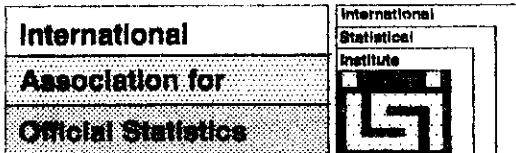


SS449 ✓



JOINT IAOS/AFSA CONFERENCE

IAOS/AFSA/IS.4/10

Addis Ababa, 22-24 May 1995

*Managing Development in the 1990s
and Beyond: New Trends In Statistics*

INVITED SESSION 4: DATA FOR DIVERSE AREAS; THE GEOGRAPHIC DIMENSION IN DATA

NEED FOR DATA THAT IGNORE BOUNDARIES
DEVELOPMENT IN SMALL AREA ESTIMATION AND THE PROBLEMS OF
PROTECTING CONFIDENTIALITY

By

Asmerom KIDANE

Department of Statistics
Addis Ababa University
Addis Ababa, Ethiopia

**ADDIS ABABA
MAY 1995**

Joint IAOS/AFSA Conference
Addis Ababa, Ethiopia 22-24 May 1995

Need for Data that ignore boundaries
Development in small area estimation and the problems of
protecting confidentiality

By

Asmerom Kidane
Department of Statistics
Addis Ababa University
Addis Ababa, Ethiopia

Abstract

The paper summarizes problems encountered in designing surveys among nomadic populations. These populations are always mobile and they are difficult to trace. Because of this there is the difficulty of obtaining and preparing sampling frames. The sampling method also becomes complex. In this paper we identify possible solutions to such problems; we also consider the advantages and disadvantages of the methods suggested. We conclude by suggesting that more reliable results may be obtained from survey data provided that information obtained from individuals or households is kept confidential and reported only in broad generalities.

1. Introduction

In many African countries there are non sedentary nomadic population and conducting surveys among these groups is proving to be difficult. The difficulties are above and beyond those problem that prevail in many African countries. The problem of conducting scientific survey in nomadic areas may be due to

- i. Unavailability as well as the difficulty of formulating sampling frame.
- ii. Lack of cartographic or mapping material
- iii. Even if a cartographic and mapping materials are available nomadic population are so mobile that preparation of Enumeration Areas (AEs) are likely to yield marginal results.
- iv. In nomadic areas there are few major urban centres with fewer satellite towns. Most of the urban dwellers in nomadic areas become sedentary and may not represent the entire nomadic population

2. Sampling Frame

A major bottleneck in conducting population or household survey among the nomadic population is lack of appropriate sampling frame. Little or non is available and at the same time it is almost impossible to prepare one within a short time and with relative ease. Sample frame may be available for purposes of population census. But as Kirigiera (1982) has pointed out such frames become relatively obsolete as a result of rural urban migration. In nomadic areas sampling frames designed for purposes of census or survey have never been reliable.

Tax records may be another source of sampling Frame. Nomadic population pay some form of tax. The taxes are usually not based on land ownership, but on ownership of cattle camels and other animals. To try to use tax records as a frame could significantly

underestimate the population; population below and above certain age groups may not be represented at all. even within certain age groups the frame may not be reliable as some households may not own cattle and may not be in the records. The use of tax records as sampling frame depends on the subject matter under investigation. For example for demographic surveys where a household or a woman is a unit of observation tax records cannot and should not be used as a sampling frame. On the other hand if the purpose of the survey is meant to assess the extent of wealth (cattle) ownership, tax records could be used as a frame.

A particular area may be classified into different administrative districts and information may be sought from each district; this information could be another source of sampling frame. Again because of the mobile nature of the population the use of district level information as sampling frame may not be satisfactory. For nomadic population both administrative district as well as international boundary do not have any meaning. Moreover administrative districts in some African countries seem to be under perpetual changes.

Other sources of sampling frame may also be considered. These include information from Land Commission; the latter is expected to have a list of land owners by various categories. Most nomadic population own land not privately but communally; the latter may be for grazing purposes only. Naturally this information cannot be a source of Sampling Frame.

The above discussion suggest that alternate method of survey sampling ought to be designed for nomadic population. This is briefly considered below.

3. Alternate Methods

It will be inappropriate and probably misleading to prescribe a standard procedure of preparing sampling frame or to choose

appropriate sampling technique that is applicable to all nomadic population. Some nomads live in different ecological zones while others live in different topographical setting. Some nomads are organized on a clan or tribal basis while others are simply individuals or households. In some nomadic communities only a household head may be mobile while other members are not; in other areas the entire clan or tribe may be on the move. Some nomads are on the move only during a particular season while others move throughout the year. The method of preparing a sampling frame as well as the method of sampling suggested below will be applicable to nomads who are primarily cattle herders, nomads who are on the move in search of water and green pasture and where the entire household or clan is on the move. These category of nomads make up the majority of nomadic population in African countries.

The suggested methods of identifying such nomads for purposes of either preparing a sampling frame, for choosing appropriate sampling methods or direct selection can be one or some or all of the following combinations:

- a. Water Point Method
- b. Market Place of Market day
- c. Use of Clan or tribal chiefs

Other methods may also be suggested when appropriate and feasible.

When nomads are on the move in search of water or green pasture, they usually camp at lake shores, river basin, desert oasis or other water points. Once they arrive at such points they usually stay for several days. Survey statisticians may use this opportunity to prepare sampling frames, to identify the appropriate method and to select respondents and conduct interviewers. This is usually cost effective and time saving procedure . Alternately one may have to use market places or market days that take place regularly - usually once or twice a week. Both water point and

market place/day have three possible disadvantages namely

- a) Incompleteness of coverage
- b) Age and Sex Selective Representation
- c) Possible duplication

The first disadvantage is straight foreword. Not all nomads arrive at a water point or a market place. Thus the method may not be satisfactory. The second item suggests that a population of a particular age group or a given sex may predominate. This problem may be solved by making appropriate adjustment and weighted selection method such as probability Proportional to size (PPS). For example in any society the ratio of males to females is around 103 males for every 100 females. This ratio may not be observed at the sites; for example there may be 150 males for every 100 females at a particular site. In such a case three females may be selected for every two males. Similar procedures may be followed if there is over representation of certain age group such as those between 20 and 30 years. The age distribution of a nomadic population is similar to that of other LDCs and follows a certain known pattern; in other words over representation of 20 to 30 years age group is at the expense of those less than 20 or more than 30 years. Thus the selection process should take this into consideration. The above adjustments are made whenever the unit of observation is the individual. Similar adjustments may also be made if the unit of observation is a household. The average size of household may be about five; if the majority of households, present at the sites is less than five then more could be selected from those more than five etc.

When one uses water point or market place methods and when the survey covers several days and several sites the possibility of duplication of respondents may cause a serious problem. Survey statisticians should develop methods of avoiding duplication.

The third method of estimating the population or household size of

nomadic population in a particular area is through clan chiefs. This is potentially a promising approach the major advantage being that the total number of households are likely to be estimated and listed with relative ease. The disadvantage is that while the total size may be estimated the problem of gathering individuals or households at a particular site may not be easy; there may be a high percent of no shows. Besides, this method assumes that all nomads may be clan or tribe based; this may not always be the case.

4. Methods of Sampling

In household surveys of human population the most common and preferred method is stratification. Such stratification may be done on the basis of ecological zones, administrative districts, ethnic and other non overlapping method of stratification. In the case of nomadic population stratification may not yield the desired result. Except in the unusual case where the method of stratification is by sex or age group, other methods of stratification may not be readily available. The former is not recommended as the number of strata is only two. The suggested methods of preparing sample frame or selection are not amenable for stratification. The preferred method will thus be cluster Sampling. In other words a certain group of nomads may arrive at a water point or market centre at a particular time. This may form one cluster. The disadvantage of Cluster Sampling over stratified Random Sampling with regard to the degree of precision is known to survey statisticians, but the fact that it is cost effective and appropriate to nomadic population is self evident. Alternately all nomads at water point or at market place or through clan chiefs may be listed and Systematic Sampling could be adopted. However the inherent weakness of systematic Sampling over Cluster Sampling may result against adopting this approach. A better way of sampling would be a combination of the two. Non scientific sampling such as Purposive or Quota sampling could also be utilized if Cluster or Systematic Sampling did not yield the desired degree of precision.

5. Protection of Confidentiality

The need to protect confidentiality should be within the declaration of professional ethics as enumerated by the International Statistical Institute. The declaration which was adopted on August 21, 1985 recognizes that statisticians work within a variety of economic, cultural, legal and political setting and should be governed by customs and practices prevailing in their environment. Statisticians also work within different disciplines such as sociology, economics, medicine and another biological sciences. The latter have their own ethical declaration with respect to protecting confidentiality.

With regards to the protection of confidentiality regarding small area surveys one can identify two major obligations namely obligation to society as well obligation to the subjects under investigation. When considering the former a statistician should consider possible consequences of collecting and disseminating various types of survey related to empirical information. More specifically he should make sure that information is not misused, misinterpreted or abused. Naturally one may not be able to develop specific instruction, guideline or formula for assessing the likely risk to society of a particular statistical inquiry. Even though a society has a right to know about its overall characteristics, strengths and weaknesses from a survey data, this should not be at the expense of individual's freedom or privacy. Certain findings may affect a particular group of people. In such a case a statistician should employ various techniques so as to open the scope of the statistical inquiry and communicate those benefits that occur to as many members as possible. Statisticians should also protect themselves from undue influence, pressure not only from their own finding but also through indirect techniques such as withholding some information, postponing the official release of findings. In other words statisticians should uphold their professional integrity and maintain their objectivity without fear. If this is upheld then the protection of confidentiality of survey

results could be upheld. In survey sampling objectivity may include methods of sampling (Cluster versus Stratified) data collection, selection of appropriate method of analysis as well as the interpretation of the results. In some instances findings may depend on the method used for analysis. The involvement of statisticians in the interpretation of results should be as minimal and as empirical as possible. They may not be involved in analyzing the policy implication of their empirical finding.

The second issue with regard to the protection of confidentiality deals with the obligation to subjects. It should be noted that statisticians have no special entitlements to conduct any empirical inquiry. In other words the pursuit of empirical information cannot be taken as a sole justification for overriding existing social and cultural values. To avoid unnecessary and unwarranted intrusion survey investigation should be conducted after getting the consent of members or their representatives, in other words contact may have to be sought with subjects before undertaking a particular inquiry. Also information should not be sought of from third party without the consent of the subjects under study.

The acquisition of informed consent depends on the amount of information required. At one extreme there may be the need to require comprehensive knowledge from the subjects; this is both time consuming and may involve questions that are too personal. On the other hand the inquiry may be too general and obvious. In general the information sought should be made clear to the subjects or their representatives. In the declaration of professional ethics of statisticians, the following items regarding the survey may have to be explained to the subjects or their representatives.

- i. Purpose of the study
- ii. Identity of the funder
- iii. Anticipated uses of the data
- iv. Identity of the interviewer

- v. How the subject has been chosen
- vi. Subject's role in study
- vii. Possible harm or discomfort to subject
- viii. Degree of anonymity and confidentiality
- ix. Data storage arrangements or degree of security
- x. Procedures of and duration of interview
- xi. Whether participation is voluntary or compulsory.

In general the acceptability of survey data depends not only on technical considerations but on the willingness of statisticians to respect the subject as well as to acquire their confidence.

6. Some Concluding Remarks

The preceding discussion shows that conducting surveys among nomadic population is quite a difficult undertaking. Before one is to conduct survey among these group the method of preparing sampling frame should be considered. The method depends on the subject matter under investigation. In other words methods that appropriate for one topic may not be appropriate for another. The methods suggested earlier are by no means ideal; a better method can be developed only after carefully studying the customs, practices and the general way of life of the survey population- the nomads. If one clearly knows the nomadic population then a better technique of preparing a sampling frame, an appropriate method of sampling could be prescribed. Like in other societies protecting the confidentiality of respondents is a prerequisite for a successful survey investigation in nomadic areas.

7. References

1. International Statistical Institute (1985) Declaration of Professional Ethics.
2. Kidane, A. Sampling Errors from Socio Economic Baseline Survey SSRR V1:1pp 69-78
3. Kireggera, B. (1982) On Sampling Frames in African Census and Surveys *The Statistician* 31:2 pp 153-168.

4. Kish, L. (1965) Survey Sampling, John Wiley, New York
5. O'Muircheartaigh, C.A. (1977) Sample Design for Lesotho
WFS/TECH.666
6. United Nations Economic Commission for Africa (1977) Study on
Special Techniques for Enumeration of Nomads in
Africa. Addis Ababa.