Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), 149-172

Small Area Estimation – Some Applications in NSSO Surveys

A. K. Srivastava and Hukum Chandra
ICAR-Indian Agricultural Statistics Research Institute, New Delhi.

Received: 05 December 2020; Revised: 06 January 2021; Accepted: 09 January 2021

Abstract

The purpose of this article is to use small area estimation (SAE) method to produce district level estimates for some of the important indicators such as living condition, poverty incidence and working population ratio. For this purpose, data from 68th round (2011-12) of National Sample Survey Office (NSSO) pertaining to Household Consumer Expenditure Survey (HCES) and Employment and Unemployment Survey (EUS) for Uttar Pradesh has been used along with the 2011 Population Census data. The empirical results, evaluated through set of internal and external diagnostics measures, show that the district-level estimates generated through SAE approach are precise than the direct estimates. Spatial maps showing district level inequality in distribution of living condition, poverty incidence and working population ratio in Uttar Pradesh are also produced. These maps and districts level estimates are important for target oriented effective policy planning, monitoring and decision-making. In this article we deliberately consider two types of estimates *viz*. averages and proportions and use two different survey data of NSSO for producing district level estimates. We then illustrate how the existing survey data can be linked with Census data to produce reliable, timely and cost-effective district-level estimates of averages and proportions. The SAE methodology, illustration and guidelines set out in this paper can be adopted in other existing surveys for generating the disaggregate level estimates.

Key words: NSSO survey*;* Small area estimation; Precision; Living condition; Working population ratio.

0. Prologue

This paper is a tribute in honour and loving memory of Dr. Aloke Dey who had been a close friend to me all along for more than five and a half decades. Right from our student days to the entire professional career, he had been a source of strength and inspiration to all of us. His intense concern for maintaining high standards and values in research and teaching had a deep influence on his friends, colleagues, and students. Improvement in statistical system of the country was also remarkably close to his heart. This paper is an effort towards bringing in Small Area Estimation Techniques closer to application into some of the NSSO surveys – a hearty tribute from our side. - A K Srivastava.

Corresponding Author: A. K. Srivastava
E-mail: arunsrivast@gmail.com

1. Introduction

The NSSO surveys are generally conducted to generate a huge range of invaluable and crucial data, separately for the rural and urban sectors of the country, for States and Union Territories, and for different socio-economic groups. However, there is a rapidly growing demand for disaggregate level estimates (*e.g*. district or further disaggregate level) in India as the country is moving towards more decentralized system of governance. The disaggregate level estimates are also inevitable for several sustainable development goals (SDGs) related indicators. Just to mention some early attempts in India, an expert committee on small area statistics (SAS) was set up by (then) Ministry of Planning and Programme Implementation, Government of India (Government of India, 1997) under the chairmanship of Professor J. Roy. The committee deliberated upon the implications of 73rd and 74th amendments in the Constitution in view of data needs and its availability and highlighted the need for methodological studies for generating small area statistics appropriate to Indian conditions. This paper particularly concentrates on providing district level estimates for NSSO surveys where estimates are mostly generated at state-level. The SAE techniques provide a viable approach for producing estimates at smaller levels (Rao and Molina, 2015). The models used in SAE are commonly grouped as area level or unit level model. Area-level modelling is typically used when unit-level data are unavailable, or, as is often the case, where model covariates or auxiliary variables are only available in aggregate form. In this article, we motivate the SAE method based on area level small area modeling because in India the auxiliary variables are often accessible and available at aggregate (*e.g.* district) level. In this context, Fay–Herriot model (Fay and Herriot, 1979) is a widely used area level model in SAE. But this model is suited for continuous data. If the variable of interest is binary and the aim is to estimate small area proportions, then the area level generalized linear mixed model with logit link function, also referred to as the logistic linear mixed model (LLMM) is generally used (Johnson *et al*., 2010; Chandra *et al*., 2011 and Chandra *et al*., 2019). Srivastava (2007) used Fay–Herriot method of SAE to generate district level estimates for monthly per capita consumer expenditure (MPCE) using the 2004-05 Household Consumer Expenditure Survey (HCES) data of NSSO for the state of Uttar Pradesh. Srivastava (2009) further used the same data for estimating several poverty indicators at district level. Singh. *et. al.* (2005) used NSSO data for application of spatio-temporal models in SAE. More recently, Anjoy *et al.* (2020) used All India Debt and Investment Survey 2012-13 of NSSO for estimating the district-wise proportions of indebted households in rural areas of Karnataka. Chandra (2020) applied SAE method to estimate the incidence of food insecurity in different districts of rural areas of the state of Uttar Pradesh using the 2011-12 HCES of NSSO.

In this article, we consider SAE methods to produce district level estimates of the average household MPCE, the proportion of poor households and the employment rate for both rural and urban sectors for the state of Uttar Pradesh. Throughout this article, the proportion of poor households (*i.e.* proportion of households below poverty line) is also referred by poverty incidence and poverty rate (PR). The employment rate (UR) is referred as the proportion of persons employed to total persons. Alternatively, the worker-population ratio (WPR), also referred as work-force participation rate (WFPR) is defined as the number of persons employed per 1000 persons (*i.e.* WPR=1000×UR). The work force in the usual status includes the persons who worked for a relatively long part of the 365 days preceding the date of survey and the persons from among the remaining population who had worked at least for 30 days during the reference period of 365 days preceding the date of survey. The estimates of average household MPCE and poverty rate from the HCES of NSSO and the estimates of employment rate from the employment and unemployment survey (EUS) are common statistics generated by all the states and used by different departments and ministries. This article deliberates these parameters and illustrates how the existing HCES and EUS data can be used to generate precise district level estimates. We elaborate two types of estimates *viz*. average and proportion (rate) and use two different survey data (HCES and EUS) of NSSO linking with Census data for producing district level estimates. This example can also be used as guidelines for generating the district level estimates of other commonly required parameters from the other existing surveys.

The paper is organized as follows. Besides introductory part in section 1, we describe data sources, different indicator variables of interest, and choice of auxiliary variables for SAE modelling in Section 2. Section 3 briefly delineates methodology used in the applications considered in this paper. Some aspects of the methodological framework have been discussed in Srivastava (2007, 2009) and Chandra (2020). In fact, Chandra (2020) applied the approach to estimate district-wise proportion of food insecure households in rural areas of Uttar Pradesh. However, for the sake of clarity and completeness, approach is described briefly in Section 3. The empirical results including essential diagnostic measures and discussions are deliberated in Section 4. Finally, Section 5 concludes the paper with some final remarks and recommendations.

2. Data Sources and Model Selection

The small area applications reported in this paper are based on the HCES and the EUS data from 68th round (2011-12) of NSSO for both rural and urban sectors of Uttar Pradesh and the 2011 Population Census. The 2011-12 HCES data is used to estimate the average household MPCE and the proportion of poor households (*i.e.* poverty ratio or PR) at district level for both rural and urban sectors in Uttar Pradesh. On the other hand, the estimation of UR (or WPR) is based on the 2011-12 EUS data. The household MPCE and the binary variable indicating whether a household is poor or not are the target variables of interest in 2011-12 HCES data. In this application a household having MPCE below the state poverty line is defined as poor. The poverty line used in this study (Rs. 768 for rural and Rs. 941 for urban) is the same as that set by the then Planning Commission, Govt. of India, for 2011-12. The parameters of interest are the average household MPCE and the PR within each district. In 2011-12 EUS data, the parameter of interest is the UR or WPR. In 2011-12 HCES, a total of 5916 rural and 3102 urban households from the 71 districts of Uttar Pradesh were surveyed. The district sample sizes for rural areas ranged from 32 to 128 with average of 83. Similarly, the district sample sizes for urban areas varied from 30 to 128 with average of 44. On the other hand, the 2011-12 EUS enumerated 49513 persons (33738 in rural areas and 15775 in urban areas). The district level sample sizes are relatively small for generating precise district level estimates.

The 2011 Population Census has a range of auxiliary variables (covariates) which can be explored for SAE modelling. However, we identified few relevant auxiliary variables for this study. We also used Principal Component Analysis (PCA) to derive composite scores for selected groups of auxiliary variables, separately for both rural and urban areas. Using district aggregates of rural data, we did PCA for two groups of auxiliary variables, denoted as R1and R2. The first group (R1) consisted of the proportions of main worker by gender, proportions of main cultivator by gender and proportions of main agricultural labourer by gender. The second group (R2) consisted of proportions of marginal cultivator by gender and proportions of marginal agriculture labourers by gender. The first principal component (R11) for the first group explained 44% of the variability in the R1, while adding the second component (R12) increased explained variability to 69%. The first principal component (R21) for the second group explained 52% of the variability in the R2 group, while addition of second component (R22) increased explained variability to 90%. For urban areas, we further applied PCA, separately for two groups of variables, as defined in rural data, but using district aggregates of urban data. These are denoted as U1 and U2. Here, the first principal component (U11) explained 53% of the variability and addition of the second component (U12) explained 83% variability to in the U1. The first principal component (U21) for the U2 explained 63% of the variability, while adding the second component (U22) enhanced explained variability to 87%.

For both rural and urban data, we separately fitted a linear model using district-wise direct estimates of MPCE as the response variable and the PCA scores and other auxiliary variables as covariates. The final model with selected covariates was used to produce district-wise estimates of average household MPCE. The model was fitted using the lm() function in R using the district specific sample sizes as the weight. We also fitted a generalized linear model (GLM) using direct estimates of proportions of poor households versus the PCA scores and other auxiliary variables for each group of data. The model was fitted using the glm() function in R and specifying the family as "binomial" and the district wise sample sizes as the weight. We also fitted a GLM using direct estimates of employment rates versus a set of auxiliary variables. In each case, model fitting was used for selection of final model for SAE analysis. Table 1 provides list of selected covariates which were used in SAE of average household MPCE, poverty incidence and employment rate.

Table 1: Selected auxiliary variables for SAE of the average MPCE, the proportion of poor households (poverty incidence) and the working population ratio

Parameter	Rural	Urban
MPCE	SC (Proportion of scheduled caste to total population), Literacy rate, R11, R21and R22	Literacy rate and TWPR (Proportion of worker to total population)
Poverty incidence	SC, Literacy rate, R11, R21and R22	Literacy rate and TWPR
Employment rate	SC, TWPR, Number of households and Total population	SC, Literacy rate and TWPR

3. Methodological Framework

This Section briefly introduces the SAE methods applied for producing the district level estimates of average household MPCE, poverty incidence and employment rate (or WPR) and their measure of precision for rural and urban areas of the state of Uttar Pradesh. Let y_{ij} denote the value of the variable of interest for unit $j(j=1,\ldots,N_i)$ in district $i(i=1,\ldots,D)$, where N_i and D denote the population size of district i and total number of districts in the population respectively. The quantity of interest in district i is the population mean (or proportion, in case of binary variable) m_i defined as $m_i = N_i^{-1}\sum_{j=1}^{N_i} y_{ij}$. Let n_i denotes the sample size in district i, then the direct estimator of m_i is $\hat{m}_i^{Direct} = \left(\sum_{j=1}^{n_i} w_{ij}\right)^{-1}\sum_{j=1}^{n_i} w_{ij}y_{ij}$, where w_{ij} is inverse of the inclusion probability for unit j in district i. The estimate of variance of direct estimator $\hat{m}_i^{Direct}$ is $v(\hat{m}_i^{Direct}) \approx \left(\sum_{j=1}^{n_i} w_{ij}\right)^{-2}\sum_{j=1}^{n_i} w_{ij}(w_{ij}-1)(y_{ij}-\hat{m}_i^{Direct})^2$. Let $\hat{m}_i^{Direct}$ be the observed direct estimate of average MPCE for district i. Let $\mathbf{x}_i$ be the k-vector of known district level auxiliary variables, related to the population parameter m_i. Then district specific Fay and Herriot (1979) model is described as $\hat{m}_i^{Direct} = m_i + e_i$ and $m_i = \mathbf{x}_i^T\boldsymbol{\beta} + u_i$. Alternatively, this model can be expressed as

$$\hat{m}_i^{Direct} = \mathbf{x}_i^T\boldsymbol{\beta} + u_i + e_i;\ i = 1,...,D. \quad (1)$$

Here $\boldsymbol{\beta}$ is a k-vector of unknown fixed effect parameters, u_i's are independently and identically distributed normal random errors with $E(u_i) = 0$ and $Var(u_i) = \sigma_u^2$, and $e_{\alpha i}$'s are independent sampling errors normally distributed with $E(e_i \mid m_i) = 0$, $Var(e_i \mid m_i) = \psi_i$. The two errors are independent of each other within and across districts. Let $\hat{\sigma}_u^2$ denote the estimator of σ_u^2 and $\hat{\boldsymbol{\beta}}$ the empirical best linear unbiased estimator of $\boldsymbol{\beta}$. The empirical best linear unbiased predictor (EBLUP) estimate of m_i is then

$$\hat{m}_i^{EBLUP} = \mathbf{x}_i^T\hat{\boldsymbol{\beta}} + \hat{u}_i;\ i = 1,...,D. \quad (2)$$

Here, $\hat{u}_i = \hat{\gamma}_i(\hat{m}_i^{Direct} - \mathbf{x}_i^T\hat{\boldsymbol{\beta}})$, where $\hat{\gamma}_i = \hat{\sigma}_u^2\left(\hat{\sigma}_u^2 + \psi_{\alpha i}\right)^{-1}$ defines the shrinkage effect for district i. The mean squared error (MSE) estimation of EBLUP (2) follows from Rao and Molina (2015). Readers can also refer to Chandra (2013) for the expression of MSE estimate of EBLUP (2).

It is worth noting that the direct estimate of proportions (*e.g.* PR and ER) can also be modelled by Fay-Herriot model (1) and the EBLUP estimate of district level proportions can be obtained. However, the estimate of district level proportions derived from the EBLUP (2) might be inconsistent in the sense that they might not be within the [0,1] interval. We describe approach to model district-specific proportions under a LLMM to produce precise district level estimates of PR and ER. For example, for estimating PR, the binary variable y_{ij} takes value 1 when household j in district i is poor and 0 otherwise. Similarly, in case of ER, it assumes value 1 when person j in district i is employed and 0 otherwise. In this case, population parameter of interest in district i is the district level proportion. Let $y_{si} = \sum_{j \in s_i} y_{ij}$ denotes the sample count in district i, which follows a Binomial distribution with parameters n_i and π_i, *i.e.* $y_{si} \mid v_i \sim \text{Bin}(n_i, \pi_i)$, where π_i is a success probability. The model linking π_i with the covariates $\mathbf{x}_i$ is the LLMM of form

$$logit(\pi_i) = \ln\left\{\pi_i(1-\pi_i)^{-1}\right\} = \eta_i = \mathbf{x}_i^T\boldsymbol{\alpha} + v_i, \quad (3)$$

with $\pi_i = \exp(\mathbf{x}_i^T\boldsymbol{\alpha} + v_i)\left\{1 + \exp(\mathbf{x}_i^T\boldsymbol{\alpha} + v_i)\right\}^{-1}$, where $\boldsymbol{\alpha}$ is the k-vector of regression coefficients and v_i is the district-specific random effect with $v_i \sim N(0, \sigma_v^2)$. Here, the sampling information has been incorporated by replacing the "actual sample size" and the "actual sample count" with the "effective sample size" and the "effective sample count" respectively, see for example, Chandra *et al*. (2019). Assuming $N_i >>> n_i$, a plug-in empirical predictor (EPP) of proportion (*e.g*., PR or ER) m_i in district *i* is

$$\hat{m}_i^{EPP} = \exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}} + \hat{v}_i)\left(1+\exp(\mathbf{x}_i^T\hat{\boldsymbol{\beta}} + \hat{v}_i)\right)^{-1}. \quad (4)$$

The expression for the estimate of MSE of EPP (4) is given in Chandra *et al.* (2019).

4. Results and Discussions

This Section illustrates and discusses the district-wise estimates of average household MPCE, incidence of poverty and UR generated by direct and SAE methods for both rural and urban areas of Uttar Pradesh. The EBLUP (2) under FH model is used to produce the district-wise estimates of average household MPCE and the EPP (4) is applied for generating the district-wise estimates of incidence of poverty (or PR) and employment rate for rural and urban areas. The corresponding estimates of MSE are also computed to assess the reliability of estimates and also to construct the confidence interval for the estimates. The district-specific estimates of average household MPCE, PR and WPR along with their SEs and CVs generated by the Direct and SAE methods for Uttar Pradesh are provided in the Appendices (Tables A1-A6).

A set of diagnostics measures are implemented before making inferences about small area estimates. Such diagnostics measures are (i) the model diagnostics, and (ii) the small area estimates diagnostics. The model diagnostics are tested to verify the assumptions of the underlying model. For example, the small area models (1) and (3) assume that the random district specific effects have a normal distribution with mean zero and fixed variance. The district specific residuals are expected to be randomly distributed around zero if the model assumptions are satisfied. Histogram and q-q plot are also checked to inspect the normality assumption. For this study, the district level residuals are randomly distributed around zero and the histograms as well as the q-q plots also provide evidence in support of the normality assumption. In addition, we use the Shapiro-Wilk (SW) test to examine the normality of the district random effects. The other diagnostics are demonstrated to examine the level of validity and accuracy of the small area estimates. Three commonly used diagnostics measures for evaluating the validity and the reliability of the small area estimates: the bias diagnostic, the percent coefficient of variation (CV) diagnostic and the 95% confidence interval (CI) diagnostic. The first diagnostics assesses the validity and last two review the improved precision of the small area estimates level (Chandra *et al*., 2011). For bias diagnostic we plot direct estimates (*Y*-axis) vs. small area estimates (*X*-axis) and we looked for divergence of the fitted least squares regression line from the line of equality. Although results not reported here, the bias diagnostic plots revealed that the district level estimates of MPCE, poverty incidence and WPR for both rural and urban are less extreme when compared to the corresponding direct estimates. We also use a Goodness of Fit (GoF) diagnostic, which is equivalent to a Wald test, for whether the differences $D_i = \hat{m}_i^{Direct} - \hat{m}_i$ between direct estimates $\hat{m}_i^{Direct}$ and small area estimates $\hat{m}_i$ of a population parameter (m_i) are statistically different. The null hypothesis is that the direct and small area estimates are statistically equivalent. The alternative is that the direct and small area estimates are statistically different. This Wald test statistic is computed as $\mathrm{W} = \sum_i \left\{ D_i^2 \left[var(\hat{m}_i^{Direct}) + mse(\hat{m}_i) \right]^{-1} \right\}$. Assuming $\hat{m}_i^{Direct}$ and $\hat{m}_i$ are independently distributed, which is not unreasonable for large sample sizes, the value of W can be compared with an appropriate critical value from a chi square distribution with degrees of freedom *D* equal to the number of districts. For our analysis, $D = 71$, with a critical value of 91.67 at a 5% level of significance calculated using qchisq function in R. A small value (<91.67 here) of W indicates no statistically significant difference between small area and direct estimates. The results from GoF diagnostic are given in Table 2. The values of W are smaller than the 91.67, which indicates that small area estimates are consistent with the direct estimates. In general, the bias diagnostics reflect that the small area estimates are consistent with the direct survey estimates.

Table 2: Goodness of fit diagnostic

Estimate	Rural	Urban
MPCE	11.87	7.70
Poverty	28.39	13.04
WPR	26.20	44.17

We computed the CV to compare the extent to which the small area estimates of MPCE, poverty incidence and WPR improve in precision compared to the corresponding direct estimates. There is no standard, universally accepted definition of what constitutes large or small CV values. However, different organizations have different cut-offs: for instance, the UK Office for National Statistics has a cut-off CV value of 20% for acceptable estimates, while in the US the National Center for Health Statistics has a cut-off of 30% for county-level health statistics (Baffour *et al*., 2019). Figure 1 displays the district-wise values of CV for small area estimates and direct estimates in increasing order of sample sizes. The distribution of CV in Figure 1 shows that in most of the districts, the CVs of the small area estimates are significantly smaller than those of the direct survey estimates, implying that the small area estimates are less variable, and hence relatively more precise than the direct survey estimates. The improvement CV is higher for the districts with smaller sample sizes as compared to the larger sample sizes. A set of summary statistics for the direct and small area estimates along with associated standard errors and CV of the MPCE, poverty incidence and WPR over 71 districts for rural and urban area are reported in Table 3. As expected, the average values of MPCE, poverty incidence and WPR estimates generated by SAE are almost identical to those of the direct estimates but with lower variability (*i.e.* smaller values of standard deviation). For example, the standard deviations of MPCE estimates for rural area generated by the direct and the SAE methods are 50 and 24, respectively. From Table 3, it is obvious that the small area estimates of MPCE, poverty incidence and WPR are more precise and representative than the direct estimates for both rural and urban areas. We now examine the 95 % confidence interval for the direct estimates compared to the small area estimates. For more precise estimates, we expect the width of the confidence interval to be narrower. The district-wise plots of the 95% confidence intervals (CIs) for the average household MPCE, poverty incidence and WPR generated by direct and SAE methods (EBLUP for average household MPCE and EPP for both poverty incidence and WPR) are displayed in Figure 2. These plots show that the 95% CIs for the direct estimates are wider than the 95% CIs for the small area estimates for the average household MPCE, poverty incidence and WPR. We further note that in many districts the 95% CI for direct estimates are invalid (for example, negative values for poverty incidence) due to large standard errors. Finally, we examine the aggregation property of the small area district-level estimates generated by SAE methods at higher (*e.g*. State) level. Let $\hat{m}_i$ and N_i denote the estimate of an average or proportion y_i and population size for district *i*. The divisional and state-level estimates an average or proportion is then calculated as $\hat{m} = \sum_{i=1}^{D} N_i \hat{m}_i \Big/ \sum_{i=1}^{D} N_i$.

Table 4 reports the state level estimates of the average household MPCE, poverty indicator and WPR generated by direct and SAE methods. Comparing these estimates, we see that the small area estimates are close to the direct survey estimates at state level.

Table 3: Summary distribution of direct and model-based small area estimates along with their standard error (SE) and percent CV of MPCE, poverty incidence and WPR

Parameter	Statistics	Direct estimate			Small area estimate		
		Estimate	SE	CV	Estimate	SE	CV
Rural							
Average household MPCE	Minimum	774	38	3.73	791	37	3.67
	Maximum	1958	309	25.35	1558	139	14.43
	Average	1083	83	7.44	1059	68	6.40
	Std. deviation	224	50	3.61	165	24	1.96
Poverty incidence	Minimum	0.002	0.011	12.91	0.060	0.024	12.75
	Maximum	0.578	0.169	99.38	0.506	0.083	42.32
	Average	0.249	0.071	35.08	0.251	0.055	24.37
	Std. deviation	0.137	0.025	17.44	0.107	0.014	6.72
WPR	Minimum	224	22	6.80	280	20	5.86
	Maximum	507	79	21.68	430	36	10.16
	Average	338	37	11.03	337	26	7.65
	Std. deviation	57	10	3.25	33	3	0.99
Urban							
Average household MPCE	Minimum	791	61	6.08	796	60	6.02
	Maximum	6453	609	22.42	4762	482	19.39
	Average	1623	185	10.97	1569	172	10.61
	Std. deviation	835	123	3.83	637	99	3.34
Poverty incidence	Minimum	0.003	0.003	11.98	0.023	0.010	11.54
	Maximum	0.736	0.136	105.41	0.667	0.122	56.94
	Average	0.281	0.077	36.73	0.278	0.061	26.57
	Std. deviation	0.183	0.032	18.47	0.161	0.024	10.89
WPR	Minimum	222	28	7.51	260	18	5.75
	Maximum	487	57	21.51	393	27	7.72
	Average	313	41	13.39	312	20	6.51
	Std. deviation	51	7	2.73	25	2	0.49

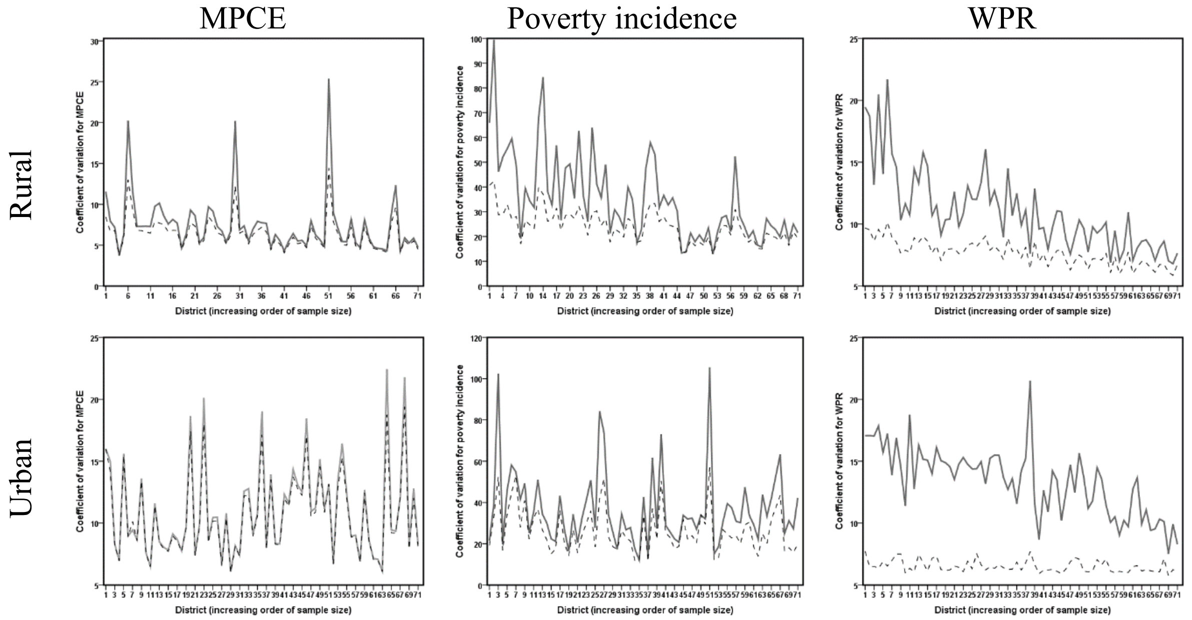

Figure 1: District-wise coefficient of variation (%) for the small area estimates (solid line) and the direct estimates (dash line). Districts are arranged in increasing order of sample sizes.

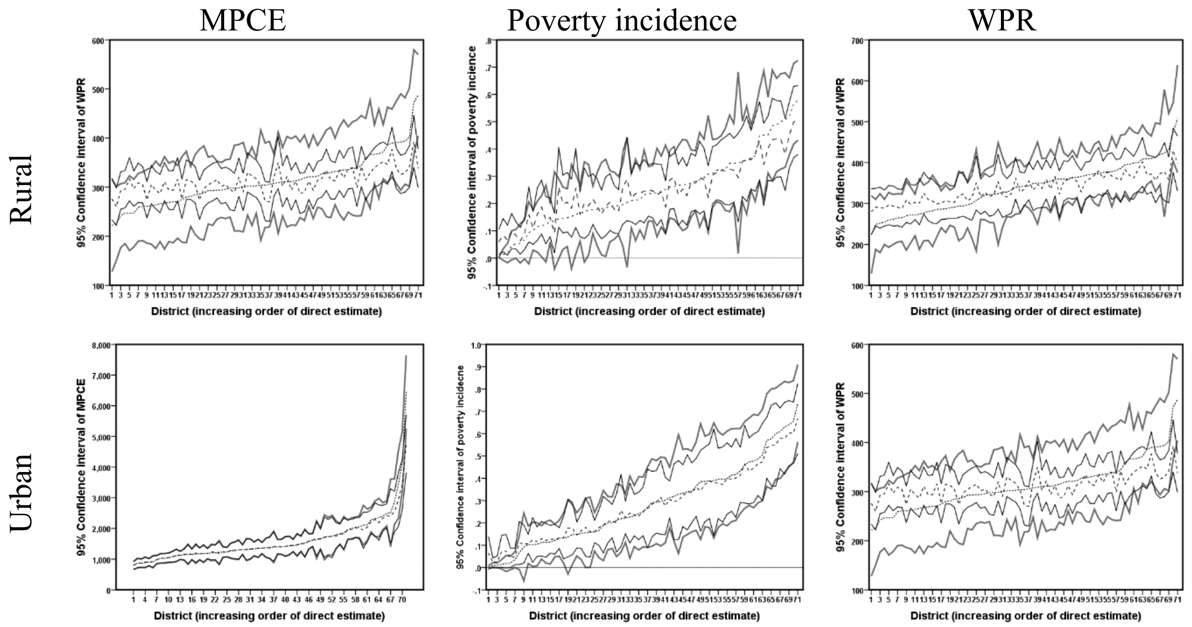

Figure 2: District-wise 95 percentage nominal confidence interval (95% CI) for the direct (solid line) and small area (thin line) estimates. Direct (dotted point) and model-based estimates (dash point) are shown in the 95% CI.

The district-specific estimates of average household MPCE, poverty incidence and WPR along with their CVs and and 95% CIs generated by the direct and SAE methods are provided in the Appendices (Table A1-A6). The diagnostics measures clearly demonstrate that the small area estimates are more efficient, precise, and representative than the direct estimates. Consequently, statistical inferences and conclusions based on the small area estimates of MPCE, poverty incidence and WPR are expected to offer better and effective policy decisions. Therefore, hereafter in discussion we focus on the estimates of MPCE, poverty incidence and WPR generated by SAE methods. Figures 3-5 provides maps showing spatial distribution of MPCE, poverty incidence and WPR estimates respectively at district level for rural and urban areas of Uttar Pradesh produced from the SAE methods. Darker areas of the maps correspond to the areas with high values of estimates. These maps supplement the district-wise estimates along with CVs and 95% CIs set out in Appendices (Table A1-A6).

Table 3: Aggregated level estimates generated by direct and SAE methods

Parameters	Rural		Urban	
	Direct	SAE	Direct	SAE
MPCE (Rs)	1073	1050	1942	1934
Poverty rate (%)	25.8	25.7	19.2	19.4
WPR	338	336	317	310

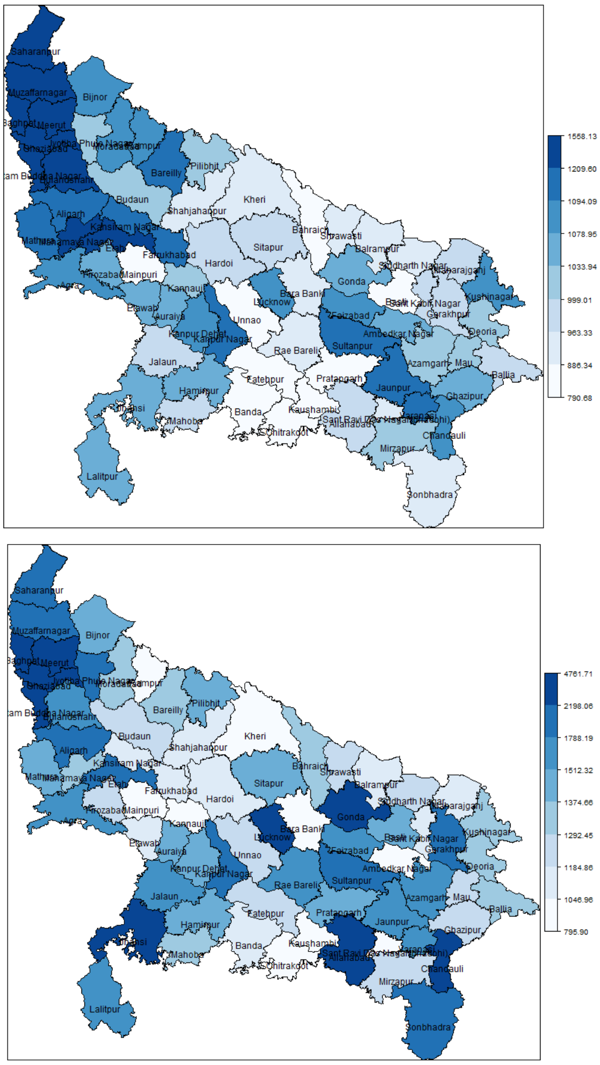

Figure 3: District-wise mapping of MPCE for rural (left) and urban (right) areas in the state of Uttar Pradesh generated by small area estimation method, 2011-12

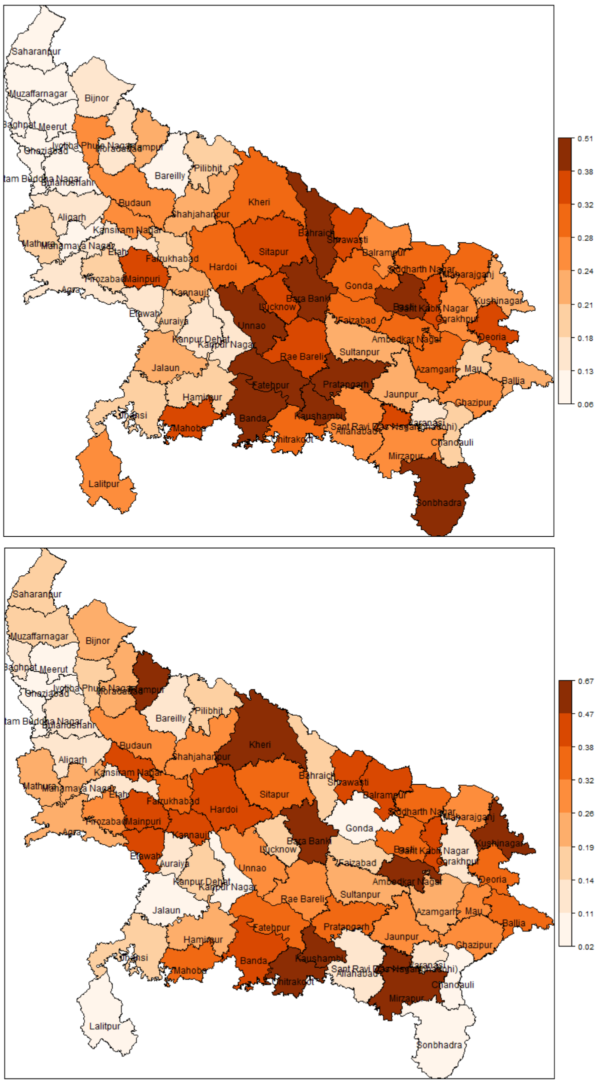

Figure 4: District-wise mapping of poverty incidence for rural (left) and urban (right) areas in the state of Uttar Pradesh generated by small area estimation method, 2011-12

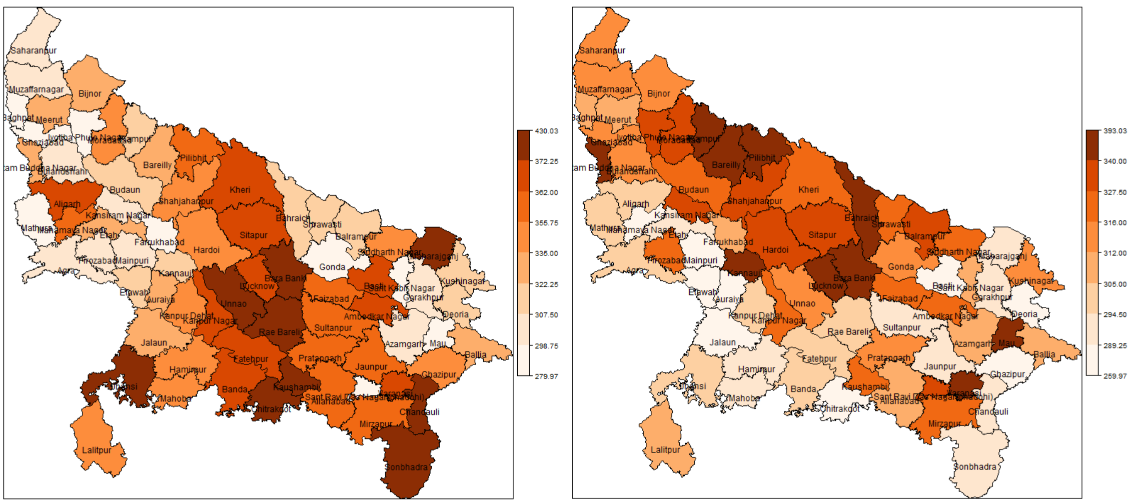

Figure 5: District-wise mapping of worker population ratio for rural (left) and urban (right) areas in the state of Uttar Pradesh generated by small area estimation method, 2011-12

5. Concluding Remarks and Recommendations

In India, Censuses are usually limited as they tend to focus mainly on the basic socio-demographic and economic data and are not available for every time - period. On the other hand, country is fortunate to have regular NSSO surveys for generating number of socio-economic indicators. The NSSO surveys are aimed to generate estimates at national and state level. They do not provide sub-state level statistics. There is no regular flow of estimates at districts and further disaggregate levels. It is known that state and national estimates usually mask variations (heterogeneity) at the sub-state or district level and render little information for micro level planning and allocation of resources. Recently, there has been a pressing demand for disaggregate level sustainable development goals (SDGs) related indicators in various departments in central and state governments and United Nations agencies in the country. Therefore, need for SAE has again achieved momentum. Despite the importance and urgent requirements, there are several virtual reasons for this topic not being implemented in the system. To the best of our knowledge and understanding, one such reason is technicality involved in SAE method. For example, SAE is combination of statistical modelling and survey estimation and there is no unique solution for all type of problems encountered. In order to develop a team of personnel with technical knowledge and experience in the field, adequate stability of the staff needs to be ensured.

This article demonstrated application of SAE approach to generate district level reliable and representative of the average household MPCE, poverty incidence and working population ratio for rural and urban areas of Uttar Pradesh by linking the latest round of 2011-12 HCES and 2011-12 EUS data of NSSO with the 2011 Population Census. The diagnostic measures clearly confirm that the estimates generated by SAE have reasonably good precision. The SAE method has also generated reliable estimates for the districts with smaller sample sizes. The district level estimates, and spatial mapping can provide useful information for the purpose of better strategic decision and policy planning. For example, many programmes are launched by Government of India with an objective to uplift the socio-economic condition of masses. NITI Aayog requires values of some socio-economic parameters for the backward districts, which they have identified, to see the impact of policy interventions and for future planning in these backward districts. NITI Aayog has identified 114 backward districts in rural India and 112 backward districts in urban India. They are monitoring some indicators related to socio-economic parameter on a continuous basis and thus providing district level estimates is very much apt for these districts. Further, the district level estimates are likely to be advantageous for allocating budget to target welfare interventions through recognizing the districts or regions with low average MPCE (or high poverty rate) and working population ratio. The indicators chosen here are based on HCES and EUS Surveys. In fact, NSSO conducts several other important Household Surveys such as Health Surveys, Education Surveys, Situation Analysis Surveys, AIDIS Surveys besides Establishment Surveys. There are several well identified indicators of interest for each of these surveys. The methodology and application presented in this paper can be used as guideline for producing reliable, timely and cost-effective estimates using survey data from different sectors.

Acknowledgments

The authors would like to acknowledge the valuable comments and suggestions of the Chair Editor, and the reviewer. These led to a considerable improvement in the paper.

References

Anjoy, P., Chandra, H. and Parsad, R. (2020). Estimation and spatial mapping of incidence of indebtedness in the state of Karnataka in India by combining survey and census data. *Statistics and Applications*, **18**(**1**) (New Series), 21-33.

Baffour, B., Chandra, H. and Martinez, A. (2019). Localised estimates of dynamics of multidimensional disadvantage: an application of the small area estimation technique using Australian survey and census data. *International Statistical Review*, **87**(**1**), 1-23.

Chandra, H. (2020). District-level estimates of extent of food insecurity for the state of Uttar Pradesh in India by combining survey and census data. In*: Special Proceeding of the 22nd Annual Conference of SSCA held at Savitribai Phule Pune University, Pune, during January* 02-04, 2020*,* 25-38.

Chandra, H. (2013). Exploring spatial dependence in area level random effect model for disaggregate level crop yield estimation. *Journal of Applied Statistics*, **40**, 823-842.

Chandra, H., Chambers, R. and Salvati, N. (2019). Small area estimation of survey weighted counts under aggregated level spatial model. *Survey Methodology*, **45**(**1**), 31-59.

Chandra, H., Salvati N. and Sud U. C. (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.

Fay, R. E. and Herriot, R. A. (1979). Estimation of income from small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association,* **74**, 269-277.

Government of India (1997). Report of the expert committee on small area statistics. *Department of Statistics, Ministry of Planning and Programme Implementation, Government of India, New Delhi*, April 1997.

Johnson F. A., Chandra H., Brown J. and Padmadas S. (2010). Estimating district-level births attended by skilled attendants in Ghana using demographic health survey and census data: an application of small area estimation technique. *Journal of Official Statistics,* **26** (**2**), 341–359.

Rao, J. N. K. and Molina, I. (2015). *Small Area Estimation*. 2nd Edition. John Wiley & Sons.

Singh, B. B., Shukla, G. K. and Kundu, D. (2005). Spatio-temporal models in small area estimation. *Survey Methodology*, **31**, 183195.

Srivastava A. K. (2009). Some aspects of estimating poverty at small area level. *Journal of the Indian Society of Agricultural Statistics*, **63** (**1**), 1-23.

Srivastava A. K. (2007). Small area estimation – a perspective and some applications. *Journal of the Indian Society of Agricultural Statistics*, **61** (**3**), 295-309.

Appendix

Table A1: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of MPCE for rural areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			MPCE	SE	CV	MPCE	SE	CV
Western	Saharanpur	96	1419	89	6.30	1361	78	5.76
	Muzaffarnagar	128	1366	73	5.31	1345	66	4.91
	Bijnor	96	1068	59	5.55	1087	56	5.12
	Moradabad	128	1081	51	4.71	1083	49	4.49
	Rampur	64	1092	80	7.31	1081	71	6.61
	Jyotiba Phule Nr	64	1012	74	7.28	1032	67	6.47
	Meerut	64	1958	191	9.75	1558	123	7.87
	Baghpat	32	1885	218	11.56	1542	130	8.44
	Ghaziabad	64	1454	147	10.14	1430	110	7.71
	Gautam B Nr	32	1547	123	7.93	1465	98	6.68
	Bulandshahar	96	1247	53	4.29	1247	51	4.07
	Aligarh	95	1135	87	7.66	1151	76	6.61
	Hathras	64	1546	133	8.58	1360	102	7.50
	Mathura	64	1109	84	7.59	1116	74	6.63
	Agra	96	1063	58	5.50	1080	55	5.07
	Firozabad	64	1014	83	8.14	1075	73	6.82
	Etah	64	1436	111	7.71	1338	91	6.79
	Mainpuri	64	836	39	4.68	853	38	4.46
	Budaun	96	1016	65	6.42	1018	61	5.98
	Bareilly	95	1168	55	4.67	1168	52	4.43
	Pilibhit	64	1021	65	6.33	1024	60	5.84
	Shahjahanpur	96	921	51	5.52	939	49	5.18
	Farrukhabad	64	1149	107	9.28	1146	89	7.74
	Kannauj	64	973	84	8.63	1023	75	7.33
	Etawah	64	1045	55	5.28	1045	52	5.00
	Auraiya	64	1087	65	5.97	1076	60	5.59
	Kashiramnagar	32	1230	88	7.19	1161	79	6.79
Central	Kheri	128	936	76	8.11	947	69	7.25
	Sitapur	128	1002	59	5.86	993	55	5.58
	Hardoi	128	967	46	4.80	965	45	4.63
	Unnao	96	861	48	5.59	867	46	5.32
	Lucknow	64	1130	110	9.69	1083	92	8.51
	Rae Bareli	128	930	43	4.60	930	41	4.44
	Kanpur Dehat	64	1104	101	9.15	1090	85	7.80
	Kanpur Nagar	64	1139	83	7.27	1126	74	6.53
	Fatehpur	96	777	38	4.91	791	37	4.69

Southern	Jalaun	64	993	67	6.74	987	62	6.25
	Jhansi	64	1070	58	5.39	1056	54	5.14
	Lalitpur	32	1061	40	3.73	1052	39	3.67
	Hamirpur	32	1079	67	6.25	1069	62	5.81
	Banda	64	774	52	6.71	793	49	6.22
	Chitrakoot	32	839	170	20.23	879	114	12.99
	Mahoba	32	975	114	11.68	976	92	9.40
Eastern	Mahrajganj	96	1012	81	7.99	984	73	7.38
	Pratapgarh	128	870	40	4.54	880	38	4.35
	Kaushambi	63	809	59	7.29	819	56	6.83
	Allahabad	128	991	43	4.32	999	41	4.14
	Barabanki	96	900	56	6.26	906	53	5.86
	Faizabad	64	1378	278	20.18	1080	132	12.22
	Ambedkar Nagar	96	1047	59	5.67	1041	55	5.32
	Sultanpur	128	1313	115	8.78	1197	92	7.71
	Bahraich	96	828	40	4.85	833	39	4.69
	Shrawasti	64	888	61	6.88	887	57	6.47
	Balrampur	63	892	65	7.29	895	61	6.77
	Gonda	128	1063	131	12.31	1034	100	9.68
	Siddharthnagar	96	1220	309	25.35	962	139	14.43
	Basti	96	861	78	9.02	885	70	7.86
	Sant K Nagar	64	1006	74	7.36	991	67	6.77
	Gorakhpur	128	993	42	4.25	996	41	4.10
	Kushinagar	128	1108	65	5.89	1087	61	5.59
	Deoria	96	988	70	7.11	999	64	6.45
	Azamgarh	128	1020	54	5.28	1020	51	5.00
	Mau	64	1000	55	5.51	1007	52	5.16
	Ballia	96	955	52	5.44	976	49	5.07
	Jaunpur	128	1115	65	5.86	1098	61	5.53
	Ghazipur	128	1051	50	4.71	1050	47	4.50
	Chandauli	64	1092	76	6.98	1087	69	6.31
	Varanasi	96	1136	61	5.41	1152	57	4.99
	Sant Ravidas Nr	64	873	69	7.91	929	63	6.83
	Mirzapur	96	1023	84	8.18	1024	74	7.21
	Sonbhadra	64	946	73	7.74	928	67	7.21

Table A2: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of poverty incidence for rural areas of Uttar Pradesh in 2011-12

Region	District	Sample Size	Direct			SAE		
			Poverty Incidence	SE	CV	Poverty Incidence	SE	CV
Western	Saharanpur	96	0.068	0.0362	53.23	0.089	0.0297	33.33
	Muzaffarnagar	128	0.052	0.0272	52.31	0.083	0.0257	30.95
	Bijnor	96	0.165	0.0524	31.78	0.160	0.0417	26.07
	Moradabad	128	0.131	0.0371	28.36	0.149	0.0365	24.48
	Rampur	64	0.240	0.0825	34.36	0.232	0.0574	24.76
	Jyotiba Phule Nr	64	0.268	0.0850	31.71	0.253	0.0589	23.28
	Meerut	64	0.002	0.00	0.00	0.060	0.0237	39.44
	Baghpat	32	0.138	0.0909	65.86	0.092	0.0374	40.67
	Ghaziabad	64	0.054	0.0367	68.04	0.063	0.0251	39.84
	Gautam B. Nr	32	0.018	0.0179	99.38	0.066	0.0279	42.32
	Bulandshahar	96	0.103	0.0377	36.59	0.110	0.0308	28.02
	Aligarh	95	0.181	0.0861	47.56	0.160	0.0452	28.23
	Hathras	64	0.013	0.0110	84.27	0.084	0.0311	37.08
	Mathura	64	0.179	0.0684	38.22	0.192	0.0504	26.25
	Agra	96	0.192	0.0628	32.73	0.179	0.0443	24.73
	Firozabad	64	0.252	0.0824	32.70	0.187	0.0496	26.52
	Etah	64	0.126	0.0716	56.79	0.139	0.0437	31.44
	Mainpuri	64	0.451	0.1198	26.57	0.343	0.0772	22.51
	Budaun	96	0.233	0.0829	35.57	0.245	0.0591	24.11
	Bareilly	95	0.047	0.0272	57.88	0.099	0.0329	33.20
	Pilibhit	64	0.177	0.0844	47.71	0.192	0.0546	28.43
	Shahjahanpur	96	0.270	0.0820	30.38	0.229	0.0549	23.96
	Farrukhabad	64	0.184	0.0907	49.27	0.190	0.0553	29.11
	Kannauj	64	0.308	0.1089	35.37	0.221	0.0601	27.19
	Etawah	64	0.093	0.0582	62.61	0.160	0.0516	32.23
	Auraiya	64	0.148	0.0540	36.51	0.205	0.0555	27.07
	Kashiramnagar	32	0.161	0.0744	46.19	0.257	0.0740	28.80
Central	Kheri	128	0.295	0.0717	24.30	0.288	0.0588	20.42
	Sitapur	128	0.324	0.0634	19.57	0.321	0.0570	17.76
	Hardoi	128	0.260	0.0579	22.26	0.287	0.0550	19.18
	Unnao	96	0.566	0.0756	13.36	0.499	0.0669	13.41
	Lucknow	64	0.347	0.0903	26.02	0.326	0.0670	20.55
	Rae Bareli	128	0.367	0.0604	16.46	0.360	0.0548	15.21
	Kanpur Dehat	64	0.152	0.0971	63.89	0.183	0.0543	29.68
	Kanpur Nagar	64	0.115	0.0473	41.16	0.161	0.0488	30.30
	Fatehpur	96	0.520	0.0722	13.88	0.453	0.0618	13.64
Southern	Jalaun	64	0.213	0.0760	35.69	0.236	0.0578	24.49
	Jhansi	64	0.117	0.0573	48.95	0.187	0.0506	27.06
	Lalitpur	32	0.144	0.0751	52.15	0.260	0.0752	28.91
	Hamirpur	32	0.169	0.0939	55.54	0.201	0.0659	32.78
	Banda	64	0.486	0.1038	21.35	0.434	0.0773	17.82
	Chitrakoot	32	0.204	0.1213	59.47	0.290	0.0780	26.89
	Mahoba	32	0.349	0.1693	48.51	0.295	0.0828	28.08

Eastern	Mahrajganj	96	0.354	0.0757	21.38	0.354	0.0652	18.42
	Pratapgarh	128	0.451	0.0711	15.77	0.403	0.0604	14.99
	Kaushambi	63	0.450	0.0863	19.17	0.430	0.0734	17.07
	Allahabad	128	0.244	0.0664	27.22	0.242	0.0517	21.35
	Barabanki	96	0.501	0.0895	17.86	0.437	0.0711	16.28
	Faizabad	64	0.287	0.0884	30.81	0.295	0.0677	22.97
	Ambedkar Nagar	96	0.310	0.0639	20.60	0.303	0.0535	17.65
	Sultanpur	128	0.210	0.0512	24.37	0.221	0.0449	20.34
	Bahraich	96	0.488	0.0873	17.90	0.437	0.0716	16.39
	Shrawasti	64	0.359	0.1008	28.09	0.364	0.0797	21.89
	Balrampur	63	0.196	0.0773	39.42	0.257	0.0678	26.39
	Gonda	128	0.274	0.0625	22.82	0.277	0.0544	19.64
	Siddharthnagar	96	0.263	0.0616	23.44	0.295	0.0583	19.77
	Basti	96	0.578	0.0746	12.91	0.506	0.0645	12.75
	Sant Kabir Nagar	64	0.325	0.0765	23.53	0.320	0.0634	19.81
	Gorakhpur	128	0.283	0.0563	19.89	0.275	0.0504	18.33
	Kushinagar	128	0.214	0.0567	26.52	0.238	0.0511	21.47
	Deoria	96	0.347	0.0746	21.49	0.322	0.0597	18.56
	Azamgarh	128	0.322	0.0585	18.16	0.315	0.0517	16.40
	Mau	64	0.146	0.0582	39.88	0.198	0.0539	27.24
	Ballia	96	0.267	0.0738	27.62	0.232	0.0564	24.31
	Jaunpur	128	0.177	0.0443	25.01	0.216	0.0473	21.91
	Ghazipur	128	0.236	0.0509	21.56	0.248	0.0477	19.25
	Chandauli	64	0.190	0.0663	34.87	0.205	0.0530	25.86
	Varanasi	96	0.192	0.0546	28.43	0.170	0.0415	24.40
	Sant Ravidas Nr	64	0.506	0.0880	17.40	0.380	0.0673	17.71
	Mirzapur	96	0.237	0.0522	22.05	0.247	0.0509	20.60
	Sonbhadra	64	0.375	0.0854	22.77	0.382	0.0698	18.27

Table A3: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of MPCE for urban areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			MPCE	SE	CV	MPCE	SE	CV
Western	Saharanpur	64	2118	262	12.39	2047	247	12.06
	Muzaffarnagar	64	2057	184	8.95	2012	179	8.87
	Bijnor	64	1405	127	9.02	1397	125	8.94
	Moradabad	64	1363	94	6.89	1360	93	6.85
	Rampur	32	988	69	6.99	988	69	6.96
	Jyotiba Phule Nr	32	2108	328	15.58	1945	300	15.44
	Meerut	96	2401	225	9.35	2334	215	9.2
	Baghpat	32	2290	205	8.97	2219	198	8.91
	Ghaziabad	96	4180	504	12.06	3439	416	12.1
	Gautam Buddha Nr	32	6453	609	9.44	4762	482	10.12
	Bulandshahar	64	1803	229	12.68	1756	218	12.43
	Aligarh	64	2009	173	8.61	1968	168	8.55
	Hathras	32	1335	116	8.67	1336	114	8.57
	Mathura	64	1445	103	7.14	1446	102	7.06
	Agra	96	1714	373	21.77	1715	333	19.39
	Firozabad	64	1229	87	7.07	1236	86	6.99

	Etah	32	2354	320	13.6	2191	294	13.4
	Mainpuri	32	1026	78	7.61	1030	78	7.54
	Budaun	32	1234	80	6.49	1229	80	6.48
	Bareilly	64	1311	80	6.12	1313	80	6.08
	Pilibhit	32	1419	164	11.58	1410	161	11.39
	Shahjahanpur	32	1175	100	8.49	1175	99	8.42
	Farrukhabad	32	1150	92	8.04	1157	92	7.92
	Kannauj	32	1027	80	7.82	1035	80	7.72
	Etawah	32	1118	102	9.13	1130	101	8.96
	Auraiya	32	1401	122	8.7	1412	120	8.53
	Kashiramnagar	32	1158	90	7.77	1158	89	7.72
Central	Kheri	32	894	87	9.72	902	86	9.56
	Sitapur	32	1400	261	18.64	1410	246	17.43
	Hardoi	32	1046	78	7.45	1051	77	7.37
	Unnao	32	1273	126	9.88	1285	124	9.65
	Lucknow	128	2318	296	12.79	2296	276	12.02
	Rae Bareli	32	1742	350	20.11	1756	316	18
	Kanpur Dehat	32	1499	129	8.62	1509	127	8.43
	Kanpur Nagar	128	1956	162	8.29	1966	159	8.07
	Fatehpur	32	1214	127	10.45	1229	125	10.17
Southern	Jalaun	32	1659	174	10.47	1659	169	10.18
	Jhansi	64	2507	562	22.42	2407	451	18.74
	Lalitpur	32	1620	108	6.66	1629	107	6.55
	Hamirpur	32	1437	155	10.78	1457	152	10.41
	Banda	32	1120	68	6.08	1127	68	6.02
	Chitrakoot	32	791	65	8.18	796	64	8.1
	Mahoba	32	1179	87	7.39	1184	87	7.31
Eastern	Mahrajganj	32	1328	167	12.58	1335	163	12.21
	Pratapgarh	32	1458	186	12.78	1477	181	12.23
	Kaushambi	32	867	79	9.11	878	79	8.95
	Allahabad	63	3436	564	16.41	2940	450	15.3
	Barabanki	32	911	99	10.83	923	98	10.59
	Faizabad	32	1632	310	19.01	1668	286	17.13
	Ambedkar Nagar	32	868	70	8.03	875	69	7.92
	Sultanpur	31	1847	277	15	1832	260	14.17
	Bahraich	32	1313	183	13.93	1313	178	13.52
	Shrawasti	30	1224	196	16	1190	190	15.97
	Balrampur	32	1076	90	8.36	1077	89	8.29
	Gonda	32	2488	207	8.32	2414	199	8.25
	Siddharthnagar	32	1178	145	12.33	1186	143	12.02
	Basti	32	1371	159	11.62	1375	156	11.34
	Sant Kabir Nagar	32	1153	165	14.34	1173	161	13.74
	Gorakhpur	64	1820	172	9.45	1820	168	9.21
	Kushinagar	32	1376	180	13.09	1368	175	12.79
	Deoria	32	1306	163	12.5	1306	160	12.23
	Azamgarh	32	1734	320	18.44	1719	293	17.03
	Mau	32	1210	132	10.93	1235	130	10.54
	Ballia	32	1348	151	11.23	1361	148	10.89
	Jaunpur	32	1522	231	15.15	1513	220	14.56
	Ghazipur	32	1280	143	11.13	1288	140	10.87
	Chandauli	32	2875	377	13.13	2552	336	13.15
	Varanasi	96	1572	127	8.11	1585	126	7.93
	Sant Ravidas Nr.	32	902	61	6.72	905	60	6.67
	Mirzapur	32	1169	157	13.41	1201	153	12.77
	Sonbhadra	31	2039	169	8.3	2021	165	8.17

Table A4: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of poverty incidence for urban areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			Poverty Incidence	SE	CV	Poverty Incidence	SE	CV
Western	Saharanpur	64	0.153	0.0603	39.43	0.165	0.0407	24.69
	Muzaffarnagar	64	0.166	0.0621	37.43	0.179	0.0407	22.76
	Bijnor	64	0.199	0.0614	30.85	0.206	0.0490	23.78
	Moradabad	64	0.231	0.0698	30.21	0.235	0.0473	20.14
	Rampur	32	0.576	0.1144	19.86	0.547	0.0918	16.78
	Jyotiba Phule Nr	32	0.148	0.0678	45.83	0.175	0.0595	34.00
	Meerut	96	0.035	0.0184	52.68	0.050	0.0187	37.42
	Baghpat	32	0.068	0.0395	58.08	0.110	0.0492	44.72
	Ghaziabad	96	0.010	0.0063	63.25	0.023	0.0100	43.48
	Gautam B Nr	32	0.010	0.0055	54.77	0.026	0.0138	53.02
	Bulandshahar	64	0.104	0.0493	47.40	0.116	0.0339	29.23
	Aligarh	64	0.113	0.0381	33.70	0.129	0.0392	30.42
	Hathras	32	0.226	0.0938	41.51	0.244	0.0683	28.01
	Mathura	64	0.201	0.0587	29.18	0.208	0.0382	18.37
	Agra	96	0.222	0.0555	25.00	0.224	0.0373	16.64
	Firozabad	64	0.314	0.0696	22.18	0.305	0.0424	13.91
	Etah	32	0.103	0.0507	49.22	0.124	0.0508	40.96
	Mainpuri	32	0.417	0.1053	25.24	0.401	0.0896	22.35
	Budaun	32	0.219	0.0764	34.89	0.256	0.0828	32.35
	Bareilly	64	0.114	0.0498	43.68	0.127	0.0316	24.90
	Pilibhit	32	0.156	0.0796	51.00	0.180	0.0663	36.81
	Shahjahanpur	32	0.296	0.1018	34.39	0.314	0.0841	26.78
	Farrukhabad	32	0.333	0.1003	30.13	0.322	0.0707	21.96
	Kannauj	32	0.433	0.0970	22.40	0.415	0.0625	15.07
	Etawah	32	0.474	0.1001	21.12	0.433	0.0756	17.47
	Auraiya	32	0.132	0.0570	43.19	0.134	0.0487	36.33
	Kashiramnagar	32	0.399	0.1081	27.10	0.399	0.0815	20.44
Central	Kheri	32	0.630	0.1039	16.50	0.579	0.0822	14.19
	Sitapur	32	0.385	0.1321	34.30	0.344	0.0923	26.83
	Hardoi	32	0.486	0.0999	20.57	0.466	0.0727	15.61
	Unnao	32	0.293	0.0948	32.36	0.294	0.0713	24.24
	Lucknow	128	0.160	0.0437	27.31	0.161	0.0253	15.71
	Rae Bareli	32	0.329	0.1356	41.22	0.302	0.0904	29.95
	Kanpur Dehat	32	0.158	0.0802	50.75	0.160	0.0575	35.96
	Kanpur Nagar	128	0.102	0.0430	42.17	0.106	0.0212	20.01
	Fatehpur	32	0.365	0.1012	27.74	0.359	0.0663	18.48
Southern	Jalaun	32	0.092	0.0775	84.20	0.105	0.0438	41.73
	Jhansi	64	0.149	0.0497	33.36	0.146	0.0295	20.20
	Lalitpur	32	0.021	0.0155	73.77	0.043	0.0221	51.48
	Hamirpur	32	0.243	0.0834	34.33	0.231	0.0541	23.43
	Banda	32	0.414	0.1063	25.67	0.396	0.0736	18.59
	Chitrakoot	32	0.600	0.1041	17.34	0.551	0.0957	17.36
	Mahoba	32	0.291	0.1002	34.43	0.284	0.0752	26.47

Eastern	Mahrajganj	32	0.386	0.1033	26.77	0.373	0.0852	22.84
	Pratapgarh	32	0.395	0.1098	27.80	0.377	0.0804	21.34
	Kaushambi	32	0.609	0.1071	17.59	0.579	0.0692	11.95
	Allahabad	63	0.121	0.0370	30.59	0.127	0.0341	26.82
	Barabanki	32	0.736	0.0882	11.98	0.667	0.0802	12.03
	Faizabad	32	0.188	0.0801	42.62	0.192	0.0643	33.51
	Ambedkar Nr	32	0.654	0.0932	14.25	0.604	0.0697	11.54
	Sultanpur	31	0.212	0.0834	39.32	0.201	0.0667	33.19
	Bahraich	32	0.137	0.0845	61.68	0.177	0.0656	37.05
	Shrawasti	30	0.460	0.0991	21.54	0.463	0.0887	19.16
	Balrampur	32	0.388	0.1064	27.43	0.387	0.0859	22.20
	Gonda	32	0.015	0.0110	73.03	0.073	0.0367	50.33
	Siddharthnagar	32	0.340	0.0982	28.88	0.337	0.0856	25.39
	Basti	32	0.395	0.1015	25.69	0.372	0.0856	23.00
	Sant Kabir Nr	32	0.477	0.1067	22.36	0.445	0.0785	17.64
	Gorakhpur	64	0.113	0.0475	42.07	0.122	0.0387	31.75
	Kushinagar	32	0.504	0.1044	20.71	0.473	0.0884	18.70
	Deoria	32	0.390	0.1317	33.76	0.382	0.1217	31.87
	Azamgarh	32	0.249	0.0796	31.95	0.252	0.0559	22.17
	Mau	32	0.301	0.0975	32.40	0.276	0.0660	23.92
	Ballia	32	0.341	0.0922	27.04	0.335	0.0685	20.44
	Jaunpur	32	0.271	0.0917	33.84	0.264	0.0901	34.11
	Ghazipur	32	0.318	0.1020	32.07	0.295	0.0868	29.42
	Chandauli	32	0.003	0.0032	105.4	0.065	0.0370	56.94
	Varanasi	96	0.131	0.0407	31.10	0.132	0.0239	18.09
	Sant Ravidas Nr	32	0.640	0.0964	15.06	0.602	0.0745	12.38
	Mirzapur	32	0.571	0.1061	18.58	0.537	0.0699	13.01
	Sonbhadra	31	0.018	0.0184	102.4	0.073	0.0382	52.34

Table A5: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of worker population ratio for rural areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			WPR	SE	CV	WPR	SE	CV
Western	Saharanpur	557	287	29.65	10.33	307	23.02	7.50
	Muzaffarnagar	785	291	25.08	8.62	303	20.49	6.76
	Bijnor	557	321	30.99	9.66	327	23.66	7.24
	Moradabad	768	343	24.19	7.05	335	20.49	6.12
	Rampur	407	321	34.27	10.68	317	24.70	7.79
	Jyotiba Phule Nr.	365	260	34.02	13.09	296	24.49	8.28
	Meerut	354	349	37.80	10.83	333	25.69	7.71
	Baghpat	184	224	48.55	21.68	280	28.46	10.16
	Ghaziabad	392	272	34.56	12.71	298	24.49	8.22
	Gautam Buddha Nr.	171	360	50.61	14.06	323	28.98	8.97
	Bulandshahar	500	272	26.15	9.61	303	21.21	7.00
	Aligarh	536	399	38.81	9.73	363	26.65	7.34
	Hathras	348	413	42.74	10.35	357	28.28	7.92
	Mathura	412	249	31.07	12.48	291	23.24	7.99
	Agra	582	297	29.04	9.78	307	22.14	7.21
	Firozabad	379	286	34.30	11.99	299	24.49	8.19
	Etah	346	366	42.11	11.51	328	27.02	8.24
	Mainpuri	387	274	43.95	16.04	301	27.20	9.04
	Budaun	539	323	35.71	11.06	322	25.30	7.86

	Bareilly	515	341	33.11	9.71	328	24.90	7.59
	Pilibhit	349	432	44.95	10.41	358	28.64	8.00
	Shahjahanpur	541	355	39.04	11.00	335	26.65	7.95
	Farrukhabad	400	252	36.51	14.49	290	25.69	8.86
	Kannauj	379	328	43.88	13.38	320	27.57	8.62
	Etawah	334	294	43.13	14.67	320	27.57	8.62
	Auraiya	315	294	39.09	13.30	323	27.57	8.54
	Kashiramnagar	205	282	44.25	15.69	306	27.39	8.95
Central	Kheri	686	363	34.33	9.46	364	26.83	7.37
	Sitapur	701	359	28.38	7.90	366	24.08	6.58
	Hardoi	727	343	27.67	8.07	355	23.24	6.55
	Unnao	470	475	36.25	7.63	430	27.57	6.41
	Lucknow	350	373	36.53	9.79	369	28.11	7.62
	Rae Bareli	700	377	26.45	7.01	377	22.58	5.99
	Kanpur Dehat	283	350	51.15	14.61	347	29.66	8.55
	Kanpur Nr	299	396	46.09	11.64	369	29.15	7.90
	Fatehpur	445	340	33.62	9.89	362	26.27	7.26
Southern	Jalaun	349	277	34.91	12.60	325	26.08	8.02
	Jhansi	286	409	42.21	10.32	402	30.66	7.63
	Lalitpur	134	324	63.01	19.45	353	34.21	9.69
	Hamirpur	160	308	63.02	20.46	350	33.47	9.56
	Banda	390	378	44.12	11.67	367	28.98	7.90
	Chitrakoot	145	507	66.83	13.18	398	34.35	8.63
	Mahoba	145	423	79.24	18.73	378	36.06	9.54
Eastern	Mahrajganj	549	360	32.45	9.01	355	24.90	7.01
	Pratapgarh	765	369	30.08	8.15	356	23.02	6.47
	Kaushambi	308	373	40.06	10.74	380	29.33	7.72
	Allahabad	745	351	30.60	8.72	359	24.70	6.88
	Barabanki	595	395	37.95	9.61	383	27.20	7.10
	Faizabad	345	386	41.20	10.67	357	27.39	7.67
	Ambedkar Nr.	546	381	28.85	7.57	362	22.80	6.30
	Sultanpur	708	371	26.19	7.06	361	21.68	6.01
	Bahraich	488	281	36.15	12.86	310	26.65	8.60
	Shrawasti	317	270	42.64	15.79	308	27.75	9.01
	Balrampur	310	300	43.37	14.46	314	27.93	8.89
	Gonda	705	255	27.93	10.95	289	22.36	7.74
	Siddharthnagar	541	350	30.69	8.77	337	23.87	7.08
	Basti	534	410	32.60	7.95	371	24.29	6.55
	Sant Kabir Nr.	394	265	30.10	11.36	291	22.80	7.84
	Gorakhpur	777	286	23.12	8.08	301	20.25	6.73
	Kushinagar	727	314	27.07	8.62	317	22.14	6.98
	Deoria	584	328	30.45	9.28	316	22.80	7.22
	Azamgarh	896	290	22.18	7.65	304	20.49	6.74
	Mau	374	261	31.35	12.01	294	24.90	8.47
	Ballia	630	333	33.79	10.15	324	24.70	7.62
	Jaunpur	796	376	26.43	7.03	361	22.14	6.13
	Ghazipur	816	364	24.76	6.80	354	20.74	5.86
	Chandauli	395	420	37.56	8.94	373	26.65	7.14
	Varanasi	636	410	28.27	6.89	372	22.58	6.07
	Sant Ravidas Nr.	467	265	29.50	11.13	288	23.45	8.14
	Mirzapur	566	360	27.94	7.76	358	22.80	6.37
	Sonbhadra	347	400	36.46	9.11	383	27.02	7.05

Table A6: District-wise sample size, direct estimates (Direct) and small area estimates (SAE) along with their standard error (SE) and percentage coefficient of variations (CV) of worker population ratio for urban areas of Uttar Pradesh in 2011-12

Region	District	Sample size	Direct			SAE		
			WPR	SE	CV	WPR	SE	CV
Western	Saharanpur	298	332	34.90	10.51	313	18.71	5.98
	Muzaffarnagar	322	294	30.04	10.22	315	18.97	6.02
	Bijnor	354	281	31.08	11.06	313	19.49	6.23
	Moradabad	349	302	29.97	9.92	330	20.25	6.14
	Rampur	171	473	54.68	11.56	393	26.83	6.83
	Jyotiba Phule Nr.	163	336	44.27	13.18	337	20.98	6.22
	Meerut	470	296	30.56	10.32	305	18.44	6.05
	Baghpat	126	268	47.82	17.84	311	19.75	6.35
	Ghaziabad	402	321	30.59	9.53	312	18.97	6.08
	Gautam Buddha Nr.	117	309	52.71	17.06	350	27.02	7.72
	Bulandshahar	289	322	32.38	10.06	314	19.24	6.13
	Aligarh	339	247	31.45	12.73	296	19.24	6.50
	Hathras	194	315	39.27	12.47	287	20.74	7.23
	Mathura	336	294	28.56	9.71	302	18.17	6.02
	Agra	509	298	30.20	10.13	295	20.98	7.11
	Firozabad	390	303	28.53	9.42	317	19.49	6.15
	Etah	159	321	43.45	13.54	294	19.24	6.54
	Mainpuri	151	306	44.39	14.51	270	20.25	7.50
	Budaun	194	247	38.66	15.65	316	22.36	7.08
	Bareilly	315	390	35.20	9.03	377	23.02	6.11
	Pilibhit	183	368	46.58	12.66	346	21.45	6.20
	Shahjahanpur	170	334	45.43	13.60	334	20.98	6.28
	Farrukhabad	199	261	35.81	13.72	305	18.71	6.13
	Kannauj	191	386	39.50	10.23	346	20.49	5.92
	Etawah	161	261	38.46	14.74	274	18.97	6.92
	Auraiya	148	226	38.16	16.88	260	19.49	7.50
	Kashiramnagar	186	391	42.69	10.92	338	20.98	6.21
Central	Kheri	190	302	40.70	13.48	323	19.75	6.11
	Sitapur	160	367	56.16	15.30	329	20.00	6.08
	Hardoi	151	394	44.86	11.39	329	19.75	6.00
	Unnao	158	313	46.56	14.88	312	19.24	6.17
	Lucknow	656	367	30.46	8.30	340	20.49	6.03
	Rae Bareli	152	304	56.98	18.74	303	19.24	6.35
	Kanpur Dehat	159	339	50.02	14.75	299	19.24	6.43
	Kanpur Nr.	548	308	30.54	9.92	320	20.00	6.25
	Fatehpur	174	247	35.13	14.22	297	18.44	6.21
Southern	Jalaun	157	275	44.29	16.10	276	20.25	7.34
	Jhansi	272	311	35.65	11.46	302	21.45	7.10
	Lalitpur	163	262	39.92	15.24	305	19.75	6.47
	Hamirpur	134	270	42.39	15.70	294	20.25	6.89
	Banda	125	288	49.17	17.07	297	19.24	6.48
	Chitrakoot	153	286	46.58	16.29	274	20.49	7.48
	Mahoba	143	264	45.44	17.21	294	19.24	6.54
Eastern	Mahrajganj	162	298	44.59	14.96	290	19.24	6.63
	Pratapgarh	187	330	46.89	14.21	315	19.75	6.27
	Kaushambi	154	311	47.11	15.15	322	21.91	6.80
	Allahabad	267	259	35.20	13.59	309	20.49	6.63
	Barabanki	154	344	51.86	15.08	342	20.98	6.13
	Faizabad	166	335	51.88	15.49	319	20.25	6.35

	Ambedkar Nr.	201	328	36.72	11.19	318	19.24	6.05
	Sultanpur	176	269	40.93	15.22	294	20.00	6.80
	Bahraich	161	338	48.61	14.38	340	21.45	6.31
	Shrawasti	155	270	37.74	13.98	316	22.58	7.15
	Balrampur	168	328	44.89	13.69	335	21.45	6.40
	Gonda	125	280	47.65	17.02	315	20.49	6.51
	Siddharthnagar	152	392	50.02	12.76	326	19.75	6.06
	Basti	162	304	43.75	14.39	272	20.49	7.53
	Sant Kabir Nr.	158	285	41.50	14.56	305	18.97	6.22
	Gorakhpur	344	242	33.07	13.66	293	19.24	6.56
	Kushinagar	179	359	41.52	11.57	315	20.98	6.66
	Deoria	177	222	47.75	21.51	276	21.21	7.69
	Azamgarh	202	293	34.59	11.80	305	18.71	6.13
	Mau	168	402	51.15	12.72	342	21.45	6.27
	Ballia	192	293	36.41	12.43	310	19.49	6.29
	Jaunpur	205	311	45.03	14.48	294	20.74	7.05
	Ghazipur	192	266	39.16	14.72	287	19.75	6.88
	Chandauli	166	300	46.43	15.48	290	19.24	6.63
	Varanasi	541	372	27.93	7.51	352	20.25	5.75
	Sant Ravidas Nr.	179	487	42.23	8.67	338	20.00	5.92
	Mirzapur	157	299	44.96	15.04	320	20.00	6.25
	Sonbhadra	144	319	44.26	13.88	290	20.25	6.98