a survey with a complex sample design to ensure that sampling errors are calculated and reported appropriately. The problem of appropriately taking into account design features when estimating sampling errors has been greatly simplified by the fact that several available analysis packages will do those adjustments. (See Chapter 10.)

Finally, the appropriateness of any sample design feature can be evaluated only in the context of the overall survey objectives. Clustered designs are likely to save money both in sampling (listing) and in data collection. Moreover, it is common to find many variables for which clustering does not inflate the sampling errors very much. Oversampling one or more groups often is a cost-effective design. As with most issues discussed in this book, the important point is for a researcher to be aware of the potential costs and benefits of the options and to weigh them in the context of all the design options and the main purposes of the survey.

HOW BIG SHOULD A SAMPLE BE?

Of the many issues involved in sample design, one of the most common questions posed to a survey methodologist is how big a survey sample should be. Before providing an approach to answering this question, perhaps it is appropriate to discuss three common but inappropriate ways of answering it.

(1) One common misconception is that the adequacy of a sample depends heavily on the fraction of the population included in that sample—that somehow 1%, or 5%, or some other percentage of a population will make a sample credible. The estimates of sampling errors discussed above do not take into account the fraction of a population included in a sample. The sampling error estimates from the preceding equations and from Table 3.1 can be reduced by multiplying them by the value $(1 - f)$, where f = the fraction of the population included in a sample.

When one is sampling 10% or more of a population, this adjustment can have a discernible effect on sampling error estimates. The vast majority of survey samples, however, involve very small fractions of populations. In such instances, small increments in the fraction of the population included in a sample will have no effect on the ability of a researcher to generalize from a sample to a population.

The converse of this principle also should be noted. The size of the population from which a sample of a particular size is drawn has virtually no impact on how well that sample is likely to describe the population. A sample of 150 people will describe a population of 15,000 or 15 million with virtually the same degree of accuracy, assuming that all other aspects of the sample design and sampling procedures are the same. Compared to the total sample size and other design features such as clustering, the impact of the fraction of a population sampled on sampling errors is typically trivial. It is most unusual for it to be an important consideration when deciding on a sample size.

(2) A second inappropriate approach to deciding on sample size is somewhat easier to understand. Some people have been exposed to so-called standard survey studies, and from these they have derived a typical or appropriate sample size. Thus some people will say that good national survey samples generally are 1,500, or that good community samples are 500. Of course, it is not foolish to look at what other competent researchers have considered to be adequate sample sizes of a particular population. The sample size decision, however, like most other design decisions, must be made on a case-by-case basis, with the researchers considering the variety of goals to be achieved by a particular study and taking into account numerous other aspects of the research design.

(3) A third wrong approach to deciding on sample size is the most important one to address, for it can be found in many statistical textbooks. The approach goes like this: A researcher should decide how much margin of error he or she can tolerate or how much precision is required of estimates. Once one knows the need for precision, one simply uses a table such as Table 3.1, or appropriate variations thereon, to calculate the sample size needed to achieve the desired level of precision.

In some theoretical sense, there is nothing wrong with this approach. In practice, however, it provides little help to most researchers trying to design real studies. First, it is unusual to base a sample size decision on the need for precision of a single estimate. Most survey studies are designed to make numerous estimates, and the needed precision for these estimates is likely to vary.

In addition, it is unusual for a researcher to be able to specify a desired level of precision in more than the most general way. It is only the exception, rather than the common situation, when a specific acceptable margin for error can be specified in advance. Even in the latter case, the above approach implies that sampling error is the only or main source of error in a survey estimate. When a required level of precision from a sample survey is specified, it generally ignores the fact that there will be error from sources other than sampling. In such cases, the calculation of precision based on sampling error alone is an unrealistic oversimplification. Moreover, given fixed resources, increasing the sample size may even decrease precision by reducing resources devoted to response rates, question design, or the quality of data collection.

Estimates of sampling error, which are related to sample size, do play a role in analyses of how big a sample should be. This role, however, is complicated. The first prerequisite for determining a sample size is an analysis plan. The key component of that analysis plan usually is not an estimate of confidence intervals for the overall sample, but rather an outline of the subgroups within the total population for which separate estimates are required, together with some estimates of the fraction of the population that will fall into those subgroups. Typically, the design process moves quickly to identifying the smaller groups within the population for which figures are needed. The researcher then estimates how large a sample will be required in order to provide a minimally adequate sample of these small subgroups. Most sample size decisions do not focus on estimates for the total population; rather, they are concentrated on the minimum sample sizes that can be tolerated for the smallest subgroups of importance.

The process then turns to Table 3.1, not at the high end but at the low end of the sample size continuum. Are 50 observations adequate? If one studies Table 3.1, it can be seen that precision increases rather steadily up to sample sizes of 150 to 200. After that point, there is a much more modest gain to increasing sample size.

Like most decisions relating to research design, there is seldom a definitive answer about how large a sample should be for any given study. There are many ways to increase the reliability of survey estimates. Increasing sample size is one of them. Even if one cannot say that there is a single right answer, however, it can be said that there are three approaches to deciding on sample size that are inadequate. Specifying a fraction of the population to be included in the sample is never the right way to decide on a sample size. Sampling errors primarily depend on sample size, not on the proportion of the population in a sample. Saying that a particular sample size is the usual or typical approach to studying a population also is virtually always the wrong approach. An analysis plan that addresses the study's goals is the critical first step. Finally, it is very rare that calculating a desired confidence interval for one variable for an entire population is the determining calculation in how big a sample should be.

SAMPLING ERROR AS A COMPONENT OF TOTAL SURVEY ERROR

The sampling process can affect the quality of survey estimates in three different ways:

- If the sample frame excludes some people whom we want to describe, sample estimates will be biased to the extent that those omitted differ from those included.