

47193



UNITED NATIONS
ECONOMIC
AND
SOCIAL COUNCIL



Distr.
LIMITED

E/CN.14/ASPP/L.9
E/CN.9/CONF.3/L.9
8 October 1962

Original : ENGLISH

ECONOMIC COMMISSION FOR AFRICA
SEMINAR ON POPULATION PROBLEMS IN AFRICA
29 October-10 November 1962
Cairo, United Arab Republic

MAJOR APPLICATIONS OF SAMPLING IN POPULATION CENSUS PROGRAMMES

62-22282

E/CN.14/ASPP/L.9
E/CN.9/CONF.3/L.9

Page 2

This Seminar has been organized by the secretariat of the Economic Commission for Africa in co-operation with the United Nations Bureau of Social Affairs, Statistical Office and Bureau of Technical Assistance Operations, and the Government of the United Arab Republic as host.

MAJOR APPLICATIONS OF SAMPLING IN POPULATION CENSUS PROGRAMMES

Prepared by the Statistical Office of the United Nations

CONTENTS

	<u>Paragraphs</u>
I. INTRODUCTION	1-2
II. PRE-CENSAL AND CENSUS APPLICATIONS OF SAMPLING	3-29
A. Sampling at pre-enumeration stage: (experimental or pilot census)	3-6
B. Sampling in enumeration	7-13
C. Sampling in data processing	14-29
1. Quality control	14-19
2. Advance (provisional) tabulations	20-24
3. Sampling for additional tabulations	25-29
III. POST-CENSAL APPLICATIONS OF SAMPLING	30-51
A. Designing sample surveys	33-40
1. Frames	34-37
2. Stratification	38-39
3. Weighting	40
B. Types of post-censal inquiries	41-51
1. Verification of completeness and accuracy of the regular census	41
2. Verification of completeness and accuracy of civil registers of birth and death	42-43
3. Obtaining supplementary information	44-47
4. Up-dating information obtained from censuses	48-50
5. Exploiting census results	51
IV. THE USE OF SAMPLE SURVEYS WHEN A CENSUS IS NOT FEASIBLE	52-63

I. INTRODUCTION

1. The purpose of this paper is to show how sampling methods can be utilized as an integral part of the census programme in planning, executing and analyzing the results of a population census, how it may be employed to evaluate and supplement the data obtained from a census, and how it may provide interim estimates of basic data of the kind traditionally obtained by means of a census.

2. Sampling, as a statistical technique, is a tool which can be applied at every stage of census taking and utilization. Its proper employment can not only help improve the census results but it can effectively contribute to the rational use of resources including inter alia funds, equipment, personnel and time. The possibility of applying sampling techniques should, therefore, be considered at every stage of the census, that is, at the pre-enumeration stage; during the enumeration; in the processing stage; in connexion with tabulations; and in planning post-censal surveys. The salient features of each of these applications will be discussed below, with special attention to the planning of censuses and surveys as interrelated parts of a concerted programme of data collection and research.

II. PRE-CENSAL AND CENSUS APPLICATIONS OF SAMPLING

A. Sampling at pre-enumeration stage: (experimental or pilot census)

3. In planning a census of population, one is confronted, among other problems, with determination of the questions to be asked; the manner of phrasing each question; the most effective format for presenting the questions;

the selection of the most adequate respondent; the method of training enumerators and a host of other procedural questions. Good census practice suggests that these questions of procedure and operations be settled on the basis of pre-testing the census schedule in a sample of enumeration areas.

4. The nature and extent of the pre-testing, which may need to be undertaken in advance of the enumeration proper, depends on the amount of relevant experience already at hand from previous census projects. Countries with no previous census experience to rely upon, and those which expect to expand the scope of their census rather substantially or radically change their methods, may have to provide for quite an extensive pre-testing programme. Sampling, since it offers a scientific basis for such an approach, can make a valuable contribution toward providing reliable answers to a number of important questions. The use of sampling in the pre-testing stage can therefore help to ensure that the final census plan is the one best suited to the problem at hand and to conditions prevalent in a given country.

5. Among the African countries, the one which has probably made the most use of a pre-test or experimental census is the Sudan in 1953. The purpose of this pilot inquiry was to pave the way for the 1955-56 census. Inasmuch as sampling was expected to constitute an integral part of any subsequent census, an attempt was made to get some notion of the variability of the sampling material which would enable the sample size to be determined. The census test proved very valuable by appraising the census officials of people's superstition and objections to being counted, and improved the chances of eventual success by disseminating news about census plans.

6. Census tests have also formed part of census plans in Liberia; for a more detailed discussion of this and the Sudan experience see the Handbook of Population Census Methods, ^{1/} Chapter VII.

B. Sampling in enumeration

7. Modern experience in the use of sampling techniques has confirmed that it is not necessary to gather all desired demographic information on a complete, 100 per cent coverage, basis. The rapidly growing needs for extensive and reliable demographic data have made the modern census of population an undertaking of great magnitude. Census-type operations are expensive and, if all needs were to be met through complete enumeration, the cost in time and organization required would necessarily place a severe strain on the resources of most countries. To lower the cost and to reduce the operations of census taking, and at the same time to obtain the maximum amount of desired data of the best quality, sampling methods can successfully be introduced at the enumeration stage.

8. Sampling has been applied to the enumeration phase of census taking in two distinct ways. One way is to select a sample of the population to whom all questions will be addressed - a method which is, in reality, a sample census or survey discussed in part II of this paper. The second application consists of utilizing the sampling method to broaden the scope

^{1/} Handbook of Population Census Methods - Volume I: General Aspects of a Population Census. United Nations, Statistical Office, ST/STAT/SER.F/5/Rev.I, Sales No: 58.XVII.6 (Vol. I).

of the enumeration by directing selected census questions to a sample of the population only, rather than to all persons in the enumeration area, and in doing this concurrently with the enumeration proper.

9. In the latter case, two methods of enumeration may be employed for the sample questions: the first allows the regular enumerator to ask the sample questions as well as the basic inquiries. This technique is used in cases where the supplementary information can be obtained fairly easily by conventional enumeration techniques and where reducing the cost of the census is the principal objective.

10. The second method of enumeration is based on the assumption that ordinary enumeration techniques will not elicit accurate information of the type required, and therefore use is made of a especially trained crop of enumerators. These special enumerators may obtain the information during the regular enumeration or a short time afterwards but their activities are considered an integral part of the enumeration and not independent of it in any way. It is this characteristic of non-independence which distinguishes such sample enumeration activities from post-censal sample survey inquiries described in paras. 44-47.

11. Topics which can be and are investigated in a supplementary sample enumeration include questions on fertility, income, internal migration, economic characteristics and the like for which intensive questioning is required. In general, however, the decision as to which particular questions may be reserved for the sample enumeration depends not only on the desired accuracy of the results, but also on the needs for statistics of small

geographic areas and for cross-tabulations of these topics. Information collected by sample enumeration is not always adequate for tabulation by small civil divisions or population groups, or for providing significant cross classifications. It may therefore fall short of satisfying present or future needs. This is particularly important if the census results are to serve as a frame for future sample surveys, in which case basic data for small aggregates is essential for stratification purposes (see paras. 38-39).

12. Account must be taken of all these factors before questions are selected for sample, rather than complete, enumeration, but it can generally be concluded that items of information which have no statutory force, items which need not be tabulated for small geographic areas and items for which detailed cross tabulations are not required can be secured by means of a sample operation conducted concurrently with the population census enumeration. Such items should, however, be as few as possible, and should be confined to those that are time consuming or require better trained investigators.

13. Examples of the use of sample enumeration as a supplement to complete enumeration are wide spread.^{2/} In Africa, supplementary information was obtained for Comoro Islands from 25 per cent of the persons 14 years of age and over enumerated in the census of 1958. In Morocco, the 1960 census plan called for only 1/10th of the rural population to fill in the questionnaire

^{2/} Ibid. pp. 127-131.

utilized on a 100 per cent basis in the urban areas. Similarly, the 1956 census of Sudan covered 10 per cent of the rural population and 100 per cent of the urban. Gabon, in its 1960/61 census, collected data on natural movement of the population from 10 per cent of the households, concurrently with the regular census.

C. Sampling in data processing

1. Quality control

14. Sampling can appreciably assist in controlling the quality of data processing, beginning with the editing of the census schedules and ending with verification of computations.

15. During recent years, there has come into being a new attitude with regard to statistical data and census results in particular. This attitude is based on the recognition that statistics are useful even if they are not 100 per cent accurate, provided the margin of error is objectively assessed and found to be within reasonable limits. The use of sampling methods to assist in the measurement and control of the quality of census-processing operations is one aspect of the new philosophy.

16. "Data processing", as used here, comprises editing of schedules for consistency and credibility of response; coding of replies; verification of coding; transcription of codes to spread sheets, punch cards or tape; verification of transcription; scrutiny of tabulations for credibility; and computation of rates and ratios. Even in countries of moderate population size, these operations may involve millions of schedules and cards, and a far greater number of columns of figures. Hence, the magnitude of the

processing operations of a modern census makes achievement of complete accuracy in the results of each stage, a practical impossibility. Moreover, the scale of the process makes 100 per cent verification equally impossible. Such considerations suggest the adoption of sampling methods of the type used in industry to control the quality of operations.

17. The aim of statistical quality control of census data-processing is limited primarily to ensuring that the over-all frequency of error in the various processing operations remains small enough to make corrections or adjustments unnecessary and, subsidiarily, to avoid the inclusion in the results of those work lots which are of extremely poor quality. This type of continuous sample control actually results in a much lower census cost than if the verification had been performed on a 100 per cent basis.

18. It should be noted, however, that the economies which may be achieved by substituting statistical quality control for the traditional method of repeating each operation on a 100 per cent basis, depend on the relative costs of manpower and equipment. Such savings will be relatively less in economically less-developed countries where wages tend to be lower and mechanical equipment more expensive as well as more difficult to obtain.

19. Examples of the application of statistical quality control appear to be limited to the United States and Canada and to European countries, such as France, Norway and Yugoslavia. No information is available for Africa.

2. Advance (provisional) tabulations

20. Another aspect of data processing where sampling can be, and often has been, used in connexion with the census of population is the important one of providing provisional results as early as possible after enumeration is completed, that is, tabulations prepared in advance of the final detailed cross-classifications for meeting urgent needs.

21. A complete national census is a huge undertaking. If all the returns are completely processed before tabulation begins, several years may elapse before results are made available and the utility of the census for many purposes may thus be seriously impaired. The probability that such a lag will occur is especially great in the economically less-developed countries where resources are limited and few or no funds can be devoted to modern tabulation equipment. If, on the other hand, a sample of returns is processed and tabulated in advance of the main processing, the provisional results obtained may be utilized as estimates of the final returns. A carefully executed sample tabulation plan, based on a random sample of returns, can make available statistics on age, employment status, industry, etc. required for many administrative decisions, in time for them to be of maximum utility. If sampling has been used to broaden the scope of the enumeration (see paras. 7-13) the same sample may be used for advance tabulations. Advance tabulations may also provide a convenient tool for verifying the credibility of the complete tabulations.

22. Naturally the selection of a sample of census returns for advance processing of data has certain disadvantages as well as benefits. One of the potential disadvantages is the possibility that the imposition of the sampling procedures may cause a delay in the regular census processing. The danger also exists that concentration on the preparation of sample tabulations may even result in elimination of the corresponding complete tabulations. Since the sample tabulations will usually provide only national totals, the geographic detail for the tabulations in question will never become available.

23. Many types of sampling plans are employed by countries for obtaining results in advance of the complete tabulations, but the simplest is one which uses the enumeration district as the sample unit, and processes all returns for each unit selected. A second plan used frequently is a two-stage design in which the enumeration district is the first stage sample unit and the household is the second. In this plan, only the households selected in the sample are processed. A third method relies on a systematic sample of census schedules.

24. Information on the application of this method to African censuses is not comprehensive, but it is known that advance tabulation of a 10 per cent sample of returns to secure early estimates of the principal results was used in the 1956 census of Rhodesia and Nyasaland. The results of the 1960 censuses of Morocco and Ghana also were tabulated in advance using a sample of returns.

3. Sampling for additional tabulations

25. The complete tabulation of all the information collected in a population census requires considerable time and money. Consequently, to enlarge the scope of the tabulation programme within the limits of available resources, the tabulation programmes of most population censuses provide for the complete tabulation of only the legally prescribed items and those elements of demographic and economic information which are regarded as of importance for small administrative areas, while the remainder are made on the basis of a sample of the returns.

26. Demographic data needed only for large areas and for the country as a whole which can be provided by sample tabulations tend to fall into two classes: (a) detailed tabulations of certain characteristics such as age by single years, or occupation or industry classifications by detailed codes, and (b) tabulations of multiple variables such as marital status classified by age and religion. Sampling enables one to obtain these tabulations for large areas with reasonably small sampling errors at a much reduced cost and in a shorter time than is needed for tabulations on a complete basis. Even for detailed cross-tabulations prepared on a sample basis, the cost is relatively not high.

27. The sampling considerations involved in the present case are somewhat different from those in the preparation of advance tabulations referred to in para. 21. Whereas advance sample tabulations have a provisional

status, being prepared for meeting urgent needs and scheduled to be replaced eventually by a complete tabulation, here one is concerned with sample tabulations and cross-tabulations which are not intended to be prepared on a complete basis at any later date. This important difference highlights the need to exercise the utmost care in the designing and executing of the sample with a view to attaining the highest accuracy within permissible costs. It may be mentioned however, that in some countries the sample selected for advance tabulations is also utilized for the additional tabulations designed to enlarge the scope of the programme.

28. No information is available as to the application of this procedure in Africa, but numerous examples of its application in other areas are set forth in the Handbook of Population Census Methods.^{3/}

29. In addition, it is advisable to retain a small sample of the census punched cards or tapes after the tabulation programme has been completed, for future use in preparing additional tabulations for purposes of analysis and comparisons with future census data, though the need for such tabulations may not be foreseen at the time of census processing. It is advisable to keep the size of such a sample small so that the tabulations can be carried out quickly with minimal expense.

^{3/} Ibid. pp. 147-150.

III. POST-CENSAL APPLICATIONS OF SAMPLING

30. It is becoming more and more evident that periodic sample surveys have an important place among the traditional methods of data collection. In view of the fact that for their proper execution, sample surveys depend largely on the information provided by censuses, the two techniques must be viewed as interrelated parts of a concerted programme of data collection and research. It is imperative therefore that links be established between the sample survey and the census at the planning stage.

31. It goes without saying, of course, that one must avoid the danger of over-optimism about the potentialities of sampling. As in the case of any other scientific tool, sampling can be exploited to the maximum extent only by appreciating the limitations of the technique. But within the limitations of the method, sample surveys can provide more detailed information and investigate more complex relationships than would be feasible in a census. This is true because although it is, like a census an interview-type procedure, the sample survey can utilize fewer and better trained enumerators, it can allow more time for obtaining adequate replies and use probing questions, it can select the respondent more carefully, and generally carry out the interviewing procedure at a higher level of efficiency.

32. Post-censal sampling has several demographic objectives all of which are based on the fact that the census can be of assistance in designing the sample and also in indicating areas of research where more intensive investigations in depth are required. The salient features of each of these applications will be set forth below.

A. Designing sample surveys

33. The use of the census of population for designing subsequent sample surveys rests on two facts: first, sample units for a survey operation in single or multi-stage sampling have to be selected from a complete list of units known as a sampling frame and normally such a frame is obtained if possible only from the census of population. Second, for the efficient estimation of parameters for the population as a whole, the principles of stratification should be introduced and the use of this principle depends, to a large extent, on basic information provided by the population census.

1. Frames

34. Sampling frames may be composed of one of a variety of materials, including geographic sub-divisions of a country, small administrative areas, housing units, households, groups of households such as compounds or villages, individuals or any other identifiable unit whose dimensions or boundaries can be precisely defined, which together cover the universe to be sampled, which have permanency and stability of location, and which are operationally convenient. The criterion of operational convenience is especially important inasmuch as the survey interviewer should be able to conduct the survey operations within a short fixed period of time, and widely scattered units of observation would defeat this objective.

35. The preparation of a good sampling frame or frames can be an arduous task especially in under-developed countries. It is particularly difficult when sample surveys have to be carried out subsequent to the census, rather than concurrently with it, or immediately following the completion of enumeration. In the latter cases, the census itself provides directly a satisfactory sampling frame. When sample surveys are planned to take place subsequent to the census, the frame provided by the census becomes increasingly obsolete as time elapses after the census.

36. Therefore, before a sample selection is undertaken it is necessary to ensure that the sampling frame is free from such defects as obsolescence as well as inaccuracy, incompleteness, duplication, and inadequacy. Inaccuracy in a frame may arise not only from wrong information about the units listed in it or defined by it but also from listing of units which are in actual fact non-existent. If some units of the material are omitted entirely, the frame is called incomplete; if some units of the material are included more than once, the frame is said to be subject to duplication. A frame used for a particular survey may sometimes exclude certain parts of the material which the survey is expected to cover; such a frame is called inadequate for the survey under study. As noted in para. 35 a frame though accurate, complete and free from duplication at the time it was constructed, may no longer be so at the time it is required for use. Such a frame is called obsolete and defects of the first three of the above types may be introduced through the use of an obsolete frame.

37. Since each of the five types of defects in a frame introduces defects of a different type into the sampling procedure, it is very desirable that population census data which are to be used as the frame or frames for a proposed sample survey be carefully examined with this in mind. Of even greater importance is that census planners should try to anticipate the fact that census records (including maps, lists of enumeration districts and schedules) may be used as a frame for future sample surveys and, in view of this possibility, make every effort to keep these records in a form which can readily be used for sample survey work; a clear indication of their deficiencies when used as a sampling frame should also be available. It is not possible of course to anticipate all the types of sampling inquiries that are likely to be undertaken and the kinds of frames that may be required, but efforts should be made to provide data in geographic detail sufficient to allow the size of each unit as well as its location be used to build suitable sampling frames.

2. Stratification

38. Populations can be classed as homogeneous or heterogeneous depending on the degree to which the characteristics of each population vary from one sample unit to another. If there is little variation, a good estimate of the average for the characteristic in question can be obtained from a small sample of the homogeneous population. But if the population is heterogeneous, it will need to be sub-divided into more homogeneous groups for the efficient estimation of parameters for the population as a whole. The sub-groups are known as strata and the division into strata is stratification.

39. For purposes of stratification and multi-stage designs there is need first for the census to provide tabulations of elementary characteristics of the population by very small areas, such as city blocks and the smallest division of the rural areas. The geographic tabulation will not prove troublesome, but the selection of the characteristics to be cross-tabulated is often dependent on national characteristics and on the purpose of the survey. For example, if religious communities are well demarcated in the population, this characteristic will need to be the basis of tabulation for small areas. Other attributes to be considered might be literacy, urban/rural residence, language spoken, locality size, nativity, ethnic groups, and so forth. If the census of agriculture helps provide data required for the sampling frame, certain agricultural elements might need to be employed in stratification, for example, tenure, size of farm, principal crop and so forth. The housing census might also be helpful in deciding on the frame; it could provide tabulations of houses by type and size, for example.

3. Weighting

40. In addition to providing the sampling frame, the census results also produce data which are used as weights for application to the data obtained from the subsequent survey in making national estimates. Whenever stratification is employed in selecting the sample, it is necessary to take into account the relative value of the various elements on which the strata are based in order that the national estimate may reflect these values. For example, the results for individuals from a family living survey among various types of families would need to be weighted by the number of families of corresponding type found in the census.

B. Post-censal inquiries

1. Verification of completeness and accuracy of the regular census

41. Post-censal field checks of the census in which the coverage and content of the original census are verified by a repetition of the census on a sample scale, have been discussed in detail in another document. ^{4/} As indicated in that paper, a post-censal field check must meet three criteria, namely, (1) be independent of the original census; (2) be representative of the whole country and of all population groups; and (3) involve one-to-one matching of records to produce an identical sample from each investigation. The salient features of a survey designed to meet these criteria, as well as examples of such surveys which have been carried out, are set forth in the document referred to above.

2. Verification of completeness and accuracy of civil registers of birth and death

42. Checking of the coverage and content of civil registers of birth and death by means of sample household surveys is another type of post-censal inquiry which may utilize the census for design of the sample.

43. As noted in the document mentioned above, a field-check on vital statistics derived from civil registers would have to be based on a probability sample selected from a frame consisting of local civil registration districts.

^{4/} E/CN.14/ASPP/L.10.

Since registration districts are, insofar as possible, co-extensive with minor civil divisions, results of the census of population could be used to obtain the population size of at least most of the units in the frame. Some indications of the methods of carrying out such an evaluation are set forth in part B of the other document.

3. Obtaining supplementary information

44. Topics on which data can usefully be collected by the regular census are limited by certain levels of cost and time, by the available supply of trained personnel, and by the feasibility of the inquiry, i.e., the anticipated public reaction to certain questions. Information on the variety of other other subjects relative to measurement of levels of living and the changes therein required for planning economic and social development cannot normally be collected most efficiently and economically through the traditional population census. Thus, supplementary information on such topics as levels of wages and earnings; under-employment; internal migration; income and expenditure; educational characteristics; fertility; mortality; health; housing, food consumption and nutrition, which are constantly required, have to be provided by other methods, among which is the sample survey of households. ^{5/}

45. Surveys designed as single-subject or multi-subject inquiries are usually household inquiries carried out as a data collection operation

^{5/} The wide applicability of this method of data collection has initiated the preparation by the United Nations of a Handbook of Household Surveys: A Practical Guide for Inquiries on Levels of Living. It is anticipated that the Handbook may be available in 1963.

independent of the census operation. They are linked to the census through the sampling frame obtained from the census and they are also, if possible, linked by making use of the same basic definitions and concepts as pertained to the census. For example, the household, which is the unit of enumeration, should be defined in the survey exactly as it was in the census of population which preceded it. The same concepts of residence should also be adopted, as well as the same definitions of urban and rural and characteristics of the population surveyed should be defined in accord with the census definitions. Classifications employed should be coordinated with those used in the census. Strict coordination is especially important because of the fact that data obtained from sample surveys is designed to supplement that obtained from the census and they must, therefore, be capable of being related to the analytical results of the census proper. An equally important reason for adopting identical definitions is efficiency, since the survey interviewers will most likely be selected from the census enumerators and their performance will benefit from being able to make use of their previous training and experience.

46. In organizing an independent sample survey, it is possible to use a frame consisting of households enumerated at the census or an entirely independent area sample, in each of which every household will be interviewed. If the sample survey is part of a post-censal field check ^{6/} the list of households mentioned above will not be adequate as a frame because of the

^{6/} See E/CN.14/ASPP/L.10.

need for complete independence. In any case, for reasons of efficiency, it is better to utilize area sampling in order to minimize travel between interviews. Further, if enumeration districts are the sampling units, the selection can be done in a central office well in advance of the enumeration, thus saving interviewer time in the field. Data processing would also be simpler since the enumeration districts are often the smallest units used in preparing tabulations.

47. Examples of sample surveys utilizing the census districts as a frame in Africa, include the following: Nigeria has adopted this technique in connexion with its post-censal field check, to obtain additional information on marital status, fertility, economic characteristics, migration and disability. Ghana also has used the 1960 Census as a frame for subsequent surveys. In the 1959 census of Uganda, African population and non-African plus Africans residing with them were enumerated separately. The African census took the form of a group count of the members of households, asking only age group and sex. This was followed by a "sample census" asking more detailed information. The same technique was used in Tanganyika. In Togo, the 1958-60 census was the basis for a national Demographic Sample Survey conducted in 1961. In Dahomey, a sample survey was carried out in May-Sept. 1961, based on the census of September 1960. Tunisia carried out a survey in 1959/60 making use of the census of 1956. Morocco is organizing a multi-subject sample inquiry to supplement the 1960 census. Additional examples of the wide application of the use of the census data as a frame for subsequent sample surveys may be found in the Handbook of Population Census Methods.^{1/}

^{1/} Op. cit.

4. Up-dating information obtained from censuses

48. Very closely related to the use of sample survey for obtaining supplementary information is their applicability for bringing up-to-date the information obtained in a census. Censuses are usually conducted at intervals of ten years, and in the intervening period between two censuses some of the data might become grossly out of date. This is likely to be the case with respect to data on the labour force, information on which is obtained in the population censuses of many countries. Similarly, in areas subjected to intense public health programmes or to migration, the rates of population growth as estimated from previous census data may no longer be applicable and current estimates of population size and characteristics may be required. Thus, the most common subjects for investigation in surveys designed to up-date the census are total population size, population growth rates and characteristics such as age, sex and economic characteristics.

49. Insofar as the sample survey brings up-to-date data collected at the preceding census with a view to appraising changes which have occurred since, it constitutes a micro-census. For practical reasons, only a fraction of the data collected at the census can be brought up to date by a survey, but as was noted in para. 45 it is essential that these items be defined in a manner identical with the census and that the population to which they refer be defined in the same way. It is also recommended that in the survey the questions be put precisely in the same terms as were used in the census so that comparisons will be feasible.

50. Monthly surveys of population and labour force are being carried out in many countries of the world, but, for understandable reasons, few of these are found in Africa, especially in the demographic field. For a review of these and other sample inquiries, see Sample Surveys of Current Interest.^{8/}

5. Exploiting census results

51. Still another use of sample surveys in connexion with censuses may be sited. Through analysis of the census data, of questions and hypotheses which require more intensive investigation by means of sample survey may be formulated. In this, the census acts as a clue to subject areas requiring investigation in depth. For example, different rates of population growth in different areas of the country may indicate need for sample inquiries relating to mortality, fertility, or migration to explain these differences. In the same way, the census may indicate particular geographic areas or special groups of persons which should be studied intensively by the sample survey technique. For these inquiries, the census can provide a list of households having certain characteristics from which a sample may be drawn.

^{8/} United Nations publication Statistical Papers Series C Nos. 1-10.

IV. THE USE OF SAMPLE SURVEYS WHEN A CENSUS IS NOT FEASIBLE

52. The applications of sampling method described in the preceding sections are those used as an integral part of complete enumeration or as a device to broaden the scope of the census by obtaining supplemental information by post-censal, independent sample surveys. This section on the other hand, will discuss the use of sample surveys where a full, nation-wide census has never been taken or where one is not feasible. This includes most of Africa.

53. As has been pointed out in the Handbook of Population Census Methods,^{9/} every country must sooner or later take a complete census even if it is of imperfect quality or of limited scope in terms of information collected. A complete census is indispensable not only because it forms the basis of numerous subsequent statistical inquiries, but also because it provides detailed geographic data. It is true that in the short run basic demographic information can be effectively obtained through sampling, since sampling techniques are flexible and so are usually adaptable even to conditions which, from the viewpoint of census taking, are very difficult. Nevertheless, a sample survey should not entirely replace a complete enumeration of some kind except under conditions which make a complete enumeration impossible. Even when a country resorts to sampling in the place of a census, the sampling scheme used may best be considered an experimental census, paving the way to a complete census to be executed in the near future.

^{9/} Op. cit.

54. The nature and extent of the difficulties which prevent countries from conducting a complete enumeration vary greatly from country to country, depending on the degree of social and economic development. The sample survey has served in these cases as a useful method of obtaining current information, but it should be noted that the very factors which have prevented the carrying out of a census, such as lack of experience and trained staff, also affect the conduct of sample inquiries. Among the fundamental obstacles to all schemes for measuring quantitative aspects of social and economic life in these countries are (1) the ignorance, illiteracy and superstition of a large part of the population; (2) inadequate transport and communication facilities; (3) nomadic population; (4) lack of qualified canvassers; (5) lack of interest in a census or survey.

55. Any sample survey instituted as a substitute for a complete enumeration under the circumstances mentioned above has necessarily to possess the following three features. First, the field work must be designed to win and maintain the confidence and assistance of the people. Secondly, the range of information covered by each survey should be limited to a very few subjects. It is desirable to cover a field of different topics by repeated sample inquiries rather than by one inquiry covering the entire field. The case for a sample survey of limited scope is particularly strong when de facto demographic statistics are required, since a survey must obtain the needed information in the shortest possible time. Thirdly, whenever possible two or more interpenetrating samples should be selected in every administrative

district or any other smallest domain of study. The use of interpenetrating samples has two important advantages. It enables the work of field officers to be organized in such a way as to facilitate the examination of each officer's performance and provides a means of discarding unreliable estimates. The second advantage is that an alternative estimate is available for almost any domain of study in the event of unforeseen contingencies (such as accidents to field officers) that may arise while the survey is in progress. The practice of using interpenetrating samples yielded highly satisfactory results in Southern Rhodesia.

56. Where no previous population census exists, the search for a suitable sampling frame is beset with many difficulties, not the least among these being the complete absence of maps for large areas. Most countries maintain lists of taxpayers but the natural tendency to evade taxation results in incomplete registration. Registers are of varying quality in different districts of the same country, some including misleading book entries, (e.g. persons working in towns entered in certain rural areas for administrative convenience), fictitious entries, etc. In using registers for frames it is preferable to employ the register of villages as the frame and consider the village as the sample unit since the opportunities for evasion of the registration of whole villages are more remote than for tax evasion. In countries which have no satisfactory registers, the only practicable method would appear to be a prior complete count of heads accompanied by the sampling of every nth village for the measurement of the more detailed quantitative characteristics.

57. When deciding on the size of the sample in an area for which no prior census exists, it is rarely possible to fix on a definite sampling error in advance to work out the sample size. The sample size is more frequently determined by the personnel available to analyse the results, the available transport, and the money which a particular country is prepared to spend. Expenditure is usually determined by the other factors mentioned. As a consequence, the accuracy of the various characteristics under study is determined also by the above noted factors.

58. The analysis and tabulation of sample results have to be accomplished speedily if these results are not to lose their usefulness, especially in view of the fact that in most countries there is usually a great urgency attached to demographic sample surveys. Again, where interpenetrating samples are taken, it is desirable to complete the statistical cross checks of the work of each officer as quickly as possible in case it should prove desirable and practicable to conduct another survey in a particular domain of study. Provision also must be made for the computation of standard errors.

59. Examples of sample surveys carried out in the absence of a complete census as a bench mark are found primarily in Africa. The 1948 sample inquiries in Uganda, Kenya and Tanganyika are called "sample censuses" but they may equally well be considered as important sample surveys in which data on place of birth, sex, age, marital status, occupation, education, literacy and fertility were collected. Sample areas were selected, and from each, a cluster of 200 huts were chosen for investigation.

60. Another demographic survey of the African population was undertaken in 1948 in Southern Rhodesia. The frame for the sample was administrative sub-divisions listed in the registers of Kraals and plots maintained by the native commissioners. Complete enumeration was then made within the selected units; including total persons, sex and age distribution and fertility of women. A second survey was carried out in 1953 and a third in the main urban areas in 1958-60.

61. In Northern Rhodesia, a survey of African population was conducted in 1950 to obtain estimates of the geographical distribution, age and sex composition, the birth, death and infant mortality rates and fertility. In this case, insuperable obstacles were encountered in the way of conducting a complete enumeration and consequently a sample survey was the only feasible means of obtaining the needed information. This survey is described in some detail in the Handbook of Population Census Methods. ^{10/}

62. In Sudan, conditions ruled out a complete enumeration, so a 10 per cent sample census of the rural and nomadic population supplementing a 100 per cent enumeration of 68 towns was carried out in 1956. The sample was designed on the basis of the 1953 experimental census (see para. 5) which, therefore may be considered an example of a sample survey in an area which had had no previous census of population. The experimental census is described in the Handbook of Population Census Methods. ^{11/}

^{10/} Op. cit., pp. 151-154.

^{11/} Op. cit., p. 126.

63. Other African countries which have never had a full-scale nation-wide census of population but which have nevertheless completed demographic sample surveys include the following:

Central African Republic - National demographic sample survey, covering 100,000 persons, was held in 1959-1960.

Congo (Brazzaville) - National demographic sample survey, covering 800,000 persons, was completed in 1961.

Guinea - National demographic sample survey, covering 300,000 persons, was held in March 1955.

Ivory Coast - National demographic sample survey was carried out November 1957-November 1958.

Mali - National demographic sample survey, covering approximately 120,000 persons, was taken June 1960 - September 1961.

Niger - Demographic sample survey of the sedentary population of 399 villages was carried out October 1959-March 1960. Survey of nomads, who comprise about 20 per cent of the population, is planned for 1962.

Senegal - National demographic sample survey was carried out in 1960-1961.

Upper Volta - National demographic sample survey covering 185,000 persons, was held March-November 1960.

These surveys in general provided an estimate of the total population and its geographic distribution, the composition of the population in respect of age, sex, marital status and so forth, and estimates of the rate of population growth as revealed by the crude birth and death rates. Similar surveys are either underway or scheduled to be undertaken in a number of other African countries including Somalia (1962); Chad (1961-1964); Ethiopia (1962-1965); Madagascar (beginning 1961); Mauritania (1961-1962) and Cameroun (beginning 1961).