## SURVEY SAMPLING AND IMPLEMENTATION FOR DEVELOPMENT PROGRAMS

Roe Goodman, Bureau of the Census

The main considerations in this paper are those relating to the <u>implementation</u> of sample survey in underdeveloped countries, hence this term is to be regarded as the key word in the title. Since certain factors are important in relation to all phases of sample surveys, the discussion on the whole is broader than that of survey sampling alone. However the illustrations are taken from the field of sampling.

It appears that the use of sample surveys for development is currently enjoying a wholesome growth. Various factors are tending to stimulate the use of this statistical tool in underdeveloped countries. Outstanding examples of current sample surveys for fact-gathering purposes are agricultural censuses based on samples which are now in progress in several Middle and Far Eastern countries, as a part of the FAO-sponsored World Census of Agriculture. Expansion of survey research designed specifically for policy making or analytical purposes can now be anticipated.

Considering the intrinsic difficulties in carrying through sample surveys in underdeveloped countries there is no doubt that the motivation of responsible survey officials is crucial for the success of this activity. More than this it is my observation based on experience in several countries that adequate motivation of survey officials cannot be taken for granted. To quite an extent the duties and responsibilities required from survey statisticians and field directors has little precedence in these countries. And only those who can see sufficient reasons for high standards of execution can really be expected to implement surveys well.

A major contention of this paper then is that in order to develop and maintain adequate motivation for survey work in underdeveloped countries great emphasis must be placed on the scientific character of sample surveys. I am stressing this point here in order to facilitate an orderly presentation by first showing the utility of this approach and second showing its necessity. The persons whose motivation could be affected by emphasis on scientific work are in general those who are college graduates, a substantial part of the staff of sample survey organizations being in this category.

To amplify a little on the notion of science and surveys, the idea is to emphasize that work on sampling and sample surveys is scientific work and that those participating in it are a part of the large body of persons engaged in scientific undertakings in the world today. This emphasis does not imply that the results from a survey are necessarily going to be used scientifically (or even used at all). Rather it is that the scientific approach permits those actively involved to feel that they are capable of uncovering new information of unquestioned validity. They can thus identify themselves with something that has prestige now and is likely to be even more promising for the future.

An emphasis upon science fits well with the idea expressed at the beginning of this paper — that proper implementation is a key consideration in survey work in underdeveloped countries. Obviously, if survey work is to be called scientific the sampling and other phases as well must be carried out strictly according to plan. Since in underdeveloped countries there has been little tradition to see to it that instructions are followed such an orientation becomes particularly appropriate.

When it comes to sampling, the textbooks usually state that the practicability of the design of a probability sample is a cardinal consideration. This matter, however, is often dismissed without specific amplification. In underdeveloped countries especially it must be borne in mind that practicability means that the plan must produce samples which can be identified correctly and easily by the field worker.

What has been said is merely that, in order to be considered scientific, the sample as implemented in the field must be a true probability sample. The very extensive work that has been done on sampling efficiency is also highly scientific. However, an efficient sample which makes great demands on the field worker for its proper implementation may be carried through so badly that the undertaking can no longer be called scientific. In my work in underdeveloped countries I have found that an insistence upon practicability often requires a sacrifice of sampling efficiency. The emphasis on science highlights the need for adhering to a thoroughly practicable, even if inefficient, plan at all costs.

A few examples may serve to clarify what has just been said. The first example, though beginning with an inefficient sample, is one in which the efficiency of the sample can be improved over time. The case is one of repeated surveys, twice a year, in order to estimate the acreages under various crops in East Pakistan. For this purpose a sample of fields is selected by a multi-stage sampling. Data on land use of sample fields are then obtained and recorded by Agriculture Department employees on the basis of personal observation. For the beginning years the idea was to select fields in clusters of about ten contiguous

fields within the villages selected at the preceding stage. Good maps, though not up to date (often forty years old), were available for the use of the field workers and it had been found repeatedly that randomly selected fields could be located on the spot with the aid of the maps alone provided sufficient care were used. In order to more or less force the field worker to stay in the proper vicinity for enough time to verify the field identification the use of the clusters of fields was decided upon. Of course, based on analyses carried out in many parts of the world there could be no doubt that the selection of fields in smaller clusters or scattered individually throughout the village would have been much more efficient statistically. But sober reflection led to the use of the relatively large clusters anyway in order to provide greater guarantee that the field work would be done strictly in accord with the plan. Naturally it is intended in cases of this kind that steps would be taken to increase the efficiency of the sample by gradually reducing the size of clusters. This would be done year by year as evidence accumulates that the field workers are performing well and that fields would be located and reported upon properly in smaller clusters. Although these surveys generally use the same sample from one survey to the next, there is no reason why a transition cannot be safely accomplished from clusters to individual fields within sample villages over time.

The second example is of an inefficiency of a different type in that it involved extra work at one step in the sampling process but finally led to the selection of a sample of normal efficiency. In this case (a family expenditure survey in Santiago, Chile) blocks had been selected with probabilities proportionate to the number of segments, the segments having been allocated to the blocks on the basis of estimated numbers of households in each block. The particular step involved the dividing of the selected blocks into the pre-assigned numbers of segments. For this purpose the field workers were asked to proceed around the blocks, going into and out of 'cites' and 'pasages' as they came to them and listing the addresses of all dwelling places one by one. Subsequently the listings were utilized in the office to establish readily identifiable points of division for the segments and one of the segments thus defined was selected at random within each block. An alternative would have been to have sent persons out to determine and note down suitable points of division within blocks without the preparation of any listing. The Director-General of the National Statistical Service, however, believed that the second plan would have led to the using of many unsatisfactory division points between segments. In this connection it is worthy of note that the original listing did

not eliminate the necessity for a subsequent visitation to the families within selected segments, which was an additional preliminary to the selection of individual families for the family expenditure survey. The detailed listing of 100 or more addresses per sample block had the sole purpose of making sure that the segments as defined could always be accurately identified in the field and was an extra task carried out for this purpose.

As a final example, there is the case of the sample of villages which was selected in Pakistan for enumeration in the country's 1960 (first) Census of Agriculture. As selected and used this sample was one of the largest samples ever taken anywhere in the world but it could have been much smaller with a more efficient design. In planning the sample it was concluded that the only possible means of doing sub-sampling within selected villages would be to place complete reliance upon officials within the different districts (1) to see that accurate lists of households were obtained, and (2) to make random selections from these lists. At the same time it was well known that to the district officials the. idea of sampling was relatively new and the necessity for really rigorous work in such matters was entirely unrecognized. The inescapable conclusion, therefore, was that if sub-sampling were to be attempted it could not be expected that the households finally to be enumerated would be a true probability sample of the rural households of Pakistan. Accordingly, the design that was decided upon was a stratified random sample of whole villages. The entire work of sample selection was done at a central point for each of the two Provinces, a total of about 12,000 villages being selected in all. Instructions were then issued that within each selected village all cultivators (farmers) were to be enumerated, which meant that although the Census was based on a sample the work of an enumerator within a village was in all major respects exactly the same as it would have been if a complete Census were being taken. The sample was, of course, a grossly inefficient one (however, a much more 'efficient' operation than a complete Census would have been) but the rigorous adherance to a scientifically selected sample seemed a goal worthy of the price which was paid.

To put these examples of inefficiencies in their proper light it should be mentioned that obtaining the necessary funds to do good survey work, however inefficient the sampling design, does not constitute a major problem. In fact, the examples given are instances in which nationals in the countries either suggested or at least agreed wholeheartedly with the procedures used.

If the problem is not one of money where then does the problem lie? Fundamentally the problem is that the type of motivation among survey workers which is generally effective in the U.S. is not present in sufficient measure in underdeveloped countries. It is probably pointless to go into extensive discussion of these matters especially since a number of the factors involved are rather negative in character. However, it may be useful to mention the points which I believe have the greater relevance. The enumeration of these problems does not imply that persons within the countries are not aware of them or that no progress is being made in resolving them.

Some of the critical points are:

(a) Those at the highest levels who have authorized a survey may not really expect to see useful results emerge (but have approved the undertaking for a variety of other reasons) and consequently they do not go to any particular pains to see that personnel, equipment and so forth are available at exactly the right time. Moreover, they often underestimate the difficulties and more or less assume that progress will be made in due time once approval has been given.

(b) Officials at a high level may be keenly interested but due to lack of professional training they are not especially impressed with the need for rigorous methods and consequently they do not support the decisions of their well-trained subordinates and do not make adequate provision for necessary personnel or other resources.

(c) An enthusiasm for the work which might develop on the part of technical staff is stifled by knowledge that no matter how much they may be interested their superiors will not give them any particular encouragement or credit for a job well done.

(d) Professional workers, and nonprofessional too, have good reason to believe that no matter how hard they apply themselves to the tasks at hand they will receive no advantage in the way of promotion over those who take their responsibilities lightly. It does not matter that a new day may be dawning and they may in fact receive faster promotion; what is controlling is that they do not think they will.

(e) Officials at higher levels may have ordered the inclusion of too many questions or too difficult questions and the professional worker may not dare to advise his chief of the impossible character of the assignment given to him.

(f) Rigidity in personnel policies often means that relatively unsuited persons are placed in responsible positions in survey organization (perhaps due to seniority) and their better suited colleagues are therefore aware of the inability of the organization as a whole to function properly under the existing arrangements.

(g) Administrative ineptness either in the way of delayed approval of expenditures or inopportune transfers of personnel may place serious pitfalls in the way of the normal progress of a sample survey and require that the well-trained official do his work under particular difficulties.

Isolated instances can be found of even more extreme obstacles which rear themselves in the way of successful surveys. From one point of view all these obstacles may be seen as inefficiencies which will be overcome in due time just as sampling efficiencies can and will be overcome also. The effects of the administrative deficiencies taken all together may, however, be sufficiently extensive to create a feeling of futility and discouragement among professional workers.

What is therefore needed is a survey design which will enable those who want to make scientific progress to achieve limited objectives which will meet the test of scientific scrutiny. Even if a large part of an early survey is inevitably a failure, for some of the reasons mentioned, key professional workers can with sufficient imagination salvage information on a few items or gain methodological insights which point the way to future progress. It is necessary primarily that the desire to follow scientific procedures be firmly established in the minds of the professional staff at all stages of the work. Required is the discovery of procedures which are sufficiently simple to permit scientific standards to be upheld.

The manner in which the relatively inexperienced but well-trained workers in many countries around the world are striving to do good scientific work in what to them are pioneering days is worthy of whole-hearted support. There is an ever-growing number of people in these countries who have received good training at home or abroad, which is a portent of constructive development for the future in the area of sample surveys.