

On Utilisation of Sample Registration Areas for Improving the Quality of Agricultural and Allied Statistics¹

K.C. Seal²

SUMMARY

Various statistical schemes designed to strengthen the quality and timeliness of agricultural and allied statistics in the country have been discussed. Suggestions are given to strengthen the current national data system by introducing regular collection of longitudinal data generated through Sample Registration Area (SRA) pertaining to key agro-socio-economic variables in a cost effective manner. Intensive use of SRA for collection of 'quality data' and organisational set up for the purpose are discussed in detail.

Key words : Longitudinal data, Agro-socio-economic variables, Sample Registration Areas (SRA), Quality data.

1. Introduction

Soon after gaining independence India made rapid stride in revamping the national statistical system to ensure proper formulation of social and economic development plans covering various sectors of the economy. Prof. P.C. Mahalanobis, who was appointed as Honorary Statistical Adviser and Member, Planning Commission implemented a large number of new country-wide statistical schemes which helped the country to improve the statistical data base. A number of important research and official institutions were established to carry out both theoretical and applied research in Statistics and to improve the quality of official statistics in the country. India was recognised as one of the statistically developed countries in the world. United Nations Statistical Office launched a world wide project called 'National Household Survey Capability Programme' in 1977 to assist the developing countries to carry out national sample surveys in their respective countries on the lines of National Sample Surveys which were launched in 1950 in India under the technical guidance of the Indian Statistical Institute. Such sample surveys were recognised as the most cost-effective method to generate essential data base required for formulation of meaningful development plans in the

1 The paper was presented at the Dr. P.V. Sukhatme Memorial Session held at Rajkot on 7 December, 1997 during the 51st Annual Conference of the Indian Society of Agricultural Statistics.

2 Present Address: D-135, Saket, New Delhi - 110017

country. Many statistical schemes were designed during sixties and seventies to strengthen the quality and timeliness of agricultural and allied statistics in the country. Prof. P.V. Sukhatme played a crucial role in establishing an Indian Agricultural Statistics Research Institute to help the Indian Council of Agricultural Research (ICAR) to use modern scientific methods for deriving dependable estimates of key agricultural and allied parameters required for planning and policy making. With the departure of Prof. Mahalanobis from the statistical scene in mid-seventies further development in the national statistical system became more and more difficult with the passage of time primarily due to paucity of funds in the Five Year and Annual Development Plans allocated to Statistics. The changing scenario after Mahalanobis has been summarised in [2].

In recent years there is a growing concern amongst many professional statisticians and economists in the country about the deteriorating quality of official statistics pertaining to several vital sectors including agricultural ([1], [2], [3]). The causes of deteriorating quality of agricultural and rural development statistics have been critically examined recently in [1]. Remedial measures for strengthening the national statistical system in general and agricultural and rural development statistics in particular have been elaborated in [1], [2]. Unfortunately most of the suggestions for improvement are yet to be implemented. There has been no significant improvement in agricultural and allied data system in recent years. The agricultural production figures are often doubted. The statistical activities which are assigned to land revenue officials are not often given due importance by the land revenue officials as compared to their more 'prestigious' land revenue work which generally counts towards their future career. Unless suitable rewards and penalties for good and indifferent work in statistical activities are introduced in the present data collection system of acreage and land use statistics which are maintained by the village patwaris, the quality of data would be a suspect in many parts of the country. It would be desirable to introduce an independent sample check of key agro-socio-economic statistics which are currently produced by various data collection agencies in the country on a regular basis. In this paper some suggestions are given to strengthen the current national data system by introducing regular collection of longitudinal data pertaining to key agro-socio-economic variables in a cost-effective manner.

The national sample surveys which are being conducted in the country more or less on a continuing basis since 1950 deal only with cross-sectional data. On the other hand, measurement of change in complex socio-economic variables over time can be assessed more accurately from longitudinal data. The Sample Registration Areas which are currently used by the Registrar General, India (RGI) only for deriving dependable estimates of vital rates (birth rate, death rate and the consequent natural growth rate) can be utilised more

intensively by the official agencies to generate meaningful longitudinal data elaborated in the next section.

2. Sample Registration Areas

The Civil Registration System (CRS) existing in the country has been found to be so defective in most parts of the country that the vital rates derived from CRS are still considered unacceptable for formulation of a dependable development plan for the country. The Sample Registration System (SRS) was introduced by RGI in 1969-70 to remove the aforesaid deficiency. A short background note on Sample Registration Areas (SRA) is given at the Appendix for the benefit of those agricultural and other scientists who are not familiar with the above system.

The SRA's constitute an 'area sample' spread all over the country. It enables the statisticians and demographers to derive dependable estimates of vital rates. Concentrated efforts are made by the resident part-time enumerators and regional staff of RGI to collect information on each and every vital event occurring within an SRA. These are thereafter used to get estimates of vital rates at six-monthly interval.

The SRA's are being replaced by RGI only at an interval of about 10 years on the basis of population statistics generated by the decennial population census. Though it is admitted that this is expected to introduce some 'conditioning effect' in the sampled area on account of repeated use of the same set of households residing in the sampled area over a long period, the replacement of the 'area sample' could not be attempted by RGI before the 10 years period on account of practical difficulties. Rotation sample of area sample every 3 years or so would be desirable to take into account of the 'conditioning effect' and introduction of some unknown bias in the estimates based on these SRA's.

It is suggested that the SRA's should be used more intensively not only to drive reliable estimates of vital rates but also:

- A. To generate additional data which could serve as post-enumeration sample check for similar data collected under a complete census such as decennial population (conducted by RGI), quinquennial livestock census (conducted by the Directorate of Economics & Statistics, Ministry of Agriculture), quinquennial economic census (conducted by Central Statistical Organisation, Ministry of Planning) and for checking the accuracy of the number of operational holdings and land utilisation statistics (which are maintained by the village patwaris on a complete census basis for the Land Record States);
- B. For arriving at more reliable statistics on many agro-socio-economic variables such as cropwise agricultural production and acreage, livestock

number and employment, poultry statistics, animal feed as well as other important agricultural variables for which the currently available data are considered generally unreliable;

- C. For working out dependable estimates of many complex socio-economic variables like gender specific time use of household members in economic activities which would normally require repeated visits to the same household over a fairly long period (to take into account the seasonal variations in the household activities); and
- D. For deriving more dependable estimates of the direction of change in the level of living, income and expenditure, chronic unemployment and under employment amongst the poorest segment of the population, at regular intervals of say 3-5 years.

The above points are further elaborated in the next section.

3. Intensive Use of SRA's for Collection of 'Quality Data'

A. Sample Check

During a complete census the SRA's are also covered. Soon after completion of the main census a post-enumeration check on a few key agro-socio-economic variables covered under the main census should be carried out in all or on a sample of the usually resident households within the SRA. Matching the results of the main census with those of the post-enumeration check is expected to shed valuable light on the quality of main census data. Whenever wide divergence is noticed between the two sets of figures pertaining to the same SRA further probing in the concerned SRA would be necessary to find out the reasons of the divergence and suitable remedial actions should be taken.

B. Probing into Quality of Doubtful Statistics

Quality of data on the number of operational holdings, acreage under various crops, livestock number and employment in that sector, animal feed, poultry statistics etc. is generally considered unsatisfactory under the current data collection system. Data on these doubtful agricultural variables can be checked within the SRA from the longitudinal data on the aforesaid variables that can be collected by a trained resident enumerator within the SRA. The doubtful figures are expected to become apparent from the serial data that can be derived for the selected SRA.

The SRA's can also be used to get an independent plausible estimate of agricultural production for different major crops by directly interviewing all the farmers residing within the SRA to get their estimate(s) of cropwise

production soon after the harvesting of the crops. The part-time enumerators currently engaged by RGI to collect only information on vital events in the SRA may not find it feasible to take up this additional work without neglecting their vital registration work. The statistical activities within the SRA need to be rationalised and suitably integrated by reorganisation of the present statistical system. Some suggestions in this regard are given in Section 5. In this connection, it should be noted that recent research experiments carried out in several countries in Africa [4] as well as in a state in India tend to suggest that estimate of agricultural production based on the aforesaid interview method is quite close to actual agricultural production and sometimes better than the estimates derived under the current methodology using crop acreages and yield rates derived from crop cutting experiments. The estimate of agricultural production within the SRA's should normally be fairly close to the estimate derived from the current methodology. Wide divergence between the two sets of estimates would tend to suggest a large margin of uncertainty in the estimates of agricultural production derived under the current methodology.

C. Estimates of Complex Socio-economic Parameters

A resident field enumerator in an SRA would be able to canvass without much difficulty a schedule on time use by selected members of a household through repeated visits to the same household as well as by collection of related details through 'participant observation' method (often used in social anthropological surveys), whenever considered necessary.

For carrying out the assigned tasks satisfactorily it is essential that the resident enumerator gets adequate training for the purpose. The field enumerator/investigator may have to reside within the SRA for a fairly long period to collect 'quality data' on such complex socio-economic variables.

D. Repeat Surveys within SRA

Repeat surveys amongst the usually resident households within an SRA at regular intervals of 3-5 years can be used to collect longitudinal data on the change noticed in key socio-economic characteristics such as the number of persons within the SRA who are 'visibly deprived' of basic human needs, chronic unemployment etc. at regular intervals. The changes observed during the intervening period of 3-5 years are expected to yield fairly dependable estimate of the direction of change in the level of living of resident population who are 'visibly deprived' of minimum human needs covering all the principal components of level of living (such as food, health and nutrition, shelter, fuel, literacy, drinking water, sanitation and so forth). Extreme poverty might be defined as deprivation of the minimum needs pertaining to most of the aforesaid principal components. Population 'visibly deprived' of basic human needs can

be assessed in terms of a small number of fairly simple socio-economic indicators, most of them should preferably be non-monetary and cover all the principal components of levels of living. An illustrative list of such socio-economic indicators for identifying the extremely poor rural population is given below:

Socio-Economic Indicators	
Food	(i) Families not having two square meals a day.
Shelter	(ii) Families living in thatched houses without permanent roof.
Fuel	(iii) Families using only cowdung/dried leaves as fuel for cooking.
Literacy	(iv) Families having at least one adult illiterate.
	(v) Families having children, none attending any school.
Employment	(vi) Families have not a single adult gainfully employed.
Sanitation	(vii) Families without a household latrine.
Drinking Water	(viii) Families with no access to safe drinking water in the vicinity.
Asset	(ix) Families not having any land.
	(x) Families not having any livestock.

The above indicators would need precise definition and further refinement so as to ensure comparability of data pertaining to different SRA's that would be collected by different field investigators. Suitable action plan for ameliorating the condition of 'visibly poor' will be easier to evolve on the basis of multi-dimensional definition of 'visibly poor' in different SRA's as advocated in [5].

The improvement/deterioration in the magnitude of deprivation of basic human needs in an SRA over time can be easily assessed by conducting repeat surveys at intervals of 3-5 years. If the trend revealed by the repeat surveys in the SRA's is found to be contradictory to the trend arrived at by following the methodology currently used by the Planning Commission based primarily on cross-sectional data of NSS, further probing regarding suitability of the current methodology adopted by the Planning Commission would be desirable.

To avoid possible controversies pertaining to the quality of official statistics and the claims made by policy makers/politicians regarding improvement or deterioration in the 'quality of life' of people living in different parts of the country it is necessary to give priority to introduce the aforesaid panel study in the SRA's on a regular basis as soon possible.

4. Remarks

The above suggestions for panel study using longitudinal data generated from the SRA's for assessing the direction of change in complex socio-economic variables like 'level of living' are expected to provide better insight to the

administrators, planners and policy makers for formulation of more realistic action plan to fight poverty in the country. The nature and degree of deprivation as revealed by socio-economic indicators suggested earlier are expected to help the policy makers in deriving appropriate strategies and modalities to eliminate such malaise in general and reorient poverty alleviation programmes in particular. Similar study is possible with other complex variables like chronic unemployment and underemployment.

In this connection it is noteworthy that SRA's can provide estimates both on 'de jure' or 'de-facto' concepts (which are commonly used in collecting demographic statistics in the country). While estimates of acreage, crop production, livestock and poultry statistics pertaining to the entire sampled area can be determined on the basis of 'de facto' concept, by asking each and every household residing in the SRA, estimate of change in the percentage of visibly poor families should appropriately be confined to only those who are usual residents in the SRA, i.e. on the basis of 'de jure' concept. Since the estimates of 'poverty line' are derived from the cross-sectional data collected by NSSO on the basis of 'de facto' concept, the estimates of poverty based on (i) panel data with in SRA's and (ii) cross-sectional data on consumption and expenditure collected by NSSO would not be strictly comparable. Nevertheless if the direction of change over a period of 3-5 years as suggested earlier is found to be contradictory for the aforesaid two concepts, deeper probing for the causes of the contradictory trend is considered necessary.

There is bound to be in-and out-migrations of members of households pertaining to an SRA. For getting estimates based on 'de jure' concept, the number of usually resident households is expected to get diminished with longer period of repeat surveys; this would affect the precision of estimates derived from panel data.

5. Suggestions on Organisational Set Up

The task of collecting additional data by resident enumerators keeping in view the urgent need of our economy has to be entrusted to the Central Department of Statistics at the centre and to the Directorates of Economics and Statistics/State Statistical Bureaus at the State level. While deciding on the detailed procedure to be followed for collection of reliable data at the SRA level it has to be ensured that additional responsibility for collection of data other than on vital statistics should not impair the quality of vital statistics which are being utilised by RGI for deriving dependable estimates of vital rates. The additional data to be collected within an SRA may relate to several sectors of our economy; for instance, for data collection on agricultural production, number of operational holdings, acreage under various crops etc. the Directorate of Economics & Statistics-Ministry of Agriculture should be closely associated to ensure collection of quality and comparable data in the SRA; data on time-use

survey through repeated visits to selected households in the SRA should be carefully planned in consultation with the Central Department of Statistics and other interested agencies like Social Welfare Division of Planning Commission at the Central level. Efforts should be made to collect minimum statistics to ensure collection of quality data by trained female resident enumerators under the close supervision of local level supervisory staff belonging to SSB or Central NSSO. For ensuring collection of meaningful data local resident enumerators (preferably female) and regular supervision of data collection, it is essential that the detailed planning of the survey is done in consultation with nodal agency viz. CSO at the Centre and SSB/DES at the State level. Feasibility of the establishment of a plausible data collection agency at the Centre and the States should be carefully considered by carrying out a pilot study in selected SRA's. Experience gained in introducing a meaningful development monitoring service in certain selected areas in Kerala [6] should be taken into account before introducing the new system of data collection in selected SRA's. As mentioned earlier, adequate training has to be imparted to the field enumerators/investigators before they are engaged for collection of longitudinal data pertaining to complex socio-economic characteristics. Technical details for collection of longitudinal/panel data should be decided by the Central Department of Statistics/State Statistical Bureaus in consultation with RGI and other administrative departments interested in the findings of the study. The data collection work on different complex agro-socio-economic variables in the same SRA needed by different administrative agencies have to be suitably harmonised by an experienced data collection agency like NSSO at the Central level.

In depth analysis of collected data and preparation of technical reports pertaining to newly emerging issues like gender specific time-use in economic activities within the sampled households, studies on magnitude and direction of change in deprivation of basic human needs in the SRA's or other complex socio-economic characteristics should preferably be entrusted on contract basis to some experienced research institutions of repute who are fully conversant with the subject. The overall responsibility for undertaking special studies based on additional data collected in the SRA's should be entrusted to the Central Department of Statistics.

6. Concluding Remarks

- (a) The present number of sample registration areas spread all over the country has been based on the desired precision of vital rates at State (rural and urban areas separately) level. The sample size may not be adequate for many socio-economic variables that might be decided to be studied.

- (b) The representativeness of SRA's for the country is likely to be affected due to 'conditioning effect' on SRA's on account of continuing use of the same area for several consecutive years.
- (c) The SRA's can be used by and large as social observatories to study the change in complex socio-economic characteristics over time amongst the usually resident population, more or less, as case studies.
- (d) Additional data collection on a continuing basis in an SRA has to be done by resident enumerators in the concerned SRA. Repeat surveys can, however, be carried out by trained field investigators visiting the SRA at regular intervals. The supervision of field work of resident enumerators or field investigators has to be organised by the regular staff of the available data collection agencies such as NSSO under the Central Department of Statistics or by the supervisory field staff of the State Statistical Bureaus in different States or by the supervisory field staff of the concerned Central Department which is vitally interested in the findings of the study.
- (e) The organisational set up that would enable meaningful panel studies in the SRA's should be decided on the basis of experience to be gained during some pilot studies in different demographical and socio-economic situations in the country. Such pilot studies in selected SRA's would provide better insight into the practical problems that are likely to be encountered in introducing the new system of longitudinal data collection in the SRA's on a regular basis.

REFERENCES

- [1] Ghosh, Arun, 1997. India's crumbling data system. *Eco. & Pol. Weekly*, June 21, 1997, 1439-1441.
- [2] Rao, V.R., 1993. Statistics for rural development—An Agenda for the Nineties. *J. Ind. Soc. Agril. Statist.*, 46, 1-47.
- [3] Seal, K.C., 1994. Quality of official statistics and statisticians-The changing scenario after Mahalanobis. *Proceedings of Triennial International Symposium on Probability & Statistics*, Organised jointly by Calcutta Statistical Association and Calcutta University, Dec. 1994-Jan. 1995.
- [4] Seal, K.C., and Bansil, P.C., 1993. Poverty and agriculture—Issues and problems of measurement. Annual Conference of *J. Ind. Soc. Agril. Statist.*, Bhubaneswar.
- [5] Scott, Wolf and Mathew, N.T., 1983. A development monitoring service at the local level. *United Nations Research Institute for Social Development*, Geneva.

Appendix

Background Note on Sample Registration Areas

Sample Registration Area (SRA) used in this article stands for the sample unit under the Sample Registration System (SRS) which was introduced in 1969-70 by the Registrar General, India (RGI) to derive reasonably good estimates of Birth Rate (BR), Death Rate (DR) and the consequent Growth Rate (GR) at the State level separately for rural and urban areas in the country. Technical details of the SRS are well-documented and need no detailed elaboration. Broadly speaking, SRS is a large scale demographic sample survey based on the dual recording system of births and deaths in the country. The objective is to provide dependable estimates of birth rate, death rate and growth rate in different states (rural/urban) which are vitally needed by the development planners and policy makers. Based on data on vital statistics collected on a continuing basis, SRS also provides various other measures of fertility and mortality often required by demographers for making various analytical studies. These are taken up by RGI and many other research workers in the country.

The field investigation under SRS consists of continuous enumeration of births and deaths in a sample of villages/urban blocks by a resident part-time enumerator appointed by RGI in consultation with the concerned state administration. Generally the enumerators are school teachers. RGI also organises an independent six-monthly retrospective survey by a full time supervisor selected by RGI. Data obtained independently through these two sources are matched. The unmatched and partially matched events are re-verified in the concerned SRA to get an unduplicated count of correct vital events in the SRA. This matching provides an insight of the sources of distortion in the two sets of records and enables RGI to improve the quality of performance of the resident field enumerators and their supervisors who generally belong to the office of Registrar General, India.

The part-time resident field enumerator in an SRA is paid a small honorarium for the work entrusted to him. The full-time supervisors stationed at the concerned state headquarters have a complement of staff necessary for planning and organising various field operations, training of the field staff, effecting proper and timely supervision and control, ensuring regular flow of returns from the field, forwarding of various returns to the office of the Registrar General and for undertaking certain minimum tabulations. One supervisor is usually assigned a set of 10 SRA's for conducting half-yearly surveys.

Initially about 3,700 SRA's were selected from 1961 Population Census frame both in rural and urban areas. Another 1700 SRA's were added during 1978-79 from the 1971 Population Census frame. After the 1981 Population Census, 6022 SRA's were selected afresh from the 1981 Population Census houselist frame covering the whole country (except Mizoram). After the 1991

Population Census the number of SRA's was further increased to 6,613 (4,420 rural and 2,193 urban) covering a total population of 5.9 million (4.7 million rural and 1.2 million urban) spread all over the country (excluding J & K and Mizoram).

The sample design adopted is a unistage stratified simple random sampling without replacement. In rural areas, each district within a state has been divided into two strata, namely stratum 1 - villages with population less than or equal to 1,500 and Stratum 2- villages with population more than 1,500. In order to cover the village by a part-time resident enumerator, villages belonging to the second stratum (having population of more than 1,500) were broken down into two or more segments of approximately equal size. A simple random sample of villages and segments has been selected from each of the two strata, without replacement in each State/Union Territory. In urban areas stratification has been done on the basis of size class of the towns/cities. The towns/cities were grouped into five size classes, namely (a) towns with population below 20,000 (b) towns with population of 20,000 and more but less than 50,000 (c) towns with population of 50,000 and more but less than 1,00,000 (d) towns with population of 1,00,000 and more but less than 5,00,000 (e) cities with population of 5,00,000 and more but less than 10,00,000 and (f) each city with population of 10,00,000 or more, treated as a separate stratum. The SRA in urban area is a census enumeration block. A simple random sample of these enumeration blocks has been selected without replacement from each of the size classes of towns/cities in each State/Union Territory. Statement 1 at the end shows the statewise distribution of the number of SRA's and population covered in 1994 separately for rural and urban areas of all the States and Union Territories where SRS is being implemented.

The number of SRA's under SRS has been by and large, based on the desired precision of vital rates at State (rural and urban separately) level. For socio-economic variables other than the vital rates, the current sample size may or may not be adequate depending on the variability of the socio-economic variable under study. However, one advantage of SRA's is that it remains fixed for a period of almost 10 years till a fresh sample is drawn on the basis of the next decennial census. Collection of data on vital events by resident enumerator enables the administrators and planners to check the accuracy of the collected data which are available on a longitudinal basis. The national sample surveys cannot collect longitudinal data as their survey design adopted up till now cannot accomodate such panel studies/longitudinal data collection on a regular basis.

Reference : V.S. Swamy, 1992. Sample Registration System in India—A Perspective, Occasional Paper No. 4 of 1992, Office of RGI, Vital Statistics Division, Ministry of Home Affairs, Government of India, New Delhi.

Statement 1 : Number of sample units and population covered India, States and Union Territories — 1994

No.	States/UT	No. of Sample Units			Population Covered ('000)		
		Total	Rural	Urban	Total	Rural	Urban
	India	6613	4420	2193	5933 ¹	4668 ¹	1265 ¹
	<i>Major States</i>						
1.	A.P.	310	210	100	340	274	66
2.	Assam	254	200	54	215	184	31
3.	Bihar	500	400	100	506	449	57
4.	Gujarat	300	200	100	315	262	53
5.	Haryana	180	110	70	197	148	49
6.	Karnataka	375	260	115	348	275	73
7.	Kerala	250	150	100	327	267	60
8.	M.P.	400	300	100	344	287	57
9.	Maharashtra	375	200	175	344	237	107
10.	Orissa	405	300	105	280	222	58
11.	Punjab	200	120	80	175	127	48
12.	Rajasthan	350	250	100	302	255	47
13.	Tamil Nadu	375	200	175	394	278	116
14.	Uttar Pradesh	650	450	200	648	546	102
15.	West Bengal	475	300	175	438	334	104
	<i>Smaller States</i>						
1.	Arunachal Pradesh	56	50	6	38	33	5
2.	Goa	70	40	30	60	43	17
3.	H.P.	190	140	50	96	69	27
4.	J & K	166	100	66	Not Available		
5.	Manipur	150	100	50	123	90	33
6.	Meghalaya	120	100	20	54	45	9
7.	Nagaland	35	25	10	22	18 ²	4
8.	Sikkim	60	50	10	48	41	7
9.	Tripura	75	60	15	83	75	8
	<i>Union Territories</i>						
1.	A & N Islands	45	30	15	30	21	9
2.	Chandigarh	30	5	25	24	6	18
3.	Dadra & Nagar Haveli	20	20	—	21	21	—
4.	Daman & Diu	17	15	2	19	18	1
5.	Delhi	130	10	120	95	14	81
6.	Lakshadweep	10	5	5	11	7	4
7.	Pondicherry	40	20	20	35	21	14

1. Excludes Jammu & Kashmir and Mizoram. The field work could not be undertaken in Jammu & Kashmir due to disturbed conditions and in Mizoram due to some administrative reasons.

2. Relates to 1993 due to incomplete returns for 1994.

3. No sample.