

Gender issues in the measurement of paid and unpaid work

Expert Group Meeting on
Methods for Conducting Time-Use Surveys
23-27 October 2000
New York

**Sampling issues in time use survey—
Indian experience**

Country Paper for the Expert Group Meeting on Time Use Survey to be held at New York during 23-27 October, 2000

Sampling Issues in Time Use Survey- Indian Experience

*Dr. R. N. Pandey, Joint Director,
CSO, Ministry of Statistics and Programme Implementation, New Delhi*

INTRODUCTION

To estimate properly the contribution of women to the national economy and to study the gender discrimination in household activities, a pilot Time Use Survey was conducted in 18620 households spread over six selected States, namely Haryana, Madhya Pradesh, Gujarat, Orissa, Tamil Nadu and Meghalaya during the period July, 1998 to June, 1999. As this survey was first of its kind in India and even among the developing countries, a number of sampling issues were to be resolved before launching the survey. In this paper an attempt has been made to present the details of the sampling procedures adopted in the Indian Pilot Time Use Survey. The details presented here may be useful to many other developing countries, which are planning to conduct the Time Use Survey in their own countries.

DETAILS OF SAMPLE SELECTION FOR THE TIME USE SURVEY

(a) Need for Using Sampling

As no previous survey was conducted on this topic and concepts and methodologies to be used were not firmed up, it was decided to conduct this survey on a pilot basis. However, to ensure the use of data collected in the pilot survey and to draw meaningful conclusions from the data, a proper sampling procedure was followed. The sampling procedure adopted were to ensure the reliable data at national and state level.

(b) Selection of States:

India has lot of socio-economic, demographic, geographic and cultural diversities. To ensure that all aspects of diversities are captured, six States were selected purposively to give geographical presentation to each regions of the country. Haryana, Madhya Pradesh, Gujarat, Orissa, Tamil Nadu and Meghalaya were chosen to represent northern, central, western, eastern, southern and north-eastern regions respectively. The location of the States selected for the survey are shown on the Indian map in Annex- I.

(c) Distribution of Sample in the Selected States:

The total sample size of 18600 households was first distributed among the States in proportion to the total number of estimated households, as per the 1993-94 survey of the National Sample Survey Organisation. The further distribution of the allocated

sample in the States in the districts, villages and towns were as per the following procedures,

(d) Selection of Districts

The number of districts to be selected in each State was also decided purposively. It was decided to select 1/3 of the total number of districts in five States. In the sixth State of Meghalaya, 1/2 of the districts were selected because the total number of districts were only 7 and variation was likely to be more because of its specific geographical location.

For the selection of the requisite number of districts in a state, all the districts were grouped in four strata using the criteria of density and proportion of schedule tribe population as follows:

Stratum 1. Districts with population density less than median density and proportion of schedule tribe population less than median schedule tribe population	Stratum 2. Districts with population density less than median density and proportion of schedule tribe population more than median schedule tribe population
Stratum 3. Districts with population density more than median density and proportion of schedule tribe population less than median schedule tribe population	Stratum 4. Districts with population density more than median density and proportion of schedule tribe population more than median schedule tribe population

(e) Determination of Number of First Stage Units (Villages/ Urban Blocks)

No. Of first stage units (villages and sample blocks) were determined using the initial sample size to be allocated to each state and by assuming that in each f.s.u., 12 households will be surveyed. The number of f.s.u. so arrived at was adjusted to be multiple of 8 so that at least 2 f.s.u. can be surveyed in each of the 4 sub-rounds. As the sample was selected in the form of two interpenetrating sub-sample, it was also required to select 2 f.s.u.

The no. Of districts to be selected (fixed earlier) in each state were distributed in 4 strata in proportion to the total population in each stratum. If more than one district was to be selected from a stratum then all the districts were arranged in the decreasing order of population density and proportion of tribal population. Selection of more than one district was done using circular systematic sampling with equal probability.

For getting the district – wise allocation of the f.s.u., the total f.s.u. sample was distributed among the selected districts in proportion to the population of the districts.

The rural (villages) and urban (sample blocks) sample in a particular district was worked out by distributing the total number of f.s.u. in rural and urban sample in proportion to the rural and urban population in the district.

(f) Selection of villages

All the villages in the selected district were grouped in 3 categories namely large (population above 1200), medium (population between 400 to 1200) and small (population less than 400). The total rural sample was distributed in three strata in proportion to the population in the three strata. In case any stratum was not applicable in a particular district, the allocated sample was distributed in the remaining stratum only. If more than one village was to be selected in the particular stratum, then villages were selected using circular systematic sampling with probability proportional to the population. If all the three strata were present then minimum sample size allotted in each stratum was 2.

(g) Selection of urban sample blocks

All the towns in the selected district were grouped in 3 categories namely large (population more than 2 lakhs), medium (population between 50000 to 2 lakhs) and small (population less than 50000) . The total urban sample was distributed in three strata in proportion to the population in the three stratum. In case any stratum was not applicable in a particular district, the allocated sample was distributed in the remaining stratum only .if more than one sample block was to be selected in the particular stratum , then ufs blocks in each of the towns were presented by investigator unit and ufs blocks no. The requisite number of UFS blocks were then selected by using circular systematic sampling with equal probability. If all the three strata were present then minimum sample size allocated in each stratum was 2. Due to this, in some cases, overall urban sample size allotted in a particular district might have increased.

(h) Distribution of villages/urban blocks in sub-sample and sub-rounds

The allocated sample in a particular stratum of rural or urban area of a particular district was first distributed in two interpenetrating sample of equal size i.e. both the sample were drawn independently using circular systematic sampling. The sample size for a particular sub-sample was then evenly distributed in 4 sub-rounds. If there were only 2 sample units then they were covered in two sub rounds only. But the selected sub-rounds were not contiguous to take care of the seasonality effects. In case of odd number of sample size, sub-round wise allocation was adjusted by increasing the sample size in one sub-round and decreasing in another.

(i) Selection of Households in the Selected Village/ Urban Blocks

Sampling design was so prepared to ensure the selection of all types of households viz., rich and poor in the sample. After listing all the households in the selected village, the rural households were grouped into six sub-strata viz. those owning land of 7 acres or more, 2-7 acres and less than 7 acres and those not owning land as self-

employed in non-agriculture, rural labour and others. Similarly the urban households in the selected sample blocks were grouped into six sub-strata viz., self-employed households : Monthly Per Capita Consumption Expenditure (MPCE) \leq Rs.1200, MPCE $>$ Rs.1200 : regular wage/salary households : MPCE \leq Rs.1200, MPCE $>$ Rs.1200 : casual labour households and other households. A sample of 12 households was selected from each village/urban blocks. This sample size of 12 was allocated amongst six strata in proportion to the number of total number of households in the respective sub-strata.

(j) Tabular Presentation of the Sampling Details

Details of the sampling procedure presented in the above paragraphs are presented in the following Table

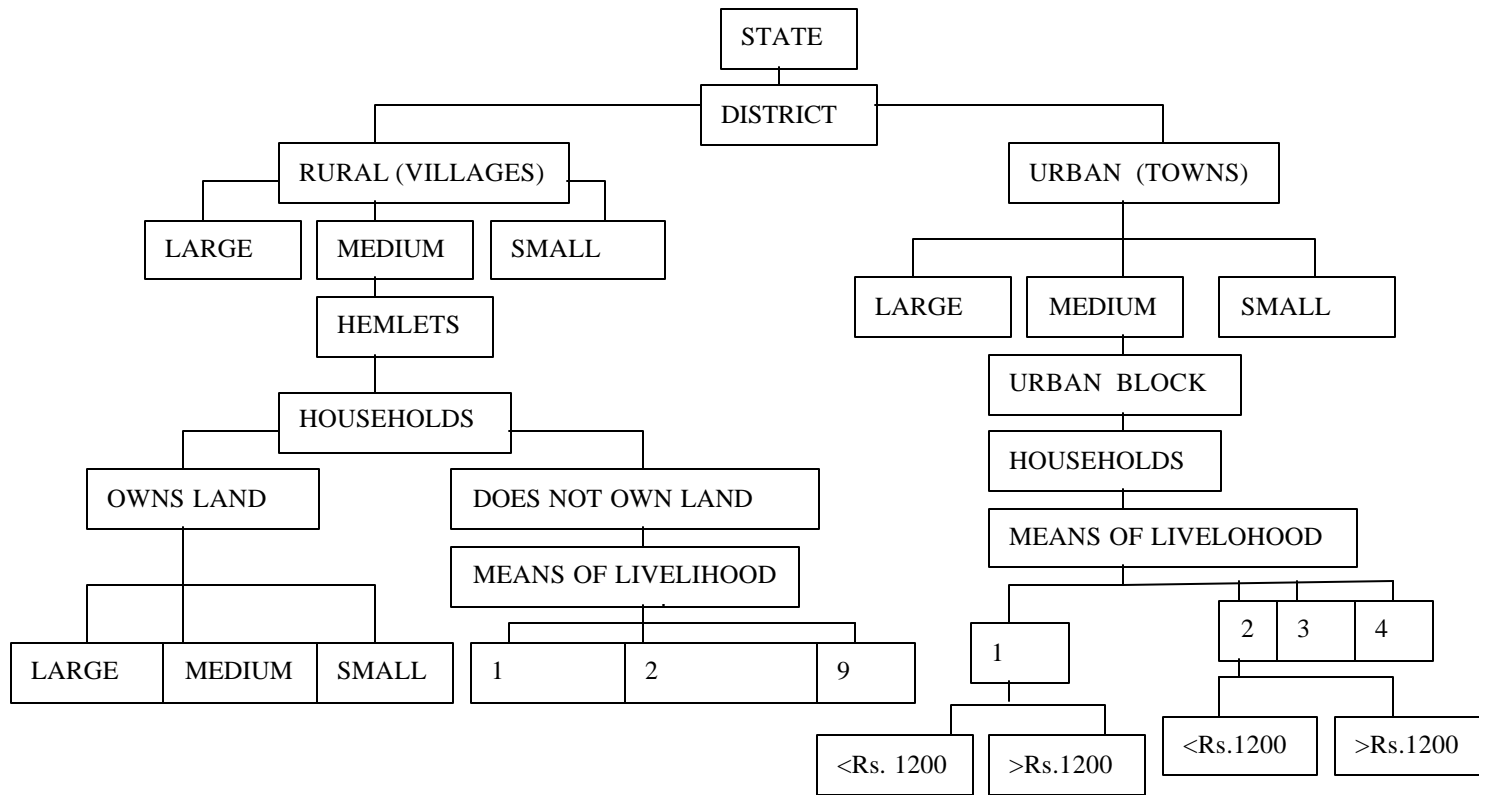
State	Total No. Of Distts.	Total sample initially allotted	No. Of distts. in the sample	No. Of first stage unit			No. Of selected house-holds		
				Rural	Urban	Total	Rural	Urban	Total
Madhya- Pradesh	45	5027	15	318	106	424	3816	1272	5088
Gujarat	19	3174	7	140	124	264	1680	1488	3168
Meghalaya	7	520	4	34	14	48	408	168	576
Tamilnadu	29	5588	10	304	168	472	3648	2016	5664
Haryana	17	1320	6	82	30	112	984	360	1344
Orissa	30	2758	10	188	46	234	2256	552	2808
Total		18387		1066	488	1554	12792	5856	18648

(k) Sampling of Days

To catch the variation in the activity pattern, time use data was collected for all the individuals aged 6 years and above for three type of days, namely normal, abnormal and weekly variant. After selecting the household, the investigator visited the selected households to find out the details of three types of date from different members of the selected households. Based on the information the investigator prepared the work programme for the next one week. If the normal date for an individual was Monday, the investigator will visit that member on Tuesday to find out his time used pattern on Monday. This arrangement worked quite well for normal days but there was some problem, particularly in the urban areas for weekly variant because in urban areas, Saturday and Sunday were generally reported as weekly variant and the investigator has to cover all the selected households on Sunday and Monday. The information was

collected within a reference period of one week for only one day each of normal day, weekly variant and abnormal days.

(k). Diagrammatic Presentation of the Sampling Procedure Adopted in the Survey



Rural

- 1 Self Employed in Non-Agriculture
- 2 Rural Labour
- 3 Others

Urban

- 1 Self Employed
- 2 Regular Wage/ Salaried
- 3 Casual Labour
- 4 Others

(I) Estimation Procedure

Notations :

- S : staratum D: total number of districts d : number of sample districts
- i: subscript for i-th sample district r/u : subscripts for rural/urban
- j : subscript for j-th class(j=1,2,3) k : subscript for k-th sample village/blocks
- m : subscript for p-th sample household z : size of a stratum x sample district x rural/urban x class
- n: number of sample villages/blocks available for tabulation
- Z: village/block size (equal to village population as per frame and z = 1 for each urban
- L: total number of hg's/sb's formed in a village/block
- H: total number of households listed
- h: number of households surveyed and available for tabulation
- y: sample value of any characteristic under estimation
- \hat{Y} : estimate of the population total of the characteristic y

FORMULA FOR ESTIMATION

$$\hat{Y}_{sr} = \frac{D_s}{d_s} \sum_{i=1}^{d_s} \sum_{j=1}^3 \left[\frac{Z_{sirj}}{n_{sirj}} \sum_{k=1}^{n_{sirj}} \frac{1}{Z_{sirjk}} \left\{ \sum_{m=1}^6 \left[\frac{h_{sirjkm}}{h_{sirjkm}} \sum_{p=1}^{h_{sirjkm}} y_{sirjkmp} \right] \right\} \right]$$

The same formula may be used to obtain \hat{Y}_{su}

$$\hat{Y}_s = \hat{Y}_{sr} + \hat{Y}_{su}$$

Estimate of total of y at the state level may be obtained by summing $\hat{Y}_s, \hat{Y}_{sr}, \hat{Y}_{su}$ as the case may be State level estimate of aggregates (Y) may be added to obtain estimate of the total of y for all states combined.

The above formulae may be applied to obtain sub-sample wise estimates first where n will denote number of sample villages / blocks available for tabulation in the concerned sub-sample .Pooled estimate based on both the sub-samples may be obtained as the simple average of the two sub-sample estimates.Ratios of the form R=Y/X may be estimate at the last stage at any desired level by taking the ratio of estimate of Y to the estimate of X.

Standard Error Estimates :

The standard error estimates may be calculated on the basis of sub-sample wise estimates of stratum totals. The estimates of the variance of Y and R are given by :

$$\widehat{V}(Y) = \frac{1}{4} \sum_s (\widehat{Y}_{s1} - \widehat{Y}_{s2})^2$$

and

$$\widehat{V}(R) = \frac{1}{4\widehat{X}^2} \left[\sum_s (\widehat{Y}_{s1} - \widehat{Y}_{s2})^2 - 2\widehat{R}(\widehat{Y}_{s1} - \widehat{Y}_{s2})(\widehat{X}_{s1} - \widehat{X}_{s2}) + \widehat{R}^2(\widehat{X}_{s1} - \widehat{X}_{s2})^2 \right]$$

where suffixes 1 and 2 denote the sub-samples 1 and 2 respectively.

These formulae may be applied to obtain estimates of variances for rural as well as urban areas also if required by using rural and urban estimates of Y, X and R.

Standard errors of estimates may be obtained by taking square roots of the estimated variances. Relative standard error of an estimate is the standard error expressed as percentage of the estimate.