

CHAPTER 10

SAMPLING FRAMES FOR AGRICULTURAL CENSUSES AND SURVEYS

One of the important aims of the census of agriculture is to provide a sampling frame for the agricultural survey programme. This chapter explains what a sampling frame is, and how a census can be used to create a sampling frame. It also discusses the use of population and agricultural censuses to create sampling frames for the agricultural census supplementary modules and the programme of agricultural surveys. Both single-stage and multi-stage sampling frames are considered, and the problems of keeping sampling frames up-to-date are highlighted.

What is a sampling frame?

10.1. In a census, each unit (such as person, household or holding) is enumerated, whereas in a sample survey, only a sample of units is enumerated and information provided by the sample is used to make estimates relating to all units. In an agricultural production sample survey, for example, a sample of agricultural holdings is enumerated, and information from the sample holdings is used to make estimates of total agricultural production.

10.2. In a sample survey, the sample of units to be enumerated must be selected using strict statistical procedures. A method known as random sampling is used. Random sampling is the process of selecting units for inclusion in the sample in such a way that each unit has a known, though not necessarily the same, chance (or probability) of selection. The simplest type of random sample is one selected by “lottery”, where all units have the same chance of selection in the sample; for example, in an agricultural survey, each agricultural holding would have the same chance of selection. Usually, sampling schemes are more complex than this, with units having differing probabilities of selection in the sample. In an agricultural survey, for example, large holdings may be given more chance of selection than small holdings; some very large holdings may even be completely enumerated.

10.3. To select a random sample for a sample survey, one needs, first, to clearly define which units are within the scope of the survey. Some specific surveys and the in-scope units are:

- National agricultural survey: all agricultural holdings in the country.
- Agricultural survey in Province A: all agricultural holdings in Province A.
- National rice production survey: all rice producers in the country.
- National aquacultural production survey: all aquacultural producers in the country.

10.4. Having determined the scope of the survey, a means of identifying all the in-scope units is needed, so that each unit can be given the required chance of selection in the sample. This is called a sampling frame. A sampling frame could be a list of units (such as households or holdings), areas (such as EAs), or any other materials (such as maps), and may also include information about each unit, such as their size, to help with the sample selection or survey estimation (FAO, 1989, pp 32–41; UN, 1986).

10.5. The best type of sampling frame is a list of all units within the scope of the survey. For example, for a national agricultural survey, the sampling frame would be a list of all agricultural holdings in the country. Here, the sample of holdings would be selected directly from this list, by giving each holding on the list the appropriate chance of selection in the sample. The sampling frame must provide a complete and up-to-date list of holdings, without omissions or duplications, and without including any units other than holdings. Holdings missing from the frame would have no chance of selection in the sample, while duplicate holdings would have more chance of selection than they should. In these circumstances, the sample would no longer be a random sample, which would bias the survey results.

10.6. Often, a list of units within the scope of the survey is not available. For example, an agricultural census identifies all agricultural holdings at the time of the census, but this would not provide an accurate list of holdings for a survey conducted some time after the census. Lists of households or agricultural holdings from a census quickly become out-of-date because:

- New households are created.
- Households cease to exist because people die or because of changes in family circumstances.
- Households move from one place to another.
- New agricultural holdings are created as households become agricultural producers.
- Agricultural holdings cease to exist as households stop being agricultural producers.
- The management structure of an agricultural holding changes, with the result that the holding is split into two, or combined with another holding.

10.7. Often, updating lists of households or agricultural holdings from a population or agricultural census for sampling frame purposes is too difficult or expensive. Instead, a sampling technique known as multi-stage sampling is used. In multi-stage sampling, random sampling is carried out in stages, as opposed to single-stage sampling where the sample is selected directly from lists of households or holdings. Thus, for an agricultural survey, a sample of EAs could be selected first, and then a sample of agricultural holdings selected in each sample EA. In multi-stage sampling, sampling frames are needed for each stage of sampling: in the example above, a list of all EAs in the country to select the sample of EAs, and lists of agricultural holdings in each sample EA to select the sample of holdings.

10.8. Multi-stage sampling is widely used for agricultural surveys, especially for the household sector. Its main advantage is that it is cheaper and easier to create lists of holdings just in the selected areas, rather than for the whole country. Data collection is also cheaper because the sample holdings are concentrated in the selected areas, rather being spread around the whole country. However, sampling errors are higher because of the “clustering” of sample in selected areas. Sometimes, multi-stage sampling is used in conjunction with single-stage sampling (see paragraphs 10.25–10.27).

10.9. Note that, for an agricultural survey, it is not necessary to have any information about agriculture on the sampling frame. A sampling frame of households, rather than agricultural holdings, is often used for the household sector as part of a multi-stage sampling scheme. Often the population census, rather than the agricultural census, is used for this purpose (see paragraphs 10.31–10.33).

10.10. Another type of sampling frame often used in agricultural surveys is an area sample frame. In area sampling, the unit being sampled is a physical piece of land, called a segment. A sample of segments is selected and data are collected in respect of the agricultural activities of each sample segment. Here, the sampling frame consists of all the segments making up the whole area within the scope of the survey. For example, for an agricultural survey in Province A, the sampling frame could be a map of Province A divided into clearly defined segments.

10.11. A variety of sampling techniques – such as stratification, systematic sampling, and probability proportional to size sampling – can be used to improve the efficiency of the sample design. Techniques such as ratio and regression estimation may also be used to improve the reliability of survey data. A description of these techniques is outside the scope of this publication.

Sampling frames for census supplementary modules

10.12. In the agricultural census, the supplementary modules are undertaken at the same time as, or soon after, the core module. Thus, the core module provides up-to-date lists of holdings for use as sampling frames for the supplementary modules. Examples of supplementary modules and the relevant sampling frames from the core module are:

- Crop supplementary module: list of agricultural holdings with temporary crops in Item 0011 or permanent crops in Item 0012.

- Livestock supplementary module: list of agricultural holdings with livestock in Item 0013.
- Agricultural practices module: list of all agricultural holdings.

10.13. One way to carry out a supplementary module based on the core census module is to use single-stage sampling. For example, a livestock module could be carried out as follows:

- Conduct the core census module by enumerating all holdings.
- During the core census enumeration, identify all holdings with livestock, to be used as a sampling frame for the livestock module.
- Select a sample of holdings with livestock based on this sampling frame, in accordance with the required sampling scheme, and enumerate those holdings for the livestock module.

10.14. This method may be difficult to implement for a supplementary module carried out at the same time as the core module, because enumerators would need to do the sample selection in the field. A multi-stage sampling approach is more commonly used as follows:

- Divide the country into EAs for the purpose of organizing the enumeration of the core census module.
- Prior to the census enumeration, select a sample of EAs for the livestock module.
- Conduct the core census module by enumerating all holdings in all EAs.
- During the core census enumeration, identify all holdings with livestock in the sample EAs. Special holdings with livestock, such as large units, would also be identified in the non-sample EAs.
- For the livestock module, enumerate all holdings with livestock in the sample EAs and all special holdings with livestock in the non-sample EAs.

10.15. The advantage of the multi-stage approach is that the sample selection of EAs can be done by technical staff prior to the fieldwork, rather than requiring each enumerator to do the sample selection. This makes the census field operations easier. A convenient way to organize the census enumeration would be to assign the best enumerators to the sample EAs, to interview each holding for the core module, and, if the holding is within the scope of the supplementary module, ask further questions for the supplementary module. All other enumerators would be assigned to the non-sample EAs to collect core data only. Senior field staff could enumerate the special holdings.

Sampling frames for the programme of agricultural surveys

10.16. Some examples of agricultural surveys and the applicable sampling frames from the core census module are shown below:

- Rice production survey: list of holdings growing rice in Item 0011.
- Pig production survey: list of holdings with pigs in Item 0013.
- Gender survey: list of holdings with sex of agricultural holder as female in Item 0003.
- Survey of young farmers: list of holdings with age of holder less than 25 years in Item 0004.

10.17. From a sampling point of view, an agricultural survey is like a census supplementary module, except that it is not carried out as part of the agricultural census, but some time later. This has implications for the sampling methodology. For example, a list of holdings with pigs from the core census module would not be accurate for a pig production survey conducted some years after the agricultural census. Even a gap of several months could result in serious shortcomings in the sampling frame.

10.18. Where a list of in-scope units from the core census module is deemed to be acceptable as a sampling frame for an agricultural survey, the frame can be established in a similar way to that for a census supplementary module, using either single-stage sampling (see paragraph 10.13) or multi-stage sampling (see paragraph 10.14). The single-stage approach is usually better suited to an agricultural

survey than a census supplementary survey, because the survey is carried out some time after the agricultural census, and sample selection can be done by technical staff prior to the survey, rather than by the enumerators as in a census supplementary module. However, the multi-stage approach is often still preferred because the “clustering” of sample reduces data collection costs (see paragraph 10.8).

10.19. Where a list of in-scope units from the core census module is not accurate enough to be used as a sampling frame for the agricultural survey, a different sampling approach is needed. Some alternatives are discussed below.

(a) Update the sampling frame of in-scope units: single-stage sampling

10.20. The best approach is to maintain an up-to-date list of agricultural holdings to provide a sampling frame for agricultural surveys undertaken at any time. Some countries maintain a register of holdings, containing basic information about each holding such as main agricultural activity and size. If specific crop production surveys are required, information about specific crops grown would also need to be provided on the register.

10.21. Keeping registers of holdings up-to-date is difficult and expensive. Often, it can be done by making use of information from government regulatory agencies, producers’ associations, telephone directories, or other administrative sources. Sometimes, the results of surveys and other statistical activities can be used to update the register. Many countries are unable to make available the necessary resources to do this work.

(b) Multi-stage sampling

10.22. Even if the list of holdings from the agricultural census core module is not good enough to directly select the sample for the agricultural survey, the core census data can still be useful for the sample design and fieldwork in a multi-stage sample design. A common sampling method, using a wheat production survey as an example, is given below:

- Select a sample of agricultural census EAs, using sampling techniques such as stratification and probability proportional to size sampling, based on wheat data from the core census module. A typical design would sample important wheat growing areas more heavily than other areas.
- Prepare a list of wheat producers in each sample EA, by updating lists of units available from the agricultural census core module.
- Select a sample of wheat producers in each sample EA, and enumerate those units for the survey.

10.23. Multi-stage surveys may not be efficient for agricultural surveys because they do not allow large holdings to be sampled more heavily than small holdings. This is of particular concern where there are a few dominant holdings, especially in the non-household sector. For this reason, a combination of single- and multi-stage sampling is preferred for agricultural surveys (see paragraph 10.25–10.27).

10.24. Another problem with multi-stage surveys of this type is that there can be changes in the administrative structure on which the EAs are based. Low-level administrative units such as communes and villages may change frequently, and this makes it difficult for enumerators to identify EAs in a survey conducted some years after the agricultural census. Special field procedures are needed to deal with this problem.

(c) Combination of single- and multi-stage sampling

10.25. Another approach is to use a combination of single- and multi-stage sampling. One way to do this is to create an up-to-date list of holdings for certain types of holdings only, with single-stage sampling used for those units and multi-stage sampling used for all other units.

10.26. For agricultural surveys, single-stage sampling is often used for the non-household sector or for large or otherwise important units, using a sampling frame of agricultural holdings obtained from the

agricultural census and updated using business registration records. Multi-stage sampling is then used to cover the household sector.

10.27. Another way is to use a list of holdings from the agricultural census as a sampling frame for those units present at the time of the census, with a small supplementary multi-stage sample used for other units.

Special sampling frame situations for agricultural surveys

(a) Periodic surveys

10.28. Some agricultural surveys are conducted on a regular basis; for example, a cassava production survey may be needed once a year to measure annual cassava production. For such surveys, the sampling frame must be updated each time the survey is run. Where multi-stage sampling is used, the list of units – for example, cassava production holdings – in each sample area must be updated for each round of the survey.

(b) Longitudinal surveys

10.29. A longitudinal survey is a special type of periodic survey aimed at studying changes in the behaviour of a particular group over time; for example, to assess how a group of maize farmers change their farming practices over time. Unlike normal ongoing surveys, longitudinal surveys are not designed to provide aggregated data, such as national maize production. In a longitudinal survey, the sample of units is selected at the start of the study and those units are followed up in each round of the survey. Longitudinal surveys only need a sampling frame at the beginning of the study. An agricultural census provides an ideal frame for a longitudinal study beginning soon after the census.

(c) Census supplementary modules as sampling frames for agricultural surveys

10.30. Usually, the sampling frame for an agricultural survey is obtained from the core module of the agricultural census. However, a supplementary module can also provide a good sampling frame, as shown in the following example for a survey of organic farmers.

- Conduct the agricultural census core module as normal.
- As part of the agricultural census, carry out a census supplementary module on agricultural practices, based on the sampling frame of agricultural holdings from the core module.
- Form a list of organic farmers from the agricultural practices module, and use this as a frame to select the sample of organic farmers for the survey of organic farming. This frame would need to be updated if the survey is conducted some time after the agricultural census.

(d) Use of the population census as a household sampling frame for agricultural surveys

10.31. As noted in paragraph 10.9, a household sampling frame from the population census can be used to select the sample for the household sector of an agricultural survey. Non-household units would be covered by a separate sampling frame. A common household sampling method is:

- Select a sample of population census EAs.
- In each sample EA, form a list of population census households, updated as necessary if the survey is conducted some time after the population census. Select a sample of households in each sample EA, based on these household lists.
- Interview each sample household to ask screening questions to find out if the household contains an agricultural holding within the scope of the survey.
- Enumerate each holding identified in this way to collect data for the survey.

10.32. A population census sometimes provides a better sampling frame than an agricultural census, because it covers all households in the country, rather than just agricultural holdings. It can therefore be

used as a sampling frame for agricultural surveys with a wider coverage than agricultural holdings. For example, a household food security survey usually covers all rural households or even all households in the country. Also, a farm labour survey should cover not only agricultural holdings, but also other households with members working in agriculture.

10.33. Sampling error issues need to be considered when using the population census as a frame for agricultural surveys. The smaller the proportion of households covered by the survey, the higher will be the sampling errors. Thus, the household sampling approach might be suitable for a chicken production survey, but would be less satisfactory for a goose production survey.

(e) Master sampling frames

10.34. A master sampling frame is a general purpose sampling frame created from a census, for use in selecting samples for different surveys or different rounds of a periodic survey. The frame is usually maintained by the national statistical office, and is updated on an ongoing basis so that it is available for any survey carried out at any time.

10.35. A master sampling frame has several benefits. It is quick and easy to conduct surveys of any kind, because a ready-made frame is already available. The cost of preparing sampling materials and selecting samples is also reduced. Master sampling frames also make it easier to relate data from different surveys and to control the reporting burden on survey respondents.

10.36. Master sampling frames suitable for agricultural surveys may be available from either the population census or the agricultural census.

- Population census. The population census master sampling frame is a database of small geographical units, such as villages or EAs, containing key data about each unit, such as the population and the number of households. This can be used to select samples for any type of household survey, including agricultural surveys (see paragraphs 10.31–10.33). A supplementary frame may be needed for non-household holdings and any other large holdings.
- Agricultural census. Agricultural census master sampling frames are of two types. One type is similar to a population census master sampling frame in that it is a database of small geographical units, such as villages or EAs, containing key agricultural data about each unit, such as the area of major crops and the number of livestock. Another type of agricultural census master sampling frame is a database of agricultural holdings, created and maintained following an agricultural census, containing key agricultural data about each holding, such as the crops grown and the livestock raised. Often, the master sampling frame is a combination of the two frame types, with a database of all important agricultural holdings complementing the EA frame. A master sampling frame from an agricultural census can be used for any type of survey of agricultural holdings.

10.37. A similar form of sampling frame is a register of businesses, containing information such as management structure, industry, employment, and turnover for each business unit. This can be used as a sampling frame for any economic survey. Usually, such frames include agriculture, sometimes for the non-household sector only, and can also be used for agricultural surveys.

Sampling frames for a sample-based agricultural census core module

10.38. So far, this chapter has looked at sampling frames from the agricultural census in situations where the core census module is based on a complete enumeration of agricultural holdings. This section considers sampling frame issues where the core module is carried out on a sample basis.

10.39. In a sample-based core census module, a sampling frame is needed for the core module itself. For the household sector, this is normally based on the population census. The sampling methods are similar to those for any other agricultural survey. Sometimes, the population census provides information on

households engaged in own-account agricultural production for use in the sample design and selection for the core module. For the non-household sector, a frame from administrative sources is often used.

10.40. A sample-based core census module can still provide a sampling frame for the census supplementary modules and the programme of agricultural surveys, even though the core module itself is only a sample of agricultural holdings. This can be done by sub-sampling core module holdings. This is illustrated in Figure 10.1, using an aquacultural supplementary module as an example.

10.41. The main proviso in the use of a sample-based core census module as a sampling frame for the census supplementary modules is that the core census sample must be big enough to provide sufficient sample for the supplementary modules. Thus, a sample of 100,000 holdings for the core census module might yield 10,000 holdings with aquaculture, more than sufficient for an aquacultural module. However, it might yield only 50 holdings growing potatoes, which would be insufficient for an in-depth potato survey. Countries should plan their agricultural census and survey programme at the outset to ensure that data requirements can be met.

Figure 10.1: Household sampling frames from complete and sample enumeration core modules

Step	Complete enumeration core module	Sample core module
1. Core census module frame	List all households by EA from the population census.	List all households by EA from the population census.
2. Core census module enumeration	Enumerate all households in all EAs to identify agricultural holdings; do core census module for all agricultural holdings.	Select a sample of EAs and then a sample of households in selected EAs; enumerate sample households to identify agricultural holdings; do core census module for sample holdings.
3. Frame for aquacultural supplementary module	List all holdings in the country with aquaculture, based on the core census module.	List holdings with aquaculture as enumerated in the sample core census module.
4. Sample for aquacultural module	Select a sample of holdings with aquaculture from the aquacultural frame.	Select a sample of holdings with aquaculture from the aquacultural frame (sub-sample of core module holdings).

