UNITED NATIONS
ECONOMIC
AND
SOCIAL COUNCIL

Distr.
LIMITED

E/CN.14/ASPP/L.10
E/CN.9/CONF.3/L.10
4 October 1962

Original : ENGLISH

ECONOMIC COMMISSION FOR AFRICA
Seminar on Population Problems in Africa
29 October - 10 November 1962
Cairo, United Arab Republic

FIELD CHECKS ON ACCURANCY OF POPULATION AND VITAL STATISTICS

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, the Statistical Office and the Bureau of Technical Assistance Operations, and the Government of the United Arab Republic as host.

/...

FIELD CHECKS ON ACCURACY OF POPULATION AND VITAL STATISTICS

Prepared by the Statistical Office of the United Nations

CONTENTS

CONTENTS (cont'd)

/...

I. INTRODUCTION

1. During recent years, increasing emphasis has been attached to the need to evaluate and publicize the accuracy of statistical series. This relatively new philosophy of admitting the possibility of error and the need (1) to quantify it in order to indicate the degree of reliability of the results, (2) to pin-point faulty procedures for correction when the collection procedure is repeated, and (3) to correct or adjust the totals in some cases, permeates all fields of statistics. It arises from the recognition that, because of their nature, statistics can not, nor need they be, 100 per cent accurate, and provided the margin of error is objectively assessed and found to be within reasonable limits, the data may be of considerable use for many purposes.

2. Application of this principle of evaluation to population and vital statistics is not easy because of the nature of the collection processes involved. For example, the successful taking of a census or national demographic sample survey depends on a complex of highly variable components. It is, first of all, dependent on a nation-wide administrative organization which covers every part of the country. It is dependent on the quality of performance of a large body of mobile enumerators recruited on a temporary basis. It is dependent on the design of a schedule or questionnaire which determines the items on which information is collected. It is dependent on the accuracy of a respondent's reply to those questions. Finally, it is dependent on the quality of the data-processing procedures utilized in the compilation. Similarly, achievement of complete (100 per cent) registration of births, deaths, marriages and divorces and the compilation of reliable vital statistics from these records is also a complex undertaking. It is dependent on the existence of a nation-wide network of local registration offices. It is dependent on each office being staffed by a competent registrar who interviews the informant. It is dependent on the ability of this registrar to collect information on certain topics or items which characterize the vital event. It is dependent on the accuracy of the respondents' answers to the questions. And finally it is dependent on the adequacy of the data-processing procedures utilized in producing the vital statistics. Similar considerations apply to data on births and deaths obtained from household sample surveys - a

procedure of data collection which has been followed in many parts of Africa. Errors are to be anticipated at every stage of enumeration, registration and statistical processing since each is a product of human endeavour; therefore procedures must be devised to eliminate, minimize or at least measure these errors at every stage.

II. TYPES OF ERRORS TO BE EXPECTED

3. Errors to be expected in population and vital statistics can be classified into two main categories. The first category comprises errors of coverage, that is, quantitative errors which have a bearing on the reliability of the total count of persons or events and on the corresponding frequencies for the geographic subdivisions of the country. These are errors of under-enumeration or under-registration of population and vital events, that is, deficiencies in geographic or ethnic coverage characterized by failure to include whole groups of population, inadvertent omission of households or individuals and failure to register or report all vital events. Over-enumeration or over-registration is also a quantitative error. It may be caused, inter alia, by failure by the enumerator to restrict himself to his particular areas of enumeration or to follow instructions for enumerating de facto or de jure population as the case may be, as of a specified date; or by deliberate or inadvertent registration of a vital event in two separate registration areas.

4. The second principal category of errors may be defined as errors of content, that is, mistakes in reporting and recording information concerning households, individuals or events. These are qualitative errors, as contrasted with quantitative, which may affect the accuracy of the distribution of population, births, deaths, marriages and divorces by such characteristics as age, sex, economic activity, and the like.

5. Good census survey and registration procedures should provide built-in safeguards designed to minimize both quantitative and qualitative errors. To help ensure complete geographic coverage in the population census, for example, the territory to be enumerated should be well demarcated, with precise geographic boundaries easily identifiable in the field with the help of clear maps. Accurate cartographic materials will materially assist in avoiding the possibility of gaps

or duplication in enumeration. Omission of households can be minimized by a system of pre-numbering dwellings and checking on coverage by dwelling number. Receipt of census schedules should be controlled by strict supervision at the enumeration district level. Similarly, clear delimitation of local registration areas will assist in control of registration completeness. Good indexes will help to avoid duplicate registration. Receipt of statistical reports on each registered event from each local register should be strictly controlled. Control of the quality of the performance of the enumerator or the registrar by on-the-spot verification of an interview will help improve accuracy. Editing and querying of census schedules and vital statistics reports for internal consistency, omission of responses, illegibility and the like can improve the accuracy of the data. Despite these safeguards, which are part of the normal conduct of a statistical inquiry, the population and vital statistics will reveal inaccuracies which have their source in the enumeration or registration process. Methods for detecting and measuring these are set forth below.

III. METHODS OF EVALUATING ACCURACY OF ENUMERATION AND REGISTRATION

6. Methods of evaluating the accuracy of statistics may vary in detail according to the purposes to be achieved and the resources and degree of statistical development of the countries and territories involved, but basically they may be classified into two types, designated for convenience as the direct and the indirect methods of evaluation.

7. The indirect method of detecting and, if possible, assessing the extent of omissions or duplications in a census, survey or register is the analytical procedure which consists in scrutinizing the statistical results tabulated from the census schedules or the records in the register, for plausibility, comparing these with corresponding numerical data from another time period or from a similar geographic area, and examining the various distributions for consistency one with the other and for conformity to classic relationships which are known to exist between different population characteristics from the same universe and between these and a theoretical model.

8. The direct method is the more fundamental evaluation procedure designed to detect errors of coverage and quality. It consists of checking each item on the

individual census or survey schedule or those on the entries in the civil register against corresponding data from other current records for the same individual obtained by means of an independent data-collection mechanism. Since either set of records may be in error, discrepancies must be resolved by further investigation.

9. It must be emphasized that both the direct and the indirect methods should be employed in evaluating any type of statistics. Suggested analytical assessment methods have been set forth in document E/CN.9/CONF.3/L.11; the direct method of "field checks" is described below. Applications to population data are presented in Part A, and those applicable to vital statistics are given in Parts B-C.

A. Direct Evaluation of Population Census Results

10. In this section, the problems of direct evaluation of population statistics are discussed with reference to results of an attempt at 100 per cent enumeration of the population. To a large extent the methods discussed here are also relevant to the problem of evaluation of population data obtained from sample surveys though some modifications are required for the latter purpose.

11. As noted above, direct evaluation of the census results consists in matching each census report on an individual to a similar record from another source and assessing the degree of comparability observed. This comparison can take place immediately after, or long after, the census date so long as it utilizes lists of persons presumed to be living in the area as of the time of the census. It is obvious that lists originating months or years following the census will make matching difficult or impossible because of population mobility.

12. "Independent records" for this purpose can come from two sources: the commonest and most easily obtained are the lists or registers of the population or of segments thereof which have their origin in an administrative procedure not necessarily statistical in nature. A second source of independent records is a sample survey, that is, the repetition of the census process on a sample basis and under conditions which seek to ensure that the errors committed in the original enumeration will not be repeated and other errors will be minimized.

13. The relative adequacy of either of the sources mentioned above can be measured by its ability to provide (1) an estimate of the completeness of enumeration of households and persons for both the country as a whole and for selected geographic

subdivisions and (2) a measure of the accuracy of the characteristics recorded for individuals in the population. This criterion will be applied first to registers, to show that they are rarely adequate for the comprehensive evaluation envisaged here and that the field check procedure is likely to be the best method.

1. Registers

14. Registers which have been used to check census schedules are lists of children enrolled in schools; lists of registered voters, i.e. electoral lists; registers of births and deaths; families on population registers; lists of taxpayers (real estate tax, income tax, etc.); registers of military recruits, social security beneficiaries and so forth.

(a) Partial lists

15. It will be apparent at once that, with the exception of the population register, all of those mentioned are partial lists in that they cover only limited segments of the population such as school children, voters, taxpayers, new-born children and so forth. Because of this limitation, comparison of census schedules with these cannot provide an estimate of the completeness of the census enumeration for the country as a whole or for its major civil divisions, the first prerequisite for adequacy. Nor can an evaluation of quality of response be obtained in most cases because such registers do not usually supply information on an individual's personal characteristics. The completeness with which children under one year of age have been enumerated may be evaluated by collating the infants reported in the census with these inscribed in the birth register and the reliability of response in respect of topics related to infants can be checked, but this is of limited value.

16. In addition to their limitations in terms of coverage, it must be borne in mind that the completeness and accuracy of the partial register itself is likely to be questionable. In countries where school enrolment is low, the list of school children will be deficient. Birth registers tend to be incomplete in the countries which are likely to have poor census enumeration. Electoral registers are notably incomplete in many countries. Tax registers are likely to be deflated while food

ration registers tend to be inflated.[1] For this and the other reasons given above, it is clear that partial registers are not entirely adequate for assessing either the completeness or accuracy of national census results.

(b) Population registers

17. Population registers, being the most comprehensive of the registers mentioned, can serve as a better independent source against which to check the census records, and indeed they are so used in some European countries. But even in this case, when discrepancies are found in two records, there is no way of knowing which set of data is correct without locating the individual concerned and re-interviewing him. This is usually the procedure adopted. For example, after the count of 1950 in Norway, all census schedules were sent to the office of the local population register to be verified. Discrepancies were followed up in the field and as a result close to 40,000 persons were added to the census schedules while some 20,000 were deleted. Thus, a net addition of about 1/2 per cent of the total domiciled population was made, a correction of little significance in relation to the whole, but of importance for breakdowns such as age and occupation. The limiting factor on the use of this source of records is, of course, the relative rarity of reliable population registers.

2. Sample surveys

18. A second type of independent record which can be used for checking the completeness and accuracy of a census can be obtained from a sample survey. Such a check depends on matching the documents relating to the same individual obtained from the enumeration and the survey. Four groups of documents should result from this procedure: one group comprises persons enumerated in each of the two investigations; a second group consists of persons enumerated in the census but not found in the survey households; a third group is the reverse of the second, i.e., persons enumerated in the survey households but not found in the census; a hypothetical fourth group is persons missed in both the census and the survey.

1/ For example, Japan, in addition to independent post-censal sample field checks, also used the food ration register to check both the 1950 and 1955 censuses of population. A surplus of 0.5 per cent was found in the register for 1950 and 0.66 per cent in 1955. However, these percentages could not be considered equivalent to under-enumeration because of the possibility that they represented duplicate registration in the food register.

for which some allowance is usually made. When frequencies for the first three groups are determined, re-examination or re-interviewing of those in groups 2 and 3 must be undertaken to determine whether or not these persons should have been enumerated or surveyed, as the case may be, according to the definitions of the two investigations.

19. Theoretically, any independent random sample survey of households should yield records adequate for the purpose provided the sampling frame is such that the unit of enumeration is the same in each type of investigation, and that households and individuals which failed to be enumerated in one or the other can be identified. However, experience with the use of labour force and other surveys for this purpose has been disappointing, primarily because of differences in the basic definitions used in the two inquiries. Canada, for example, utilized the May 1956 Labour Force Survey (an area sample of about 1 per cent of the population) to check on the June 1956 census of population. Documents for the same household from each investigation were matched and those households where differences in composition were found were re-enumerated. Analysis of the resulting data produced an estimate of about 1.1 per cent under-enumeration at the census[2/] but it was also established that differences between the two investigations were due in part to lack of comparability in the definition of "household". Difficulties of matching may be posed also by the adoption of slightly different concepts of the population to be enumerated in the household, for example, the enumeration of population usually residing in the household at the time of the census and that present at the time of the sample inquiry.

20. The experience of Japan in using the Labour Force Survey as a mechanism for checking the 1950 population census exemplifies another disadvantage of using a sample survey designed for a purpose other than evaluation. Following the 1950 census of Japan, households comprising the 10 per cent sample of census returns used for preliminary tabulations were matched to records from the monthly Labour Force Survey.[3/] Net under-enumeration was found to be 0.74 per cent, but since the Survey was concerned only with persons 14 years of age and over, this percentage

---

2/ Edward, D. and Yablonski, E. "Population Quality Check of the 1956 Census of Canada". Ottawa, 1958.

3/ Morita, Yuzo, "An appraisal of the population census statistics". Bulletin de l'Institut International de Statistique, tome 34, 3ème livraison, 1954, pp. 189-190.

could not be extrapolated to the population as a whole since it did not reflect under-enumeration throughout the early ages.

21. Still another example of an unsuccessful attempt to use a survey not specifically designed for the purpose can be found in the experience of the United States of America. In 1950, an experiment was planned to study differences arising from different enumeration techniques, training programmes and questionnaire design as well as to measure, if possible, completeness of coverage, age reporting errors, income reporting, and so forth. It was intended to match returns from the April 1950 Current Population Survey (CPS) to the schedules from the 1 April 1950 census of population, but due primarily to procedural difficulties, the experiment was abandoned in favour of a post-censal sample survey designed specifically for that purpose.

3. Post-censal sample field check defined

22. Since sample surveys designed to obtain current data of one type or another may not be strictly comparable to census enumerations in matters of geographic subdivisions, units of enumeration, type of population included or population coverage, they are not ideally suited for checking the general population census for completeness of coverage or accuracy of response. A more efficient method is replication of the census under intensive conditions of precision, in a post-censal field check defined as the independent re-enumeration of a representative sample of the population in such a way as to provide a reliable measure of (1) the number of households and persons who were omitted from, or erroneously included in, the original census count and (2) the magnitude and nature of response errors.

23. A true post-censal field check must meet three specific criteria, namely, (1) be independent of the original census; (2) be representative of the whole country and all population groups; and (3) involve one-to-one matching of records to produce an identical sample from each investigation. Independence is an especially important criterion and one which failed to be achieved in many of the so-called "field checks" carried out in connexion with several recent censuses. In Ethiopia, for example, a post-censal sample survey, covering every 15th household reported in the census of Addis Ababa, was conducted in 1961. In the course of the original enumeration, the households had been identified by the use of stickers affixed to the living quarters of each household. The supervisor of each enumeration district

was subsequently given a list of sample-household numbers for each of the districts under his supervision and the re-enumeration was limited to these households. The procedure followed could not uncover omitted households and in this respect fell short of the independence criterion for a post-censal field check.

24. Similarly, after the 1953 census in Ceylon, a 1 per cent sample verification survey was made to test the completeness of enumeration but this test, also, was based on a list of households already reported in the 1953 census; it could not, therefore, be considered independent of the census.

25. Representativeness in respect of the whole geographic area and all population groups is also an important criterion and one which has also failed to be achieved by several inquiries which were undertaken as field checks. For example, a post-censal survey was undertaken in Costa Rica following the 1950 census but it was confined to those parts of the country where enumeration was thought to be inadequate Re-enumeration was carried out in 135 of the 316 enumeration districts but because of the nature of the recheck these were not uniformly distributed throughout the country. The results showed 6,878 persons omitted in the 135 districts but since this appeared insignificant, the additional schedules were simply added to those originally enumerated, revising the count upward by 0.86 per cent. Since it was not comprehensive, or unbiased, this is not considered a true post-censal field check.

26. Similarly, a "post-censal field check" of the 1950 census of Mexico was designed to cover Mexico City and 18 municipios in the rest of the Republic where illiteracy was high and school attendance and number of teachers low.[4/] The selected geographic areas were not representative of the country. Moreover, lacking detailed maps, the selection of households in the municipios was made from a list of the original census schedules. Hence the test was not independent of the original census, except in Mexico City and it thus fails on two counts as a true post-censal field check.

27. In India, the frame for the sample field check on the 1951 census was the National Register of Citizens which had been transcribed from the original census schedules. Naturally, with this alone it would be impossible to find a household

4/ Informe sobre los censos de población de México, by Oscar Morales Cabanas. (UN document E/CN.9/CONF.1/G.10).

which was omitted in the original enumeration, though the test could give an indication of the extent of over or under-enumeration of persons within the households. To locate households which had escaped the original enumeration the re-enumeration was supposed to investigate not only the selected household but three adjacent dwellings, to see if all had been caught in the first enumeration. This procedure goes some way toward independence but does not ensure representativeness.

4. Features of a post-censal sample field check

28. True post-censal sample field checks have been carried out or attempted in relatively few countries during recent years; these are described briefly in paragraphs 63-73. The techniques used vary in detail, but the salient features of the re-enumeration, common to all, are outlined below. These in effect might constitute a guide for planning a true post-censal sample field check.

(a) Scope of re-enumeration

29. Re-enumeration of the entire population is not desirable nor is it necessary. It is not desirable because it would be excessively expensive and the cost, together with the burden imposed on the public in two successive censuses, makes 100 per cent re-enumeration prohibitive. It is not necessary because modern sampling techniques make it possible to re-enumerate only a fraction of the total population in order to obtain an estimate of the possible error involved. This fraction should however be a random sample of the population, in order that quantitative measures of under and over-enumeration may be derived from the results. Non-random or purposive samples might perhaps be acceptable where quantitative measures for the entire country or population are not of crucial importance, but the utility of maintaining a sample of the population for the study of future problems indicates the desirability of selecting a representative sample to begin with.

(b) Timing of the re-enumeration

30. To be most effective, the re-enumeration should take place as soon as possible after the conclusion of the original census of the areas which fall into the sample to be re-canvassed. The longer the time interval between the two investigations,

the greater will be the number of differences resulting from births, deaths, marriages and migrations which occurred in the interval, and the greater the chance of introducing response errors in respect of those items involving retrospection. The more differences there are between the two sets of records, the more time will be consumed in the subsequent matching and resolution process. Re-enumeration shortly after the census will tend to minimize discrepancies. A secondary, though not unimportant, reason for having the post-censal field check soon after the original enumeration is that by so doing one may take advantage of the field organization already in existence for the original census.

31. Although it should take place soon after the census, the re-survey must not interfere with the original census procedures. Neither should it be carried out as an integral part of the first enumeration, with the same enumerator asking both sets of questions. Re-interrogation of the informant by the same enumerator is good census practice, which will tend to enhance the accuracy of the original responses, but it does not constitute an independent check on the accuracy of the enumeration.

32. Some countries have found it possible to organize a sample survey at about the same time as the complete census, in the course of which additional items of information, not included in the general questionnaire, are collected on a separate schedule, on a different day, by an entirely independent group of enumerators. When such a sample survey is designed to cover the entire territory and with the same definitions as the complete census, it may also be used to give an assessment of the completeness of enumeration and the accuracy of reporting at the complete census. It can also be used to assess the cost, speed, adaptability, efficiency and accuracy of sampling methods in comparison with complete enumeration.[5/] For most countries, however, it will be easier to organize a relatively simple sample field check to follow the original enumeration and to have as its sole objective the checking of the accuracy of the census.

(c) Type of sample

33. Theoretically it is possible to select a purposive sample of areas about which information is particularly desired but, for a number of reasons mentioned below,

---

5/ Field checks of this type have been carried out by several countries. In France, the post-censal field check of the 1954 census was coupled with a periodical survey of the employment situation.

it is recommended that the post-censal field check be carried out in a well-designed random cluster sample, small enough to minimize cost but large enough to give at least an evaluation of the enumeration in the country as a whole and in its major civil division since determination of quantitative measures of under- and over-enumeration is one of the important objectives of the procedure. Moreover, the random cluster sample will provide a scientific basis for future investigations, and for this reason alone it would be desirable to employ random sampling. Should there be suspicion of great regional variability in the quality of enumeration, 100 per cent re-enumeration in the areas suspected of large errors might have to be carried out separately, but results of such intensive checks might then be combined with those of a small sample of the rest of the country to evaluate over-all accuracy.

(d) Sampling unit

34. In organizing the independent sample survey, the ultimate sampling units for re-enumeration should be clearly defined geographic areas, operationally convenient, preferably the enumeration districts used in the original census. This point is particularly important. It means that the sample should be chosen not from a list of housing units, households or names enumerated at the census being evaluated, but from a comprehensive frame consisting of geographic area units, the probable population-size of which is already known and within which every household will be re-canvassed. This would imply the availability of good maps showing enumeration districts; a list of such districts with boundary descriptions is not adequate for the purpose. If good maps of local areas are not available, sampling units of defined size and shape could be indicated on a large-scale map by utilizing natural boundaries such as roads and streams, for example, and all households falling into the demarcated area could be enumerated. In such cases, where the population-size of the selected areas is not known precisely, more care will need to be exercised in order to control the size of the sample. Another solution applicable to Africa might be to use the village as the unit of sampling, because the boundaries of villages are always precise and their identity is safeguarded by a unique name on the village lists. Moreover, in Africa, the village may well have been the enumeration district for purposes of the original

/...

census. Since matching of census schedules from the survey and the census is the ultimate goal, the closer the identity of the areas being re-enumerated with the original areas, more easily can matching procedures be carried out.

(e) Size of the sample

35. The usual practice is carefully to evaluate the precision desired in the estimates of error (at least in the most important items being investigated) and the geographic detail required and then to choose the sample design which would provide this accuracy at least cost. The sample should not be a large one because this will tend to increase the non-sampling errors, but it should be such as to afford estimates of known reliability for at least major geographic sub-divisions of the country and major categories of the population.

(f) Enumerators

36. Since the sample field check is itself a micro census, it is also subject to errors of coverage and content, just as was the original census. Special precautions, therefore, need to be taken to keep each source of non-sampling error to a minimum, and the first precaution is to select the most highly trained and experienced corps of enumerators possible. This corps can be much smaller than that employed in the original census and it therefore can be subjected to more intensive training and special supervision. It can be paid on a daily rate basis rather than on a piece rate and because of this it can take sufficient time to obtain valid replies to questions and to verify its work by cross-checking and repeat visits. In many of the recent post-enumeration sample surveys, the census supervisors who received training and gained experience in connexion with the total enumeration were used as survey enumerators.

(g) Informant

37. In most 100 per cent censuses, the enumerators are allowed to receive information for each member of a household from one acceptable adult informant. In the sample field check, there can be more selectivity because the enumerator in the sample field check operation will be able to make several revisits in order

.../...

to interview each individual adult member of the household who falls into the sample; this should, presumably, work to improve the quality of the replies to survey questions.

(b) Contents of test schedule

38. The questions to be asked during the sample field check should preferably be fewer than those included in the original census. A simple short list of items which will provide the means of identifying the household and the individuals who are members of it, and questions on the items selected for verification of response (see para. 40), will prove most efficient, because of the resources and time required for a comprehensive inquiry and the possible antipathy of the public to repeated questioning. This is especially true in countries which do not have a long tradition of census taking and where, consequently, the public may be distrustful and unco-operative at best.

39. Naturally, errors are possible in connexion with every item on which information is collected at the census, and in theory, it would be desirable to measure the degree of reliance one might give to each. But since it is clear that the re-examination of every item is not feasible or desirable, some criteria must be used to choose the items to be verified.

40. Many criteria for selection might be set up, but the most effective appear to be (1) the relative importance of the potential error in connexion with the use to which tabulations on each item might be put, and (2) the stability of the information. On this basis, one might hazard a guess that the items with priority for re-measurement would be sex, age, tribal affiliation or ethnic group, martial status, and the economic characteristics of the population, since these are more or less fundamental. Another contender for verification might be the fertility item, that is "number of children born to this woman". High priority might also be given to place of birth and place or residence, especially if tabulations on these items are to be used to estimate internal migration.

41. In summary, it may be said that selection of the items to be verified must necessarily be made on a national basis, taking into account the special enumeration problems in the country (such as large numbers of aliens or non-residents, personal

speaking foreign languages, the extent of literacy, extent of under-registration of births and marriages and the like) as well as the subsequent uses to which the data may be put and the manner in which they have been collected originally.

42. Where it is necessary to design the sample survey both as a post-enumeration check and as a means of collecting additional information as is being done in Nigeria and Ghana (see paras. 64 and 69), care should be taken to keep these additions to a minimum so as not to overburden the survey unduly. Inclusion of a large number of questions may in the end bring into question the validity of the results of both the post-enumeration check and of the additional inquiry.

43. Whatever the content, the questions on the post-censal schedule could be phrased more carefully than was possible on the original, and probe questions designed to help the respondent give more accurate replies could be introduced. Such probe questions will probably be needed in connexion with "age" where this characteristic is not well known, with "marital status" when consensual rather than legal marriage is prevalent, and with "economic characteristics" in all cases because concepts in this area tend to be confused. It is assumed that experience gained in the original census will have resulted in more effective interviewing methods and ways of structuring questions.

(4) Testing procedures

44. When completed schedules from the field check are assembled, the next step is to search for the corresponding record in the original census files for the sample area. As noted in para. 18, the result of the search should yield three groups of schedules: (1) matched schedules; (2) schedules present in original census but not found in the field check; (3) schedules present in field check, but not found in the original enumeration. Discrepancies in matching, that is groups (2) and (3), are revisited to verify that they were properly enumerated and that the difference is real. The percentages derived from the three groups which remain would constitute meaaures of under- or over-enumeration of households and persons, on the assumption that the field check as corrected through re-verification of the two lists is more reliable than the original census.

45. It should be noted in this connexion that name-to-name matching is not an easy task in any country and in cultures where names lack uniqueness, it may be difficult. Even the first step of identifying the household might be particularly

/...

difficult in Africa where street names and house numbers are practically non-existent. A practice which has been found useful in this respect is the affixing of an identifying household number in advance of the original enumeration. Such a number will assist in identifying the household if it should appear in the sample check.

46. Using the matched schedules, comparisons are made between the responses obtained to each question being verified, and discrepancies can be recorded to evaluate the accuracy of the original replies. In each case the replies to the sample field check, after investigation of discrepancies by re-interview to the extent considered necessary and feasible, are assumed to be correct and the basis for evaluating the original response. A different reply on the original census is assumed to be erroneous unless it were patently impossible.

5. Errors revealed by a post-censal sample field check

(a) Errors of coverage

47. Errors of coverage revealed by the matching procedure have a bearing mainly on the total population figure, though, as will be mentioned below, they may also, in some cultures, involve the accuracy of certain characteristics. The term usually used to designate errors of omission in coverage is under-enumeration; that used to designate errors of duplication is over-enumeration. Over-enumeration errors seem to be less frequent than under-enumeration errors, but the possibility of both must be considered because each has a direct bearing on the accuracy of the total population figure for the country and its administrative divisions.

(1) Under-enumeration

48. Under-enumeration, or the omission of persons who should have been included in the enumeration, can be caused by failure to cover all the area to be enumerated, or by failure to enumerate all the persons within that area.

49. Considering first failure to cover the entire area, it will be obvious that if a country is large, with scattered population and poor transportation facilities, it will be difficult to reach every household. The field check, on the other hand, being based as it most often is on a sample of the area, will be more likely to result in complete coverage because the distances to be covered will be smaller. This is especially true for countries which have almost inaccessible territory due to natural barriers such as dense forests, mountains, deserts, swamps and so forth.

50. Omission of indigenous or nomadic population groups, like omission of territorial units, is usually due to distance and inaccessibility. The field check on a sampling basis, with better trained enumerators able to give more effort and care, will be more likely to cover them.

51. Other groups of population likely to be overlooked in the census are those who live in unconventional housing, as for example in forest huts, caves, boats, tents, cellars or other non-residential housing units. Again, the sample field check will be more likely to take account of these. Institutional population, lodgers, persons living in hotels and transients also constitute potential omission, and the incidence of these omissions will tend to be reduced in a sample.

52. Discussion of potential omissions would not be complete without reference to the under-enumeration which occurs almost universally among certain age groups of the population, but especially among infants and young children. New-born babies and children under 5 years of age are easily overlooked at the census. An intensive re-interview will be more likely to uncover these omissions through probing questions and cross-checking.

53. In some cultures, there can also be serious under-enumeration of women - especially is this so in Moslem countries where women remain secluded and the male canvasser is not allowed to interview them. When omission of this type is suspected, female enumerators could be employed in the field checks to try to remedy the deficiency. Sudan experienced this problem at its 1956 census and remedied it with an intensive re-enumeration, using female enumerators.

(ii) Over-enumeration

54. Special mention must be made of the errors of over-enumeration produced most often by failure to interpret correctly the effective date of the census when the operation is carried out over a period of time. Persons born just after the specified census data are likely to be included in the enumeration while persons who died just before may also be included. This type of error is especially subject to verification by the field check technique, since the questioning can be more thorough.

55. Double counting of certain persons, particularly when the enumeration is on a residence basis, may also be uncovered by the field check. As a matter of fact, the the re-enumeration may also provide a means of obtaining both types of count, i.e., the present-in-area figure as well as the resident population.

(b) Errors of content or response

56. As distinct from coverage errors which affect census totals primarily, errors in reporting or recording data can jeopardize or invalidate the subsequent detailed tabulations by personal characteristics. This type of error is much more difficult to detect and correct than are those of coverage, but every effort must be made to reduce them to a minimum since they can so seriously impair the use of the data.

57. First among the reporting errors might be mentioned the possible incorrect designation of a person's place of residence - his geographic attribute - and hence his inclusion in the wrong enumeration district. This type of error would be particularly important in so far as matching of records is concerned but it would not ordinarily affect to any degree the totals for the major territorial subdivisions, or the total for the country as a whole. It could, however, produce large errors in small area statistics. The results of the coverage check of the 1961 population census of Greece, for example, showed little change in the total population as a result of the application of the more detailed and careful procedures of the post-enumeration sample survey, but it did reveal that the census enumerators made both types of errors, missing people at addresses where they should have been counted and enumerating people at places where they should not have been counted in accordance with census rules, resulting in an estimated 67,000 erroneous inclusions and 46,000 omissions at listed addresses.

58. The most important errors of reporting are those made in describing the personal characteristics of the population, such as age, sex, marital status and the like. Errors of this type may arise from ignorance, a misunderstanding of the question, or from deliberate misstatement and falsifications. Irrespective of the reason, the results are unreliable distributions of population, which must be evaluated in order that the data may be useful.

6. Use of the results of the field check

59. The methods used in evaluating the completeness of the census results and the accuracy of response should be fully explained in the final census publications. If the enumeration coverage is deficient or over-counted to a relatively small and statistically insignificant degree, the results of the check should simply be

announced as guides to the user of census data. If the errors are of medium magnitude, they also should be announced and published in the report on the census, together with the method of evaluation but, in addition, it might be advisable to correct the total result of the census - in other words, to adjust the total upward or downward in accord with the test result.6/ It will not ordinarily be possible to adjust the distributions, but in any case, correction factors should be computed for use in analytical studies and in connexion with future censuses.

60. If the errors found are very large, the only recourse is to invalidate the census and carry out a new enumeration unless, as mentioned above, the sample field check itself is adequate to provide useful aggregates.

7. Limitations of the sample field check

61. Despite the fact that every effort is made to achieve exceptionally high quality of performance in the sample field check and although a sampling error can be established, it must be borne in mind that in so far as methodology is concerned, this method in reality involves a new census or survey. Errors of coverage and content will be present in the check census just as they were in the original census, though presumably on a much smaller scale, considering the efforts made to eliminate them. However, the continued presence of some error means that the differences observed between the results of the original census enumeration and those of the sample field check do not necessarily represent total error. In the United States, for example, a smaller scale field check was made to test the accuracy of the regular post-censal field check. The results indicated that the post-enumeration check had under-estimated the number of persons omitted from the original count. In view of this, the results of a sample field check must be considered as an estimate of the true situation. It is not a sufficiently sensitive tool to be used alone; it should be complemented by analytical checks which can confirm or indicate the possibility of error. These analytical methods of evaluating census results have been presented in document E/CN.14/ASPP/L.11; E/CN.9/CONF.3/L.11.

6/ This is not always feasible, however. In some countries, the tabulated results of the census enumeration have a legal status which cannot be altered even though known to be incorrect.

62. Another limitation on the use of the sample re-enumeration, which perhaps should be mentioned, is the matter of expense. Since a sample field check is in reality a sample census of small scale, it will require expenditures for staff and equipment similar to that of the census. The cost of a properly planned check, however, can be very modest in comparison with that of the census itself and it can do much to enhance the value of the results.

8. Examples of true post-censal field checks

63. Even in countries with a long census tradition, post-censal sample field checks of the type described above have been conducted in relatively few countries. Details of the re-enumeration checks for which information is available as well as information on checks which did not meet the criteria of "independence" mentioned in para. 23 are described in the Handbook of Population Census Methods, Vol. 1.7/ A brief reference to those which are thought to have met the criteria for a true post-censal check are made below.

(a) Ghana

64. To check the reliability of the main census of 20 March 1960, and to inquire into various social and economic problems such as population growth, employment and under-employment, migration, etc., a post-enumeration sample survey was carried out in June 1960. The survey covered about 5 per cent of the population. The sample design frame of the first stage and the house or compound for the second stage. Additional topics were included in the survey, e.g., religion, marital status, migration, literacy, mortality, etc. The results of the survey are in the process of analysis.

(b) Greece

65. True post-censal sample field checks were carried out after the October 1950, 1955 and 1960 censuses of population. Households and individuals in a sample of 233 enumeration districts were re-enumerated on 15 December 1950. According to this test, under-enumeration was estimated to be between 0.1 per cent and 0.7 per cent. This may be compared with the results of the comparison of census

7/ United Nations document ST/STAT/SER.F/5/Rev.1, pp. 108 and 131.

records with (1) food ration lists and (2) Labour Force Survey records (see paras. 19 and 20 which give net under-enumeration of 1.1 per cent and 0.74 per cent.

66. A sample of 840 enumeration districts from the October 1955 population census were re-enumerated 15 November 1955. Estimates of under-enumeration were derived in the same way as in 1950 and with similar results.

67. The post-enumeration survey for the 1960 population census was taken as of 5 November 1960. The enumeration districts selected at random numbered 659 for cities and 243 for counties. The households were selected at the ratio of 1/5 from all the households located in the sample enumeration districts. Personal characteristics were checked for accuracy of response, whereas the items on economic activities will be collated with the entries in the questionnaire of the current Labour Force Survey.

(c) Liberia

68. The first census of population of Liberia was taken in April-May 1962. A post-enumeration survey was planned to be carried out shortly after the census. Details of the survey are not yet available.

(d) Nigeria

69. Immediately on conclusion of the Federal Census of Nigeria, May 1962, a post-enumeration survey was conducted. The survey covered about 1,000,000 persons, or roughly 2 1/2 per cent of the estimated total population and was designed to supplement as well as evaluate the census. Topics covered in the survey included marital status, fertility, economic characteristics, migration and disability. Results are not yet available.

(e) Tunisia

70. In April 1956, Tunisia undertook a post-censal field check with a view to assessing both the qualitative and the quantitative accuracy of the 1956 census of population. Difficulties of distance and inaccessibility necessitated the exclusion of the rural areas as well as certain communes, reducing coverage to about 27 per cent of the total for Tunisia. Further difficulties in execution were encountered, including problems of recruitment of qualified personnel, resistance of the public to accord interviews, difficulties in demarcating sampling units and absence of street names and house numbers. None the less, the survey was useful as a pilot experiment.

(f) United States

71. After completion of the 1950 census of population, a post-censal sample field check was undertaken to evaluate coverage and content errors. The first-stage sampling units were enumeration districts designed to represent the entire land area of the United States of America; the second stage was "dwelling units". Intensive re-interviews were conducted with the object of obtaining accurate information on the characteristics of the population.

72. The post-censal field check itself was subjected to evaluation by a small-scale field check. Results of this re-check indicated that the post-censal field survey erred in the direction of underestimating omissions.8/

(g) Yugoslavia

73. A post-censal field check following the 1953 population census in Yugoslavia was designed to (1) check completeness of enumeration and (2) appraise the quality of the answers. All persons in selected enumeration districts were re-enumerated one week after census day, and the re-enumeration schedules were matched to the original census schedules. Differences were investigated and resolved. More details on this survey are given in the Handbook of Population Census Methods.9/

9. Plans for post-censal sample field checks in connexion with 1960 censuses

74. While numerous demographic, socio-economic, agricultural and other sample surveys are being taken or are planned in African countries, only little information is available as to any detailed census plans for post-enumeration surveys. Information as to the experience gained in connexion with the post-enumeration surveys of Ethiopia, Ghana and Nigeria is still in the process of analysis. A post-enumeration survey is planned to be carried out following the 1962 population census of Kenya.

8/ For additional details, see the Handbook of Population Census Methods, pp. 109 and 135 and Methods of Appraisal of Quality of Basic Data for Population Estimates, Manual II, ST/SOA/Series A, No. 23, pp. 16, 17.

9/ Ibid., page 134.

B. Direct Evaluation of Vital Statistics from Civil Registers

75. The direct method of evaluating the accuracy of birth, death, marriage and divorce statistics is essentially comparable to the direct method of evaluating population census results. It depends on the one-to-one matching of vital events from two independent records with the objective of assessing the extent of omissions or duplications (or, in other words, the proportion of events which have failed to be recorded or have been recorded twice) and also the accuracy of the response to questions on the fact, date of occurrence and characteristics of the birth, death, marriage or divorce.

76. The discussion in this section refers to methods of direct evaluation of birth and death statistics from civil registers. The special problems of evaluating registration data on marriages and divorces are not considered here, as the priority need is for checking birth and death records. In Africa at present the problem of evaluating birth and death statistics from civil registers has somewhat limited immediate relevance inasmuch as civil registration in the majority of African countries either does not exist or is limited to urban areas or minor ethnic groups. Nevertheless, such evaluation is important for those areas where civil registers are maintained. Moreover, to a considerable extent, the methodological principles of such evaluation are also applicable to the problem of evaluating birth and death statistics obtained from household sample surveys - a source of data which has been developed recently in an increasing number of African countries where satisfactory registration records are lacking. The need for evaluating the quality of vital statistics from whatever source they may be obtained cannot be overemphasized, especially where, as in most of Africa at present, comprehensive systems of record-keeping and current statistical reporting in this sphere have not yet been firmly established. In this situation no data should be accepted at face value without the most thorough verification possible in the circumstances.

77. As is the case with post-censal field checks of census returns, the independent records for checking the completeness and accuracy of the registration of livebirth and death may come from two types of independent sources: (1) other registers or lists and (2) records obtained from a sample survey.

/...

1. Registers

78. Registers or lists against which the live birth register may be checked include any more or less comprehensive lists, the most useful of which are (1) notifications of birth, which originate within hospitals or with the medical attendant; and (2) church or mission records of baptisms. For deaths, secondary registers which can be used for checking the completeness of the register include inter alia (1) notifications of deaths originating in the hospitals or in records of the medical attendant; and (2) lists of interments.

79. All of the lists mentioned constitute partial registers of births and deaths. They will not likely contain more entries than the official birth and death registers but they may include different events. However, since they are not comprehensive, they can only serve as indicators of deficiencies in the official registers; for evaluation of the extent of the deficiency and its nature, recourse must be had to an independent list of births and deaths generated by a household survey.

2. Census of population

80. The most comprehensive household survey which has been used for checking a civil register is the population census. A number of countries have in the past utilized the enumeration to carry out a birth registration test.[10] In such a test, a special schedule (called an "infant card") is filled out by the enumerator for infants born during a selected period (usually three or four months) preceding the census. Each of these cards, representing an infant under 3 or 4 months of age, is checked to the birth register to ascertain whether a corresponding birth record is on file.

81. The cases in which a birth record is found constitute a set of matched records; a second group will consist of births registered but not enumerated; a third group will consist of infants enumerated for whom no birth record could be located; and a fourth group which escaped registration and enumeration can be assumed.

10/ For a description of these, see United Nations Handbook of Vital Statistics Methods, Studies in Methods, Series F, No. 7, pages 204-208.

82. After taking account of deaths in the interval between birth and the census date, and with some allowance for group 4, the second group may be indicative of under-enumeration of young infants at the census. The third group, with some assumptions for group 4, constitutes the extent of under-registration. The procedure is theoretically sound as a device to obtain an independent list of births for evaluating the birth register but it must be noted that it is expensive and time consuming and may result in delaying the normal enumeration and tabulation. Moreover, it can be done only when a census is taken, which is most often decennially. A more flexible method utilizes the sample household survey technique for generating the independent list of births and deaths. Dependence on a sample of the total population rather than on the population of the whole country materially reduces the amount of work involved.

3. Sample surveys

83. The technique consists of inquiring of each household in a particular area (1) whether there has been a birth to any member of the household during the previous 12 months; and (2) whether any member of the household had died during the previous 12 months. In addition to the fact of birth or death, enough information on characteristics of the newborn, and the decedent, is collected to allow the event to be matched with an entry in the birth or death register as appropriate. The degree to which the births and deaths reported in the survey are not able to be found in the registers constitutes the degree of under-registration, suitable account having been taken of various problems mentioned in paras. 104-108 below.

84. As was the case with post-censal field checking, it is theoretically possible to utilize any household survey for testing birth and death registration completeness. For example, the questions on the incidence of births and deaths could be added to a sample household survey of population or labour force, thus producing a multi-subject survey. However, there is no reason to believe that this would be any more successful in the case of births and deaths than it was in the case of population (see paras. 19-21) and for the same reasons. A sample designed to produce births and deaths for checking a civil register must be drawn from a frame of civil registration area; most sample surveys utilize other types of frame and hence would not be appropriate for the purpose under discussion.

/...

4. Features of a vital statistics field check

85. So far as is known, checking of civil registers by data obtained from sample surveys of households has not been carried out on a very wide scale to date. The sample survey technique has been used rather extensively in Africa to obtain estimates of the crude birth and death rate where the traditional civil registration system is lacking or unreliable[11] but it appears that evaluation by the survey method is not so prevalent. This is not unusual, inasmuch as registration is so deficient in Africa that the survey will almost surely turn up more cases of births and deaths than the register. Nevertheless some salient features of the method are described in some detail below on the theory that it should form part of any procedure for checking the civil registers now in existence and certainly be included in any experiment or system of civil registration adopted in the future. It should be pointed out that schemes[12] for the continued observation and recording of vital events in a sample of areas, supplemented by an initial and subsequent periodic demographic survey of the population under observation, involve inter alia checking the registered births and deaths against the events recorded in the periodic surveys. It is essential, therefore, that the salient features of the field inquiry be clearly understood.

(a) Scope of survey

86. The inquiry envisioned is a sample survey of the areas covered by the registers being evaluated. Modern sampling methods enable an index of completeness to be calculated from results obtained from a fraction of the total population, so long as it is a probability sample. In the case of the birth-registration test utilizing the census of population described above, the scope was nation-wide and based on the schedules for all the households. However, to bring the project into workable proportions, the births during only the three previous months were recorded. The sample survey scheme is more limited in population coverage but more extensive in the time period of reference (see para. 92).

11/ Central African Republic, Congo(Brazzaville), Congo(Leopoldville), Dahomey, Gabon, Ivory Coast, Guinea, Mali, Niger, Senegal, Upper Volta.

12/ See Problems in African Demography. Bulletin of the Union Internationale pour l'Etude Scientifique de la Population. Paris 1960, Chapter IX, pp. 23-26.

(b) Timing of the survey (period of observation)

87. A field check on the accuracy of vital statistics can take place at any time, since it depends on a method completely independent of the civil registers. But, as in the case of a post-censal field check, the longer the interval between the registration and the survey, the more differences there will be between the register and the survey records because of deaths and migration. Extension of the interval also increases the chance of omission in the survey due to memory lapse and name changes. On the other hand, the events reported in the survey for the time period immediately preceding will certainly exceed those registered because of the time lag in registration. A statutory time period is allowed by law but this registration time period is seldom strictly adhered to. In practice, the period is often as much as a year.

88. The most effective timing of the survey should, therefore, provide the most complete register and survey list for matching, taking into account the statutory and actual period during which registrations of birth and death should and are usually made and also the factor of memory lapse. The exact period will need to be decided for each country but a period of six to twelve months is indicated.

(c) Type of sample

89. The field check on the completeness of civil registers should be carried out in a probability sample of civil registration areas, small enough to minimize the cost of the operation but large enough to give an evaluation for determination of the degree of completeness for the parts of the country where civil registration is compulsory.

90. The importance of designing a probability sample for checking purposes is further emphasized by the fact that the field check will be the source of demographic information independent of the registers and these quantitative results must originate in a random sample rather than a purposive one in order that they may be used as valid estimates of the true birth and death rates.

(d) Sampling unit

91. In order that the records from the survey can be matched to records in the civil register, it is imperative that the ultimate sampling unit be the registration district. Thus the sample will be chosen from a list of basic

registration districts, and every household in the district will be canvassed. Since the ultimate objective is to match to the civil registers, no other type of sample will suffice. The survey area must coincide with a defined portion of the registration area. In Africa, this will probably result in the sampling unit being a village.

(e) Size of the sample

92. The size of the sample depends on the nature of the inquiry, and the length of the period of observation. The longer the observation period, the smaller the sample may be, but in deciding the balance between these two elements, account must be taken of the potential increase in non-sampling errors when the period of reporting is extended unduly.

(f) Interviewers

93. As is the case in any survey, the accuracy of the results depends in large part on the skill and ingenuity of the interviewers. For obtaining information on births and deaths, the corps of interviewers should be selected and trained with care. They will need to be aware of customs, language, environment and taboos of the population they will interview. They should also belong, if possible, to the same ethnic group as the persons being interviewed. In some areas, it may be necessary to employ only women interviewers inasmuch as in a survey of the type envisioned responses will normally be sought from the women of the household, as well as from the head.

(g) Informant

94. Information on births must be obtained from each woman in the household who is in the child-bearing age. This is important because the nominal head of the household will not necessarily have knowledge of births which may have occurred to women members who have joined the household within the period under observation. Questions on the number of births will also be put to the head of the household, but only for checking purposes.

95. Information on the deaths which have occurred among the household members as constituted twelve months previously will normally be obtained from the head of the household. However, in the case of any member reported as "widowed" a question should be put concerning the circumstances of the death of the spouse, so as to make sure it has been reported if appropriate to the survey.

(h) Contents of the test schedule

96. The questions to be asked in a field survey designed to check civil registers should include information which will identify the event in the register. Items required for this purpose in case of a birth include the following:

Name of child
Name of father
Name of mother
Age of father
Age of mother
Place of birth
Place of registration
Date of birth
Date of registration
Sex of child
Ethnic characteristic of child
Is child still living

97. For a death the following items of information must be available:

Name of decedent
Age of decedent
Sex of decedent
Date of death
Place of death
Date of registration
Place of registration
For infants: Name of father
Name of mother

98. The items set forth above are necessary for matching reports of birth and death to the civil registers because the unit of matching in this case is an individual rather than a household as it was in the case of the post-censal field check. Since names change, this piece of information will not be unique and must be supplemented by other identifying data. Information on the place where the birth and death occurred is especially important because of the fact that

populations are mobile and vital events are normally registered in the place where they occurred. Hence, some births and deaths reported in the survey may refer to infants and deceased persons whose birth and death were recorded in a registration district outside the one chosen in the sample. Similarly, some registered events will not turn up in the survey because the household involved is now living in another registration district. Unless a national index of births and deaths is available, these will have to be excluded from the test results before any conclusions are drawn from them.

99. Each of the substantive items of information included could be checked for accuracy when the match is made to the civil registers.

(i) Testing procedures

100. When the completed schedules from the survey are assembled, the next step is to search for the events recorded in the relevant civil register. Sample survey events which occurred outside the civil registration district being tested should first be excluded; the remaining events, which allegedly occurred and were registered in the sample civil registration district, would then be sought in the official register, using the items of information recorded for identification of a match. As in any matching process, rules will need to be established to determine what constitutes a "match" because the reliability of the results depends in large part on the precision of the matching operation, but these will need to be set up in accord with local conditions.

101. The results of the matching operation should yield three groups of certificates comparable to those mentioned in para. 18: (1) matches; (2) births or deaths found in the civil registers but not reported in the survey; and (3) births or deaths reported in survey but not found in the registers. A fourth group of events which escaped both registration and surveying is assumed. Events in groups two and three are then further investigated to eliminate any which should not be considered as a mismatch. Group two of births is checked to the death register to take account of liveborn infants who died in the interval since registration. For deaths not turned up in the survey, attempts are made to determine if the death eliminated or removed the corresponding household from the district, thus making it impossible for the survey to reflect the death. When the cases in groups two and three have been minimized by procedural checking, the percentages derived from the

three groups which remain may be used to estimate under- or over-registration of births or deaths. The survey results combined with the register may also provide an estimate of the rates of natality and mortality.

102. The matching of a survey record to the birth and death register will be even more difficult than the identification of the post-censal field check records. This is so because the search is not limited to one enumeration district within which every household and individual is supposed to have been enumerated. It involves events which are not common to all households, information about which has been recorded in completely different circumstances with different informants in most cases, and different types of interviewers in all cases. The problems of matching are thereby intensified. For this reason it may not be possible to match records and in this case, the test will be reduced to a comparison of the birth and death rates derived from the sample survey with those obtained from the civil registers. This minimum may show a deficiency in the registers but it provides no information on the reasons for the difference observed, and thus does not indicate the possible means for improvement in the registers.

103. In the matter of response error as related to various questions, comparison of the responses made to a number of questions will be automatically made in the process of matching, since a number of important items are used for identification. This verification will also be eliminated if matching cannot be carried out.

5. Limitations of the field check of vital statistics

104. Despite its apparent simplicity, the survey method of generating an independent list of births or deaths for matching to the civil registers in order to evaluate their completeness is subject to many sources of error. The overriding disadvantages are those of any household survey, namely dependence (1) on the willingness of the informant to give information; (2) on his knowledge of the events; and (3) on his ability to remember well enough to place the event correctly in time and space.

105. Willingness to divulge information is especially important in connexion with death surveys. There has long been a feeling that, even in modern societies, there is a reluctance to mention or discuss a death which had occurred in the household and hence there would be considerable deliberate under-reporting. Indications from

small-scale experimental inquiries in industrialized countries have tended to disprove this thesis but it cannot be denied that in some societies there are likely to be strong superstitions which prevent the revealing of a death, especially of young and unnamed children. This is one of the reasons why the list of deaths generated by a survey will often be only a partial list, as is evident from the low death rates estimated by this technique in countries where civil registration of death is not established. However, it is likely that reluctance to mention a death to an interviewer would be correlated with failure to register it with the proper authorities. Hence, this particular disadvantage of the survey method is probably more damaging to its use as a means of measuring mortality levels, where civil registration does not exist, than it is in connexion with checking completeness of the death registers. These unreported deaths would likely form part of the hypothetical "not reported/not registered" fourth group mentioned above which may constitute a larger proportion of the total than is generally assumed. Unfortunately, there is no perfectly reliable technique for estimating the size of this group.

106. Even if the informant has no basic inhibitions about discussing the birth or death, the success of the survey technique is dependent on his ability to remember the event and to place it correctly in time and space. It is probably true that memory lapse, especially memory of events in their proper time, accounts for most of the omissions in survey results. One remedy might be to shorten the time period covered by the inquiry, but when a household demographic survey is conducted on a sample basis, either the sample size has to be increased unduly or the period of inquiry must be kept long in order that an adequate number of births and deaths may fall in the sample. Increasing the size of the sample would increase costs; prolongation of the time period for reporting tends to increase memory lapse and so thereby decrease precision.

107. Failure to place the birth or death correctly in time will not completely invalidate this technique for checking purposes. It will increase the difficulty of locating the event in the register and, if the survey-reported events are used as measures of natality and mortality, it may also have the effect of inflating the rates for the time period in question. But failure of the survey to place the birth or death correctly in space presents a more serious problem. In a

demographic survey, the informant is asked to give information about births and deaths which occurred among residents of the household, while the civil register, on the other hand, records births and deaths which occurred in the area, irrespective of habitual residence. Failure to obtain information on the place where the survey-reported birth or death occurred may result in non-productive searching of the registers and consequent understatement of registration completeness. This is so because if a birth or death occurred outside the jurisdiction of the register being examined, it would normally have been registered in another area. Similarly, failure to find a matching survey-reported birth or death for each event on the registers may be due to the fact that the mother or the decedent, as the case may be, was a non-resident or a prior resident of the area. Either result will to some extent impair the utility of the name-matching technique to measure the quantitative completeness of coverage of birth and death registers.

108. A more important problem in using the household survey technique in measuring completeness of death registration is that often, the occurrence of the death of an individual eliminates the single-person household which he constituted or breaks up the household of which he was head. In the first, there remains no one who has knowledge of the fact to report it in a survey; in the second case, there is the likelihood that the broken family group will be dispersed outside the registration area under study and the death will not be reported in the survey.

6. Examples of field checks of vital statistics from civil registers

109. So far as is known, the only example of the use of the results of a sample survey of births and deaths to check a civil register (and also of civil registration data to check the results of the sample survey) is that of the Mysore experience[13/] and this was limited in coverage to selected areas of Mysore State, India, because it was a pilot study. The objectives of the experiment were inter alia to indicate the levels of the birth and death rates, and to test a method of obtaining data on births and deaths; but, in the course of this, the data from the survey were cross-checked with independent data to assess accuracy and one of the independent sources was the civil register. For the results of the experiment, reference should be made to the Report.[14/]

13/ The Mysore Population Study - A Co-operative Project of the United Nations and the Government of India. Document ST/SOA/SERIES A/34 (Sales No.61.XIII. 3).

14/ Ibid.

110. Field checks which made use of the census of population to gather independent data on the occurrence of births and deaths include the 1950 tests of birth registration in the United States and the 1953 test of vital registration in Ceylon. No such test has been carried out in Africa.

111. The use of the census of population to evaluate the accuracy of response to items occurring in the civil registers of birth and death is exemplified by the tests undertaken by the United States of America and United Kingdom in connexion with the 1950 and 1951 censuses of population respectively. In the United States of America, one test was designed to measure comparability between the reports on occupation in the death register and in the census, in the course of which the accuracy of the age reporting was evaluated.[15/] Another United States test was a comparison of the birth register with information obtained at the census to measure the accuracy with which information on "birth order", "age of mother", "race" and "place of residence" was recorded. The United Kingdom test was an attempt to match selected information for each of 9,864 deaths registered between 1 and 7 May 1951 to that on schedules from the 1 April 1951 population census.[16/] The item of major interest was occupation, although reporting of age and marital status were also the object of study.

C. Direct evaluation of household sample survey data on births and deaths

112. Since household surveys are the main source of data on births and deaths at present in a number of African countries where comprehensive systems of registration of vital events are lacking, it is a matter of special importance in this region to develop adequate methods for direct evaluation of birth and death data obtained from such surveys. Where a registration system is in operation, even if it covers only a part of the country or functions imperfectly, direct checks by matching registration lists with household survey returns along the general methodological lines indicated above may be useful in evaluating the survey

15/ The Comparability of Reports on Occupation from Vital Records and the 1950 Census. Vital Statistics-Special Reports, Vol. 53, No. 1, June 1961. U.S. Department of Health, Education and Welfare, Washington 25, D.C.

16/ Census 1951. England and Wales; General Report. General Register Office, London; Her Majesty's Stationery Office, 1958.

/...

data. Where such registration records are lacking or inadequate for the purpose, it is important to consider other possible means of evaluating the survey data. Among the possible means might be a sample field check, perhaps built into the over-all plan of the main survey, and general principles similar, with necessary modifications, to those indicated above for field checks on completeness of enumeration of population and accuracy of reporting on characteristics in censuses. As yet, no such direct evaluations of sample survey data on births and deaths appear to have been attempted in Africa. This is an important field for methodological development and experimentation.
