

## AREA FRAME SAMPLING IN AGRICULTURE

by Michael E. Craig\*

1981

I will present the concept of Area Frame Sampling in Agriculture from several viewpoints. First, an explanation of area frame sampling in general is presented. Then the advantages of this methodology are discussed for developing countries. Finally, I will show some specific examples of its use in various countries. These examples all come from experiences of Economics and Statistics Service personnel and include some of the countries represented at this seminar. The ESS travel and work on these projects is funded by the United States Agency for International Development (AID), both from Washington, D.C. and the individual AID missions in participating countries.

### Area Sampling Frame Methodology

Basically, any sampling frame is a complete listing of the entire population of units to be sampled. In area sampling frame methodology, the units to be sampled are areas of land. Our procedure is to divide the total land area of a country into small areas of defined size using natural boundaries. These small geographic areas must be located, listed, and measured. This listing, along with its size measurements, gives us our Area Sampling Frame (ASF). Aerial photography, topographic maps, other map type products, and LANDSAT satellite images are used as control data for breaking out these areas and for classifying them into groups with similar characteristics. A random sample of these small areas, referred to as the selected segments, are then visited by enumerators who interview the farmers and landowners found in the segment. Data obtained by these interviews is used to make statistical estimates about agricultural commodities. These interviews may also be used as ground data for remote sensing analysis.

The samples we select are designed to minimize the "variance", or error due to sampling, in the final estimates taking in to account the costs of the samples. Above, I mentioned the classification of the geographic areas into similar groups. This process, called stratification, is a technique used to reduce the sampling error. We list the units by their similar groups, called strata, as if each stratum was a separate population. This permits the focusing of the resources for collection of data on those geographic areas where the gain in terms of improvement to the overall estimate will be the greatest. Other presentations at this seminar will give more detailed descriptions of the methodology involved in area sampling.

In the United States, small agricultural areas have been used as sampling units as far back as the late 1940's. By the early 1960's, an area sampling frame had been developed for all states. For many years, ESS has been using sample survey data to estimate area of crops, yield, production, livestock numbers, prices, agricultural wage rates, farm numbers, and other related items in US agriculture. This information is mainly based on a probability sample of segments enumerated in late May and again in early December of each year. States' area frames, which provide the basis for this sample, are based on standardized land use definitions in each stratum. Strata definitions used refer to the percentage of land cultivated, the number of dwellings per unit area, land in open range or woodland, and land in non-agricultural uses.

\*PRESENTED AT CONFERENCE ON AREA FRAMES FOR AGRICULTURE,  
SAN JOSE, COSTA RICA, APRIL 1981.

Another procedure found useful in area sampling is replication. Replication is the process of drawing several smaller sized, independent random samples in each stratum rather than just one simple random sample in each stratum. This procedure allows us to change the sample size in a specific stratum by adding or deleting replications, without having to redraw or otherwise disturb the remaining replications. Other benefits of replication include the ease in rotation of segments, with one or two replications at a time being replaced by new replications. Segment rotation is important to avoid overburdening the same farmers survey after survey. Also, with replication, different questionnaire designs could be tested in different replications to compare their utility.

#### Background and Advantages

In the developing countries, the agricultural sector still dominates the entire economy. This sector usually produces more than one-half the total economic output, and utilizes more than two-thirds of the labor force. High potentials exist in most developing countries for agricultural development and in their pressing needs for more agricultural products. Thus we see higher priority for agriculture in national plans and in international economic aid efforts. Also many believe that small farmers can make major contributions to more productive agriculture, as well as to an acceptable pattern of economic growth and development.

Widespread concern over world food prospects, malnutrition, and periodic famines in some countries have generated support for focusing more attention on agriculture. More rapid changes in agriculture will increase the needs for additional, more reliable, and more timely basic information about agriculture. Data on production of crops, livestock products, herd inventories, area planted and harvested, yields, prices, and incomes are essential. Information is especially needed on potential area which could be planted, labor force size and composition, migration, farm family living conditions, etc. Area frame samples are a well developed, efficient technique for collecting agricultural and other related information.

The responsibility for collection and distribution of basic information about agriculture falls clearly on the government. Farmers and private enterprises have neither the resources nor the incentives to develop and operate such statistical systems on a wide basis. Data from these systems is needed for decisions in both private and public sectors. In countries that are very poor agriculturally, decisions must be made concerning the amount of food grains to import. Some countries that are large exporters of agricultural commodities must make decisions about the level of exports so as not to effect the available supplies for their own consumption. Also, in countries where large quantities of exports do not go through official channels, some form of objective measurement of these items must be used.

Agricultural statistics are collected both by censuses, which require enumeration of the total population of interest, and by samples requiring enumeration of only a small part of the population. Statistical methods have made clear that the enumeration of small samples can greatly reduce the cost of the collection of agricultural statistics while increasing their accuracy. A well designed sample, for which the data are carefully collected, can provide much cheaper statistics than a census and provide more timely information on current conditions. Sampling procedures provide the means for

controlling both sampling and non-sampling errors to meet the specific requirements for estimates, permitting the allocation of resources to those areas and those statistics where the needs for more and improved data are the greatest. The ability to tailor the accuracy of estimates to specific need is especially important to developing countries which have very limited resources to apply to the collection of agricultural data.

The Area Sampling Frame is a tested and well developed method which uses a small sample of land areas and employs stratification for increasing the sampling efficiency for agricultural information. For example, in the US, a sample of approximately 16000 segments is drawn for the June survey, covering only one-half of one percent of the total land area. The uses of area sampling are unlimited. Survey populations could be composed of reporting units that are households, persons, farms, plants, animals, cotton gins, sugar cane processors, suppliers of fertilizer, tracts of land, grain storage facilities, or rice processing plants. Many needs for information have been filled where area sampling was the only method available for selecting a probability sample.

In concept, an area sampling frame is always current and complete with regard to any definition of a reporting unit. This can not be said for list sampling frames which are another very common method of obtaining agricultural information. An area sampling frame does not become out-of-date in terms of coverage of a population, unless the population extends to areas not covered by the frame. Changes in land use, or number and location of reporting units, have a bearing on the sampling variance but do not introduce bias into the estimate. Parts of an area sampling frame can be updated as is necessary.

Two other uses of area sampling frames include multiple frame sampling and ground data for remote sensing analysis. Remote sensing analysis will be discussed at a later presentation. Multiple frame sampling combines the best features of both list and area frames for sampling purposes. Since a list frame is seldom, if ever, complete, the area frame can be used to supplement the list where it is not complete. For example, you may have a list of the 100 largest livestock producers but not of the smaller ones. A sample could be drawn from this list and enumerated to estimate the inventories of the largest 100 operators. To estimate the smaller operators, the area frame estimate is used with the data that belongs to any one of the 100 list frame producers edited to zero if found in a sampled segment. This combination of frames gives coverage of the population of livestock operators, and is both complete and more efficient than just sampling one frame or the other.

#### Current Area Sampling Frame Projects

Here I will mainly discuss a project funded by AID/Washington to transfer the ASF methodology to developing countries around the world. Most of the projects represented at this seminar have been funded by the individual countries and their AID missions although three; Costa Rica, Ecuador, and the Dominican Republic; have been at least partially funded by this AID/Washington project. Under this project AID has contracted for the expert technical knowledge and guidance of the Economics and Statistics Service of the US Department of Agriculture in assisting certain developing countries in establishing and using ASF procedures. We feel that the ESS offers a unique capacity for supplying this methodology because we have more experience with the ASF than any other group in the world.

Under Participating Agency Service Agreements (called PASA's) with AID, ESS has provided assistance in frame construction, methods for conducting field surveys, and for training activities to transfer the methodology to personnel in the countries involved. Let me note here that our objective is not to perform these surveys and gather information in the participating countries, but to teach the procedures to our various counterparts so that they can be as self sufficient as possible.

This project is organized into four phases for each participating country. The actual construction of the area frame has been designated as Phase I. Phase II includes the actual field survey for such data as crop area, yields, production, livestock inventories, farm improvements, and irrigation practices. Phase III concerns the computerized classification of agricultural areas using digital LANDSAT data. The last step, Phase IV, is defined as agricultural yield modeling using weather and other related data. ESS was expecting to use experiences from Phase I and Phase II in various countries to prepare feasibility study for Phases III and IV. However, at this time it is still uncertain whether the study or Phases III and IV will be carried out for any specific country or project.

ESS has thus far assisted 12 countries in the preparation and use of area sampling frames under this specific project. National level area frames have been constructed in Costa Rica, Jamaica, Tunisia, and the Dominican Republic. These were originally mission funded programs supported in the late stages by this project. Construction of national area sampling frames has been initiated, but not completed, in Bolivia, the Philippines, Indonesia, Morocco, Sierra Leone, Liberia, Sudan, and Ecuador. One other country, Paraguay, was considered for this project but it was decided that area frame sampling and construction were not feasible at this time. Intensive training in area frame construction has been given to personnel in the above mentioned countries. Note again that these projects are mainly AID/Washington funded and by no means include all countries using this methodology. These projects have varying amounts of progress to date.

Probably the most advanced projects to date are in Tunisia, the Dominican Republic, and Jamaica. The Tunisia Project began in 1974 and they have been publishing official estimates at the national level based on the ASF since 1976. Currently we are working there to update the existing frame for better accuracy using LANDSAT imagery for restratification. Tunisia also has implemented a North Star mini-computer system (discussed in a separate presentation) for the editing and summarization of survey data. The Dominican Republic, which is represented here, has been using the ASF methodology at a national level for quarterly surveys since 1972 when a Coffee and Cocoa survey was conducted. Current efforts there include rotation of segments and updating the stratification with aerial photography. The area sampling frame for Jamaica was completed in 1979, and a national level crop production survey conducted in October of that year. From this first survey, only sugar cane had an acceptable level of sampling error, with the coefficient of variation being between 5 and 15 percent. Design modifications are being planned to improve the estimates of a number of other crops. Jamaica has also implemented the North Star mini-computer system for survey data.

In Costa Rica, the ASF project has progressed very well. The pilot survey was completed in 1976 and currently the staff is preparing for the

first national level survey. The country of Thailand has a complete frame for 6 of its provinces and has been using it for estimation. They have implemented not only crop area and production surveys but a system of objective measurements for crop yield estimation. Several countries have completed pilot surveys. These countries include Ecuador, the Philippines, Indonesia, and Bolivia (although the enumeration is complete the editing and summarization is not). Frame construction has begun in several countries, including Morocco, Sudan, Liberia, Zaire, Tanzania, and North Yemen. I am not very familiar with the progress of the countries at this seminar not covered by the AID/Washington project and will be very interested in hearing your presentations.