Agricultural Economics Research Review 2020, 33 (1), 23-33
DOI: 10.5958/0974-0279.2020.00003.8

Disaggregate-level disparity in the incidence of poverty in Chhattisgarh, India

Priyanka Anjoy* and Hukum Chandra

ICAR-Indian Agricultural Statistics Research Institute, New Delhi-110 012, India

*Corresponding author: anjoypriyanka90@gmail.com

Abstract Poverty is widespread in India, especially in far-flung rural areas, and disparities exist among and between states and social groups. To design programmes that alleviate poverty, policymakers need disaggregated data. The small area estimation (SAE) technique using the hierarchical Bayes model generates representative, micro-level estimates of poverty incidence. The results of a study in Chhattisgarh show that the SAE-based estimates are precise, and can help the government formulate micro-level anti-poverty strategies.

Keywords Sustainable Development Goal (SDG), National Sample Survey Office (NSSO), hierarchical Bayes model, small area estimation (SAE) technique, poverty

JEL codes C810, C890

India's economic growth looks strong, but poverty is still deeply rooted and persistent. In 2011–12, the poverty ratio was 25.7% in rural areas and 13.7% in urban areas; the poor numbered about 269 million people, and 216.5 million (around 80%) lived in rural areas(https://niti.gov.in/planningcommission.gov.in/docs/news/pre_pov2307.pdf, Planning Commission). A disaggregated analysis finds that urban poverty is caused by urban settlements and the assimilation of migrants (Dubey and Tiwari 2018). Rural poverty may be caused by hunger, undernourishment, education, health, or unemployment. The distribution of, or reduction in, poverty varies by socio-economic category (Thorat and Dubey 2012).

If policy is to alleviate rural poverty, it needs concrete, disaggregated statistics on rural poverty that lays out disparities by region and group. National and state-level statistics mask local variations, and with the recent emphasis in India on decentralizing governance and economic planning, the need and demand has increased for micro-level, disaggregated statistics on socio-economic conditions, infrastructure, and institutions. Policymakers need such statistics to target social and spatial heterogeneity at higher levels of spatial disaggregation, monitor and evaluate parameters within and across local administrative units, and design programmes and strategies to alleviate inter-personal and inter-regional inequalities.

In the social hierarchy, Scheduled Tribes (ST) and Scheduled Castes (SC) are the most marginalized and deprived; Other Backward Classes (OBC) are in between the SCs and STs and the General category. Using the countrywide Household Consumer Expenditure Survey (HCES) of the National Sample Survey Office (NSSO), researchers have tried to capture the poverty level in India by estimating the macro-level poverty ratio or head count ratio (HCR). Chaudhuri and Gupta (2009) use the HCES 2004–05 data of the NSSO to estimate poverty by district, but a problem in their approach is that sample sizes generated large standard errors in the estimates of several districts.

Chauhan et al. (2016) study the intra- and inter-regional disparities in poverty and inequality using three quinquennial rounds of HCES data of NSSO over two decades (1993–2012), and Mohanty et al. (2016) report the district-level estimates within the region. Both studies use the traditional direct estimation and synthetic estimation approach in measuring the poverty proportions at disaggregate or local levels, while estimating the poverty indicators by fitting the regression-based fixed effects model, and therefore their estimates fail to represent the dissimilarities across areas.

The small area estimation (SAE) technique considers the random area-specific effects and potentially explores the variability between areas (Anjoy et al. 2018). This study uses the SAE technique to estimate disaggregate-level poverty in the state of Chhattisgarh. We chose Chhattisgarh because it is the poorest state in India: 40% of the people live below the poverty line (World Bank 2016); large parts of the state are conflict-affected, or they lie in remote areas, and are excluded from the benefits of development (Gebert et al. 2011); and the progress of the large numbers of the socially marginalized poor is limited.

Little reliable data is available from districts or smaller administrative units on poverty or its impact (Chaudhuri and Gupta 2009). Measuring poverty at the micro or local level using the SAE technique will help policy analysts draw acceptable inferences for representative small areas, target interventions and develop policy, and decentralize planning and monitoring for reducing poverty.

Theoretical background

The SAE technique links the target variable from a survey with the auxiliary information available for small domains or areas from other data sources (such as the record of a census or other administrative source). The technique tries to improve domain predictions by borrowing strength from related small areas under a mixed modeling framework, and its estimates are much more precise than that of the traditional direct estimation technique (Chandra et al. 2018a; Rao 2003). Based on the availability of unit-level auxiliary information, small area models are classified into area-level models, which use aggregated auxiliary and target information, and unit-level models, which are based on unit-specific variable information.

Obtaining unit-level auxiliary information may involve additional expense, therefore, small area models are mostly considered. The Fay–Herriot model is most widely used to estimate a small area mean or total (Fay and Herriot 1979). The Fay–Herriot model and its various extensions have been applied to solve small domain estimation problems under both the frequentist and Bayesian paradigms (Chandra et al. 2011; Chandra 2013; Portar et al. 2014; Pratesi and Salvati 2016; Chandra et al. 2017; Chandra et al. 2018b; Anjoy et al. 2018).

In most practical applications, however, the target variable is either binary or count, rather than continuous (Chaudhuri and Gupta 2009; Chandra et al. 2011), and the aim is to estimate small area proportions or count. It would not be appropriate to apply the Fay–Herriot model in these cases, and a potential alternative is the generalized linear mixed model, which is the logistic normal mixed model for binary data and the log normal mixed model for count data. Again, estimation problems involving small area predictors for proportion or count often ignore the underlying sampling mechanism, but small area models that do not allow to incorporate the available survey information may produce biased estimates (Liu et al. 2014; Hidiroglou and You 2016; Anjoy et al. 2018).Therefore, we prefer modeling the survey-weighted estimator. Specifically attempt is to model survey-weighted proportions—poverty proportions (HCR)—through the logistic normal mixed model to draw needful small domain inferences. In this paper, we delineate two alternative models of the logistic normal mixed model under the Bayesian framework, and we apply these to produce reliable estimates of poverty among social groups in the rural districts of Chhattisgarh by linking data from the HCES 2011–12 of NSSO and the Population Census 2011.

Assume U denotes the finite population of interest of size N, which is partitioned into M disjoint small areas, and a sample s of size n is drawn from this population with a given survey design. The set of population units in area i is denoted as U_i with known size N_i, such that $U=\cup_{i=1}^{M} U_i$ and $N=\sum_{i=1}^{M} N_i$. Where N_i and n_i respectively being the population and sample size from small area i (i=1,…,M), the units making up the sample in small area i are denoted by s_i, so that $s=\cup_{i=1}^{M} s_i$ and $n=\sum_{i=1}^{M} n_i$. Let, y_{ij} be the binary response for unit j (j=1,…,n_i) in small area i. We wish to estimate the small area proportions, $P_i=(N_i)^{-1}\sum_{j=1}^{N_i} y_{ij}$. The "small areas" or simply "areas" refer to the districts classified by the social groups of Chhattisgarh. The direct survey estimator for the proportion of poor households in area i (i=1,…,M) is defined as,

$$p_{iw} = \left(\sum_{j=1}^{n_i} w_{ij}\right)^{-1} \sum_{j=1}^{n_i} w_{ij} y_{ij}$$

where w_{ij} denotes the survey weight attached to the sampling unit (household) j in the area i. The estimate of variance of p_{iw} can be expressed as,

$$\hat{var}(p_{iw}) = \left(\sum_{j=1}^{n_i} w_{ij}\right)^{-2} \left\{\sum_{j=1}^{n_i} w_{ij}(w_{ij}-1)(y_{ij}-p_{iw})^2\right\}$$

The basic structure of the logistic normal mixed model has two components: the sampling model for the direct survey estimates, and the linking model to incorporate area-specific auxiliary variables through a linear regression framework. For estimating the small area proportions P_i, the sampling model for p_{iw} is

$p_{iw} = P_i + e_i$; i = 1,...,M

where, e_i's are independent sampling error assumed to have zero mean and known sampling variance σ^2_{ei}. The linking model of P_i can be written as,

$\text{logit}(P_i) = \mathbf{x}_i^T\boldsymbol{\beta} + v_i$; i=1,...,M

where $\mathbf{x}_i$ represent matrix of area-specific auxiliary variables, $\boldsymbol{\beta}$ is the k component regression parameter vector and v_i being the area-specific random effect, independent and identically distributed as $E(v_i) = 0$ and $var(v_i) = \sigma^2_v$. The random area-specific effects are included in the linking model to account for between-area dissimilarities. Two random errors v_i and e_i are independent of each other within and across areas (districts).

Study area and data description

The nationwide HCES of the NSSO is based on properly designed, stratified, two-stage random sampling with districts as strata; villages are the first stage units and households are the ultimate stage units. The HCES 2011–12 surveyed 1,435 households in 18 districts of Chhattisgarh. We use the monthly per capita expenditure (MPCE) data of the HCES to estimate the poverty numbers.

The estimates of state-wise poverty lines for both rural and urban areas as well as percentage and number of persons those below poverty line are also specified based on MPCE. The variable of interest at the unit (household) level is binary; if a household is poor, the value is 1 or 0, and if a household is not poor the value is 0. A household is categorized as poor if its MPCE is below the state's poverty line. The Planning Commission (http://planningcommission.nic.in/news/press_pov2307.pdf) set the poverty line for rural areas in Chhattisgarh at INR 738 for the year 2011–12.

The parameter of interest is the HCR, or proportion of poor households, by district and category. Hereabouts Table 1 presents the sample sizes of social groups in the HCES 2011–12. The sample size ranges from 8 to 160, and the average of districts is 80; further categorization into social groups makes the sample sizes too small, even zero for certain caste groups. Evidently, the direct survey estimation approach, which is based on only domain-specific sample data, fails for such districts. The SAE technique lets us obtain precise small area estimates not only for districts with negligible sample sizes but also for non-sample districts, where the direct estimation approach is typically inapplicable.

The auxiliary variables used in the study are drawn from the Indian Population Census, 2011. We carried out a preliminary data analysis to select the appropriate covariates for the SAE modeling. We followed the steps laid out in Anjoy et al. (2018) in examining the correlation of each of the covariates with the target

Table 1 Social groups and their sample sizes in Chhattisgarh

Sample Sizes	Social Group				
	All	ST	SC	OBC	Others
Minimum	8	1	0	0	0
Average	80	27	11	36	5
Maximum	160	89	30	108	15
Total	1435	485	197	656	97

variable (direct estimate of poverty proportion), and we retained the variables with reasonably good correlation for further analysis.

In the first step, we found eight such covariates: the proportion of SC population, proportion of ST population, female literacy rate, gender ratio (gender ratio = female ÷ male), main working population ratio, marginal working population ratio, the proportion of main female agricultural labourers to total female cultivators, and the proportion of marginal female agricultural labourers to total female cultivators.

In the second step, we carried out principal component analysis for a group of variables: main working population ratio, marginal working population ratio, the proportion of main female agricultural labourers to total female cultivators, and the proportion of marginal female agricultural labourers to female cultivators. The first principal component (G1) of the group explained 88.05% of the variability; adding the second principal component (G2) increased it to 96.60%.

In the third step, we fitted a generalized linear model using the direct survey estimates of poverty proportions as the response variable; as the potential covariates we used the proportion of SC population, the proportion of ST population, female literacy rate, gender ratio, G1, and G2.

Finally, we included two variables that significantly explained the model—proportion of ST population and G1—for small area estimation.

Hierarchical Bayes framework for measuring poverty incidence

We used the hierarchical Bayes method considering Gibbs sampling approach to estimate small area poverty proportion. Several researchers from the frequentist perspective use a plug-in empirical predictor under the generalized linear mixed model to estimate small area proportions, but this approach requires an analytical expression of the measure of precision that is based on the same approximation (Rao 2003). The Bayesian approach has a strategic advantage: it is implemented by simulation, in which posterior inferences can be summarized meaningfully even after complicated transformation (Jiang and Lahiri 2006; Gelman 2006; Ghosh et al. 2009). The central feature of inferences using the hierarchical Bayes method is the quantification of uncertainty; in particular, a parameter is estimated by the posterior mean, and the posterior variance is taken as the measure of the error or uncertainty of the estimates (You 2008; Lee et al. 2015).

The Bayesian paradigm lets us set up even complex multiple parameter models simply. The hierarchical Bayes method can effectively deal with complex small area models using the Markov Chain Monte Carlo method, which overcomes the computational difficulties of high-dimensional integrations of posterior densities (Liu et al. 2014).

We use the hierarchical Bayes method to explore two cases of the logistic normal mixed model with known sampling variances of the survey-weighted proportions (the HB1 model) and unknown sampling variances of the survey-weighted proportions (the HB2 model).

In the HB1 model, the sampling model is $p_{iw}|P_i \sim N(P_i, \sigma^2_{ei})$ and the linking model is $\text{logit}(P_i)|\beta, \sigma^2_v \sim N(x_i'\beta, \sigma^2_v)$

where σ^2_{ei} is the sampling variance term, which is assumed to be known, generally replaced by direct variance estimate $\hat{\text{var}}(p_{iw})$ calculated using available survey data.

Prior for the hyper-parameters (β, σ^2_v) are set as, β has $N(0, 10^6)$ prior and $\sigma^2_v \sim IG(a_0, b_0)$, where (a_0, b_0) are known positive quantity, usually taken as very small to reflect vague knowledge about σ^2_v (Rao and Molina 2015).

Let, $\tilde{P} = (p_{1w}, \ldots, p_{Mw})^T$, $\mathbf{P} = (P_1, \ldots, P_M)^T$, $\mathbf{X} = (\mathbf{x}_1^T, \ldots, \mathbf{x}_M^T)^T$, $x_i^T = (x_{i1}, \ldots, x_{ik})$, $b = (b_1, \ldots, b_k)^T$ where k represents number of auxiliary variates and x_{i1} is taken to be 1 i=1,...,M. The full conditional distribution for HB1 are given as below,

$$1)\ P_i|\boldsymbol{\beta}, \sigma^2_v, \tilde{\mathbf{p}}\ \frac{1}{P_i(1-P_i)\sqrt{\sigma^2_{ei}\sigma^2_v}} \exp\left(-\frac{(p_{iw}-P_i)^2}{2\sigma^2_{ei}} - \frac{(\text{logit}(P_i) - x_i^T\beta)^2}{2\sigma^2_v}\right);$$

$$2)\ \beta|\mathbf{P}, \sigma^2_v \sim N\left(\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T \text{logit}(\mathbf{P}), \sigma^2_v\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\right)$$

$$3)\ \sigma^2_v|\beta, P_i \sim IG\left(a + \frac{M}{2}, b + \frac{\sum_{i=1}^{M}\left(\text{logit}(P_i) - \mathbf{x}_i^T\boldsymbol{\beta}\right)^2}{2}\right).$$

In the HB2 model, the sampling model: $p_{iw}|P_i \sim N(P_i, \xi_i)$ and Linking model: $\text{logit}(P_i)|\beta, \sigma_v^2 \sim N(x_i'\beta, \sigma_v^2)$

Here, in the sampling model instead of known σ_{ei}^2, an unknown variance function ξ_i is used which can be approximated as $\xi_i = \frac{P_i(1-P_i)}{n_i} \text{deff}_i$. Given survey design of HCES, design effect is approximately $\text{deff}_i = n_i \left(\sum_{j=1}^{n_i} w_{ij} \right)^{-2} \sum_{j=1}^{n_i} w_{ij}^2$ (Anjoy et al. 2018).

The choices for priors for hyper-parameters remain the same as HB1. One of our objectives is to see how much postulating unknown sampling variances in the HB version of GLMM structure improves precision as the direct survey estimator of the sampling variance is not reliable in domains with negligible sample sizes.

The full conditional distribution for HB2 is

1) $P_i | \boldsymbol{\beta}, \sigma_v^2, \tilde{\mathbf{p}} \; \frac{1}{P_i(1-P_i)\sqrt{\xi_i \sigma_v^2}} \exp\left(-\frac{(p_{iw}-P_i)^2}{2\xi_i} - \frac{(\text{logit}(P_i) - \mathbf{x}_i^T \boldsymbol{\beta})^2}{2\sigma_v^2} \right)$;

2) $\beta | \mathbf{P}, \sigma_v^2 \sim N\left((\mathbf{X}^T\mathbf{X})^{-1} \mathbf{X}^T \text{logit}(\mathbf{P}), \sigma_v^2 (\mathbf{X}^T\mathbf{X})^{-1} \right)$

3) $\sigma_v^2 | \beta, P_i \sim IG\left(a + \frac{M}{2}, b + \frac{\sum_{i=1}^{M} (\text{logit}(P_i) - \mathbf{x}_i^T \boldsymbol{\beta})^2}{2} \right)$.

We use the hierarchical Bayes method to compute the small area proportion estimates for these models using the Metropolis–Hastings algorithm and drawing random samples from the full conditional distributions of posterior quantities (You 2008; Liu et al. 2014; Anjoy et al. 2018). In particular, we implement the Gibbs sampling method with three independent chains, each of length 10,000. We deleted the first 5,000 iterations as "burn-in" periods. The empirical results estimating the poverty proportions by socio-economic group are obtained using R and WINBUG software.

Results and discussion

This study uses the SAE technique, based on the hierarchical Bayes method, to generate the small area estimates of poverty proportion, or poverty incidence, in the districts of Chhattisgarh. We developed estimates of social groups in the socio-economic hierarchy by district. Table 2 presents the summary of the percentage coefficient of variation (%CV) across social groups and districts generated using the direct survey estimates (DIR) and the small area estimates using the HB1 and HB2 models of the SAE technique.

The results reported in Table 2 make clear that the SAE technique is superior to the traditional direct survey estimation technique. In certain social groups, where the district-specific sample size is less than 5—that is, the sample size is negligible—the estimates generated by the direct survey technique are not at all reliable, as the estimates are either approximated to 1 or close to 0; consequently, the %CV is even more than 100 for such domains or too small to be reliable.

Table 2 Summary of percentage coefficient of variation (%CV) generated by different methods

Social Group	Values	Min	Mean	Q1	Median	Q3	Max
All	DIR	15.21	28.59	19.83	25.23	32.24	70.04
	HB1	6.21	7.67	6.47	7.43	8.70	10.78
	HB2	4.00	5.87	4.45	5.39	7.23	9.54
ST	DIR	19.05	36.30	23.85	31.22	38.77	107.33
	HB1	10.02	12.77	11.75	12.51	13.38	18.74
	HB2	7.96	10.50	8.71	10.09	11.27	18.17
SC	DIR	5.28	49.73	30.22	41.5	66.71	103.46
	HB1	14.35	21.73	17.25	21.55	25.06	30.12
	HB2	12.99	16.94	14.00	16.39	19.64	22.34
OBC	DIR	22.70	45.87	32.18	45.73	54.68	86.35
	HB1	11.25	14.75	12.43	13.29	17.56	21.22
	HB2	6.86	10.31	7.46	9.39	13.17	18.26

The sample size of the SC group in Bastar district was 2; it resulted in the poverty proportion estimate 0.96 with 5.28 %CV. Again, for domains with no sample data, the SAE technique is the only option. Thus, the %CV values reported in Table 2 are the figures from the sampled areas only. The direct estimates simply cannot be consulted for the Other category—for the extreme boundary value of poverty proportions or an unacceptable %CV—we excluded the report for the Other category.

The HB2 model outperforms the HB1 model in all categories of social groups (Table 2). It shows that the postulation of the unknown sampling variance term in the HB2 model can yield estimates that are more stable. The HB1 method is based on the assumption of known sampling variances obtained using the direct survey approach, so it may result in less precise estimates relative to the HB2, as direct variance estimates are not reliable at all. As the efficiency of the HB2 method in terms of acceptable %CV make it more amenable than the HB1 for producing official estimates at the micro level, we consider only the HB2 method of the SAE technique in our discussion henceforth.

Next, we check the model diagnostics to verify the assumptions of the residuals, specifically the assumption of normality and the independence of the residuals. To examine the normality, we perform the Shapiro–Wilk test and two graphical measures—the histogram and Q-Q plots. All these tests support the normality postulation. Here, we report only the result of the Shapiro–Wilk test (Table 3); typically, the normality of residuals is confirmed if the p-value exceeds 0.05 (5% level of significance). To check for randomness, we conduct a graphical test—by plotting the district-level residuals—and the statistical Durbin–Watson test. The results of the Durbin–Watson test statistic, in the 1.5–2.5 range, support the independence assumption of residuals. Table 3 presents the Durbin–Watson test statistic values by category.

Table 3 Shapiro–Wilk (SW) and Durbin–Watson (DW) test results for different social groups

Social Group	SW statistic	p-value	DW statistic
All	0.932	0.212	1.645
ST	0.898	0.162	1.733
SC	0.970	0.816	1.894
OBC	0.933	0.246	1.896

Hereabouts Table 4 reports the estimates of poverty incidence by district and social group, along with 95% confidence interval (lower and upper) and %CV, for the direct (DIR) and the SAE method based on the HB2 model. Hereabouts Table 5 presents the estimates for various categories; the 95% confidence interval (CI) for the HB2 estimates are more precise. In a few districts, the confidence limit of direct estimates is unacceptable and invalid; in the ST category, the lower limits of the CI of direct estimates in Mahasamund and Bijapur district are negative.

In the ST category, the upper CI of direct estimates is greater than 1 in four districts: Korba, Rajnandgaon, Dakshin Bastar Dantewada, and Bijapur. In the SC category, the lower CI is less than 0 in Raigarh, Janjgir-Champa, Kabeerdham, Durg, and Mahasamund, and the upper CI of direct estimates is greater than 1 in Korba Raipur, Uttar Bastar Kanker, and Bastar. These type of abnormalities of direct estimates are still followed in the Other category, where Surguja, Korba, Bilaspur, and Dhamtari have direct estimates with a negative lower bound.

The sample size of most of the Other districts was negligible, and the direct estimation method failed for this group. The traditional direct estimation approach cannot provide adequate or representative estimates with a precise CI, and this drawback makes the SAE method the only reliable option. The HB2 method of the SAE technique provides an interval of estimates that is more credible and meaningful than the CI estimates computed from the direct approach.

A critical analysis of the HB2 estimates shows that in Chhattisgarh STs are the most affected by poverty, the average incidence is 0.49, the average poverty proportion among SC is 0.46 (Table 4). Both of these figures are greater than mean incidence of 0.43 covering all groups and districts in the state. The long deprivation of the tribal community, traditional practices, and illiteracy has delayed the progress of these categories. In the OBC category, the average poverty incidence is 0.34. In ST category poverty incidence as high as 50% or above in Raigarh, Janjgir-Champa, Bilaspur, Kabeerdham, Rajnandgaon, Durg, Mahasamund, Dhamtari, Bastar, and Bijapur. The poverty incidence is 50% or more among SCs in Koriya, Surguja, Jashpur, Korba, Uttar Bastar Kanker, Narayanpur, and Dakshin Bastar Dantewada. However, in the OBC category, the maximum average poverty

Table 4 District and social group-wise estimates of poverty incidence (Estimate) along with lower (Lower) and upper (Upper) 95% CI and percentage coefficient of variation (%CV) generated by the direct (DIR) and the model-based small area estimation using HB2 (SAE) method

Districts	Sample size	DIR				SAE			
		Estimate	Lower	Upper	%CV	Estimate	Lower	Upper	%CV
Koriya	32	0.38	0.1	0.66	37.82	0.42	0.36	0.48	7.29
Surguja	160	0.34	0.2	0.47	20.23	0.42	0.36	0.48	7.3
Jashpur	64	0.41	0.2	0.63	26.29	0.41	0.35	0.48	7.69
Raigarh	88	0.32	0.14	0.5	28.35	0.45	0.42	0.49	4.11
Korba	64	0.53	0.33	0.74	19.46	0.39	0.32	0.47	9.54
Janjgir - Champa	96	0.28	0.11	0.46	31.35	0.45	0.42	0.49	4.00
Bilaspur	120	0.37	0.21	0.52	21.86	0.45	0.41	0.49	4.30
Kabeerdham	63	0.29	0.09	0.49	34.95	0.44	0.4	0.49	4.88
Rajnandgaon	96	0.49	0.32	0.66	17.81	0.45	0.41	0.49	4.33
Durg	160	0.32	0.19	0.44	19.69	0.45	0.41	0.49	4.37
Raipur	159	0.48	0.34	0.62	15.21	0.43	0.38	0.48	6.11
Mahasamund	64	0.19	0.02	0.36	46.64	0.44	0.39	0.48	5.27
Dhamtari	55	0.33	0.12	0.54	32.54	0.45	0.41	0.49	4.67
Uttar Bastar Kanker	64	0.52	0.3	0.74	21.54	0.42	0.36	0.48	7.06
Bastar	94	0.57	0.39	0.75	16.33	0.44	0.39	0.49	5.38
Narayanpur	32	0.49	0.2	0.79	30.42	0.41	0.35	0.47	7.59
Dakshin Bastar Dantewada	16	0.77	0.4	1.13	24.17	0.43	0.37	0.48	6.41
Bijapur	8	0.49	-0.18	1.16	70.04	0.44	0.39	0.48	5.39

Table 5 District and social group-wise estimates of poverty incidence generated by the direct and the model-based small area estimation using HB2 (SAE) method for various social groups

Districts	Sample size	DIR				SAE			
		Estimate	Lower	Upper	%CV	Estimate	Lower	Upper	%CV
				ST					
Koriya	20	0.48	0.11	0.84	38.77	0.48	0.39	0.57	9.71
Surguja	89	0.37	0.20	0.55	23.85	0.48	0.38	0.57	9.85
Jashpur	43	0.40	0.13	0.66	34.33	0.47	0.37	0.58	11.17
Raigarh	19	0.53	0.13	0.93	38.93	0.51	0.39	0.63	12.15
Korba	13	0.74	0.41	1.06	22.32	0.46	0.30	0.62	18.17
Janjgir - Champa	17	0.61	0.24	0.98	31.22	0.51	0.38	0.63	12.37
Bilaspur	37	0.46	0.18	0.74	31.53	0.51	0.39	0.61	11.31
Kabeerdham	1	1.00	*	*	*	0.50	0.40	0.60	9.75
Rajnandgaon	8	0.68	0.22	1.13	34.43	0.51	0.39	0.61	11.29
Durg	29	0.60	0.34	0.85	21.54	0.51	0.39	0.61	10.94
Raipur	29	0.63	0.36	0.91	22.21	0.49	0.41	0.56	7.96
Mahasamund	20	0.01	-0.02	0.04	107.33	0.50	0.41	0.58	8.55
Dhamtari	17	0.38	0.07	0.69	41.26	0.51	0.40	0.60	10.32
Uttar Bastar Kanker	34	0.47	0.20	0.74	29.10	0.48	0.39	0.57	9.18
Bastar	63	0.57	0.36	0.78	19.05	0.50	0.41	0.58	8.42
Narayanpur	28	0.63	0.30	0.95	26.38	0.48	0.37	0.58	11.22
Dakshin Bastar Dantewada	12	0.76	0.39	1.13	24.86	0.49	0.41	0.56	8.13
Bijapur	6	0.50	-0.18	1.17	70.02	0.50	0.41	0.57	8.48
Contd...									

Districts	Sample size	DIR				SAE			
		Estimate	Lower	Upper	%CV	Estimate	Lower	Upper	%CV
				SC					
Koriya	0	**	**	**	**	0.58	0.46	0.79	13.45
Surguja	18	0.45	0.09	0.81	41.12	0.58	0.42	0.73	14.02
Jashpur	7	0.00	*	*	*	0.61	0.44	0.78	14.44
Raigarh	24	0.13	-0.10	0.35	90.79	0.30	0.19	0.45	21.90
Korba	25	0.77	0.51	1.02	16.85	0.73	0.50	0.91	14.81
Janjgir - Champa	23	0.03	-0.04	0.10	103.46	0.30	0.18	0.44	22.34
Bilaspur	30	0.49	0.20	0.77	30.22	0.32	0.21	0.46	19.93
Kabeerdham	15	0.34	-0.01	0.70	53.02	0.36	0.26	0.48	16.39
Rajnandgaon	9	0.60	0.20	1.00	34.19	0.32	0.21	0.46	19.64
Durg	13	0.21	-0.07	0.50	66.71	0.33	0.22	0.46	19.03
Raipur	14	0.76	0.48	1.04	18.55	0.48	0.37	0.60	12.99
Mahasamund	6	0.15	-0.13	0.43	95.13	0.41	0.30	0.52	14.00
Dhamtari	9	0.46	0.01	0.91	49.63	0.35	0.24	0.47	17.53
Uttar Bastar Kanker	2	0.71	0.13	1.28	41.50	0.56	0.41	0.70	13.78
Bastar	2	0.96	0.86	1.06	5.28	0.42	0.31	0.53	13.80
Narayanpur	0	**	**	**	**	0.62	0.56	0.92	12.26
Dakshin Bastar Dantewada	0	**	**	**	**	0.50	0.39	0.78	17.12
Bijapur	0	**	**	**	**	0.41	0.26	0.49	15.84
				OBC					
Koriya	11	0.24	-0.16	0.64	86.35	0.30	0.23	0.30	13.36
Surguja	41	0.16	-0.02	0.34	56.59	0.30	0.23	0.30	13.55
Jashpur	12	0.57	0.20	0.95	33.30	0.30	0.22	0.29	14.30
Raigarh	39	0.41	0.15	0.67	32.92	0.39	0.34	0.38	6.86
Korba	25	0.20	-0.05	0.46	64.55	0.26	0.18	0.25	18.26
Janjgir - Champa	47	0.28	0.02	0.55	47.14	0.39	0.34	0.39	6.90
Bilaspur	38	0.22	-0.01	0.44	52.77	0.38	0.33	0.38	7.30
Kabeerdham	41	0.30	0.03	0.57	45.73	0.36	0.30	0.36	8.33
Rajnandgaon	76	0.44	0.24	0.64	23.52	0.38	0.33	0.38	7.41
Durg	108	0.24	0.09	0.38	31.43	0.37	0.32	0.37	7.50
Raipur	102	0.39	0.22	0.57	22.70	0.33	0.26	0.33	10.95
Mahasamund	34	0.31	0.03	0.59	46.09	0.35	0.29	0.35	9.39
Dhamtari	25	0.18	-0.11	0.47	82.12	0.37	0.32	0.37	7.89
Uttar Bastar Kanker	26	0.65	0.34	0.95	24.29	0.31	0.24	0.31	12.97
Bastar	26	0.52	0.13	0.92	38.50	0.35	0.29	0.35	9.62
Narayanpur	0	**	**	**	**	0.29	0.16	0.57	27.97
Dakshin Bastar Dantewada	4	1.00	*	*	*	0.33	0.26	0.32	11.44
Bijapur	1	0.00	*	*	*	0.35	0.29	0.35	9.46
				Others					
Koriya	1	1.00	*	*	*	0.005	0.002	0.010	45.93
Surguja	12	0.03	-0.04	0.11	108.26	0.005	0.002	0.010	46.30
Jashpur	2	0.00	*	*	*	0.004	0.001	0.008	48.87
Raigarh	6	0.00	*	*	*	0.046	0.029	0.070	23.86
Korba	1	0.00	*	*	*	0.001	0.000	0.003	61.05

Contd...

Districts	Sample size	DIR				SAE			
		Estimate	Lower	Upper	%CV	Estimate	Lower	Upper	%CV
Janjgir-Champa	9	0.00	*	*	*	0.047	0.030	0.073	23.33
Bilaspur	15	0.00	*	*	*	0.039	0.024	0.062	25.33
Kabeerdham	6	0.00	*	*	*	0.027	0.015	0.044	29.11
Rajnandgaon	3	0.00	*	*	*	0.038	0.023	0.059	25.65
Durg	10	0.14	-0.11	0.40	90.83	0.036	0.022	0.057	26.15
Raipur	14	0.15	-0.13	0.42	95.63	0.010	0.005	0.019	38.33
Mahasamund	4	0.00	0.00	0.00	0.00	0.018	0.010	0.032	32.87
Dhamtari	4	0.03	-0.03	0.09	113.93	0.031	0.018	0.049	27.7
Uttar Bastar Kanker	2	0.00	*	*	*	0.006	0.002	0.011	44.46
Bastar	3	0.00	*	*	*	0.016	0.009	0.030	33.27
Narayanpur	4	0.00	*	*	*	0.004	0.001	0.008	48.74
Dakshin Bastar Dantewada	0	**	**	**	**	0.008	0.001	0.012	44.82
Bijapur	1	0.00	*	*	*	0.005	0.010	0.032	45.93

*Standard error of DIR could not be computed because poverty proportion is either 0 or 1.
**Out of sample areas.

proportion was 0.39. Government and non-governmental organizations must target the persistent poverty among STs and SCs in these districts in designing their upliftment approach.

Conclusions

In India, the Five Year Plans encapsulate the national poverty reduction policy and strategies. The 12th Five Year Plan targets to reduce poverty by 10%; it would include, by way of decentralization, framing and formulating strategic plans to tackle the social threat of poverty. This study, which applies a method based on a small area model to generate disaggregated official estimates of poverty, is a step towards efficient micro-level planning.

The most poverty-stricken ST population is concentrated largely in the southern and northernmost districts of Chhattisgarh. The model-based SAE technique estimates the HCR for such districts with an acceptable and reliable percentage of the coefficients of variation. To improve the model-based estimates within the same Bayes modeling framework, the study can be extended to account for spatial dependence among neighbouring districts.

Poverty must be eliminated worldwide and livelihoods made sustainable. UN has set “no poverty” as its first SDG, so measuring poverty properly is crucial in implementing poverty eradication programmes. Poverty has many facets—food insecurity, and the proportions of malnourished and undernourished children, children lacking access to basic education, micro-level female illiteracy—and all these need to be measured adequately for the administration to take the appropriate actions.

Many studies have used the SAE approach in the past decades. What is needed now is real-life applications and implementation for micro-level, disaggregated efficient planning and monitoring. The SAE has been initiated in developed countries like the US, the UK, and Australia, and included as a part of their objectives in the national statistical offices. Agencies and organizations in India have felt the need for small area statistics, but little initiative has been taken.

In India, the Census is usually limited in its scope in collection of data; it focuses mainly on basic social and demographic information and that too at a decennial interval. The NSSO conducts regular surveys on a number of socio-economic indicators that generate national and state-level estimates but not on administrative units smaller than a state because the sample sizes are too small. To meet the UN SDGs, central and state governments need disaggregated indicators and they are struggling to generate disaggregated statistics. Using the SAE technique can meet the growing demand, and the Ministry of Statistics and Programme Implementation and several state governments have been exploring it, but the need for expert analysts is constraining adoption in most developing countries.

Acknowledgment

The authors would like to acknowledge the valuable comments and suggestions of the anonymous referee, which improved the paper considerably.

References

Anjoy, P, H Chandra, and P Basak. 2018. Estimation of disaggregate-level poverty incidence in Odisha under area-level hierarchical Bayes small area model. *Social Indicators Research* 144(1): 251–273. https://doi.org/10.1007/s11205-018-2050-9

Chandra, H, N Salvati, and U C Sud. 2011. Disaggregate-level estimates of indebtedness in the state of Uttar Pradesh in India – an application of small area estimation technique. *Journal of Applied statistics* 38(11): 2413–2432. https://doi.org/10.1080/02664763.2011.559202

Chandra, H. 2013. Exploring spatial dependence in area-level random effect model for disaggregate-level crop yield estimation. *Journal of Applied statistics* 40(4): 823–842. https://doi.org/10.1080/02664763.2012.756858

Chandra, H, N Salvati, and R Chambers. 2017. Small area prediction of counts under a non-stationary spatial model. *Spatial Statistics* 20: 30–56. https://doi.org/10.1016/j.spasta.2017.01.004

Chandra, H. 2018a. Localized estimates of the incidence of indebtedness among rural households in Uttar Pradesh: An application of small area estimation technique. *Agricultural Economics Research Review* 31(1): 29–44. 10.22004/ag.econ.274824

Chandra, H, N Salvati, and R Chambers. 2018b. Small area estimation under a spatially non-linear model. *Computational Statistics and Data Analysis* 126: 19–38. https://doi.org/10.1016/j.csda.2018.04.002

Chaudhuri, S, and N Gupta. 2009. Levels of living and poverty patterns: a district-wise analysis for India. *Economic and Political Weekly* XLIV (9): 94–110. http://environmentportal.in/files/Levels%20of%20living.pdf

Chauhan, R K, S K Mohanty, S V Subramanian, J K Parida, and B Padhi.. 2016. Regional estimates of poverty and inequality in India, 1993–2012. *Social Indicators Research* 127: 1249–1296. https://doi.org/10.1007/s11205-015-1006-6

Dubey, A, and S Tiwari. 2018. Economic Growth and urban poverty in India. *Environment and Urbanization Asia* 9: 18–36. https://doi.org/10.1177/0975425317748451

Fay, R E, and R A Herriot. 1979. Estimates of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association* 74: 269–277. https://doi.org/10.1080/01621459.1979.10482505

Gebert, R, A Namala, and J Kumar. 2011. Poverty Impact Assessment Report, Chhattisgarh. European Union State Partnership Programme Chhattisgarh. https://www.giz.de/de/downloads/giz2011-en-poverty-impact-assessment-report.pdf

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 1: 515-533. https://projecteuclid.org/download/pdf_1/euclid.ba/1340371048

Ghosh, M, D Kim, K Sinha, T Maiti, M Katzoff, and V L Parsons 2009. Hierarchical and Empirical Bayes small domain estimation of the proportion of persons without health insurance of minority subpopulations. *Survey Methodology* 35: 53–66. https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2009001/article/10884-eng.pdf?st=i7F6JGTR

Hidiroglou, M A, and Y You. 2016. Comparison of unit level and area level small area estimators. *Survey Methodology* 42: 41–61. https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2016001/article/14540-eng.pdf?st=e4eqSUWR

Jiang, J, and P Lahiri. 2006. Mixed model prediction and small area estimation. *Test* 15: 111–999. https://doi.org/10.1007/BF02595419

Lee, D, J Minton, and G Pryce. 2015. Bayesian inference for the dissimilarity index in the presence of spatial autocorrelation. *Spatial Statistics* 11: 81–95. https://doi.org/10.1016/j.spasta.2014.12.001

Liu, B, P Lahiri, and G Kalton. 2014. Hierarchical Bayes modeling of survey-weighted small area proportions. *Survey Methodology* 40: 1–13. https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2014001/article/14030-eng.pdf?st=s8xwhBsu

Mohanty, S K, D Govil, R K Chauhan, R Kim, and S V Subramanian.2016. Estimates of poverty and inequality in the districts of India, 2011–2012. *Journal of Development Policy and Practice* 1(2): 142–202. https://doi.org/10.1177/2455133316642338

Portar, A T, S H Holan, C K Wikle, and N Cressie2014. Spatial Fay–Herriot models for small area with estimation functional covariates. *Spatial Statistics*: 10: 27–42. https://arxiv.org/pdf/1303.6668.pdf

Pratesi, M, and Salvati, N. 2009. Small area estimation in presence of correlated random area effects. *Journal of Official Statistics* 25: 37–53. https://www.scb.se/contentassets/ff271eeeca694f47ae99b942de61df83/small-area-estimation-in-the-presence-of-correlated-random-area-effects.pdf

Pratesi, M and N Salvati. 2016. Introduction on measuring poverty at local level using small area estimation methods, *Pratesi/Analysis of poverty data by small area estimation.* John Wiley & Sons, New York. https://www.wiley.com/en in/Analysis+of+Poverty+Data+by+Small+Area+Estimation-p-9781118815014

Planning Commission, Government of India. 2013. Press note on poverty estimates 2011-12., New Delhi. https://niti.gov.in/planningcommission.gov.in/docs/news/pre_pov2307

Rao, J N K. 2003. *Small area estimation*. John Wiley & Sons, New York. Rao, J N K, and I Molina. 2015. *Small area estimation*: *2nd edition*. John Wiley & Sons, New York.

World Bank. 2016. Chhattisgarh poverty, growth, and inequality (English), India state briefs., Washington. https://www.worldbank.org/en/country/india/brief/india-states-briefs—chhattisgarh

You, Y. 2008. An integrated modeling approach to unemployment rate estimation for sub-provincial areas of Canada. *Survey Methodology* 34: 19–27. https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2008001/article/10614-eng.pdf?st=Osa5qUZR

Received: July 2019 Accepted: December 2019