

Role of Subjectivity in Preparation of Sampling Frames for Data Collection

¹Sadhan Kumar Pathak

Article history

Submitted: July 10, 2021

Revised: Oct 5, 2021

Accepted: December 2, 2021

Abstract:

Collection of data from the field on various socio-economic aspects is consisting of several stages. Stratification of the population into required numbers of stratum (if it is heterogeneous) and the selection of sample units from each stratum on the basis of various random sampling method are precondition for scientific collection of data from a particular sample unit and therefore preparation of a scientific Sampling Frame which is the basis of selection of required number of sample units is of utmost importance. The established scientific procedure needs to be followed strictly while preparing the Sampling Frame and any subjectivity may enhance the sampling error to a great extent. An attempt has been made in this study to highlight the issues relating to the preparation of sampling frame in a Multistage Stratified Socio-Economic Survey and the role of subjectivity thereof.

Keywords: Sampling Frame, Subjectivity, Sampling Error, Sample Statistic

Corresponding author

*¹Research Scholar Department of Economics, The University of Burdwan West Bengal
email: sadhanpathak.sp@gmail.com*

Introduction:

The purpose of data collection is to estimate the value of a particular parameter of the population on the basis of the sample statistic and therefore the samples which are selected needs to be the true representative of the population under study otherwise the sample statistic would not be the good estimator of the population. In order to draw representative samples from a population, specific scientific

procedures need to be followed. However, it has been observed that in many cases, during survey, subjectivity of the Investigators/Enumerators replaces the established scientific methods to a considerable extent in respect of preparation of a sampling frame. The subjectivity leads to preparation of a faulty sample frame and eventually the main purpose of drawing representative samples is not ensured which in turn produces a non-representative estimation of population parameter.

Objective:

The objective of this study is to look into the extent of deviation from the standard method of preparation of a sampling frame on the basis of actual field experiences from 50 Investigators/Enumerators engaged in various field surveys.

Methodology:

The source of data for our study was primary in nature. Data were collected from Investigators/ Enumerators engaged in data collection for any Government/Private Research work. Data so collected examined and the presence of subjectivity has been analysed.

A. The Preparation of a Sampling Frame:

For any survey the first step is to prepare the sampling frame. The sampling frame needs to be prepared on the basis of established scientific procedure and should not suffer from any subjective bias. It should be systematic and exhaustive so that any number of samples drawn from this frame should be of proper representative. When the population size is large it needs to be divided into required number of units and thereafter required number of units to be selected randomly and then from these selected first stage units final sample units to be selected for our purpose.

1. Formation and Selection of Sub-Units:**a. The Scientific Method:**

We need to list of all the houses and households who are residing in the Sample First Stage Unit (FSU). Some household information like household size, usual monthly consumer expenditure of a household, educational qualifications, land possessed and value of agricultural production of a household, etc. depending on the purpose of the study are collected in this schedule.

This secondary information is used for grouping the households into different Second-Stage-Strata (SSS). The sampling frames for selection of households are

then prepared and details of the selection of sample households are recorded in the predesigned schedule. When Sub-Units (SU) are needed to be prepared, specially relating to the formation of SUs and identification of the selected SU are also recorded in this schedule.

Under National Sample Survey field survey Sub-unit formation are resorted to in the villages with census population more than or equal to 1000 (One thousand). A village is divided into a certain number of sub-units (SUs) depending on the number of populations.

In a larger village, there usually exist a number of localities where the houses of the village are concentrated. These are called 'hamlets'. In case there are no such hamlets in the village, the census sub-divisions of the village may be treated as 'hamlets'. Large hamlets may be divided artificially to achieve equal population figure for the purpose of SU formation. The procedure for formation of SU is described below:

(i) Identification of the hamlets as mentioned above (ii) Estimation of approximate present population of each hamlet. (iii) Drawing of a notional map showing the approximate location of the hamlets and numbering them in a serpentine order starting from the northwest corner and proceeding southwards. While drawing this map, uninhabited area of the village needs to be included as part of nearby hamlet, so that no area of the village is left out. The boundaries of the hamlets may be demarcated with the help of landmarks like canals, railway lines, roads, etc., so that it would be easier to identify and locate the geographical boundaries of the Subunits to be prepared in the village. (iv) Listing of the hamlets in the order of their numbering and indicate the present population figure in terms of percentages. (v) Grouping of the hamlets into numbers of SUs. The criteria to be adopted for SU formation are equality of population figure and geographical contiguity. In case there is a conflict between the two aspects, geographical contiguity needs to be given priority. However, difference between the population of the smallest and the largest SU formed is to be kept at minimum. Formation of SUs in urban areas: If an UFS block has 250 or more households then it is necessary to form SUs. Procedure for formation of SU is the same as mentioned already.

b. The Subjectivity in Sub-Units Formation and Selection:

The standard scientific and established method of various stages of Sub-Units formation has been discussed above for a First Stage Unit say, village/block but in practice the procedures as laid above in not followed as revealed from our study.

The deviations take place in the following areas due to application of subjectivity of the Investigator/Enumerators into the process.

- **Estimation of Population of the First Stage Unit:** The decision whether sub-units to be formed or not depends on the population in the village/block and therefore it is very important to estimate the current population. The census figure along with information from local panchayat/urban bodies is necessary to get the current population. However, it has been found that current population is estimated on the basis of census population plus 1.5% to 2.00% growth without taking into consideration of the actual population growth of that particular village/urban block. This practice leads to incorrect estimation of current population.
- **Subjective Hamlet Formation:** In a large village, there exist usually a few localities or pockets where the houses of the village tend to cluster together. These are called 'hamlets'. The identification of the hamlets needs to be made as per established procedure. However, in practice, it has been noticed that identification of the hamlets is made on the basis of subjective opinion. Sufficient information is not taken as well as sufficient time is not given to make proper identification of the hamlets.
- **Subjective estimation of Population per Hamlet:** In order to form sub-units of the First Stage Unit (FSU) i.e. (village/Urban Block) current population of the FSU needs to be decided. Latest census figure along with information from local knowledgeable person(s) must be consulted in order to get the current population figure of FSU.

However, it has been noticed in practice that population of the FSU is decided on the basis of some percentage increase over the latest population census. The population of a hamlet is decided on the basis of equal distribution of the FSU among its number of hamlets and then population of a sub-unit is decided on the basis of more or less equitable distribution of population among the sub-units. The real population distribution among the sub-units therefore is not taken into consideration.

- **Serpentine Numbering of the Sub Units not Followed:** In order to give number of each hamlet and sub-units serpentine numbering system needs to be followed so that randomness of the numbering is maintained. However, in practice it is noticed that number is given more or less left to right direction following a straight line. Even in some cases number is given after the random selection of required number of units from the units so that the distance between two or three selected units are kept at minimum.

- **Subjective Selection of Sub-Units:** Once the sub-units are formed within the population next task is to select the two or three sub-units from those on the basis of random sampling. Sub-units so selected are used for preparation of sampling frames. However, it has been noticed that in most cases sub-units selected on the basis of random sampling are deliberately replaced by other sub-units. In order to avoid distance between two or more selected sub-units neighbouring units are selected deliberately and then ordering/numbering of the sub-units are manipulated to match them with the random sampling number drawn.

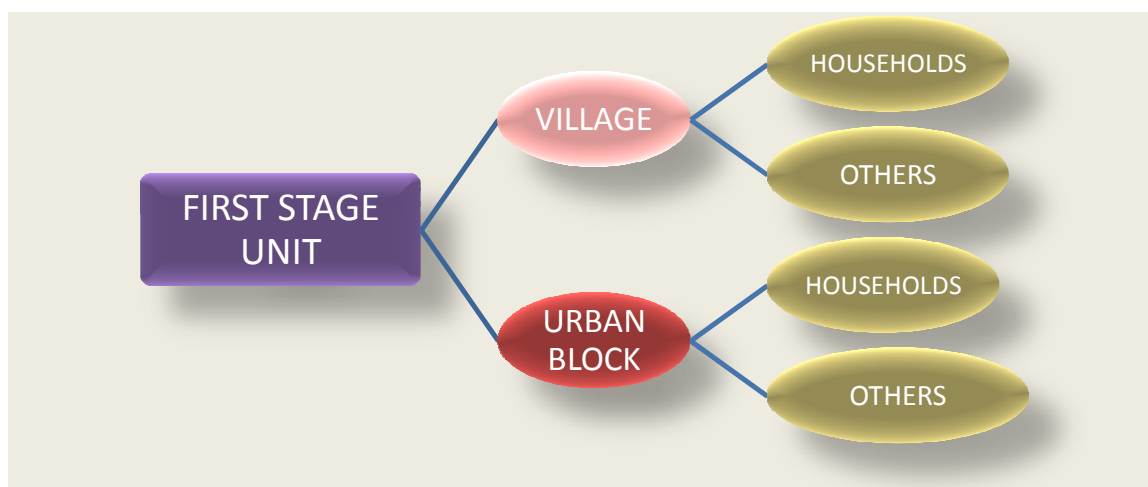


Fig I: First Stage Unit & Second Stage Unit

The Outcome: The first stage unit is the village/urban block. But due to vastness of the village/urban block it is not always feasible to survey the all the households in the village/urban block and therefore it is necessary to divide it into required number of units. Thereafter, 2 or 3 units are selected randomly and listing of all households/others second stage units are made. This listing is the sampling frame and finally we select required number of households/other units for our detail survey for getting required information for our study.

However, when selection of sub-units is not made on the basis of scientific procedure as described above but made on the basis of subjectivity of the Investigator/Enumerator the preparation of sampling frame gets deviated from the scientific procedure and outcome of the survey does not represent the true picture of the population under survey.

The Graphical Representation of Deviation: The scientific method to give number to Sub-Units is starting from the North West corner in a serpentine order as shown by the green line. The green line starts from 1 and passes through 2, 3,in a

serpentine order. However, due to subjectivity the numbering is given by the way of the line shown by orange line. When serpentine order is followed units selected differ from the subjective numbering. For instance, here unit number 1 and 9 have been randomly selected whereas unit number 4 and 6 have been selected by subjective method. It is needless to say that serpentine method is more representative than the subjective method.

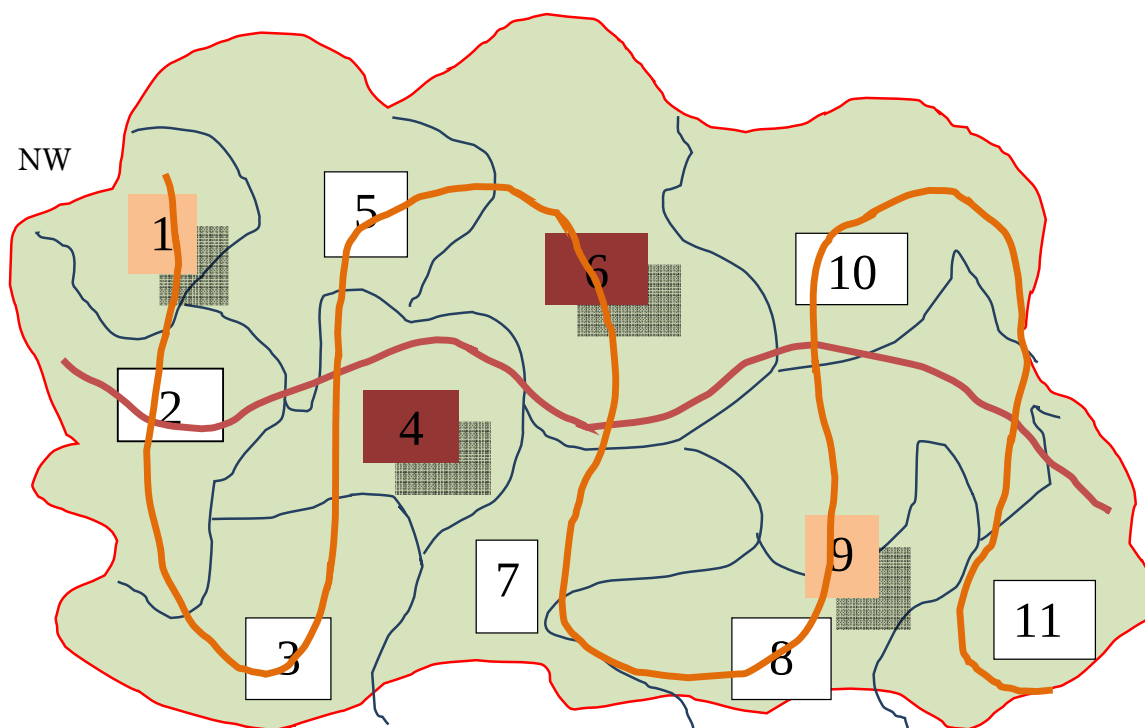


Fig II: Scientific Procedure & Subjective Procedure

Listing of Houses and Households:

Scientific Method:

It is necessary to ensure that there is no omission or duplication of any house/household. A house-to-house enquiry needs to be done to list of all the houses and households. Households, which are found locked during the time of listing (also, households which are absent in the present place for a period of less than six months during last one year) are also to be listed and included in the households frame before sample selection. After obtaining from the neighbours the details about the absentee households, attempt should be made to contact the households at the proper time.

During listing a house, the Investigator/Enumerator shall find out how many households (including the locked households) reside there and list all of them. After this, he will move to list the next house. In order to ensure complete listing of houses/households, it is better to follow a particular order for listing. The order which was followed in 2011 Population Census may be adopted, taking care that any house that has come up later is not omitted; otherwise, listing may be done in serpentine manner starting from the north-west corner and moving towards south. The names of any natural grouping of the houses like street, Mohalla, etc. and date of listing may be written at the top of the Schedule before the listing of houses starts. This will help in checking the completeness.

The 2011 Population Census house number or the number given by the local panchayat, municipality or other local bodies, may be used. The houses without such numbers need to be given a separate running serial number starting from 1 within brackets. Wherever house numbers are available, even if not for all the houses, the actual house numbers shall be recorded and no bracket to be used here. After listing of all the households associated with a house, the next house shall be listed. If the house is used solely for non-residential purposes or is vacant, the purpose for which it is used needs to be written across the line, e.g., temple, vacant structure, club, etc. For family living under a tree/bridge/ etc. (i.e., without any house), a ‘-’ may be put in the place of the Schedule.

The Subjectivity in Listing of Houses and Households:

- **The Omission of Houses:** One of the most common examples from the part of an Investigator/Enumerator is that they often omit a number of houses intentionally. For instance, if a house is located at a distant place, then they do not list that house in order to save time. Houses which are located in a row are covered but far off houses though fall in the sample sub-units are not covered intentionally.
- **The Households within the House Omission:** This is very much common. We know that household is defined as a person or a group of persons who takes their meal from a common kitchen. One house is consisting of one or more than one household. As we know that in multi storied Building many households reside. In rural areas also more than one household reside in the same house. Therefore, it is very important to list all the households reside there. However, in practice it is often observed that all the households within the same house are not listed due to improper investigation/probing on the part of the Investigator/Enumerator.
- **Locked Houses/Households not revisited:** This may so happen that some house/households are found locked at the time of listing. The

Investigator/Enumerator do not revisit those for the second time and therefore those house/households are listed.

- **Absentee Households not revisited:** There may be frequent instances of absentee households due to various reasons. Revisits needs to be made for such cases. However, in practice revisits are not generally made and information is taken from the neighboring households which often lead to incomplete information.
- **Listing of Households from a third Person:** It has also been revealed that Investigator/Enumerator makes listing of house/households with the help from a third person for the part/entire unit under survey. This is done in order to complete the survey at the cost of minimum time. This type of listing not only leads to incorrect sampling frame preparation but also strike at the very root of the purpose of the survey.

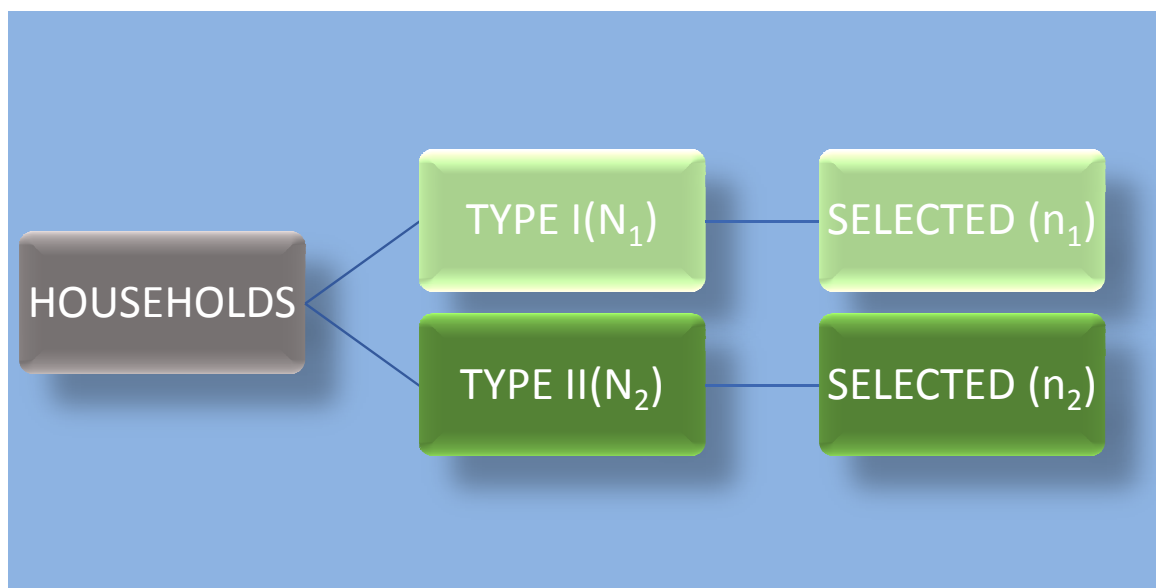
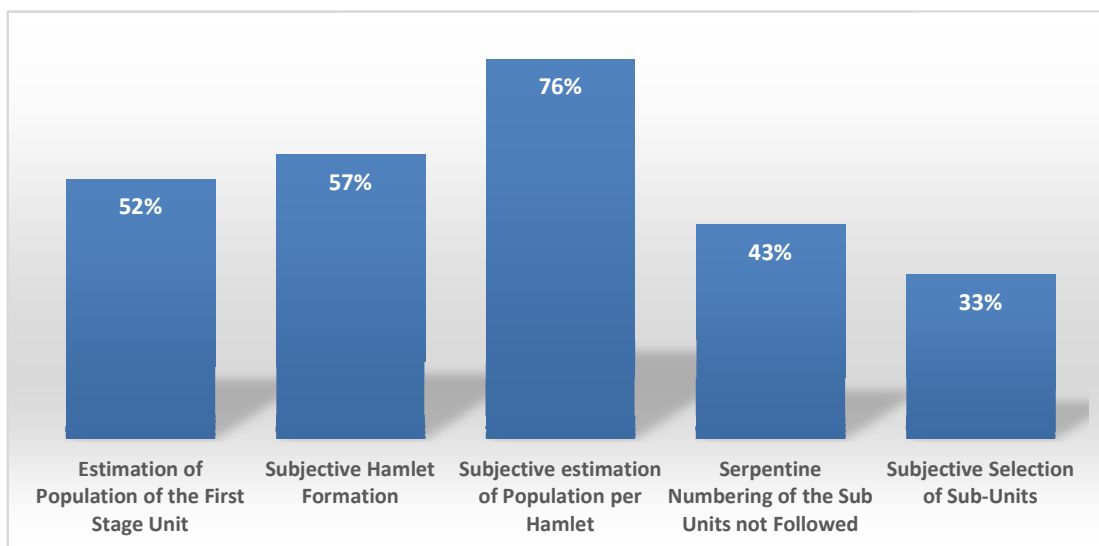


Fig III: Second Stage Stratum and Final Selection

The Outcome: When selection of the first stage unit is not proper then it will also lead to incorrect second stage units and final units. When a village is divided into number of units and out of these units 2 or 3 units are randomly taken on the basis of subjectivity then we go for listing of all households under these units. Naturally, N_1 and N_2 here will differ and this will lead to change the selection units n_1 and n_2 as well.

The degree of Deviation as revealed from our survey of the



Investigators/Enumerators:

Fig Iv: Percentage of Subjective Decision in Sub-Unit Formation and Selection
 Source: Author's own calculation based on primary data.

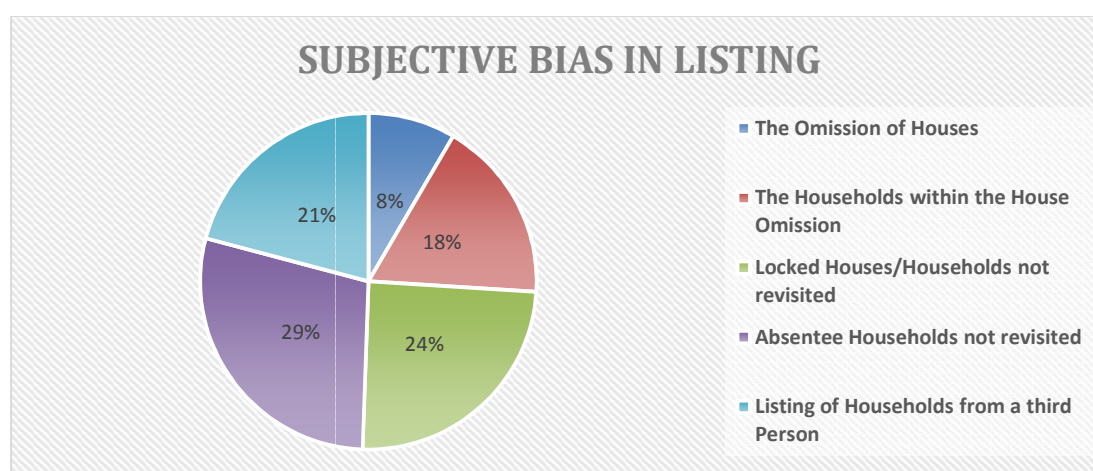


Fig IV: Percentage of Subjective Decision in Listing of House/Households

Source: Author's own calculation based on primary data.

Result:

It is observed that there was a considerable impact of subjectivity in preparation of sampling frame for a survey. The presence of subjectivity affects the preparation and deviates it from the scientific procedure and hence enhances the chances of sampling error which in turn affects the estimation of a sample statistic.

Conclusion:

The survey reveals that a considerable percentage of subjectivity is associated with the formation and selection of sub-units as well as in listing of house/households. The affect of these subjectivity on the value of sample statistic needs to be studied for a better understanding of the degree of deviation. However, we may conclude that subjective bias is noticed in the preparation of sampling frame. More is the subjectivity less is the chance of representativeness of the sample we select. The presence of subjectivity needs to be minimized for any survey otherwise it will lead to increase the sampling error and thereby the Sample Statistic will not be a true representative of the Population Parameter.

Acknowledgement

The authors are thankful to the anonymous referees of the journal for their handy suggestions to improve the quality of the article.

Declaration of Conflict of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

References:

1. Bhaumik, S K (2007): Research Methodology in Social Sciences: Methodological Issues in Field Surveys.
2. Debnath, J C (2007): Research Methodology in Social Sciences: Knowledge Sources and Research Process.
3. Hossain, A (2007): Research Methodology in Social Sciences: Social Research: The Perspectives for Theory Building.
4. Mohanty, Girisbala (2002): A Text Book of General Psychology, Kalyani Publisher.
5. NSSO, 77th Round.
6. Pathak, S K (2007): Research Methodology in Social Sciences: A Practical Guide to Data Collection: Some Issues.
7. Saravanavel, P (1991): Research Methodology