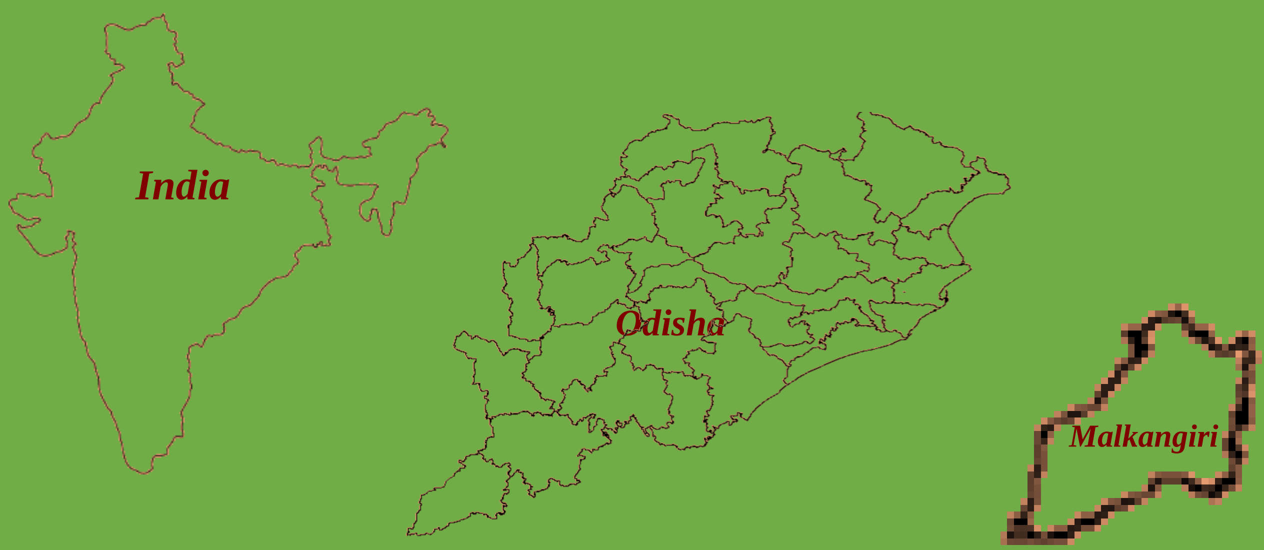
District Level Poverty Estimation for Odisha by using Small Area Estimation Technique



Directorate of Economics and Statistics,
Planning and Convergence Department
Government of Odisha
Bhubaneswar

Preface

It is a great pleasure for the Directorate of Economics and Statistics, Odisha for bringing out the Report "District Level Poverty Estimation of Odisha applying Small Area Estimation Technique" for the first time.

Measurement of poverty and its estimation has been at the center stage of the planning process in every developing country. Household surveys for consumption expenditure have been main instruments of poverty measurement.

The purpose of this paper is to provide a critical review of the main advantages in small Area Estimation (SAE) methods for poverty estimation. Poverty estimation at small area levels is a practical necessity in view of growing needs for micro level planning. Presently, estimates for number of poor as well as for poverty ratios are provided only at State level. Poverty mapping is done, based on small area level estimates. Direct estimates, based on NSSO data are likely to be less precise due to smaller sample sizes at district level. Attempts have been made through this paper to compare other aspects of poverty estimation, such as measuring the incidence, depth and severity of poverty, inequalities as well as distribution of poverty to different groups of population at district level.

I am extremely grateful to Dr. Hukum Chandra, National Fellow, Indian Agricultural Statistics Research Institute (ICAR) , New Delhi for his kind support and technical guidance for preparation of this Report.

I record my appreciation to Sri B.N.Mohanty, Deputy Director for his nice attempt for preparation of the district level poverty estimation by using Small Area Estimation Technique for the first time in Odisha.

I am also very much thankful to Dr. N.K.Singh, Expert, PHADMA and Md. Feroz Khan, Deputy Director, DES for their valuable contribution for preparation of this report.

I am also thankful to the officers and staff of the Directorate of Economics & Statistics, Bhubaneswar for their concerted effort for preparation of this report.

(Dushasan Behera)
Director

Acknowledgement

Application of Small area estimation technique helps downward estimation in spite of sample limitations. Gaining importance of poverty estimation for the district level calls for induction of the instrument in this score. The District Level Poverty Estimation of Odisha by the DES is the outcome of this attempt.

I express my deep sense of gratitude to Sri Dushasan Behera, Director, Economics and Statistics, Odisha, Bhubaneswar for his inspiration in preparing the Report on "District Level Poverty Estimation for Odisha by using Small Area Estimation Technique"..

I express my thanks to Dr. Hukum Chandra, National Fellow, IASRI (ICAR), New Delhi for his able guidance for preparation the report.

I am also thankful to Md. Feorz Khan, Deputy Director,DES and Dr. N. K. Singh, Expert PHADMA for their valuable support for preparation of this report.

And at last but not the least, I am very much thankful to the officers and staff of the Directorate of Economics & Statistics, Bhubaneswar who have extended their valuable guidance and co-operation for preparation of my dissertation paper.

I acknowledge it with sincere thanks

(Bigyanananda Mohanty)
Deputy Director
Directorate of Economics and Statistics,
Odisha, Bhubaneswar.

Contents

Abbreviation

BLUP	Best Linear Unbiased Predictor
CI	Confidence Interval
CV	Coefficient of Variance
DE S	Directorate of Economics & Statistics
EBLU	Empirical Best Linear Unbiased Predictor
FSU	First Stage Units
FH	Fay and Herriot
GP	Gram Panchayat
ICAR	Indian Council of Agricultural Research
IASRI	Indian Agricultural Statistics Research Institute
LLMM	Logistic Linear Mixed Model
MAX	Maximum
MIN	Minimum
MSE	Mean Squared Error
MPCE	Monthly Percapita Consumer Expenditure
NSS	National Sample Survey
NSSO	National Sample Survey Office
OBCs	Other Backward Classes
OLS	Ordinary Least Squares
PHADMA	Poverty & Human Development Monitoring Agency
PQL	Personalised Quasi Likelihood
SAE	Small Area Estimation
SCs	Scheduled Castes
SRS	Sample Random Sampling
SSU	Second Stage Units
STs	Scheduled Tribes
UN	United Nations
US	United States
WPR	Work Participation Rate
REML	Restricted Maximum Likelihood

Section - I
Introduction

Poverty is the main concern in most of the countries in the world.

Coexistence of poverty with population has a great resemblance in the Indian contest. Poverty is one of the main problems which have attracted attention of sociologists and economists. It indicates a condition in which a person fails to maintain a living standard adequate for his physical and mental efficiency. According to Adam Smith man is rich or poor according to the degree in which he can afford to enjoy the necessaries, the conveniences and the amusements of human life. According to the UN(1998), ***"Fundamentally, poverty is a denial choice and opportunities, a violation of human dignity. It means lack of basic capacity to participate effectively in society. It means not having enough to feed and clothes to family, not having a school or a clinic to go to not having the land on which to grow one's food or a job to earn one's living, not having access to credit. It means insincerity, powerlessness and exclusion of individuals, households and communities. It means susceptibility to violence, and it often implies living on marginal or fragile environments without access to the clean water or sanitation. According to the World Bank (2000), poverty is commonly visualized as a state of not having enough resources to meet the basic needs such as food, clothing and housing of a person.*** So poverty is a highly heterogeneous phenomenon in most of the countries of the World. Measurement of poverty and its estimations has been at the center stage of the planning process in every developing country.

Status of poverty in India according to Planning Commission of India

According to latest report by the Planning Commission of India (Tendulkar Committee), it was reported that 21.9% of all people in India fall below the international poverty line of US $1.25 per day. The number of poor is now estimated at 250 million, of which 200 million resides in rural India.

According to the release from Planning Commission, 25.7% of people in rural areas were below the so called poverty line and 13.7% in urban areas during 2010-11. This is comparable with 33.8% and 20.9% respectively in 2009-10. The poverty numbers are estimated on the basis of consumption expenditure captured in the fiver year surveys undertaken by the National Sample Survey Office (NSSO)

Poverty a great concern for Odisha.

Odisha is the tenth largest State in the Indian Union located on the eastern coast of the country surrounded by Jharkhand and west Bengal in the north, in the west by Chhattisgarh, in the south by Andhra Pradesh and in the east by Bay of Bengal. It has more than 480 kilometer long coast line of exotic attraction covers1,55,707 square kilometres with a total population of 41.97 million as per 2011 Census. The geographical boundary of the State comprises 4.74% of India's land marks, 3.58% of the country's population over 5% of the country's poor. Administratively Odisha has been divided into 3 Revenue Divisions, 30 districts, 58 sub-divisions, 317 Tahsils, 314 Blocks, and 6227 GPs, spread over 51,349 villages. The density of population in the State was 269 persons per square kilometre. The literacy rate in the State was 72.87% as against the national average of 74.04% during 2011 census. According to 2011 census, 83.31% of the population was living in rural areas of Odisha.

Odisha is a land of diversity and inhabited by different ethnic groups. About 40% of the population of the State belong to the scheduled caste (17.1%) and scheduled tribe (22.8%) communities. The geographical boundary of Odisha divides it into three regions like the coastal, the northern and the southern.

According to the Planning Commission, Odisha has been clubbed with the category of the poor states of the union. Most of these regions are either flood–prone or suffer from drought like conditions.

Inter and intra-regional disparity in Odisha causes poverty

These conditions hamper agriculture to a great extent, on which the household income of these people depends. The vagaries of nature exposed to the State's agriculture sector economy frequently causes cyclones, droughts and flash floods which substantially affect production and productivity of agriculture. The incidence of poverty in Odisha declined from 66.18% during 1973-74 to 32.59% during 2011-12.as against the national average of 54.88% and 21.92% respectively. But in rural Odisha the incidence of poverty is 35.69% as against 25.70% in all India during 2011-12.

The southern and northern regions are relatively less developed in comparison to the coastal region. There is also widening the rich poor gap across the social groups as well as the regional disparity. In Odisha, large scale disparities in economy and social development across the regions and intra-regional disparities among different communities like STs, SCs and OBCs have been major areas of concern and thus ushering regional / district level planning in the State.

Need of micro level planning in Odisha

Inputs on the different socio economic characteristics at grass root level like district, block etc. are highly essential for decentralised planning and strategies for programme implementation to handle the backward areas of the State. Certain items of information such as consumer expenditure / income data essential for grass-root planning are not covered in census schedules. Besides, conduct of decennial census in India, the National Sample Surveys Office (NSSO) carries out country wide surveys on various socio-economic parameters related to the national economy of varied topics as per demand of the Government from time to time in a regular basis in form of different rounds during the inter-censual periods. The sample sizes so designed for the surveys of the NSSO are the modest in nature and are fixed in such a way that it is possible to get some usable estimates at the national and State level.

However, due to the importance of micro level planning, in a developing country like India, where there is large scale poverty in most parts of the country, reliable estimates are being demanded by the administrators and policy planners at the small area level as per the recommendation of "Working Group on District Planning" set up by the Planning Commission of the Government of India during 1982 (Annexure-II). The Working Group in its report clearly highlighted the data requirement for planning and decision making at the district level. The sample sizes in NSSO surveys at the State level are not large enough to provide reliable direct estimate at small area level like district level, block level, community level etc. Conduct of district specific surveys with large sample also becomes expensive as well as time consuming.

Why Small Area Estimation (SAE) Technique used for micro level planning ?

Under such circumstances and in view of the demand for reliable statistics at micro level, Small Area (Domain) Estimation (SAE) techniques have been developed to produce reliable estimates for such small areas with small sample sizes by **borrowing strength** from data of other areas through *explicit and implicit models* which connects the small areas *via* supplementary data to find indirect estimates that increase the effective sample size and thus increase the precision. Because direct estimates may not provide acceptable precision at the small area level due to small sample size in small areas. The idea is to use statistical methods to link the variable of interest with the auxiliary information, from Census and Administrative sources for the small areas to define model based estimators of these small areas.

Typically, small area refers to a subset of the population for which enough information is not available from the sample survey because of limited sample size. This may be useful to a small geographical area like district, block, tehsil, gram panchayat etc. But it also includes a demographic group like a specific age, sex, group of people with a large geographical area.

Objective of the Study

The main objective of the present study is as follows

- To estimate the rural poverty in Odisha at district / regional levels based on the 68th round (2011-12) of the Household Consumer Expenditure data of the National Sample Survey (NSS) and 2011 Indian Population Census data, using direct, indirect, and small area estimation techniques.

- To compare between direct, and model based small area estimation applied for estimating the rural poverty in Odisha.

Sources of Data

In this study covariates are available at district level. Therefore, **we adopt an area level model to derive the small area estimates**.

Two types of variables are required for this analysis. These are

i. Variable of interest and

ii. Auxiliary (Covariates) variables

i. The **variable of interest** for which small area estimates are required is drawn from the Household Consumer Expenditure Survey 2011-12 of NSSO 68th round data for rural areas of the State of Odisha. The target variable used for the study was poor households. The parameter of interest is the proportion of poor households at different level. The poverty line has been used to identify whether given household is poor or not. As per Tendulkar Methodology of Planning Commission of India, the poverty line is Rs. 695.00 for rural areas has been used in this study.

ii. The auxiliary (covariates) variables are drawn from the Population Census 2011.

Matching Variables in the Survey and the Census

Before modelling, it is essential to select the list of explanatory variables to ones that exist in both the survey and the census. If the sample of the household survey is randomly selected, one can expect the distribution of the variables to be the same in the survey and in the population.

Initially, a list of common variables was constructed using both the census schedule i.e., the house list schedule and the household schedule of NSSO Consumer Expenditure survey variables.

Due to non-availability of village directory from the census , first of all convert the household level variables into village characteristics in both the census and the survey data and then convert village level data into district level and then use these generated district level variables in the regression model.

NSS data does not contain any village / district level variables. As we know that location effects captured by village / district level variables are important determinant of consumption behaviour. In order to control for location effects, we rely only on village/ district level variables that can be created from the available household level variables.

But the covariates (explanatory variables) are available at districts level not beyond that. So, the area level area model is adopted to derive the small area level estimates. These covariates are drawn from the census 2011.

To test the comparability, the relationship between variables of interest and covariates used in this study are assumed not to change significantly over the period. There were more than100 covariates available from the population census for the purpose of modelling

The sampling design used in the 68th round NSSO data for the year 2011-12 is a stratified multi stage random sampling with district as strata, villages as fist state unit (FSU) in the rural areas and household as the second stage unit (SSU). During the 68th round NSS survey, a total 2937 households were surveyed from the 30 districts of Odisha. The district wise sample size varied from 64 to 128 ***(Table 1)***

Table 1 : Distribution of districts-wise sample size

Sl No.	District Name	Rural	
		village	**household**
1	Baragarh	16	128
2	Jharsuguda	8	64
3	Sambalpur	8	64
4	Deogarh	8	64
5	Sundargarh	16	128
6	Keonjhar	16	128
7	Mayurbhanja	16	128
8	Balasore	16	128
9	Bhadrak	16	128
10	Kendrapara	16	126
11	Jagatsinghpur	12	96

Sl No.	District Name	Rural	
		village	**household**
12	Cuttack	16	128
13	Jajpur	16	128
14	Dhenkanal	12	96
15	Angul	12	95
16	Nayagarh	12	96
17	Khurda	12	96
18	Puri	16	128
19	Ganjam	20	160
20	Gajapati	8	64
21	Kandhamal	8	64
22	Boudh	8	64
23	Sonepur	8	64
24	Bolangir	12	96
25	Nuapada	8	64
26	Kalahandi	16	128
27	Rayagada	8	64
28	Nawrangapur	12	96
29	Koraput	12	96
30	Malkangiri	8	64
	Odisha	**372**	**2973**

SAE for reliable estimates

It is observed that the district level sample sizes are very small with very low values of average sampling fraction as 0.0001. So it is difficult to calculate reliable estimates and their standard errors at district level. Hence, the SAE techniques are used to solve the problem by providing reliable estimates for the districts having small sample size.

Methodology adopted

The SAE techniques has been used to estimate the poverty levels for different districts of the rural Odisha using Monthly Per Capita Consumer Expenditure (MPCE) data from the 68th round NSS conducted in 2011-12 and 2011 Population Census data of India. The poverty estimation for the district of Odisha has been carried out by two methods i.e

i. Direct estimators

ii. Estimators using mixed model approach.

The comparative pictures on assumptions and limitations of different aforesaid methods are given below :

Sl. No	Methods	Assumptions	Limitations
1	Direct estimators	• The population consists of non over lapping domains or small area. • Sample size is small • Variance is large • Unbiased estimate	• More resources are needed in the way of time, money and technical expertise for the successful completion of a survey. • It is unbiased under finite population sampling theory. • Sample size is very small at small area level. • Not giving reliable domain estimate using a direct estimator. • Direct area specific estimates may not provide acceptable precision at the small area level because sample size in small areas are seldom large enough.
2	Estimators using mixed model approach	• Area specific variability typically remains even after accounting for the auxiliary information. • It is handled by model based Small Area Estimation. • Model based methods that combine information from multiple related sources have been developed to increase the precision. • Area level models have the ability to protect confidentiality of micro data	• Area level modelling taken in to account the survey design through the use of the direct survey estimates and related design based variance estimates. • Area specific random effect assumed to be iid with mean) and variance σ^2

Limitations of the Study

The present study involves the following limitations.

Rural poverty Estimation only

1. The poverty analysis at the district level for the State of Odisha is restricted to the rural sector based on the data of the 68th round of the NSSO on Household Consumer Expenditure Survey and Indian Population Census, 2011. This restriction was adhered to because as large as about 84% of the population of the State according to 2011 census live in rural areas.

No separate Poverty line for the district

2. No separate poverty line has been used for each region or district in Odisha as the districts are only administrative units and not distinct geographical units. For the present study, the single State level official poverty line for Odisha as recommended by the Planning Commission, Government of India has been used.

3. The poverty estimation for the districts of the State has not taken into account the differential administrative efficiencies, available natural resources, infrastructural facilities, and political initiatives because of lack of quantitative indicators.

4. The poverty analysis of the NSSO consumer expenditure data is based on samples, assumed to be selected by simple random sampling (without replacement) from the population of households in the country, although the samples have been selected at different stages. This is done in order to simplify the mathematical formulae involved in the estimation.

❖❖❖❖

Section – II

Discussion on Direct Estimates

Direct Estimate

Let $U = \{ 1, 2, 3, \ldots\ldots N\}$ be the population of size N

A sample (s) of size n is drawn with sampling design $p(s)$ from the population, U.

If $\pi_j = \sum_{j\in s} p\,(s)$ is the first order inclusion probabilities then $w_j = \frac{1}{\pi_j}$ is the design weight of the element j. Under simple random sampling. $\pi_j = \frac{n}{N}$ and $w_j = \frac{1}{\pi_j} = \frac{1}{n\,/N} = \frac{N}{n}$

Assume that the population U consists of D non-overlapping small areas (domains) Ui each with population size N_i, such that $U = \bigcup_{i=1}^{D} \cup_i$, $i = 1,2,3, \ldots\ldots D$ and $N = \sum_{i=1}^{D} N_i$

Let s_i be the part of the sample s of size n_i ($n_i \geq 0$) that falls in small area i and $n = \sum_{i=1}^{D} n_i$.

Let y be the character under study.

Denote y_{ij} as the value for y of the j^{th} population unit in i^{th} small area unit. Define the population mean of i^{th} small area as

$$\bar{Y}_i = \sum_{j\in u_i} y_i \;/\; N_i$$

Direct Estimator using Sampled Data

Under the simple random sample (SRS) without replacement, a **direct estimator** of the mean $\hat{\bar{Y}}_i$ is given by $\bar{y}_i = \frac{\sum_{si} W_j y_j}{\sum W_j} = \frac{\sum y_j}{n_i}$ with variance

$Var\left(\hat{\bar{Y}}_i\right) = (1 - f_i)S_i^2 / n_i$ with $f_i = \frac{n_i}{N_i}$ and

$$S_i^2 = \frac{1}{(N_i - 1)} \sum_{j=1}^{N_i} (y_i - \bar{y}_i)^2, \quad N_i \geq 2$$

An unbiased estimator of S_i^2 is $s_i^2 = \frac{1}{(n_i - 1)} \sum_{j\in_{si}} (y_i - \bar{y}_i)^2$

Thus, an unbiased estimator for variance of $\hat{\bar{Y}}_i$ is given by

$v\left(\hat{\bar{Y}}_i\right) = (1 - f_i)s_i^2 / n_i$ when N_i is known. For unknown N_i, $f_i = \frac{n_i}{N_i}$ is replaced by $f = \frac{n}{N}$ and then the estimator for variance is $\hat{V}\left(\hat{\bar{Y}}_i\right) = (1 - f)s_i^2 / n_i$

The district level poverty estimation of Odisha has been computed by using direct estimation procedure using the 68th round NSS consumer expenditure data for rural sector. The proportion of poor, standard error, co-efficient of variation (CV) and confidence interval (CI) using the direct method are presented in ***Table 2***. It is observed that the proportion varies from 0.02 to 0.72, standard error from 0.0146 to 0.1748 and CV varies from 6.22 to 68.69 throughout the State across the districts. As per the standard, the CV more than 30% implies an unstable estimate. As the higher limit of the CV is found to be more than 68%, it signifies the unreliability of the estimate. The CVs show the sampling variability as a percentage of the estimate.

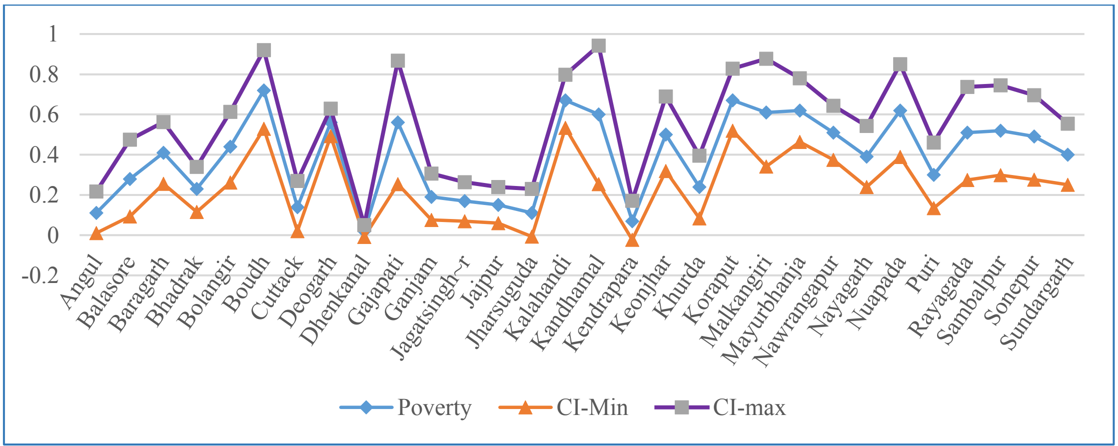
Table 2 : Direct Estimates of Poverty Ratio (District wise)

SL. No	District	Proportion	Std. Err.	CV (%)	95% Conf.	Interval
1	Angul	0.11	0.05	46.14	0.01	0.22
2	Balasore	0.28	0.10	34.26	0.09	0.48
3	Baragarh	0.41	0.08	19.01	0.26	0.56
4	Bhadrak	0.23	0.06	24.97	0.12	0.34
5	Bolangir	0.44	0.09	20.42	0.26	0.61
6	Boudh	0.72	0.10	13.72	0.53	0.92
7	Cuttack	0.14	0.06	43.74	0.02	0.27
8	Deogarh	0.56	0.04	6.22	0.49	0.63
9	Dhenkanal	0.02	0.02	68.69	-0.01	0.05
10	Gajapati	0.56	0.16	27.83	0.25	0.87
11	Ganjam	0.19	0.06	30.94	0.08	0.31
12	Jagatsinghpur	0.17	0.05	29.72	0.07	0.26
13	Jajpur	0.15	0.05	30.58	0.06	0.24
14	Jharsuguda	0.11	0.06	54.08	-0.01	0.23
15	Kalahandi	0.67	0.07	10.07	0.53	0.80
16	Kandhamal	0.60	0.18	29.24	0.25	0.94
17	Kendrapara	0.07	0.05	66.44	-0.02	0.17
18	Keonjhar	0.50	0.09	18.67	0.32	0.69
19	Khordha	0.24	0.08	33.10	0.08	0.40
20	Koraput	0.67	0.08	11.65	0.52	0.83
21	Malkangiri	0.61	0.14	22.33	0.34	0.88
22	Mayurbhanja	0.62	0.08	12.96	0.46	0.78
23	Nawrangapur	0.51	0.07	13.49	0.37	0.64
24	Nayagarh	0.39	0.08	19.72	0.24	0.54
25	Nuapada	0.62	0.12	18.93	0.39	0.85
26	Puri	0.30	0.08	27.87	0.14	0.46
27	Rayagada	0.51	0.12	23.19	0.28	0.74
28	Sambalpur	0.52	0.11	21.81	0.30	0.75
29	Sonepur	0.49	0.11	21.96	0.28	0.70
30	Sundargarh	0.40	0.08	19.10	0.25	0.55

Source : Computed from Primary data of NSSO

The ***Figure 1*** presents the district wise 95% confidence interval of the direct estimate along with value of direct estimate. This shows the degree of inequality with respect to the distribution of poverty across the districts, exemplified by the wide variations among the CI of the districts.

Figure 1 : District wise 95% Confidence Interval of the Direct Estimate



Classification of districts according to CV%

CV more than 40% in5 districts

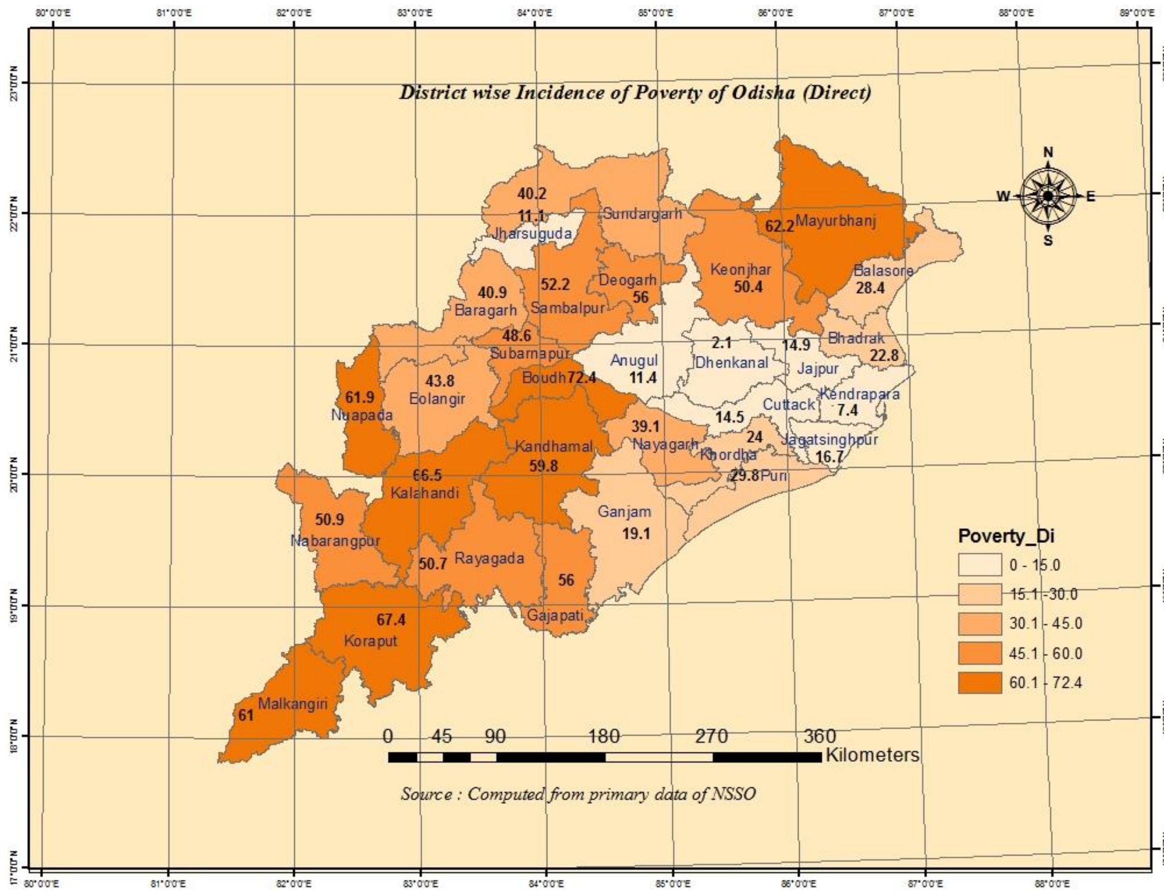
It is observed from the ***Table 3*** that out of 30 districts of Odisha, 11 districts have coefficient of variation (CV%) less than 20. The CV more than 20% in 19 districts of Odisha implies an unstable estimates (more sampling variability). Out of 19 districts, 5 districts namely, Angul, Cuttack, Dhenkanal, Jharsuguda, and Kendrapara have CV more than 40%. The spatial distribution of poverty is shown in ***Map-1***

Table 3 : Classification of the Districts according to CV of the Direct Estimate.

CV Class	No. of districts	Name of the districts
0-10	1	Deogarh
10-20	10	Kalahandi, Bargarh, Boudh, Keonjhar, Koraput, Mayurbhanj, Nabarangpur, Nayagarh, Nuapada, and Sundargarh
20-30	10	Bhadrak, Bolangir, Gajapati, Jagatsinghpur, Kandhamal, Malkangiri, Puri, Rayagada, Sambalpur and Sonepur
30-40	4	Balasore, Ganjam, Jajpur, and Khordha
40 and above	5	Angul, Cuttack, Dhenkanal, Jharsuguda, and Kendrapara.

Source : Computed from Primary data of NSSO

Map 1 : Map Showing District wise Incidence of Poverty of Odisha (Direct)



❖❖❖❖

Section – III

Discussion on Small Area Estimation (SAE)

Necessity of Small Area Estimation (SAE)

For larger sample size for each small area, the direct estimator is the most reliable one. But in practice, for most sample surveys, the sample size for each small area is usually very small in which the associated variances of these estimators are likely to be very large and unreliable. Under such circumstances, it is required to apply the estimation methods which borrow strength from the related areas. These estimators are known as the indirect estimators since they use values of survey variables (and auxiliary variables) from other small areas or times and possibly from both. They borrow information from other small areas (domains) or times or both by using statistical models based on implicit or explicit models. The usual indirect estimation techniques based on implicit models produce synthetic and composite estimators.

Mixed Models in SAE

Advantage of Mixed Models in SAE
i.Unit Level Model
ii. Area Level Model

The traditional indirect estimator assumes that all the areas of interest behave similarly with reference to the variable of interest and do not take into account the area specific variability. This will lead to severe biasness if the assumption of homogeneity within the larger area is violated or the structure of the population changed since the previous census. This limitation is taken care up by an alternative estimation techniques based on an explicit linking model named as ***mixed effect model***. Random area effects in the mixed effect model takes into account the dissimilarities among the areas. Indirect estimates based on explicit models have received a lot of attention because of the following advantages over traditional synthetic and composite estimates.

i. Model based methods make specific allowance for local variation through complex error structures in the model that links the small areas.
ii. Models can be validated from the sample data.
iii. Models can handle complex cases such as cross-sectional and time series data.
iv. Stable area specific measures of variability associated with the estimates may be obtained, unlike overall measures commonly used for traditional indirect estimates.

It is a special case of the linear mixed model and is very flexible to formulate and handle complex problems in SAE. Mixed models are used in specific situations based on data availability or the response variable of interest. These are (i) area level model which uses area specific auxiliary information and where information or response variable available only at the small area level and (ii) unit level model which uses the unit level auxiliary information and where information on the response variable is available at the unit level..

Unit Level Models

Consider a population of N units with i^{th} small areas consisting of N_i units. Let y_{ij} and x_{ij} be the unit level y-value and correlated covariate x value for j^{th} unit in the i^{th} small area. It is assumed the domain means $\bar{X}_d$ is known. Consider the one-folded nested error linear regression model.

$$y_{ij} = x_{ij}^T\beta + u_{i+}e_{ij} , j = 1, 2, \dots . N_i, i = 1,2, \dots D$$

where the random small area effects u_i have mean zero and common variance σ_u^2 and are independently distributed. Also e_{ij} are assumed to be independently distributed with mean zero variance σ_e^2 and are also independent of area effects u_i. This model was initially considered by Battese *et al.* (1988).

The model used by World Bank is a unit level regression based approach, it is very much different than the mixed model approach. For details of an exhaustive and thorough presentation of small area estimation an excellent reference is the book by Rao (2003),

Unit Level Methods

It was developed in *Hentschel et al*.(2000) amd *Elbers, Lanjouw* (2001)

i. It requires a minimum of two sets of data

ii. Household level Census data (auxiliary variable)

iii. A representative household survey corresponding to the same period as the Census.

The maximum allowable time difference will vary by the rate of economic changes in a given country,

The first step is to estimate a model of consumption-based household welfare using data from the household survey

$\log Y = \alpha + \beta_1 X + \beta_2 V + \epsilon$ is estimated using Ordinary Least Squares (OLS) where
Y = MPCE or Poverty proxy
X = Matrix of household level characteristics
V = Matrix of district/ sub-district level characteristics

The resulting parameter estimates are applied to the census data

For each household , the estimated parameter from the regression are used to compute the probability of each household in the census living in poverty.

Then household levels results can be aggregated by the districts / sub-districts by taking the mean of the probabilities for the districts.

For each household, the household level value of the explanatory variable is multiplied by the corresponding parameter estimate, which gives a predicted value of *log (MPCE)* for each household in the study area (survey)

The estimated value of the benchmark indicator is then used to find the probability of a household being a poor in terms of the given threshold

$F_{ij} = 1, \ y \log Y_{ij} < \log Z \quad E\,(F_i \,/\, X_i\beta, \sigma) = \varphi \ \frac{[\log z - X_i\,\beta]}{\sigma}$ = 0, otherwise

Where $\varphi =$ cumulative standard normal distribution.

The estimates of $\hat{\beta}$ and $\hat{\sigma}$ are obtained from the model of the benchmark indicator providing the following estimator of the expected poverty of household in the census

$$F_i^* = E\left(F_i \,/\, X_i\,\hat{\beta}, \hat{\sigma}\right) = \varphi\ \frac{\left[\log z - X_i\,\hat{\beta}\right]}{\hat{\sigma}}$$

Regional / District Poverty *F* is calculated by

$F = \frac{1}{N}\sum_{i=1}^{N} F_i$, where N = Number of households in a district or region

Expected Poverty is calculated by

$$E = (F/X, \beta\ \sigma) = \frac{1}{N}\sum_{i=1}^{N} E\ (F_i / X_i, \beta\ \sigma)$$

The incidence of poverty is calculated as the mean of probability of household being poor

$$F^* = E\left(F/X, \hat{\beta}, \hat{\sigma}\right) = \frac{1}{N}\sum_{i=1}^{N} \varphi\ \frac{[\log z - X_i\,\hat{\beta}]}{\hat{\sigma}}$$

Area Level Model

In this model, we describe the *EBLUP* estimator (*Empirical Best Linear Unbiased Predictor*), assuming a linear mixed model in which auxiliary information can be included at area level. This model was used originally by *Fay and Herriot (1979)* for the prediction of mean per capita income in small geographical areas and is given by

$$\hat{\theta}_i = \theta_i + e_i \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}$$
$$\theta_i = \boldsymbol{X_i^T \beta} + u_i \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

By combining these two equations we get,

$$\hat{\theta}_i = \boldsymbol{X_i^T \beta} + u_i + e_i \ \ldots\ldots\ldots\ldots\ldots\ldots \text{(iii)}$$

where

$\hat{\theta}_i$ is the direct survey estimate of the parameter θi (e.g. sample mean $\bar{y}_i$ etc.)

$\boldsymbol{X_i^T}$ =is the vector of covariates (area level variable) and $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots.., \beta_D)$

u_i is the model error *(iid)* $\mathrm{E}\,(u_i) = 0$, $\mathrm{v}(u_i) = \sigma_u^2$

e_i = the sampling errors which are assumed to be independent across small areas with

$E\,(e_i) = 0$ and $V(e_i) = \sigma_e^2$

Here e_i and u_i are design based and model-based random variables respectively. The models variance is σ_u^2 measure of homogeneity of the areas after accounting for the covariates x_i. Since the unknown parameters β and σ_u^2 are the same for every area, it makes sense to estimate these simultaneously across all the *D* areas. The *Fay and Herriot* (FH) method for SAE is based on area level linear mixed model and their approach is applicable to a continuous variable. But for discrete, particularly binary variable, the model linking the probability of success π_i with the covariates X_i is the logistic linear mixed model given by

$$Logit\ \pi_i = \ln\left[\frac{\pi_i}{1-\pi_i}\right] = X_i^T\beta + U_i \ \ldots\ \ldots\ldots\ldots\ldots\ldots \text{(iv)}$$

(i = 1, 2, 3… D)

Where β is the k-vector of regression coefficients often known as is fixed effect parameters and U_i is the area specific random effect that accounts for between area dissimilarity beyond that explained by the auxiliary variables included in the fixed part of the model. U_i are *iid* with mean zero and variance σ_i^2 and π_I is the probability of a poor household in area *i*, often termed as the probability of a "success". In this cases FH model is not applicable.

By definition, the means of y_{si} and y_{ri} given $u_i$under model (iv) are

$$\boldsymbol{E\,(y_{si}/u_i\,) = n_i\,\pi_i\; = n_i\,[(\mathbf{exp}\,X_i'\beta\; + \;u_i)\;(1 + \mathbf{exp}(X_i'\,\beta + \;u_i)^{-1}\,]} \text{ .... } (v)$$

$$\boldsymbol{E\,(y_{ri}/u_i\,) = (N_i\; - n_i\,)\,\pi_i = (N_i\; - n_i\,)\;[(\mathbf{exp}\,X_i'\beta\; + u_i)\;(1 + \mathbf{exp}(X_i'\,\beta + u_i)^{-1}\,]} \text{ .... } (vi)$$

Let T_i denotes the total number of poor households in the district *d*. We can write $T_i = y_{si} + y_{ri}$ where the first term y_{si} the sample count is known where as the second term y_{ri} the non-sample count, is unknown. Therefore the estimate $\hat{T}_i$ of the total number of households in area *i* is obtained by replacing y_{ri} by its predicted value under the model (i) that is

$$\hat{T}_i = y_{si} + \hat{y}_{ri} = y_{si} + (N_i - n_i)\left[\left(\exp X_i'\hat{\beta} + \hat{u}_i\right)\left(1 + \exp(X_i'\,\hat{\beta} + \hat{u}_i\right)^{-1}\right] \text{ .... (vii)}$$

An estimate of proportion of poor households p_i in small area *i* is obtained as

$$\hat{p}_i = \frac{\hat{T}_i}{N_i} = \frac{1}{N_i}\left\{y_{si} + (N_i - n_i)\left[\left(\exp X_i'\hat{\beta} + \hat{u}_i\right)\left(1 + \exp(X_i'\,\hat{\beta} + \hat{u}_i\right)^{-1}\right]\right\} \text{ ..... (viii)}$$

It is obvious that in order to compute the estimates given by equation (vii) of (viii) we require estimates of the unknown parameters β *and* u. A major difficulty in use of *Logistic Linear Mixed Model (LLMM)* for SAE is the estimation of unknown model parameters β *and* u since the likelihood function for *LLMM* often involves high dimensional integrals (computed by integrating a product of discrete and normal densities, which has no analytical solution) which are difficult to evaluate numerically. We used an interactive procedure that combines the *Penalized Quasi-Likelihood (PQL)* estimation of β *and* $u = (u_1 \ldots\ldots u_i)$ with *Restricted Maximum Likelihood (REML)* estimation of $\emptyset$ to estimate the parameters.

We now turn to estimation of mean squared error (MSE) for predictors given by equation (vii) .The MSE estimates are computed to assess the reliability of estimation and alos to construct the confidence interval (CI) for the estimates. The man squared error estimates of (vii) under model (i) is given by

$$mse\ (\hat{p}_i) = m_1\left(\widehat{\emptyset}\right) + m_2\left(\widehat{\emptyset}\right) + 2m_3\left(\widehat{\emptyset}\right) \ldots\ldots\ldots\ldots\ldots \text{(ix)}$$

The first two components m_1 and m_2 constitute the largest part of the overall *MSE* estimates in (ix). These are the *MSE* of the best *Linear Unbiased Predictor (BLUP)* –type estimator when $\emptyset$ is known. The third component m_3 is the variability due to the estimate of $\emptyset$

In this study area level models which was used by *Chandra et al. (2011)and Marteign et al. (2007)* have been applied for computing district level poverty estimate along with their mean square estimates following the mathematically tedious techniques developed by *Prasad* and *Rao (1990),* for which software packages are available. As the procedures is quite involved and need the use of software packages, we have omitted the complete elaboration here.

Selection of covariates for model-Based Estimation

First of all examined the correlation of all the available covariates with the target variable and then selected the covariates with reasonably good correlation with the target variable. After selection of covariates, the model can be estimated controlling for both household and village level effects by following step-wise regression analysis. Covariates are retained in the model according to their statistical significance. The variables with low t-values are removed. So, the five variables like household size, ST percentage, SC percentage, WPR, and female literacy rate were identified for further analysis which significantly explained the model. The $\mathbf{R}^2$ for the chosen model was 40.73 %. The consumption model used is

$$Y_i = X_i^I \beta_i + U_i$$

Where Y_i = Percentage of poverty in a district

X_i =Vector of auxiliary variables (covariate) selected from Census

β =Regression co-efficient calculated form sample survey for estimation of variables.

U_i = The area specific random effect

Diagnostic Testing of the Model

As it is very often the case with cross section data, presence of the heteroscedasticity can create potential problems. Presence and absence of hetero determines the efficient estimator for the model specification derived at the end of regression test. If heteroscedasticity is absent, then the OLS is efficient estimator.

The “hettest” has been applied and found that there is no heteroscedascity was accepted: so we estimate the models by OLS to get efficient estimates at the district level. ***Table 4: Table 5***

Post-Diagnostic Procedures on Model-Based Estimation (SAE)

- The aim of the diagnostic procedures used to validate the reliability of the *model based small area estimates vrs. direct survey estimates.*
- Generally, two types of diagnostic procedures are used in SAE, ie.
    - Model diagnostics
    - Small area estimates validation /diagnostics .
- The model diagnostics are used to verify the assumptions of underlying the model
- The second diagnostics are used to validate the reliability of the model-based SAE.
- Model-based estimates should be consistent, more precise,more stable and acceptable

Biased Diagnostics

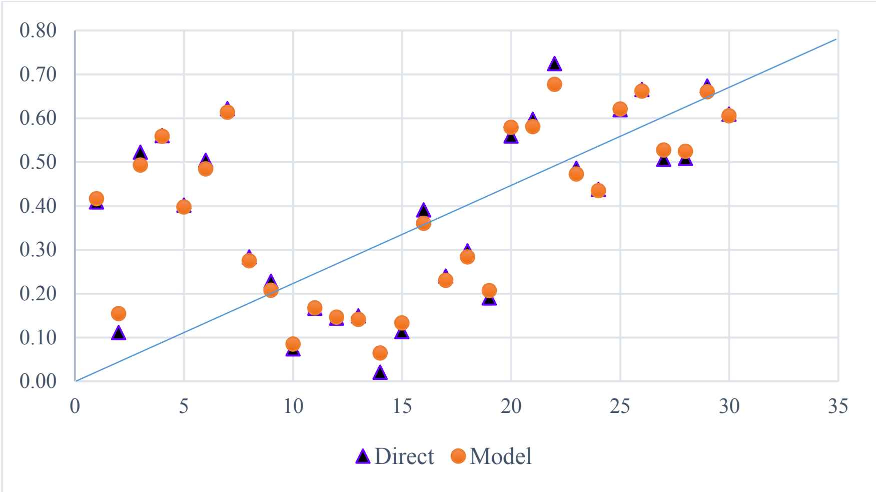
- The bias diagnostics is used to assess the deviations of the model – based estimates from the direct survey estimates.

- The model-based estimates are expected to be biased predictors of the direct estimates.

- The model-based estimates will be unbiased predictors of the direct survey estimates if the relationship between the variable of interest and the covariates have been mis-specified or mis-estimated. Where, the relationship has not been mis-estimated, a linear relationship of the is expected between the direct survey estimates and the model-based estimates.

- The ***Figure-2*** below shows the biased scatter plot of the direct estimates against the model-based with the fitted regression line and the y= x line. The plots show that the model-based estimates are less extreme as compared to the direct estimates.

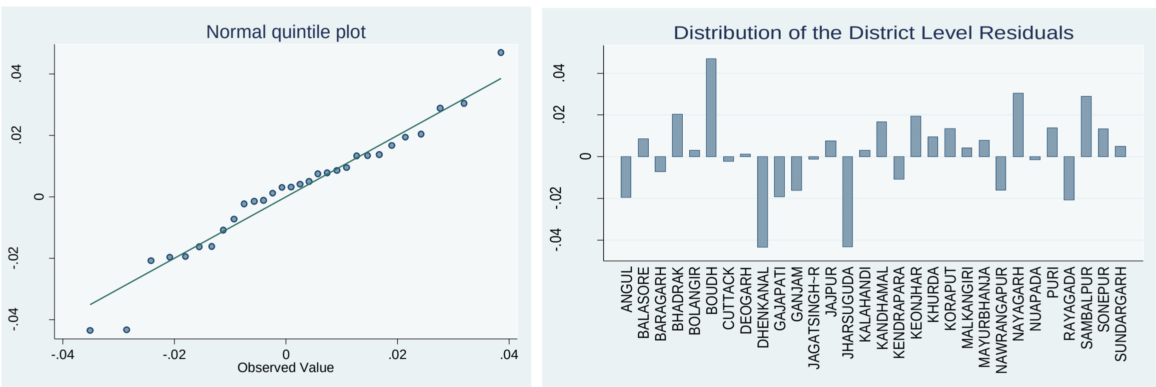
Figure 2 : Biased Scatter Plot of the Direct and Model-Based Estimates



Model Diagnostics

- The distribution of the district level residuals and q-q plots are shown in ***Figure-3*** below
- This reveals that the randomly distributed district level residuals and the line of fit does not significantly differ from the line y=0 as expected in all the plots.
- The q-q plots also confirm the normality assumption. Therefore, the model
- diagnostics are fully satisfied for the data.

Figure 3 :q-q Plot



Model based estimates is statistically accepted

Coverage Diagnostics

- Coverage Diagnostics measures the overlap between the 95 % confidence intervals of the direct estimates and those of the model based estimates.

- This diagnostics is aimed at evaluating the validity of the confidence intervals generated by the model-based procedures.

- In ***Figure-4*** the district wise 95% confidence intervals of the model- based direct estimates is pressented. The standard errors of the direct estimates are too large and therefore the estimates are un-reliable. This gives the degree of inequality with respect to distribution of poor households in different districts of Odisha.

- It is recommended that non-coverage total should not exceed 5%. In this case, there is 100% coverage between the intervals of the model-based estimates and direct survey estimates. This indicates that the method is statistically accepted.

Figure 4 : District wise 95% confidence intervals of the model-based estimates

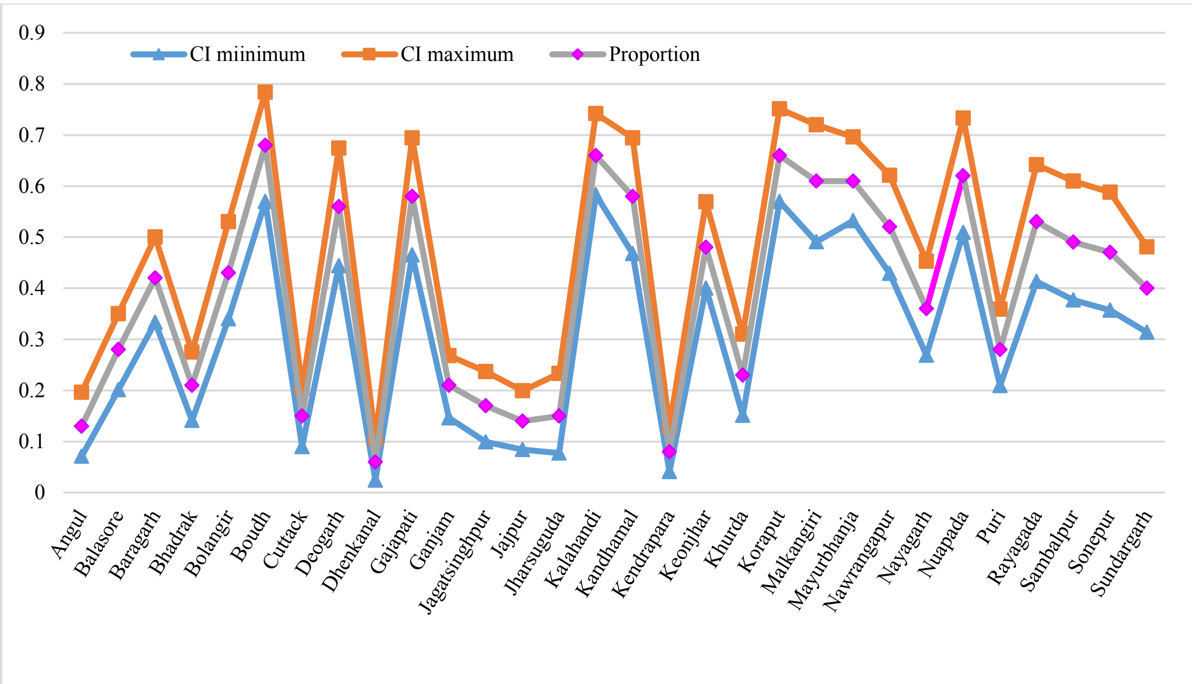


Table 4: Model-Based Estimation of Poverty (District wise)

Sl. No.	Distrcit	Proportion	Std.error	CV (%)	95% Conf.Interval	
1	Angul	0.13	0.03	23.86	0.07	0.20
2	Balasore	0.28	0.04	13.73	0.20	0.35
3	Baragarh	0.42	0.04	10.17	0.33	0.50
4	Bhadrak	0.21	0.03	16.49	0.14	0.28
5	Bolangir	0.43	0.05	11.11	0.34	0.53
6	Boudh	0.68	0.06	8.05	0.57	0.78
7	Cuttack	0.15	0.03	19.86	0.09	0.20
8	Deogarh	0.56	0.06	10.49	0.44	0.67
9	Dhenkanal	0.06	0.02	32.24	0.02	0.11
10	Gajapati	0.58	0.06	10.06	0.47	0.69
11	Ganjam	0.21	0.03	14.97	0.15	0.27
12	Jagatsinghpur	0.17	0.04	21.04	0.10	0.24
13	Jajpur	0.14	0.03	20.63	0.08	0.20
14	Jharsuguda	0.15	0.04	25.72	0.08	0.23
15	Kalahandi	0.66	0.04	6.12	0.58	0.74
16	Kandhamal	0.58	0.06	9.93	0.47	0.69
17	Kendrapara	0.08	0.02	26.26	0.04	0.13
18	Keonjhar	0.48	0.04	8.86	0.40	0.57
19	Khordha	0.23	0.04	17.66	0.15	0.31
20	Koraput	0.66	0.05	6.99	0.57	0.75
21	Malkangiri	0.61	0.06	9.65	0.49	0.72
22	Mayurbhanja	0.61	0.04	6.82	0.53	0.70
23	Nawrangapur	0.52	0.05	9.31	0.43	0.62
24	Nayagarh	0.36	0.05	13.01	0.27	0.45
25	Nuapada	0.62	0.06	9.18	0.51	0.73
26	Puri	0.28	0.04	13.45	0.21	0.36
27	Rayagada	0.53	0.06	11.10	0.41	0.64
28	Sambalpur	0.49	0.06	12.02	0.38	0.61
29	Sonepur	0.47	0.06	12.45	0.36	0.59
30	Sundargarh	0.40	0.04	10.71	0.31	0.48

Source : Computed from Primary data of NSSO

Coefficient of Variation (CV)

- The CV for the model based estimates as well as the direct estimates have been calculated to assess the improved precision of the model based estimates compared to the direct estimates **.*(Table-4)***
- The CVs show the sampling variability as a percentage of the estimates.
- Estimates with large CVs are considered unreliable (i.e smaller is better).
- There are no internationally accepted non available that allow us to judge how large is too large.
- It is observed that the CV varies from 6.12 to 32.24 throughout the State across the districts.
- Out of 30 districts of Odisha, 24 districts have coefficient of variations (CV%) less than 20.
- The CV more than 20% in 6 districts like Angul, Dhenkanal, Jagatsinghpur, Jajpur, Jharsuguda and Kendrapara of Odisha implies an unstable estimates.

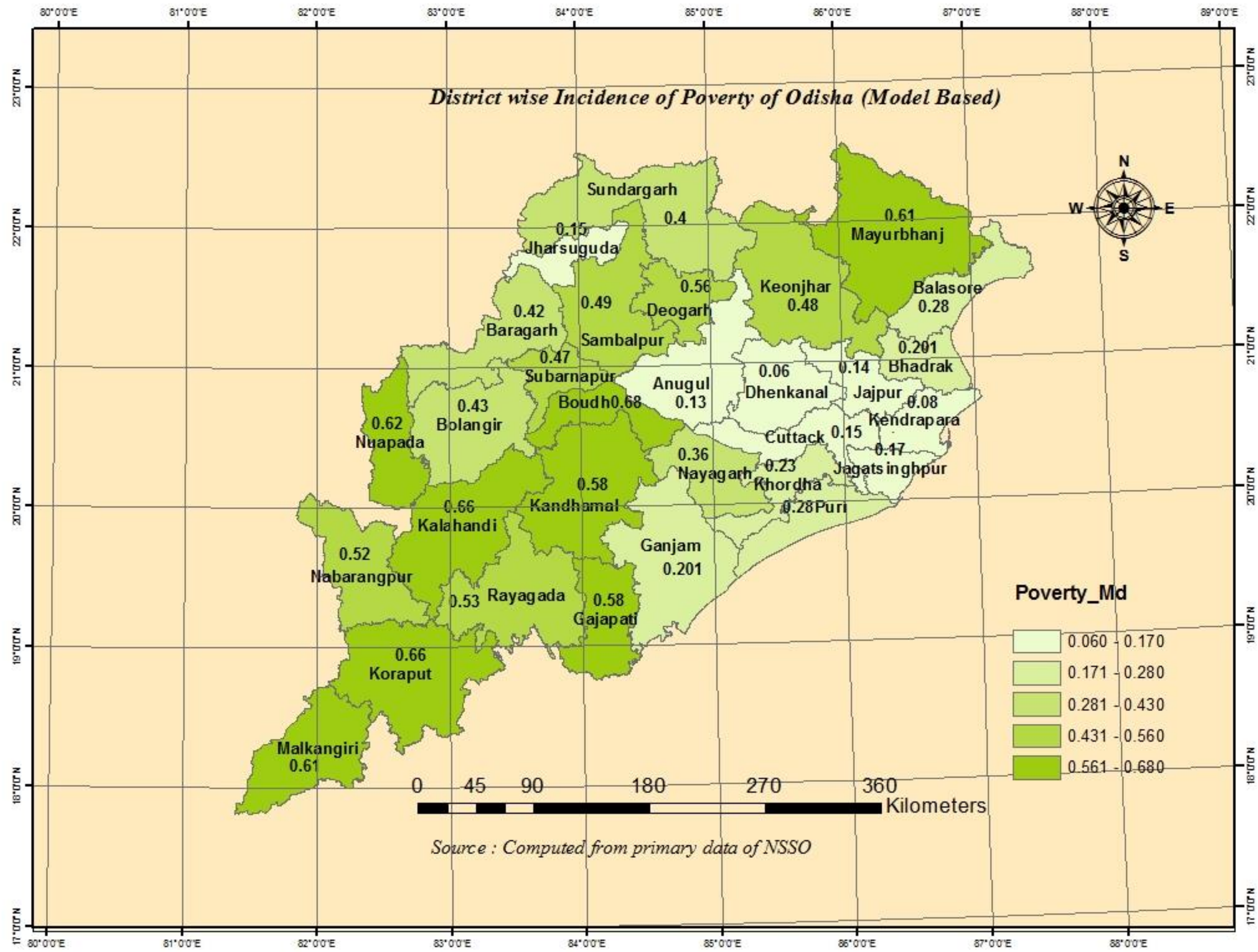
Table 5 : Classification of the Districts according to CV of the Model Based Method.

CV Class	No. of districts	Name of the districts
0-10	9	Boudh, Kalahandi, Kandhamal, Keonjhar, Koraput, Malkangiri, Mayurbhanj, Nawarangpur and Nuapada.
10-20	15	Balasore, Bargarh, Bhadrak, Bolangir, Cuttack, Deogarh, Gajpati, Ganjam, Khordha, Nayagarh, Puri, Rayagada, Sambalpur Sonepur and Sundergarh.
20-30	5	Angul, Jagatsinghpur, Jajpur, Jharsuguda and Kendrapara
30-40	1	Dhenkanal

Source : Computed from Primary data of NSSO

The ***Map-2*** shows the district wise incidence of poverty of Odisha estimated by using Small Area Technique.

Map 2 : Map Showing District wise Poverty of Odisha (Model-Based)



❖❖❖❖❖

Section – IV

Comparison of Poverty Estimate between Direct and Model Based Small Area Estimation Method

Role of SAE in generating district level poverty for grass root level planning for poor people of Odisha

The estimation of district level statistics of poor household has been carried out by using direct (head count), and model-based methods. The proportion of poor and its coefficient of variation (CV) of different methods for the districts of Odisha are presented in ***Table 8*** and ***Figure 5.*** It is observed that in many districts the lower bound of confidence interval (CI) is negative in case of all methods except model-based method. This results in practically impossible for other methods. But it is found that model-based estimates have precise CI and reasonable CV percentage which is reliable. So, small area estimates (SAE) plays an important role in generating district level poverty estimates in Odisha.

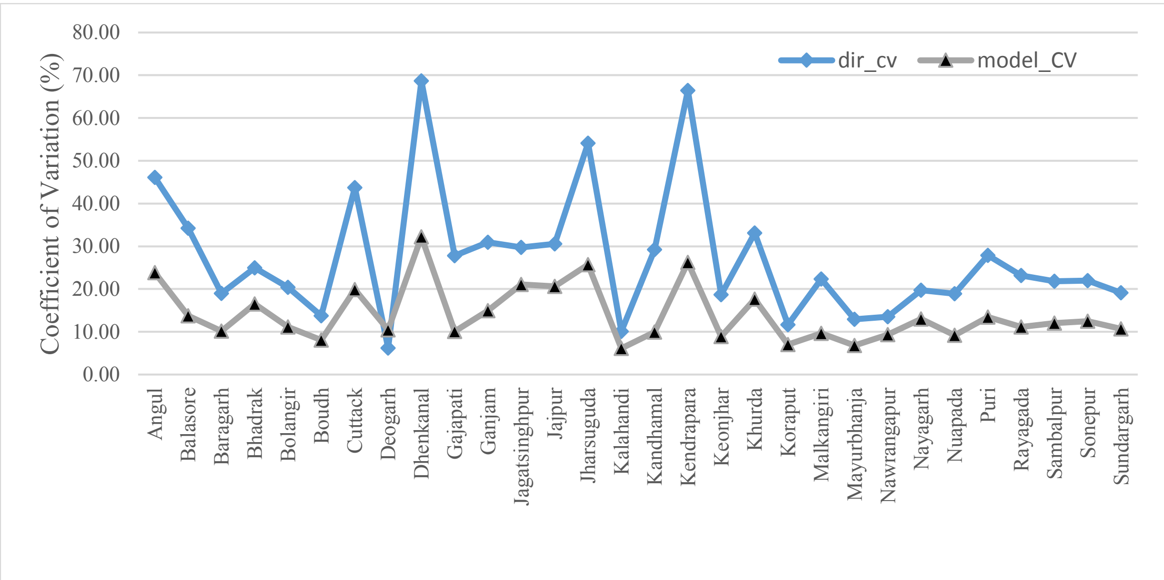
Table 6: Estimation of Poverty (District wise)

Sl. No.	Distrcit	Direct		Model-based	
		Proportion	**CV%**	**Proportion**	**CV%**
1	Angul	0.11	46.14	0.13	23.86
2	Balasore	0.28	34.26	0.28	13.73
3	Baragarh	0.41	19.01	0.42	10.17
4	Bhadrak	0.23	24.97	0.21	16.49
5	Bolangir	0.44	20.42	0.43	11.11
6	Boudh	0.72	13.72	0.68	8.05
7	Cuttack	0.15	43.74	0.15	19.86
8	Deogarh	0.56	6.22	0.56	10.49
9	Dhenkanal	0.02	68.69	0.06	32.24
10	Gajapati	0.56	27.83	0.58	10.06
11	Ganjam	0.19	30.94	0.21	14.97
12	Jagatsinghpur	0.17	29.72	0.17	21.04
13	Jajpur	0.15	30.58	0.14	20.63
14	Jharsuguda	0.11	54.08	0.15	25.72

Sl. No.	Distrcit	Direct		Model-based	
		Proportion	CV%	Proportion	CV%
15	Kalahandi	0.67	10.07	0.66	6.12
16	Kandhamal	0.60	29.24	0.58	9.93
17	Kendrapara	0.07	66.44	0.08	26.26
18	Keonjhar	0.50	18.67	0.48	8.86
19	Khordha	0.24	33.10	0.23	17.66
20	Koraput	0.67	11.65	0.66	6.99
21	Malkangiri	0.61	22.33	0.61	9.65
22	Mayurbhanja	0.62	12.96	0.61	6.82
23	Nawrangapur	0.51	13.49	0.52	9.31
24	Nayagarh	0.39	19.72	0.36	13.01
25	Nuapada	0.62	18.93	0.62	9.18
26	Puri	0.30	27.87	0.28	13.45
27	Rayagada	0.51	23.19	0.53	11.10
28	Sambalpur	0.52	21.81	0.49	12.02
29	Sonepur	0.49	21.96	0.47	12.45
30	Sundargarh	0.40	19.10	0.40	10.71

Source : Computed from Primary data of NSSO

Figure 5 : District wise coefficient of variation for direct and model-based estimates



- The Figure-5 presents the district wise distribution of the percentage CV of model-based estimates and direct estimates.

- The CV varies from 6.12 to 32.24 which is reliable in case of model-based estimates. The estimated CVs show that model-based estimates have a higher degree of reliability and stable as compared to the direct estimates.

- It is observed in direct estimate of poverty that in many districts the lower bound (Lower) of 95% confidence interval (CI) is negative which results in practically impossible and inadmissible values of CI for direct estimates.

- But, model-based estimates with precise CI and reasonable CV percentage are reliable. So, SAE plays an important role in generating micro level statistics.

- The results show the advantage of using SAE technique to cope up the small sample size problem in producing the estimates or reliable confidence intervals.

Section – V
Summary and Conclusions

Reliable estimates of poverty at district levels in Odisha are not available except the estimates based on headcount, which are accompanied by large sampling error due to small sample sizes allocated for the districts. This leads to unreliability of the poverty estimates at district levels. In this context the recently developed small area estimation pioneered techniques developed by *Rao (2003), Ghosh and Rao (1994), Saei and Chambers (2003), Manteiga et al. (2007) and Chandra et al. (2011)* have been applied to capture the district level poverty for Odisha.

The NSSO survey contributes to provide estimates on a regular basis at the State and regional level not beyond sub-State level statistics due to heterogeneity nature. This gives little information for micro level planning and allocation of resources.

The method of estimation of poverty proportion for small areas using reliable small area estimation technique is well developed and practised widely in many countries of the world. But, there is very less known application to the Indian data and no application to the valuable and informative NSSO data for Odisha. In this study much acclaimed SAE techniques have been applied to estimate the district level estimates of poor households using NSSO survey and Census data.

The present analysis using model-based method is found to be very effective for computing district level estimates of proportion of poor households in Odisha by using 68^{th} round NSSO survey rural sector consumer expenditure data and 2011 Population Census. The diagnostic procedures have confirmed that the model-based district level estimates have reasonably good precision.

A lot of emphasis is being given to micro level planning all over the World as well as in India. In India the district is an important domain for planning process in the State and therefore, availability of the district level statistics is vital to monitoring the policy and planning. This study produces reliable statistics at micro level using existing surveys and other already available auxiliary variables and may be seen as a modest attempt in this direction in Odisha. This exercise can be undertaken without conducting micro level specific survey which may involve a lot of financial burden to the State exchequer and can provide the micro level estimates for important socio-economic and demographic parameters on regular basis.

This is the beginning for application of small area technique to estimate district level poverty. Further improvement in the methodology need be explored by using different models like unit level and area level. For unit level model, availability of unit level data from Census and other socio economic studies may be ensured.

Way Forward

1. This is the beginning for application of small area technique to estimate district level poverty.

2. Further improvement in the methodology need be explored by using different models like unit level and area level.

3. For unit level model, availability of unit level data from different Census as may be ensured. Socio-economic Caste Census data of Ministry of Rural Development, Government of India is suitable for unit level model.

4. To get better estimate we may merge the unit level data of 66^{th} and 68^{th} round of NSS as these two rounds are closer to each other.

5. Broad group wise MPCE data collected through employment and un employment survey of NSS may be used to merge with consumer expenditure survey data to get reliable poverty estimate at district level.

6. To improve the methodology in the context of Odisha, a Research Committee comprising two experts from IASRI headed by Dr. Hukum Chandra, two officials of DES, Odisha and expert of PHADMA may be constituted.

❖❖❖❖

Annexure – 1

Comment on Report entitled "District Level Poverty Estimation for Odisha by using Small Area Estimation Technique" by DES Odisha

This report presents an application of small area estimation (SAE) techniques for estimation of poverty incidence (i.e. the proportion of poor households at district level in the state of Odisha). The SAE is based on 68th round (2011-12) survey data of NSS on Household Consumer Expenditure conducted during 201-12 and Population Census 2011. The SAE methodology applied in this analysis, i.e. SAE under a logistic linear mixed model with aggregated (or area) level data is appropriate. In particular, in this report SAE method is based on fitting a logistic linear mixed model, with fixed effects corresponding to covariates that are available from external sources (in this case, a census 2011), and with random effects corresponding to the districts. The exposition is clear and the methodology is correct. Results, including diagnostics, look sensible. In short, this appears to be a useful application and illustration of existing methodology. The reliable district level estimates generated in this report and also poverty mapping will definitely be useful for the state Govt in proper planning, fund allocation and also monitoring of various plans.

I appreciate the DES Odisha for this excellent initiated in generating district level estimates.

Suggestion for future and further work

DES can think of generating SAE estimates for further disaggregate level, for example, the domains can also be formed by crossing social groups with geographic districts, Different land category wise classification etc in the state of Odisha.

Model based small area estimates generated by SAE are valid and reliable in most of the districts. But, in some district still there exist scope to improve. In this spatial version of small area models can be used for further improving district level estimates.

Similar SAE methodology can be used for generating district level unemployment rate from Employment and Unemployment survey data.

It is suggested that small area estimation method should also be applied for estimation of district wise yield and area estimation using data of agriculture survey available for the state. But, in this case, appropriate SAE methodology has to be chosen.

Dr. Hukum Chandra

National Fellow & Principal Scientist

ICAR-Indian Agricultural Statistics Research Institute (IASRI)

Library Avenue, PUSA, New Delhi-110012 (INDIA)

Telephone (Office) : 0091-11-25843398

Telephone (Home) : 0091-11-47571929

Mobile : 0091 9999450020

Fax : 0091-11-25841564

E-mail : hchandra12@gmail.com, hchandra@iasri.res.in

website : www.iasri.res.in

***

Annexure – II

Recommendation of the Working Group on District Planning, under the Chairmanship of C.H. Hanumantha Rao

21.24 However, the mere disaggregation of funds to the district level and showing 'District Allocations' in the State Plan document will not amount to true decentralisation of the planning process, if the powers to plan for these funds are concentrated at the State level with the various heads of departments. The touchstone of decentralisation would be the freedom available at the district level to plan according to its local needs and local potentials. In this context, much greater stress needs to be placed upon the delegation of powers and to local resources development and the stimulation of local initiatives within the community. Governmental and administrative structures have to be better adapted to these goals.

21.25 The Planning Commission set up a Working Group on District Planning in September 1982. It has advocated a gradual, step-by-step approach towards this goal. It has suggested that Panchayati Raj Institutions (PRIs), wherever they exist, will function as the apex body to approve as well as review the implementation of District Plans. Where the PRIs either do not exist or exist without planning functions, the District Planning Boards/Councils, which already exist or are being constituted, can perform these functions. For this purpose, they will be assisted by a suitable planning machinery at that level and the Centrally-sponsored scheme of strengthening of State-level planning machinery has been extended to district level also, with effect from 1982-83.

21.26 In order to ensure effective horizontal coordination at the district level, the Working Group on District Planning has made a number of recommendations which include:

(a) Strengthening of the position of the District Collector;

(b) Placing departmental functionaries under the direct administrative control of the Collector by deeming their services to be on deputation from their departments; and

(c) Making district officials accountable to the District Planning Body.

21.27 The Working Group has also made several recommendations towards improving administrative decentralization. They bear on the procedures for administrative and technical sanctions, reappropriation measures, etc. The Working Group has recommended a participatory approach to the district planning process, such participation extending through the stages of pre-planning, Plan formulation, post-planning and implementation stages. The Planning methodology for district planning has been outlined, and a massive training programme has been envisaged for the purpose. In the `stages approach' advocated by the Working Group, Stage I will be a phases of `initiation', Stage-II will be one of `limited decentralisation', and Stage-III will be the culmination. An illustrative list of the various measures that could be taken during each stage is given in Annexure 21.1. The Working Group on District Planning has envisaged that all States in the country should reach the final phases by the year 2000 AD but as far as possible even earlier by, say, compressing Stages, I and II by the end of the Seventh Plan period itself.

21.28 During the Seventh Plan period, the decentralisation of the planning process and full public participation in development will be pursued on the lines suggested by the Working Group. These steps will seek to ensure the achievement of the twin objectives of effective implementation of the anti-poverty programmes and ensuring a balanced regional development at least in respect of the minimum needs. District Planning, as advocated in the report of the Working Group, should be vigorously pursued. Eventually the decentralisation of planning should be extended further to the block level, particularly for the more effective implementation of anti-poverty programmes.

21.29 The initiatives taken by several States on the Sixth Plan suggest that the initial resistance to decentralisation seems to have been overcome. But since for a majority of the States the decentralisation of the planning process is a new experiment, the Planning Commission propose to play a promotional and guiding role in order to impart momentum to the district planning process. This is proposed to be attempted in the following directions:

(i) Data and information system for district planning: it is proposed to update the natural resources inventory data at the district level through use of remote sensing techniques. About 100 districts in the country are proposed to be selected during the Seventh Plan period to provide a scientific basis for district planning. The institutions to undertake this work will be identified and their efforts will be coordinated,

(ii) Training for district planning: The existing scheme of strengthening the district planning machinery will be countinued during the Seventh Plan period.

(iii Training in district Planning: The task of training district-level personnel is a stupendous one. This is proposed to be achieved through massive training programmes organised by the Centre and the State Governments in selected institutions.

(iv) Technical guidance through Pilot Projects: It is also proposed to initiate a few pilot projects, including action-based research, in different States in the country to improve methodologies and procedures for decentralised district planning. This may be particularly necessary in State which are beginning the district planning exercise for the first time.

Monitoring

21.30 The essence of a good monitoring system is the speed of communication of dependable information on key result areas, the competence of the monitor to interpret the signals and the ability to lead to initervention in a constructive manner. Monitoring is not an end in itself, and, therefore, it has to be suited to the objectives. For instance, in large projects where time is of essence, monitoring has to be intensive at all levels whereas for other projects it may be intensive at the field level but selective at the higher levels. In some other activities monitoring of trends may be of more importance than that of actual details.

21.31 Monitoring has several aspects and it is necessary to identify the scope for monitoring for effective plan implementation. Broadly speaking, monitoring would cover following areas:

(a) Physical progress of implementation of projects involving civil construction, equipment erection and commissioning within time and cost schedules, e.g., irrigation canals and drains, industrial plants, power projects, etc.

(b) Quantitative and qualitative progress of implementation of programmes where physical targets are set, e.g., MNP, IRDP, NREP Hill area programme, etc.

(c) Production, productivity and profitability performance for established public sector unit in the core sector, for which key indicators specific to the units concerned may be identified.

(d) Maintenance of capital assets created to be monitored selectively so that the expenditure earmarked for this purpose in the State and Central budgets (though on the non-Plan side) is in fact utilised for e purpose.

(e) Plan expenditure-to ensure that sectoral outlays are not disturbed and outlays earmarked for specific project are not diverted for other purposes without compelling reasons.

In these areas, it will be necessary to spell out the responsibilities of the concerned enterprises, Ministries and the Planning Commission.

21.32 Monitoring will continue to be undertaken through reports, review meetings and field visits. However, the information content, channels of communication, frequency, presentational formats (including presentation through charts and graphs and other means), etc., will have to be reframed according to the nature, type, size and importance of projects, programmes and levels of monitoring, and after taking into consideration the following experience in the past:

(i) The primary responsibility of monitoring lies with the agency entrusted with the execution of the project/programme.

(ii) The farther away the monitoring level from the field, the greater is the need for selectivity in the span and items of monitoring.

(iii) Monitoring should lead to intervention, corrective as well as supportive, for resolving problems arising at site.

(iv) Over-reporting and overlapping at different moni- toring levels can cause confusion and tend to become counter-productive.

(v) Timeliness of reporting is more important than absolute accuracy, especially at higher levels of monitoring.

(vi) Data reporting has to be supplemented by direct discussions and field visits.

21.33 There are many hierarchically linked monitoring agencies and if reporting is channelled through these agencies, not only will it be delayed but the information is also likely to be distorted. There is, therefore, a need for direct communication of key information from the field level direct to each monitoring level, giving the information relevant to it. Such level jumping in monitoring can considerably reduce the time lag in reporting to higher levels.

21.34 As regard the specific steps for strengthening the monitoring and information systems, the following suggestions are made:

(a)Central Level

- Each Ministry/Department should review their existing monitoring system in order to remove the deficiencies and problem areas. For this purpose, small working groups headed by a 'nodal' officer of the Ministary should be set up to review the existing system. The main emphasis should be on simplification and usefulness of the system.

- Establishment of a computerised Data Bank within each Department/Agency should be undertaken on priority basis and the monitoring system should be linked to the Data Bank. Necessary hardware and software should be provided for these Data Banks and monitoring systems.
- It should be ensured that Monitoring Units are established and properly staffed in each Department/Organisation/Project depending on its requirement.
- Monitoring should be recognised as a specialised function requiring necessary professional skills. Intensive training efforts will be needed for this purpose.
- For selected large-scale projects, interministerial groups can be set up in the form of empowered committees with representatives from inter-linked implementation agencies, which could closely monitor the project and also take decisions for corrective action.
- For other major projects, field visits from Ministries and other concerned agencies should be regularly made for monitoring and providing assistance to project authorities.
- Maintenance of records should be improved. Where possible micro-filming should be used.
- Resource-based network systems for implementation planning and monitoring should be extended to all sectors.

(b) State Level

- For the programmes/projects executed by the State Governments (both Centrally-sponsored and State Plan) the monitoring at the Central level cannot be effective unless the basic monitoring facilities are effectively developed at the State, sub-state level on a uniform pattern with modern facilities for communication, processing, storage and retrieval of data.

- While the basic monitoring of the State Plan programmes/schemes has to be undertaken at State and/or sub-State/District level, the Centre may confine itself to monitoring the earmarked projects/programmes which are of vital importance for the country as a whole. These projects will be primarily in the Agriculture and Allied Services, Cooperation, Irrigation and Power and Minimum Needs Programme. In addition, the Central monitoring will cover the Centrally sponsored schemes. It may be helpful if a system of mid-year review of the progress under the approved Plan is introduced, where physical progress and performance under various programmes and projects could be reviewed.

- Monitoring units should be set up in important technical departments, State Public Sector Organisations and important projects. Use of computers and data processing facilities for the monitoring system should be made as widely as possible and computerised data bank and information systems should be developed. The Central scheme of providing 2/3rd assistance should cover all these aspects.

- A list of inventory of all schemes in the State Plan should be developed, sub-divided by sectors and districts. The physical aspects to be monitored in each scheme should be identified.

- Considering the large number of schemes at the State level, there is need for selectivity and differentiation in the intensity of monitoring in terms of frequency and degree of detail. The Plan schemes could be divided into, say three categories depending on investment, importance and critical linkages. For the most important category of selected schemes, the monitoring should be more frequent and intensive, while for the less important schemes, which may be large in number, monitoring may be less frequent and in lesser detail.

- The concept of implementation planning with the help of resource-based networks should also be introduced for State projects/programmes in the same manner as it is being used for the major Central Sector projects and some State power projects.

- Monitoring should be linked with the review and problem-solving mechanism so that the results of monitoring are considered by the concerned decision- making levels and used in initiating corrective action. The monitoring system should use modern techniques such as PERT/CPM and the periodical monitoring should be linked to the PERT/CPM analysis at the time of Annual Plan formulation.

- There may be an advantage in developing arrangements which ensure public participation and scrutiny in monitoring of beneficiary-oriented programmes.

- The basic records at the field level should be streamlined and simplified. The example of the Health Information System introduced in 1982 could be considered for other sectors.

(c) District Level

- The monitoring at District level is particularly important for the beneficiary-oriented schemes and those schemes requiring greater efforts at the Block and Field levels. The current scheme for strengthening and Planning machinery at the District level for which 50 per cent Central assistance is provided, covers the monitoring function but does not specify a separate functionary for this purpose. In order to ensure effective monitoring at the District level, it will be necessary to have a small monitoring machinery either in the District Planning Cell or by combining with the District Statistical Organisation. Some monitoring arrangements will also be needed at the Block level.

- There is a need for coordinated monitoring of several similar activities in the same geographical area but falling under different programmes/ schemes.

(d) Project Level

- For each major project/programmes, monitoring mechanism should be inbuilt as a part of the project implementation and specific provision should be made for monitoring unit/staff in the project estimates.

(e) General

- The Integrated Classification for planning, budgeting and accounting which has been developed by an Inter- Departmental Group set up by the Ministry of Finance and is to be introduced during the Seventh Plan, will provide a common list of major, sub-major and minor heads to be used for planning as well as accounting purposes. It will then be easier to monitor the expenditure under each Plan scheme which would be shown at one place in this classification.

- For improving data handling, including processing, storage, retrieval and communication, greater use of computers is envisaged during the Seventh Plan. It is expected that each major organisation such as Ministry/Department/Agency at the Central and State levels would develop and streamline its management information system, including a computerised data bank, located at one place or at various places, but inter- linked in such a manner that data flow is possible from one data bank to another. Later, this should help in replacing the flow of reports from one agency to another by transfer of data on computer lines, provided either by the P & T system or through satellite communication system. This requires considerable efforts in each Department/Agency in streamlining the existing data system and developing a Computerised Data Bank. The National, Regional and State Centres of the National Informatics Centre could be utilised for these data banks. It should be ensured that these data banks are compatible with each other so that the data could be easily transferred.

❖❖❖❖

References :

Battese, G.E., Harter, R.M., and Fuller, W.A. (1988): "*An Error Component Model for Prediction of Country Crop Areas Using Survey and Satellite Data*", Journal of the American Statistical Association, 83, 28 - 36.

Chandra, H., Sud, U.C. and Salvatri, N (2011):" *Estimation of District Level Poor Households in the State of Uttar Pradesh in India by combining NSSO survey and Census data"*, Journal of the Indian Society of Agricultural Statistics, 65 (I) 1-8

Fay, R.E and Herriot, R.A. (1979): "*Estimates of Income for Small Places: an Application of James Steir Procedures to Census Data*", J. Amer. Statistics Association, 85, 398-409.

Ghosh, M. and Rao, J.N.K. (1994): "*Small Area Estimation and Appraisal (with discussion)"*, Statistical Science, 9, 65-93.

Gonzalez, M.E. (1973): "*Use and Evaluation of Synthetic Estimators*", Proceedings of the Social Statistics Section, American Statistical Association, 33 - 36.

Holt, D., Smith, T.M.F., and Tomberlin, T.J. (1979): "*A Model Based Approach to Estimation for Small Subgroups of a Population*", Journal of the American Statistical Association, 74, 405 – 410.

Hentschel, J., Lanjouw, J.O., Langouw, P. and Poggi, J. (2000): "Combining Census and Survey Data to Trace the Spatial Dimensions of Poverty: A Case Study of Ecuador", The World Bank Economic Review, Vol.14, No1, 147-165.

Manteiga, G.W., Lombardia, M.J., Molina, I., Morales, D. and Santamaria, L. (2007): "*Estimation of the Mean Squared Error of Predictors of Small Area Linear Parameters under a Logistic Mixed Model."* Compute Statistical Data Analysis, 51, 2720 – 2733.

Prasad, N.G.N. and Rao, J.N.K. (1990): "*The Estimation of the Mean Squared Error of Small Area Estimators*", Journal of the American Statistical Association, 85, 163 – 171.

Rao, J.N.K. (2003): "*Small Area Estimation*", Wiley Series in Survey Methodology, Jhone Wiley and Sons Inc.

Saei, A. and Chambers, R. (2003): "*Small Area Estimation under Linear and Generalised Linear Mixed Models with Time and Area Effects*", Southampton Statistical Sciences Research Institute, University of Southampton, UK, W.P.No.M03/15.

Elbers, C., Lanjouw, J.O. and Lanjouw, P.V. (2001): "Welfare in Villages and Towns: Micro-level Estimation of Poverty and Inequality", Vrije University, Yale University and the World Bank (mimeo).

United Nations Statement, June 1998.

World Bank (2000): "The Geography of Poverty: Estimation and Analysis of Small-Area Welfare Indicators." Washington DC (mimeo)

.

❖❖❖❖❖