Research Training
for Social Scientists

21 SAMPLING STRATEGIES
IN SURVEY RESEARCH

Contributors: Dawn Burton
Print Pub. Date: 2000
Print ISBN: 9780761963516
Online ISBN: 9780857028051
DOI: 10.4135/9780857028051
Print pages: 307-319

This PDF has been generated from SAGE Research Methods Online. Please note that the pagination of the online version will vary from the pagination of the print book.
21 SAMPLING STRATEGIES IN SURVEY RESEARCH

The purpose of this chapter is to introduce some of the issues confronting research students when they have to decide how to go about the process of selecting a sample. The first issue to be addressed is whether in fact students do need to select a sample, or whether it is possible for a census to be undertaken. However, since most researchers engaged in survey work tend to select a sample, most of this chapter is devoted to understanding the various types of sampling strategies available. A number of issues are addressed including the distinction between probability and non-probability sampling; various different types of probability and non-probability sampling; and a consideration of when some sampling strategies might be more appropriate than others. Finally, the chapter will conclude by discussing some of the issues associated with sampling error in surveys and how they might be minimized.

CENSUS OR SAMPLE?

Sometimes it is possible for researchers to survey all the cases in a specific population when the number of relevant cases are small. Surveying all cases in a population is called undertaking a census. A census is what researchers should aim towards, if at all possible. For example, to survey all twenty organizations which manufacture a particular product is manageable. For practical reasons researchers are not always able to undertake a census and therefore need to take a sample drawn from a relevant population. In some respects sampling can be a better strategy than undertaking a census. Several advantages of sampling are:

- data is often cheaper to collect because of the smaller numbers involved;
- fewer people are needed to collect and analyse the data;
- sample surveys are frequently quicker to administer, analyse and process;
- having fewer cases make it possible to collect more data about each.
Sampling approaches fall into two broad categories: probability and non-probability sampling. Choosing between these approaches is a matter of evaluating the issues of validity and credibility, against a realistic assessment of alternatives both in time and effort. The choice of sampling technique is extremely important since statistical theory and tests are used in survey research to assess whether or not the null hypothesis can be rejected or upheld. Statistical tests are sensitive to sample size, sample error and sample selection. Flawed sampling can therefore seriously undermine research findings. Three of the most important sampling issues and their implications are summarized in Table 21.1. First and foremost is population definition. Researchers need to consider carefully what population is to be sampled and then ensure the sampling strategy is consistent enough to produce accurate results. Second, probability sampling techniques require that all subjects have an equal probability of being selected. If inappropriate methods are used the sample could be biased, which will undermine the research findings. Finally, the precision needed to prove or disprove a hypothesis, or theoretical assumptions needs to be given consideration.

**TABLE 21.1 Issues in sample design**

<table>
<thead>
<tr>
<th>Issue</th>
<th>Criterion</th>
<th>Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population definition</td>
<td>Consistency of target population and study population</td>
<td>Study population yields biased results by including members not in target population or leaving out members who are in target population</td>
</tr>
<tr>
<td>Sampling method</td>
<td>Sample selection equally likely to select any member of study population</td>
<td>Sample methods yield biased results if some study population members are more likely to be selected than others</td>
</tr>
<tr>
<td>Precision of estimate</td>
<td>Estimate precise enough to inform policy decision</td>
<td>All samples yield estimates, not exact figures. Lack of precision can impact on the decisions to be made</td>
</tr>
</tbody>
</table>

**PROBABILITY SAMPLING**

The distinction between probability and non-probability sampling has been succinctly summarized by De Vaus (1996: 60): ‘A probability sample is one in which each person in the population has an equal, or at least a known, chance (probability) of being
selected while in a non-probability sample some people have a greater, but unknown, chance than others of selection.’

Probability samples are often regarded as preferable by survey researchers because they are more likely to produce representative samples and facilitate estimates of sample accuracy which allow inferences to be made to a wider population. There are several different types of probability sampling techniques which are set out in Table 21.2. The rest of this section will be concerned with drawing out some of the main features of this group of sampling techniques.

**TABLE 21.2 A survey of probability sample designs**

<table>
<thead>
<tr>
<th>Type of sampling</th>
<th>Selection strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple random</td>
<td>Each member of the study population has an equal probability of being selected</td>
</tr>
<tr>
<td>Systematic</td>
<td>Each member of the study population is either assembled or listed, a random start is designated, then members of the population are selected at equal intervals</td>
</tr>
<tr>
<td>Stratified</td>
<td>Each member of the study population is assigned to a group or stratum, then a simple random sample is selected from each stratum</td>
</tr>
<tr>
<td>Cluster</td>
<td>Each member of the study population is assigned to a group or cluster, then clusters are selected at random and all members of selected clusters are included in the sample</td>
</tr>
<tr>
<td>Multistage</td>
<td>Clusters are selected as in the cluster sample, then sample members are selected from the cluster members by simple random sampling. Clustering may be done at more than one stage</td>
</tr>
</tbody>
</table>

**Simple random sampling**

Statistical theory states that the most reliable way of obtaining a representative sample is to use random sampling whereby each case, whether that be an individual, household, or organization, has an equal probability of being selected. The objective is that ‘the measurement of a particular variable can be generalized, with a calculable degree of confidence, to the population from which the sample was drawn’ (Elliot and Ellingworth, 1997: 2.1). In order to construct a probability sample, it is necessary to use a sampling frame. A sampling frame might consist of a list of all the employees in a factory, or households on the electoral register in a particular location, and so on. Once an accurate frame has been located or constructed, each case in the sampling frame
should be given a number. Cases should then be chosen at random until the required number of cases which comprise the sample have been selected. Simple random sampling is usually deemed to be sampling without replacement. This means that once a case has been selected, then it should not be returned to the pool and hence not be eligible for selection again. This is in contrast to selection with replacement, whereby the case is returned to the pool and is eligible for reselection.

Problems with random sampling occur for a number of reasons. The first relates to the precision of the sampling frame, specifically, whether in fact all possible populations which should be included are accurately represented. If there is some doubt about the adequacy of the sampling frame, the sample cannot be accurately described as a probability sample. In implementing a random sampling strategy, the researcher has no choice whom they interview. Researchers need to consider whether this strategy is practical or desirable. Selection with replacement strategies have the undesirable property whereby selecting a particular case several times can occur. Finally, in practice many large-scale surveys do not use random sampling because of the excessive costs associated with sample selection and instead opt for a compromise which often utilizes stratified, multistage, random sampling (Elliot and Ellingworth, 1997).

**Systematic sampling**

Systematic sampling refers to the process whereby the researcher knows the number of cases in the sampling frame and has chosen an appropriate sample size. The researcher then divides the number of cases in the frame by the sample size and selects every x case (known as the sampling interval). For example, if the total number of individuals in the sampling frame is 120 and the researcher wished to interview a sample of 40 individuals, the sampling interval would be 1/3 indicating that every third person in the frame would be interviewed. The starting point for the selection of cases is chosen at random.

Problems can arise in systematic sampling if the sampling frame is ordered in a way that might reflect a trend, which would then affect the characteristics in the sample: for example, a list of employees ordered by age, or income. The sampling interval might correspond to a particular characteristic of the population: for example, if every
twentieth house was sampled and they were all at the end of a street, this could be problematic if you wanted to interview people about their neighbours.

**Stratified sampling**

Stratified sampling is a modification of simple random sampling and systematic sampling techniques but includes an extra process. Stratified sampling involves identifying individuals with certain ‘target’ characteristics (for example, age, gender, ethnicity, etc.) and then drawing a sample, using simple random sampling or systematic sampling techniques, from each of the groups. This strategy is useful if the aim is to compare groups. The sampling technique requires that subjects in the sample are reflected in the same proportions as those in the population. Stratified sampling is widely used in social science research.

**TABLE 21.3 An example of multi-stage cluster sampling**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>Divide the city into a number of districts which could either be electorates or census districts</td>
</tr>
<tr>
<td>Stage 2</td>
<td>Divide each of the districts into blocks using an up-to-date street map</td>
</tr>
<tr>
<td>Stage 3</td>
<td>Draw up the list of all households in each of the blocks</td>
</tr>
<tr>
<td>Stage 4</td>
<td>Select households to include</td>
</tr>
<tr>
<td>Stage 5</td>
<td>Select people within each household to interview</td>
</tr>
</tbody>
</table>

**Multi-stage cluster sampling**

The sampling strategies already discussed in this chapter have limited use when a sampling frame does not exist. They are also not appropriate for sampling a geographically dispersed population. Multi-stage cluster sampling requires the researcher initially to draw a sample from a large population and then select out progressively smaller populations until the required number of cases are included in the sample. So for example, multi-stage sampling might be used to select a sample of households in a particular geographical area by using the stages set out in Table 21.3.
LIMITATIONS OF PROBABILITY SAMPLING

Probability sampling becomes highly problematic if the information required to construct a sampling frame does not exist. This limitation could mean that researchers might miss out on important areas of research being undertaken, for example, deviant types of behaviour (Coomber, 1997). The inability to construct a suitable sampling frame of users on the Internet is one reason why it is not a suitable vehicle for constructing probability samples of the general population (Schillewaert et al., 1998). If an adequate sampling frame needs to be constructed from scratch it can be a labour-intensive enterprise and a judgement needs to be made about whether it is worth the time and effort. Problems can also arise if a sampling frame exists but is deficient in some way. For example, it has been argued that the Electoral Register is not necessarily a good sampling frame for the British population because young people, the transient and ethnic minority groups are less likely to be included (Elliot and Ellingworth, 1997). It may also be the case that the population is so widely geographically dispersed as to make cluster sampling inefficient. Finally, if not undertaken extremely accurately, probability sampling can generate a significant number of different types of error. Bryman and Cramer (1993) suggest that response rates on sample surveys are sometimes so low that the issue of representativeness between probability and non-probability samples is often not as great as one might expect. Even where random samples have been selected, factors such as non-response may adversely affect a sample's representativeness.

TABLE 21.4 Some non-probability sample designs
A range of non-probability sampling techniques are available to researchers. Kalton (1983) has used threefold classification to summarize different non-probability techniques: haphazard, convenience, or accidental sampling; judgement or purposive sampling or expert choice; and quota sampling. Within each of these categories there are a number of designs. Henry (1990) provides a more extensive typology which is summarized in Table 21.4. In the following section of the chapter, some of the characteristics of non-probability sampling techniques will be discussed in more depth.

**Quota sampling**

Quota sampling is a widely used market research technique. We have probably all had the experience of avoiding market researchers with clipboards standing on street corners. Quota sampling attempts to approximate or represent the population characteristics by dividing the sample according to a number of specific characteristics such as gender, age and social class. The distributions of each of the groups in the population can be ascertained by using census data. Interviewers are then required to collect data which meets the *quota target*. The quota targets might be to sample 10 men and 10 women, 10 of whom are in the age range 20–30 and 10 in the range 40–50, as Table 21.5 illustrates. This example is very simple and straightforward. However, quota sampling can become very complicated when additional variables are introduced. Quota sampling techniques are examples of non-random or non-probability sampling because...
the interviewer chooses any case which fits the criteria. Therefore human judgement enters the selection process.

**TABLE 21.5 An example of quota sampling**

<table>
<thead>
<tr>
<th>Age</th>
<th>Men</th>
<th>Women</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 - 30</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>40 - 30</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>10</td>
<td>20</td>
</tr>
</tbody>
</table>

Kent (1993: 52) suggests there are a number of advantages of using quota samples:

- They are quicker, cheaper and more straightforward to administer than random samples.
- They do not require a sampling frame.
- The sample size and quota composition are usually achieved.

Quota sampling is usually quicker to implement than other strategies for two main reasons. First, for in-home quotas (calling on people in their own homes) no call backs are necessary because the researcher continues on until the requisite number of respondents have been included. For street quotas, there is no travelling time between interviews. Second, the procedures for drawing up samples are very easy because no sampling frame is required. In terms of cost per interview, quota samples are much cheaper than many other sampling techniques.

While there are certainly advantages to quota sampling, Kent also identifies a number of disadvantages:

- There is a considerable potential for bias.
- There is more variability between samples than using other techniques.
- The application of probability theory to such samples is questionable.
- The sample may not be representative of the population as a whole or the locality in which the sampling was undertaken.

Bias can be introduced in two ways: that generated by interviewer behaviour and a high level of non-response. Usually it is the researcher's responsibility to choose
respondents at the sampling point. This practice can lead to bias as interviewers may select individuals whom they believe will be the easiest to interview. These characteristics can change according to the preference of particular researchers which makes the problem of bias even more complex. The issue of non-response is largely hidden in quota research. As long as the quotas are filled, there is no need to declare the level of non-response (for example, individuals who refuse to participate) as substitutions are allowed. Levels of non-response and reasons for non-response can have important implications for research conclusions. Kent estimates that, at best, quota samples have an effective response rate of 25–30 per cent. A final problem is that accurate population proportions may not be available, which makes quota sampling extremely difficult. For example, if census data is to be used as a basis for constructing quota categories it could be out of date as the census is only undertaken every ten years. At a more micro level, localities may be subject to major social, economic and political change. The data upon which quota estimates are based may therefore not be representative of anything concrete.

**Multi-purpose sampling**

An existing survey can sometimes be used as a vehicle for reaching the target population. Multi-purpose sampling can take two forms:

- **Piggybacking** – using the sample of an existing survey.
- **Amalgam** – join forces with other researchers to identify a larger number of subjects.

Piggyback sampling would appear to be a useful way of enlisting the help of other researchers in identifying a suitable sample. It has the advantage of saving time in identifying an appropriate sample, especially when respondents might be hard to reach for one reason or another. Despite these advantages there are two main problems with piggyback sampling. First, there is the issue of confidentiality, such as giving respondents’ names and perhaps addresses to other researchers. This in turn raises ethical issues about informed consent (see Chapter 5). Second, the original survey sample may not necessarily be the ideal vehicle for a subsidiary study.
Amalgam sampling raises another set of issues about ownership of the data and publishing responsibilities and rights. There is the other possibility of researchers having different priorities and agendas for the research project which might only emerge once the project has begun. Punch (1986) describes a situation where academics nearly came to blows about disagreements concerning the priorities of a research project. Since research students are in a position where their thesis is going to be judged on originality and is a product of their own work, both of these sampling strategies need to be evaluated within that wider context. Supervisors should be able to advise on appropriate ways forward.

**Networking or snowball sampling**

In network sampling the researcher establishes contact with a suitable respondent, then asks that respondent for other contacts with the required characteristics. Network sampling is particularly useful when respondents are highly stigmatized or vulnerable, which can make them more difficult to reach. In network sampling there is a great deal of emphasis on the researcher to prove that they are a bona fide researcher and some security features might have to be built into the research design. Network sampling often requires the researcher to work through intermediaries and there is the possibility that the link person may misinterpret the aims and objectives of the research. Accurately explaining the aims and objectives to all involved is therefore crucial. Another concern is that it is often difficult to keep track of refusals if working through intermediaries. It could easily be the case that refusals might be reported when in fact no contact is made (Lee, 1993). There is also a problem of bias. Networks tend to be homogeneous in their attributes, rather than providing links to others who have different social characteristics (if that is what is required). It is therefore good practice for researchers to pace and monitor very carefully the referral chains which they generate.

**Outcropping**

Outcropping is a method of sampling the target population in a geographic location or area in which they routinely congregate with a view to surveying them. The areas in
which they might meet could be a coffee bar, a red light district, or some other place. Some disadvantages are that if the chosen location has a high turnover, data capture may be difficult. As with network sampling, there is no guarantee that a sample drawn from a particular setting is representative of the wider population beyond the sample.

**Advertising**

If all else fails, advertising is a way of obtaining a sample. However, major problems arise because of the lack of control over who responds, both in terms of representativeness and suitability. If advertising is to be used via an established route (for example, newspapers, radio, television), a press release must be constructed. Some notes of guidance may be required regarding press releases:

- Press releases should be no longer than a side of A4. They are best handled by the university press office, which will usually have names and addresses of all relevant contacts at hand. In fact, it is best if guidance from the press office is sought as early as possible. Press releases should state very clearly the objectives of the research, who the target sample is, what the research findings will be used for, and how individuals should proceed if they wish to help out.
- There is no guarantee that press releases will be taken up at all.
- Press releases can be distorted, and represented in a way not intended by the researcher. Beware!

**THE UTILITY OF NON-PROBABILITY SAMPLES**

Non-probability sampling is a useful method of sampling under some circumstances, indeed on occasions it may be the only method available. Where the researcher is specifically interested in respondents within the sample as opposed to making wider generalization or inference, the non-probability method is likely to be more appropriate. In exploratory research, a non-probability sample would also seem to be a suitable
choice and could perhaps take the form of a small pilot study with the results being integrated into a larger probability sample at a later date. Limited resources, inability to identify members of the population and the need to establish the existence of a problem all justify the use of non-probability sampling. Lee (1993) notes that sampling becomes more difficult the more sensitive the investigation, as potential respondents have more to hide. Less visible individuals are more difficult to sample. Obtaining a sample of respondents who have characteristics which are rare is more difficult and costly to sample than individuals who display frequently occurring behaviour. Sampling subjects over a wide geographical area is likely to improve representativeness, but the financial costs of doing so may be excessive.

While there are clearly advantages of using non-probability samples, there are also disadvantages. Henry (1990) notes that as a result of the subjective nature of the selection process involved in non-probability sampling, there is concern about the generalizability of the findings to the wider population. There is a risk that the findings will not be valid because of bias in the selection process. Kalton argues a similar point:

The major strength of probability sampling is that the probability selection mechanism permits the development of statistical theory to examine the properties of sample estimators. Thus estimators with little or no bias can be used, and estimates of the precision of the sample estimates can be made. The weakness of all non-probability methods is that no such theoretical development is possible. As a consequence, non-probability samples can be assessed only by subjective evaluation. Moreover, even though experience may have shown that a non-probability method has worked well in the past, this provides no guarantee that it will continue to do so in the future. (Kalton, 1983: 90)
SOURCES OF ERROR IN SURVEY RESEARCH

As with all research processes, there is potential for errors to occur in survey research. Sources of error fall into two main categories: nonsampling errors; sampling errors.

Non-sampling errors

Non-sampling errors arise as a result of decisions during data collection and are not directly related to the selection of the sample. Non-sampling errors can take a variety of forms.

*Response errors* These occur when respondents give ‘wrong’ answers. Response errors may occur due to dishonesty, poor memory, or misunderstanding, among others.

*Interviewer errors* These refer to the inaccurate recording of the response given by subjects, making mistakes, or asking questions in a non-standard fashion.

*Non-response errors* These are generated by inadequacies in the research process. Non-response errors can result from subjects who refuse to provide data, those who are asked to provide data but who are unable to undertake the required task and, finally, those whom the data collection procedures did not reach; therefore not providing individuals with an opportunity to respond. High levels of non-response can have serious implications for inference; the extent to which findings from the sample can be generalized to a wider population. For example, if a survey yields response rates of 10–25 per cent of a sample, the final sample bears little relationship to the original sample as those responding are self-selected. Whereas if a response rate of 95 per cent is achieved in a probability sample, the final sample is still very similar to the population as a whole. Fowler (1993) suggests that there are different types of bias associated with different methods of data collection. For example, mail questionnaires are more likely to be returned by individuals who are particularly interested in the project. As a result low response rates can be significantly biased in ways that are directly related to the
purpose of the research. In addition, better educated people tend to send back mail questionnaires than those who are less well educated. Bias resulting from telephone and personal surveys tends to focus on availability. For example, if a survey was undertaken on weekdays between 9am and 5pm the likely scenario is that individuals without jobs (unemployed, retired and housewives) would be over-represented in the sample.

There also tend to be lower response rates in cities than in suburbs and rural areas, which is largely a result of the disproportionate number of single people living in cities who can be difficult to contact; the nature of housing, specifically the greater number of individuals who live in flats, which makes access difficult; and the fact that some areas in the centre of cities are undesirable places to conduct fieldwork, especially at night. Researchers cannot assume that non-response is unbiased and it is not always obvious in what ways non-response is biased, which makes it difficult to counteract. The ideal solution is to reduce the non-response by amending data collection techniques such as varying methods and undertaking systematic follow-ups.

Processing errors These are coding and data entry errors. Constructing a comprehensive coding book and having in place quality control procedures should reduce such errors.

Sampling errors

Sampling errors directly relate to the sample selection process. Sampling errors can be divided into two main groups: systematic error (or bias), and random error. Hammersley and Comm provide a useful statement of the difference between the two:

Bias is generally seen as a negative feature, as something that can and should be avoided. Often, the term refers to any systematic deviation from validity, or to some deformation of research practice that produces such deviation. Thus, quantitative researchers routinely refer to measure or sampling bias, by which they mean systematic error in measurement or sampling procedures that produce erroneous results. The contrast here is with random (or haphazard) error: where
bias tends to produce spurious results, random error may obscure true conclusions. (Hammersley and Comm, 1997: 1.6)

Systematic error usually occurs as a result of one of the following: the selection process was not random, when it was intended to be; the selection of respondents was made from a list that did not cover the whole population; and non-respondents were not a cross-section of the population.

Random error relates to the error which can arise as a result of the sample size. The larger the sample, the less random error there is likely to be as the sample is more likely to be representative of the population from which it was drawn. A sample used to estimate a variable that varies widely in the population will show more random sampling error than for a variable that does not.

Quite controversially, Hammersley and Comm (1997: 1.7) argue that systematic error does not always occur by accident and that it can also include ‘a tendency on the part of researchers to collect data, and/or to interpret and present them, in such a way as to favour false results that are in line with their prejudgements and political or practical commitments’. This may include the positive tendency to report findings that promote false conclusions, or exclude from consideration other conclusions which include the truth. They argue that the abandonment of the positivist, or foundationalist's, view of error and the adoption of a more radical post-foundationalist approach leads to a much more complicated view of error with the causes of systematic error being defined in different ways. Whereas previously it was considered as relying on logic, it becomes ‘deviance from communal judgements’ about what is reasonable and unreasonable behaviour in the pursuit of knowledge, with these judgements being open to dispute and frequent revision. They acknowledge that there will always be the potential for systematic error and that some of it will be non-culpable by virtue of the fact that the researcher was unaware that the knowledge being relied upon was erroneous or dysfunctional which inadvertently led to wrong conclusions. However, it is also the case that some systematic error will be culpable, in the sense that researchers are judged to be in a position where they should have been aware that an assumption on which they were relying had an unacceptable chance of being erroneous and therefore held the possibility of leading them astray.
HOW LARGE SHOULD A SAMPLE BE?

The issue which tends to concern research students undertaking survey research more than any other is how large a sample should be. This is a very sensible question, but unfortunately there are no definitive answers. However, there are some guidelines:

2. Of particular importance are time and resource constraints. There is no point in designing a project which looks brilliant on paper, but is just not practical. It is usually best to keep projects relatively small scale, something you feel happy with, and you feel you can do well rather than being over-ambitious. The temptation to work with a large sample can result in students spending longer than anticipated collecting and processing data, rather than analysing and linking it to contemporary debates. The lack of in-depth analysis is likely to be picked up on at some stage, especially in a viva. It is therefore very important to keep your sample to manageable size.

4. As a general comment the larger the sample, the greater the degree of accuracy. In effect the sampling error (the difference between the sample and the population which are due to sampling) can be reduced by increasing the sample size. After a certain level, gains via increase in accuracy are curtailed (see Fowler, 1993; De Vaus, 1996 on this issue).

6. You need to bear in mind the issue of non-response and adjust your sample accordingly. If non-response rates are high, you will need to increase the size of your sample to compensate. Alternatively, a critical evaluation of your data collection methods will be required. You should have an indication of non-response rates from your pilot study.

8. A final issue relates to the type of analysis you wish to conduct. Specifically, how many cases are required in each sub-group. Unless you have enough cases in each of the sub-samples in which you are interested, your analysis will be limited.