Journal of Quantitative Economics
https://doi.org/10.1007/s40953-023-00338-x

ORIGINAL ARTICLE

Hierarchical Bayes Measurement Error Small Area Model for Estimation of Disaggregated Level Workers Mobility Pattern in India

Priyanka Anjoy[1]

Accepted: 16 January 2023

Abstract

Periodic Labour Force Survey (PLFS) is the major source of data on various labour force indicators in India at annual or quarterly basis which is on the field since 2017–18. It has strategically reformed the previous quinquennial Employment and Unemployment Survey of National Statistical Office, India. Mobility pattern of workers, basically in terms of commuting is one of the key information contained therein which essentially entails the workplace characteristics of the workforce. In this article PLFS 2017–18 and 2018–19 data is analysed which depicts state-wise large disparities in the commuting behaviour of workers, whereas most of the workers are out-commuting from rural areas. The potential reason behind is the rapid pace of urbanization and associated improved transportation facilities as well as search for stable non-farm employment opportunities by the rural workforce. Further, the planning of urbanization or creation of employment opportunities at rural places in each state requires within-state regional or disaggregated level information of workplaces, spatial concentration of works and workers. To pursue that, disaggregated level analysis of commuting pattern of workers is done using small area estimation approach. In particular, this article describes hierarchical Bayes (HB) measurement error (ME) small area model for binary variable of interest indicating whether individual in the workforce is commuting or not. The HBME model has been implemented to obtain district level rural commuters proportions in Uttar Pradesh state of India. This state specifically tops amongst the states in the number of rural commuters. A spatial map has been generated for visual inspection of disparity in commuting behaviour of workers, also such map is useful to the policy makers and administration for framing decentralized level plans or strategies eyeing stable mobility behaviour to persuade improvement in employment rate.

Keywords Commuting · Periodic Labour Force Survey · Small area estimation · Spatial map

✉ Priyanka Anjoy
anjoypriyanka90@gmail.com

Extended author information available on the last page of the article

Published online: 10 February 2023

Introduction

A United Nations report projects that by 2050, 68% of the world's population will live in urban settlements compared to the current 55%. Closer home, over 30% of the Indian population currently lives in urban areas. This will grow to 40% by 2030 and 58% by 2050 (United Nations 2019). The rapid pace of urbanization will be accompanied by increased motorisation and transportation facilities, creation of diverse employment opportunities, gradual change of the urban built environment resulting in attracting and retaining rural people in large numbers (Zhu et al. 2017). Existing studies suggest that, on average living standards are comparatively higher in urban populations and there is strong positive association between the living standards and income of the people. Hence, the search for a non-farm employment and better lifestyle opportunities constitute the major drivers for the mobility behaviour of rural populations. In contrast to this, there are also studies which shows fairly opposite mobility behaviour from urban to rural places. In United Kingdom, about one in ten people have changed residence annually during the last 35 years indicating that mobility in terms of change in residence or workplace is quite common (Brown et al. 2015). Unlike developing and under-developed countries, the urban to rural mobility pattern is profound in developed countries.

Structurally there are two forms of population mobility, one is migration and another is commuting. Migration is a permanent or semi-permanent change of residence of sufficient duration and distance to interrupt everyday activity patterns whether commuting is a form of population circulation that typically involves a daily journey between a permanent residence and a fixed workplace (Green 2004; Brown et al. 2015). The national as well as cross-border migration is heavily studied in most of the countries, but commuting is one of the issue which is not researched extensively (Adamson 2006; McAuliffe and Khadria 2020). Although the fact is that number of daily commuters between rural and urban areas are much more than either seasonal or permanent migration in a year (Chandrasekhar et al. 2017). Additionally, due to some of the issues like rural–urban wage differential and social inequalities, increased dependence on non-farm employment and salaried job, urbanization, improved transportation facilities and reduced cost to travel in neighbouring cities, commuting has become the far more important channel to be facilitated and understood (Bhatt et al. 2020).

In this paper Periodic Labour Force Survey (PLFS) data for the year 2017–18 and 2018–19 in India is analysed to understand the rural–urban two way commuting behaviour of workers. The National Statistical Office (NSO) in India is conducting PLFS since 2017–18 to obtain data on various labour force indicators, this survey has strategically reformed the earlier quinquennial Employment-Unemployment survey of NSO. Currently PLFS is the major source to study the worker's mobility pattern on regular basis. In both the studied years, state-wise large disparities have been found in commuting behaviour and trend while majority of the commuters are out-commuting from rural places in both the years. Uttar Pradesh shows the highest number of workers commuting from rural to urban areas whereas West Bengal constitutes highest number of urban to rural commuters. In continuation to this, attempt is to examine and measure the disaggregated level (i.e., district level) commuting pattern of particularly rural populations. Since, PLFS is designed to obtain reliable or satisfactory estimates at the national and the state level (i.e., aggregated level), the district level estimates obtained using this survey data will be poor as the sample sizes at this level are very low or even zero to provide direct estimates of reasonable precision. In this backdrop, suitable model-based small area estimation (SAE) approach is used for disaggregated level estimation of commuters proportions. The basic idea in SAE technique is to utilize existing survey or other database (i.e., census or administrative records) information in an implicit or explicit modelling framework to provide suitable small area estimates. Understanding the commuting behaviour at disaggregated or micro level of administration is essential in planning regional policies by the Central or State Governments targeting stable mobility behaviour of the workers.

The theoretical literatures of SAE techniques have technically developed towards two major directions. The First one is area level modeling, which relies on area level target and auxiliary variables to provide small domain level predictions. In contrast, unit level small area model uses unit level target and auxiliary variables for small area predictions. Amongst these, area level small area models are accepted well to account for complex survey design information which is otherwise a major problem in case of unit level models. Fay and Herriot (1979) were pioneering in forwarding the concept of area level model, popular as Fay–Herriot (FH) model. Further developments after FH model and its various extensions are duly covered in Rao (2003), Rao and Molina (2015) along with diverse practical applications. Infact, in past few years SAE techniques have become the integral component of various countries official statistics system due to its relative ability in providing quick decentralized level estimates and hence assisting in decentralized level policy reforms. In India, although SAE technique is not the part of regular NSO publications, however several inspiring applications can be found in Chandra et al. (2017, 2018, 2019), Anjoy et al. (2019), Anjoy and Chandra (2020), Anjoy and Chandra (2021), Guha and Chandra (2022). These applications are broadly disaggregated level estimation of poverty proportions, food insecurity proportions, indebtness proportions, earning inequality and crop total yield. This also induces motivation behind the present article to obtain disaggregated level estimates of worker's mobility parameters using SAE approach. In particular, to study the disaggregated level commuting behaviour area level measurement error (ME) small area model has been employed here in hierarchical Bayes (HB) framework motivated from Arima et al. (2015) and Burgard et al. (2019).

The area level FH model typically involves auxiliary variables which are available from census or another survey. It is often discussed issue that census is infrequent in nature. In India it is conducted decennially. Obtaining relevant auxiliary variables from another survey is comparatively easier and quick. In such a situation, it is necessary to take into account the variability in measuring auxiliary variables obtained from another survey. This idea has led to the ME small area model. Further, most of the ME literatures in SAE till date have been reported for continuous data as the underlined FH model is designed for continuous data (Ybarra and Lohr 2008; Arima et al. 2015; Burgard et al. 2019). In contrast to these earlier literatures, this article attempt to describe methodology of ME SAE for binary data in HB framework. Thereby, implement this HB ME small area model to estimate the small domain proportions of commuters in the workforce.

Rest of the article is organized as follows. After the introduction, a theoretical background before leading this work is presented. HBME small area model and its computational approach is explained in "Description of Methodology". In "Data Description", discussion about PLFS data and commuting pattern at aggregated level has been elaborated. In "Analysis and Interpretation", regional disparity in commuting pattern is described and small area estimates of commuters proportions for the state of Uttar Pradesh is presented. A spatial map has been generated for visual inspection of disparity in commuting behaviour of workers, also such map is useful to the policy makers and administration for framing decentralized level plans or strategies. The paper ends with relevant concluding remarks.

Background

SAE technique is known for its statistical efficiency to address the need of small domain or disaggregated level estimation. The nationwide large-scale surveys designed for the national or the state level estimation often mask the variations or heterogeneities at regional or local level. Traditional direct estimation technique is not a suitable option to derive reliable estimates at this level cause sample sizes are insufficient, negligible or even zero. Such areas named as small areas or small domains and are often of interest to the Government and private agencies in policy formulation targeting inclusive developments. The SAE methodology provides a viable and cost effective solution to address the problem of small sample sizes in small domains overcoming the drawback of direct estimation technique.

The PLFS data collected in India is based on stratified multi-stage sampling design with National Sample Survey (NSS) regions being the strata. NSS regions basically comprise of several districts within a state having similar agro-climatic conditions and socio-economic features. The rural areas of each NSS region constituted rural stratum and urban areas of each NSS region constituted urban stratum. As unplanned domains below the NSS regions (i.e., districts or lower administrative units) do not get sufficient sample sizes to produce direct estimates with acceptable accuracy and further certain domains with null sample sizes it is not even possible to get direct estimates. Hence, to deal with the situation SAE techniques are considered as promising alternative. Basically, SAE invokes the idea of borrowing strength from related areas or domains and thus improve the effective sample sizes for particular domain resulting in precise estimates for small areas or unplanned domains where direct estimation attempt may fail (Rao 2003).

Further, various literatures have discussed the issue of area level small domain estimation from either frequentist or Bayesian perspectives (Chandra 2013; Liu et al. 2014; Rao and Molina 2015; Chandra and Chandra 2015, 2020). The Bayesian approach has gained much popularity in recent years due to its flexibility in yielding quick and easier mean squared error (MSE) computation which is posterior variance; additionally, posterior mean or point estimate known to include more reasonable credible interval region (Gelman 2006). In Bayes framework estimations are described by assuming particular probability distributions, which render the opportunities to analyze the uncertainties involved in the decision process (Ghosh et al. 2009; Lee et al. 2015). The range of Bayesian methods include empirical best prediction (EBP) and HB area-level and unit level models covered in varied small area literatures (Gelman 2006; Jiang and Lahiri 2006; You 2008; Souza et al. 2009; Ghosh et al. 2009; Liu et al. 2014; Lee et al. 2015; Anjoy and Chandra 2021). The HB approach assumes particular prior distributions for the hyperparameters to obtain posterior quantities of the parameter of interest (Rao and Molina 2015). The HB approach has the flexibility to deal with complex SAE model as it overcomes the difficulties of analytical MSE estimation in frequentist set up and provides quick and easier posterior variance computation based on Markov chain Monte Carlo (MCMC) simulation (You and Rao 2002; Anjoy and Chandra 2021). In frequentist approach of SAE, unlike estimation of survey weighted linear parameters like small area means and totals, there has been comparatively little research on estimation of survey weighted small area proportions under area-level small area models. In contrast, Bayesian framework of SAE, in particular HB approach of SAE, incorporates the survey weights in estimating small area proportions under area-level small area models, see in Liu et al. (2014).

Various studies have reported ME SAE model with area or unit level covariates measured with error from either frequentist or Bayesian perspectives (Ghosh et al. 2006; Ghosh and Sinha 2007; Ybarra and Lohr 2008; Torabi et al. 2009; Datta et al. 2010; Arima et al. 2012, 2015; Burgard et al. 2019). Refer Ghosh et al. 2006; Ghosh and Sinha 2007; Torabi et al. 2009, Datta et al. 2010; Arima et al. 2012 for ME model conceptualized with unit level auxiliary variables. Ybarra and Lohr (2008) described area level ME FH model. Arima et al. (2015) has described HB version of the area level ME model of Ybarra and Lohr. Burgard et al. (2019) has discussed ME FH model from frequentist perspectives with normality assumption for measurement error and their model specifically takes maximum likelihood and restricted maximum likelihood estimation technique to obtain more robust predictor than moment-based estimation method described in Ybarra and Lohr (2008). They applied it to estimate poverty proportions in the Spanish Living Condition Survey with auxiliary information from the Spanish Labour Force Survey. The attempt in this article is to estimate disaggregate level commuting proportions of workers based on PLFS data in India. The methodological framework is based on HB approach to obtain suitable small area estimates. The next section delineates the ME SAE methods for area level data in detail.

Description of Methodology

Measurement Error Small Area Model

Let us consider a finite population U of size N which is partitioned into D distinct small areas or simply areas. The set of population units in area i is denoted as U_i with known size N_i such that $U = \bigcup_{i=1}^{D} U_i$ and $N = \sum_{i=1}^{D} N_i$. A sample s of size n is drawn from population U using a probabilistic mechanism. This resulted in sample s_i in area i with size n_i, so that $s = \bigcup_{i=1}^{D} s_i$ and $n = \sum_{i=1}^{D} n_i$. Assume, y_{ij} be the value of target variable y for unit j ($j = 1, \ldots, n_i$) in small area i with values either 1 or 0, i.e., binary response. Our aim is to estimate the small domain proportions $P_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij}$. Let, p_i be the direct survey estimator of P_i obtained from target sample data. The direct estimator can either be unweighted or survey-weighted. The expression of survey-weighted direct estimator of P_i is $p_{iw} = \left(\sum_{j=1}^{n_i} w_{ij}\right)^{-1} \sum_{j=1}^{n_i} w_{ij} y_{ij}$, where w_{ij} is the survey-weight of individual sampling units y_{ij}. The sampling model of the ME structure attaches the direct survey estimator p_{iw} to the population parameter P_i with a sampling error e_i which can be expressed as below,

$$p_{iw} = P_i + e_i; \quad i = 1, \ldots, D.$$

The sampling errors are independent and usually having normal distribution with mean 0 and known sampling variance σ_{ei}^2. Anjoy et al. (2019) has given the expression of survey variance σ_{ei}^2 computed from survey-weighted sample data. The linking model of P_i attempt to relate with area-specific auxiliary variables and random effect component.

$$g(P_i) = \eta_i = \mathbf{x}_i' \boldsymbol{\beta} + v_i; \quad i = 1, \ldots, D,$$

where the linking function $g(.)$ is logit for binary data and log for count data, $\mathbf{x}_i' = (x_{i1}, x_{i2}, \ldots x_{ip})$ represent the p-dimensional row vector of area-specific auxiliary variables, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$ is the regression coefficient vector of dimension p and v_i being the area-specific random effect assumed to be independent and identically distributed as $E(v_i) = 0$ and $\text{var}(v_i) = \sigma_v^2$. Random area-specific effects are included in the linking model to account for between areas heterogeneity.

Now, let us consider the situation in which auxiliary data are available from other surveys or are measured with error. We represent the measurement error model of ME structure as below,

$$\mathbf{x}_i = \mathbf{W}_i + \mathbf{d}_i; \quad i = 1, \ldots, D,$$

where the true auxiliary variable $\mathbf{x}_i$ being estimated by $\mathbf{W}_i$ based on other surveys with random measurement error $\mathbf{d}_i \sim N(0, \boldsymbol{\Psi}_i)$, $\boldsymbol{\Psi}_i$ is the p-dimensional variance–covariance matrix assumed to be known. We assume auxiliary variables are independent of each other, hence $\boldsymbol{\Psi}_i$ is diagonal. The combined form of the ME SAE model is expressed as,

$$p_{iw} = \mathbf{W}_i' \boldsymbol{\beta} + \gamma_i + v_i + e_i; \quad i = 1, \ldots, D,$$

where $\gamma_i = (\mathbf{x}_i - \mathbf{W}_i)' \boldsymbol{\beta} = \mathbf{d}_i' \boldsymbol{\beta}$ is the component due to taking into account the measurement error in auxiliary variables. We assume the error terms v_i, e_i and measurement error component γ_i are independent of each other. Aggregating D area level models lead to the population level version of the ME model,

$$\mathbf{p} = \mathbf{W\beta} + \mathbf{Z\theta} + \Lambda + \mathbf{v} + \mathbf{e},$$

where $\mathbf{p} = \left(p_{1w}, \ldots, p_{Dw}\right)'$ is the vector of direct survey estimates of target survey variable, $\mathbf{W} = \left(\mathbf{W}_1', \ldots, \mathbf{W}_D'\right)'$ be $D \times p$ matrix of auxiliary variates, $\boldsymbol{\beta}$ is the fixed effect parameter vector, $\mathbf{Z} = \left(\mathbf{Z}_1', \ldots, \mathbf{Z}_D'\right)'$ be the $D \times q$ matrix of auxiliary variables which are measured without error, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q)'$ is the q-dimensional regression coefficient vector associated with matrix $\mathbf{Z}$, $\Lambda = \left(\gamma_1, \ldots, \gamma_D\right)'$ incorporates measurement error due to auxiliary variables, $\mathbf{v} = \left(v_1, \ldots, v_D\right)'$ is the vector of domain random effects and $\mathbf{e} = \left(e_1, \ldots, e_D\right)'$ is the vector of sampling errors. The variance–covariance matrix of the ME model is,

$$\mathrm{var}(\mathbf{p}|\mathbf{X}) = \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_\gamma + \boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_e,$$

where $\boldsymbol{\Sigma}_\gamma = \mathrm{diag}\left\{(\boldsymbol{\beta}'\boldsymbol{\Psi}_i\boldsymbol{\beta}); 1 \le i \le D\right\}$, $\boldsymbol{\Sigma}_v = \mathrm{diag}\left\{\sigma_v^2; 1 \le i \le D\right\}$ and $\boldsymbol{\Sigma}_e = \mathrm{diag}\left\{\sigma_{ei}^2; 1 \le i \le D\right\}$.

Hierarchical Bayes Inference

Let vector of population parameters in D small domain is $\mathbf{P} = \left(P_1, \ldots, P_D\right)'$. Hereby we express ME model for binary response variable in HB framework for single auxiliary variable measured with error,

$$\begin{aligned}
&\mathbf{p}|\,\mathbf{P} \sim N(\mathbf{P}, \boldsymbol{\Sigma}_e);\\
&\mathrm{logit}(\mathbf{P})|\,\mathbf{X}, \beta, \boldsymbol{\theta}, \sigma_v^2 \sim N(\mathbf{X}\beta + \mathbf{Z\theta}, \boldsymbol{\Sigma}_v) \quad \text{and}\\
&\mathbf{W}|\,\mathbf{X}, \Psi_i \sim N(\mathbf{X}, \boldsymbol{\Psi}).
\end{aligned}$$

For univariate auxiliary variable measured with error, $\boldsymbol{\Psi} = \mathrm{diag}\left\{\Psi_i; 1 \le i \le D\right\}$. This expression of HB ME model given above has also the flexibility of extending for multivariate case.

In HB approach, the inferences about the small area parameter of interest are drawn from the posterior distribution. Particularly, posterior mean is taken as the point estimate of the parameter and posterior variance as a measure of the uncertainty associated with the estimate. The posterior density of the described HBME is,

$$\begin{aligned}
f(\mathbf{P}, \mathbf{X}, \beta, \boldsymbol{\theta}, \boldsymbol{\Sigma}_v \mid \mathbf{p}, \mathbf{W}) \propto\ & |\boldsymbol{\Sigma}_v|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\{(\mathbf{p}-\mathbf{P})'\boldsymbol{\Sigma}_e^{-1}(\mathbf{p}-\mathbf{P})\right.\\
&+ (\mathrm{logit}(\mathbf{P}) - \mathbf{X}\beta - \mathbf{Z\theta})'\left(\boldsymbol{\Sigma}_v\right)^{-1}(\mathrm{logit}(\mathbf{P}) - \mathbf{X}\beta - \mathbf{Z\theta})\\
&\left.+ (\mathbf{W}-\mathbf{X})'\boldsymbol{\Psi}^{-1}(\mathbf{W}-\mathbf{X})\}\right]\left|\frac{\partial \mathrm{logit}(\mathbf{P})}{\partial \mathbf{P}}\right|.
\end{aligned}$$

This posterior density cannot be obtained in a closed form. Hence, for implementing HB procedure, Monte Carlo Markov chain (MCMC) technique is used, which overcomes the computational difficulties of high-dimensional integrations of posterior densities to a greater extent. In particular, Gibbs sampling method is implemented to draw random samples from posterior densities. The full conditional distributions of the described HBME model under Gibbs sampler are given as follows.

$$\mathbf{P} \mid \beta, \boldsymbol{\theta}, \boldsymbol{\Sigma}_v,\ \mathbf{p}, \mathbf{X}, \mathbf{W} \propto\ |\boldsymbol{\Sigma}_v|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\{(\mathbf{p}-\mathbf{P})'\boldsymbol{\Sigma}_e^{-1}(\mathbf{p}-\mathbf{P}) + (\text{logit}(\mathbf{P}) - \mathbf{X}\beta - \mathbf{Z}\boldsymbol{\theta})'\left(\boldsymbol{\Sigma}_v\right)^{-1}(\text{logit}(\mathbf{P}) - \mathbf{X}\beta - \mathbf{Z}\boldsymbol{\theta}) + (\mathbf{W}-\mathbf{X})'\boldsymbol{\Psi}^{-1}(\mathbf{W}-\mathbf{X})\}\right]\left|\frac{\partial \text{logit}(\mathbf{P})}{\partial \mathbf{P}}\right|,$$

$$\beta \mid \mathbf{P}, \boldsymbol{\theta}, \sigma_v^2, \mathbf{p}, \mathbf{X}, \mathbf{W} \sim \text{N}\left[\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'(\text{logit}(\mathbf{P}) - \mathbf{Z}\boldsymbol{\theta}), \sigma_v^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}\right],$$

$$\boldsymbol{\theta} \mid \mathbf{P}, \beta, \sigma_v^2, \mathbf{p}, \mathbf{X}, \mathbf{W} \sim \text{N}\left[\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'(\text{logit}(\mathbf{P}) - \mathbf{X}\beta), \sigma_v^2\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\right],$$

$$\mathbf{X} \mid \mathbf{P}, \beta, \boldsymbol{\theta}, \sigma_v^2, \mathbf{p}, \mathbf{X}, \mathbf{W} \sim N(\mathbf{W} + \boldsymbol{\Sigma}^{-1}(\mathbf{p} - \mathbf{W}\beta - \mathbf{Z}\boldsymbol{\theta})\boldsymbol{\Psi}\beta,\ \boldsymbol{\Psi} - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi}\beta\beta'\boldsymbol{\Psi})).$$

$$\sigma_v^2 \mid \mathbf{P}, \beta, \boldsymbol{\theta},\ \mathbf{p}, \mathbf{X}, \mathbf{W} \sim \text{IG}\left[a_1 + \frac{D}{2},\ b_1 + \frac{(\text{logit}(\mathbf{P}) - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\theta})'(\text{logit}(\mathbf{P}) - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\theta})}{2}\right],$$

For the hyperparameters β and $\boldsymbol{\theta}$ the prior was $N(0, \sigma_0^2)$, where σ_0^2 is set to be very large. The prior choice σ_v^2 was $IG(a_0, b_0)$, (IG stands for Inverse Gamma) where very small value for a_0 and b_0 (usually $a_0 = b_0 \to 0$) is fixed to reflect lack of prior knowledge about variance parameters (Rao and Molina 2015; Anjoy and Chandra 2021). The application of described HBME model for PLFS data of NSO is delineated in the next sections. All the computations of small area models have been carried out using R and JAGS (Just Another Gibbs Sampler) software.

Data Description

In India, the PLFS has been structured majorly to meet two broad objectives. First, to measure the dynamics of labour force participation and employment status (level and change estimates) at quarterly interval for urban areas based on current weekly status (CWS) approach. Secondly, to bring out the level estimates of all important parameters annually in both rural and urban areas based on usual status (principal status + subsidiary status) and CWS approach. The usual status (ps + ss) and CWS are the activity status classification based on reference period of 1 year and 1 week respectively (Annual Report of PLFS 2019). Based on the design of PLFS, it is difficult to obtain reliable and acceptable decentralized level estimates on important labour force indicators to facilitate micro level understanding of the labour force characteristics because of the sample sizes constraint at this level. SAE techniques tend to improve the effective sample sizes for particular domain resulting in satisfactory and reliable estimates for small areas.

In this article PLFS 2017–18 and PLFS 2018–19 data have been analyzed for investigating the commuting patterns of workforce in rural and urban India. In PLFS, residence and workplace location can be tallied to get the numbers of commuters. Primarily, location of workplace is sought for industry groups 014, 016, 017 and divisions 02-99 excluding the workers engaged in farming. These industry group numbers are basically National Industrial Classification (NIC) 2008 codes corresponding to Primary, Secondary and Tertiary sectors (National Industrial Classification 2008). According to 2017–18 data, the estimated number of rural workers in urban areas were 15.8 million and estimated number of urban workers in rural areas were 2.9 million; whereas 10.3 and 9.3 million workers respectively in rural and urban places were not having any fixed place of work. In 2018–19, 18.8 million individuals living in rural areas were working in urban areas, for 2.4 million urban workers the place of work was rural while 9.8 and 7.9 million rural and urban workers respectively had no fixed place of work. In the analysis tables workplace encompasses both principal as well as subsidiary location component, although workers commuting in subsidiary capacity are considerably lower in number than the other. The gender classified as well as person wise aggregated summary in Table 1 infers that, from 2017–18 to 2018–19 the number of workers commuting from rural to urban India has increased by 3 million, on the other side number of urban workers working in rural areas has decreased by a margin (0.5 million). Combined urban workforce has increased by 5 million including urban workers as well as commuters from rural. This is also to note that, numbers of workers with no fixed workplace has decreased considerably over the year. In both rural and urban areas, proportion of male commuters (around 80% or more in both the years) are much more than female commuters.

On scrutinizing the activity status of major commuters from rural to urban as well as urban to rural workplaces, it has been found that rural residents have

Table 1 Aggregated level commuting pattern in rural and urban India from PLFS 2017–18 and 2018–19 data (in millions)

PLFS year	Residence	Workplace	Male	Female	Persons
2017–18	Rural	Rural	72.4	17.8	90.2
		Urban	14.4	1.4	15.8
		Not fixed	9.9	0.4	10.3
	Urban	Rural	2.2	0.7	2.9
		Urban	70.3	18.0	88.3
		Not fixed	8.7	0.6	9.3
Total (India)			177.9	38.9	**216.8**
2018–19	Rural	Rural	72.0	20.8	92.8
		Urban	17.1	1.7	18.8
		Not fixed	9.4	0.4	9.8
	Urban	Rural	1.9	0.5	2.4
		Urban	75.6	19.6	95.2
		Not fixed	7.4	0.5	7.9
Total (India)			183.4	43.5	**226.9**

Total number of commuters at All India level is marked as bold

Table 2 Workers commuting pattern by activity status from PLFS 2017–18 and 2018–19 data (in millions)

Activity status	Rural			Urban			All India
	Rural	Urban	Not fixed	Rural	Urban	Not fixed	Total
PLFS 2017–18							
Self employed	36.8	2.8	3.8	1.3	29.9	5.5	80.1
Regular salaried/wage	23.4	8.7	1.4	1.1	47.6	1.2	83.4
Casual wage labour	30.0	4.3	5.1	0.5	10.8	2.6	53.3
Total (India)	90.2	15.8	10.3	2.9	88.3	9.3	**216.8**
PLFS 2018–19							
Self employed	39.7	3.7	3.6	1.1	32.1	4.6	84.8
Regular salaried/wage	23.2	10.2	1.3	0.9	52.0	0.9	88.6
Casual wage labour	29.9	4.9	4.9	0.4	11.0	2.4	53.5
Total (India)	92.8	18.8	9.8	2.4	95.2	7.9	**226.9**

Total number of commuters at All India level is marked as bold

Table 3 State summary of commuting pattern in India from PLFS 2017–18 and 2018–19 data (in thousands)

Summary statistics	PLFS 2017–18		PLFS 2018–19	
	Rural to urban	Urban to rural	Rural to urban	Urban to rural
Min	0.6	0.2	1.2	0.3
Q1	15.7	5.5	20.5	10.1
Mean	440.0	82.8	523.1	66.5
Median	179.7	33.2	288.7	28.5
Q3	750.5	106.3	875.8	85.2
Max	2275.4	528.8	2835.9	356.6
Total	15,839.8	2898.3	18,830.8	2395.2

majorly moved to urban places for salaried or wage paid job followed by casual works in public or other departments. According to Table 2, around 55% rural commuters to urban places were due to salaried work followed by around 27% workers commuted for casual work. Whereas, for urban workers this commuting pattern to rural areas is not so prominently differentiated in three major activities.

Further Table 3 is the summary representing the state-wise mobility pattern of workers. The median figure of rural commuters were 1.79 lakhs and 2.88 lakhs in the consecutive years. Whereas, average number of rural commuters were much more in the consecutive years (around 4 and 5 lakhs respectively). Figure 1 is the pictorial of the share of commuters across the states which is clearly revealing the state-wise disparity in workers' mobility. The states contributing largest number of commuters from rural to urban workplaces are Uttar Pradesh (UP) followed by Tamil Nadu (TN), West Bengal (WB) and Bihar (BR). From urban to rural workplaces, the states depicting largest commuters are WB followed by UP, Maharashtra (MH) and TN.

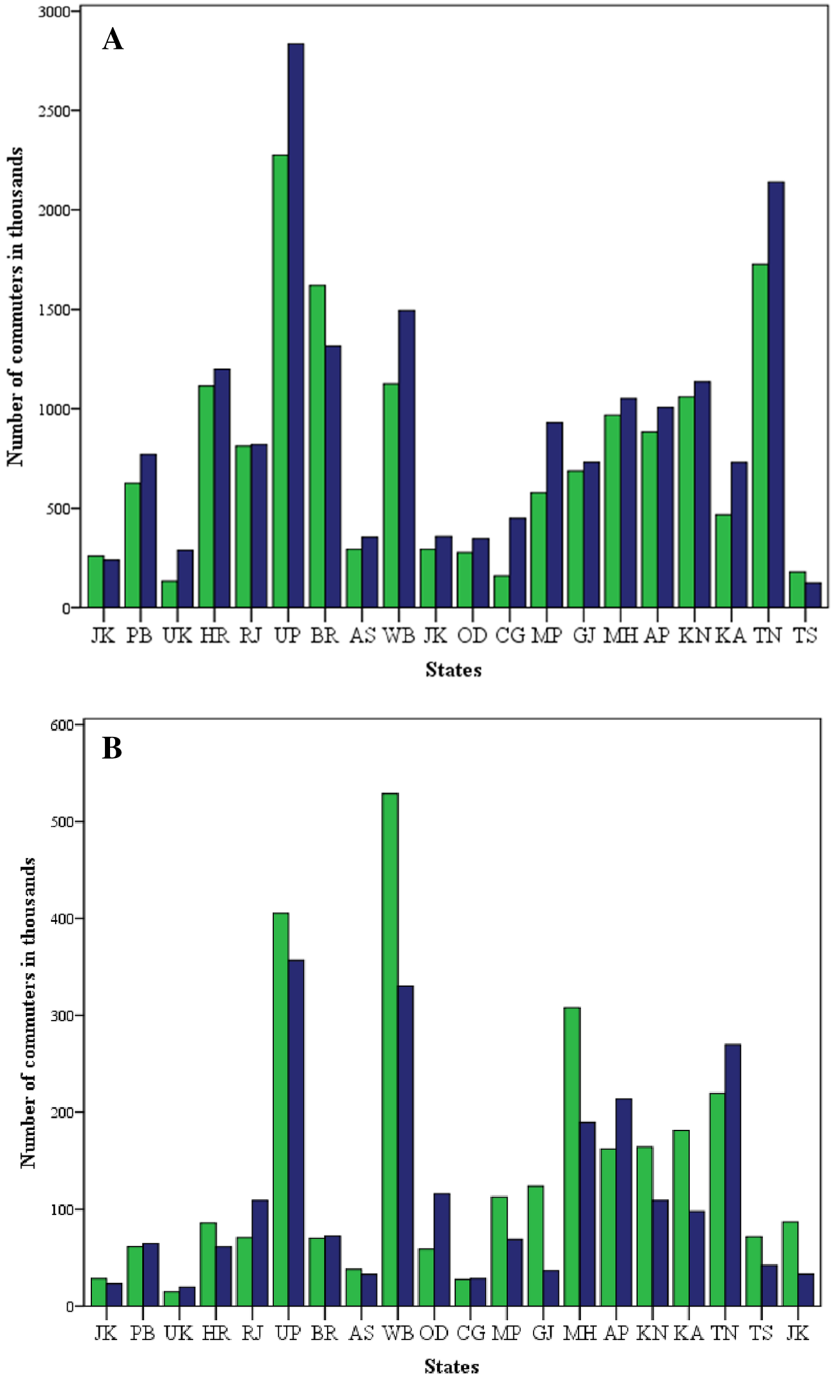


Fig. 1 **a** Distribution of rural to urban commuters across States of India (representations: green rectangle PLFS 2017–18 and blue rectangle PLFS 2018–19). **b** Distribution of urban to rural commuters across States of India (representations: green rectangle PLFS 2017–18 and blue rectangle PLFS 2018–19)

From the aggregated level picture of the number of commuters (see Fig. 1) very less idea can be obtained about the spatial concentration of commuters within states at local or regional levels. Hence, for the disaggregated level analysis of commuters behaviour Uttar Pradesh state has been chosen which was sharing largest number of rural commuters. Table 4 portray the distribution of district specific sample sizes and sampling fraction which is clearly indicating the fact of underrepresentation of samples in unplanned areas or domain. Overall, in Uttar Pradesh (rural) median sample sizes for workforce in districts are 99 only. So, we may attempt to implement SAE approach for improving estimates at small domain or district level where direct estimates cannot be acceptable. In particular, the aim is to produce proportion of rural to urban commuters (P_i) out of total workforce across the districts(rural) in Uttar Pradesh. The direct estimate of P_i is $p_{iw} = \left(\sum_{j=1}^{n_i} w_{ij}\right)^{-1} \sum_{j=1}^{n_i} w_{ij}y_{ij}$, whereas the variable y_{ij} is binary with regard to the commuting behaviour. Let p_{ij} be the selection probability attached to jth sampling unit y_{ij} in the area i. The basic design weight will be $w_{ij} = (n_i p_{ij})^{-1}$. These weights can be adjusted to account for non-response and/or auxiliary information (Hidiroglou and You 2016). The variance of the estimator p_{iw} is given by, $\sigma^2_{ei.sw} = \left(\sum_{j=1}^{N_i} w_{ij}\right)^{-2} \left\{ \sum_{j=1}^{N_i} w_{ij}(w_{ij} - 1)(y_{ij} - P_i)^2 \right\}$. The survey weighted estimator p_{iw} and its variance induce the sampling deign into HB small area modeling structure (Anjoy and Chandra 2021). Small area modeling has been done using PLFS 2018–19 data with potential auxiliary variable from PLFS 2017–18 as well as India's Population Census 2011. The small area level proportion estimates can necessarily be converted into numbers by multiplying with district wise workforce N_i or $\hat{N}_i = \sum_{j=1}^{n_i} w_{ij}$ in usual status.

Table 4 Summary of sample sizes (n) and sampling fraction (f = n/N) across regions in Uttar Pradesh from PLFS 2018–19 data

Region	Features	Min	Median	Average	Max	Total
Northern Upper Ganga Plains	n	22	109	104	211	1036
	f	0.0001	0.0002	0.0002	0.0003	0.0021
Central	n	63	137	146	297	1314
	f	0.0002	0.0002	0.0002	0.0003	0.0019
Eastern	n	48	75	106	292	2972
	f	0.0002	0.0002	0.0002	0.0003	0.0061
Southern	n	52	74	100	188	699
	f	0.0002	0.0002	0.0002	0.0003	0.0016
Southern Upper Ganga Plains	n	28	120	127	218	2156
	f	0.0002	0.0002	0.0002	0.0003	0.0037
All (Uttar Pradesh)	n	22	99	115	297	8177
	f	0.0001	0.0002	0.0002	0.0003	0.0154

Analysis and Interpretation

The auxiliary variables required for implementing HBME small area model have been obtained from PLFS 2017–18 and Population Census 2011. Since our interest is to obtain proportion of commuters commuting from rural to urban, we have taken rural unemployment rate and rural workforce participation rate as auxiliary variables. District wise proportions of rural unemployment (x) for Uttar Pradesh has been considered for the PLFS year 2017–18. The x for ith district is given as $x_i = N_i^{-1} \sum_{j=1}^{N_i} x_{ij}$, where the auxiliary variable x_{ij} is binary taking value 1 if a person in the labour force (by usual status) is unemployed and 0 otherwise. The estimate of x_i is given by $\hat{x}_i = \left(\sum_{j=1}^{n_i} w_{ij}\right)^{-1} \sum_{j=1}^{n_i} w_{ij}x_{ij}$, where w_{ij} is the survey-weight. Figure 2 shows spatial map for the district-wise proportions of rural unemployment in Uttar Pradesh from PLFS 2017–18. However, the selection of this auxiliary variable has been preceded by correlation studies with pools of auxiliary variables with the study variable as well as

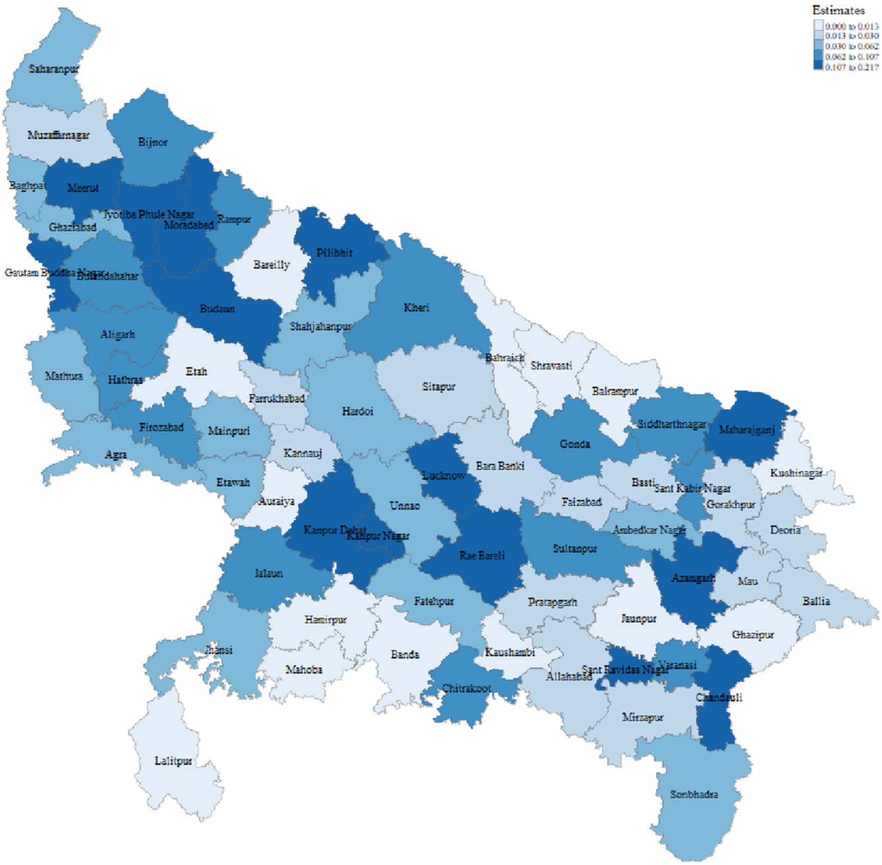


Fig. 2 Spatial map showing district-wise proportions of rural unemployment in Uttar Pradesh from PLFS 2017–18

Table 5 Comparison of direct vs. model based SAE estimates of the proportion of rural to urban commuters in workforce for Uttar Pradesh in 2018–19

Values	Summary of 63 districts		Summary of 71 districts	
	Direct	SAE	Direct	SAE
Min	0.002	0.004	0.000	0.005
Q1	0.025	0.023	0.010	0.025
Median	0.061	0.050	0.049	0.045
Average	0.089	0.081	0.079	0.077
Q3	0.117	0.107	0.116	0.099
Max	0.494	0.464	0.494	0.464

various literatures also have indicated the strong relationship between mobility vs. unemployment. The Pearson correlation coefficient between the study and this auxiliary variable was 0.30. Another auxiliary variable is rural workforce participation rate (z) from the population census. Nayka and Sridhar (2019) have suggested higher workforce participation rates generally being characterised by large number of commuters. Workforce participation rate according to census definition is proportion of total (main + marginal) workers in the population (workers + non-workers). Figure 7 in Appendix shows spatial map for the district-wise rural Workforce participation rate. Further, Figs. 8 and 9 in the Appendix showing bivariate choropleth map for the target variable with survey and census auxiliary variables respectively. Top right colour of the palette in both the figures indicates higher mobility combined with higher unemployment rate and higher workforce participation rate respectively. In Fig. 8, the districts marked in top right colour of the palette are having higher mobility and unemployment rate. Some of these districts are adjoining to big cities, for example GTB Nagar (rural), Kanpur (rural), Lucknow (rural).

We have implemented ME SAE model to obtain precise estimates for the parameter of interest at district level. For computing HB estimates of commuters proportions (P_i), we have considered prior for σ_v^2 as $IG(0.1, 0.1)$ and distribution of β and $\boldsymbol{\theta}$ has been taken to be $N(0, 10^6)$. The value of potential scale reduction factor $\hat{R}$ for each of the district was found to be close to 1, which implies the convergence success of MCMC sampler in HB method implementation. Table 5 is the summary of estimates obtained using traditional direct estimation technique vs. small area model based approach. Summary of 63 (excluding 8 districts with direct proportion value 0) and total 71 districts have been presented differently. In Fig. 3 we plot the direct and small area model based estimates across 63 districts with non-zero proportions. This plot indicates that both the estimates have coincided for majority of the districts. Wald Goodness of fit statistic is also computed to test whether there is any statistical difference between the direct estimates and model based predictor (Chandra 2018). The null hypothesis is direct and model based estimates are statistically equivalent. The Wald statistic computed based on 71 districts was 4.94 which is smaller than the critical value (91.67) from a chi square distribution with 71 degrees of freedom at 5% level of significance. This indicates that model based small area estimates are consistent with the direct estimates. The coefficient of variation % (CV%) is the criteria which

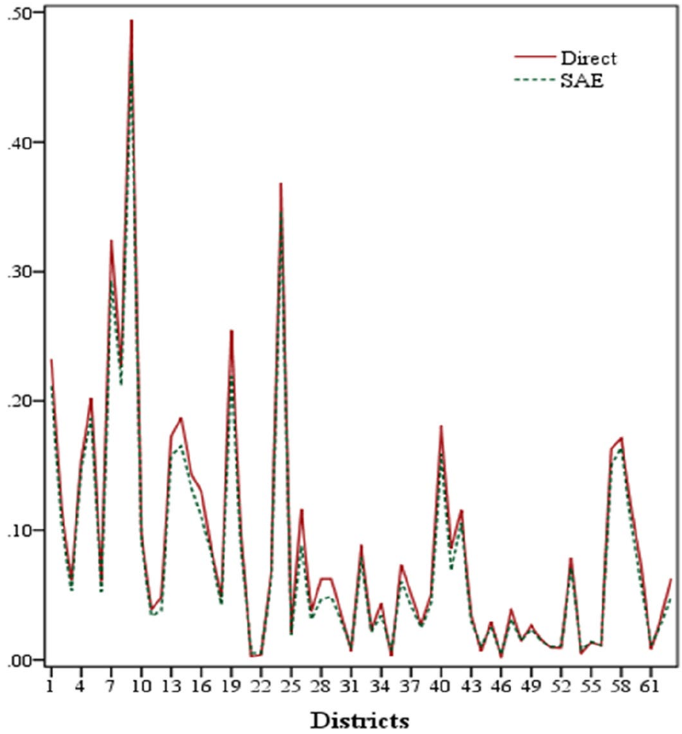


Fig. 3 Plot of direct vs. model based SAE estimates of commuters proportions in workforce

indicates the reliability of particular estimates. For districts with small sample sizes, the CV% of direct estimates may not be reliable. SAE method has also been promising for areas with very high CV% of direct estimation (see Fig. 4). For direct estimates, the maximum CV% has gone up to 233, this is the reason why we cannot rely on direct estimates of such districts with too high CV%. There were 7 such districts with CV% more than 100, whereas maximum CV in SAE estimates was 67%.

For model diagnostic of the residuals, the residuals have been plotted across districts along with histogram for normality check. As one can see from Fig. 5, residuals are symmetrically distributed along the 0-line. This infers independence assumption of the residuals have been satisfied. Histogram is suggesting that model residuals are following normal distribution. The Figs. 3 and 5 entails the fact that fitted small area model has fulfilled both the bias diagnostic and model diagnostic results.

Finally, Fig. 6 is the spatial map of rural to urban commuters' proportions across districts in Uttar Pradesh. The dark shaded regions share higher number of commuters in the workforce than the light shaded regions. In western part, districts like Gautam Budhha Nagar, Saharanpur, Baghpat, Meerut, Ghaziabad are having reasonably higher number of commuting workers, because these are near to big cities like Noida, Delhi. Again from Kanpur and Lucknow rural peoples are commuting in large numbers to the main city. In the south, Sonbhadra district is near to

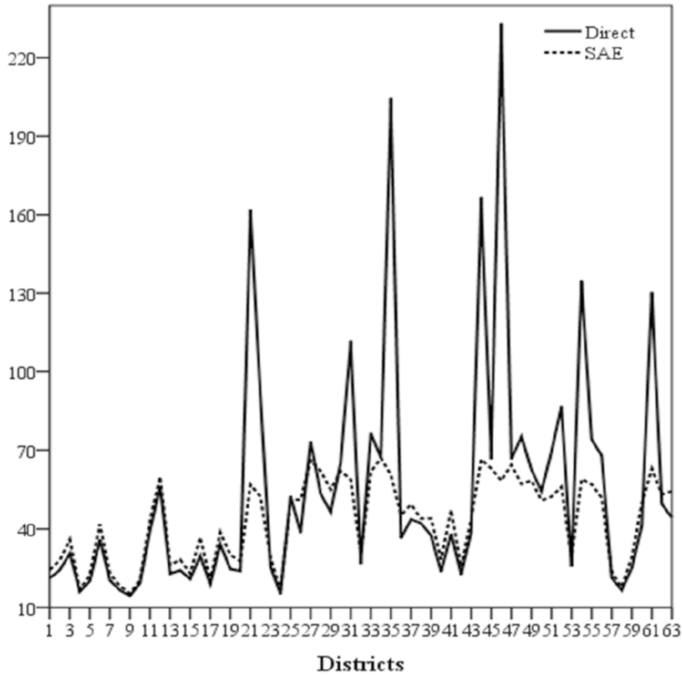
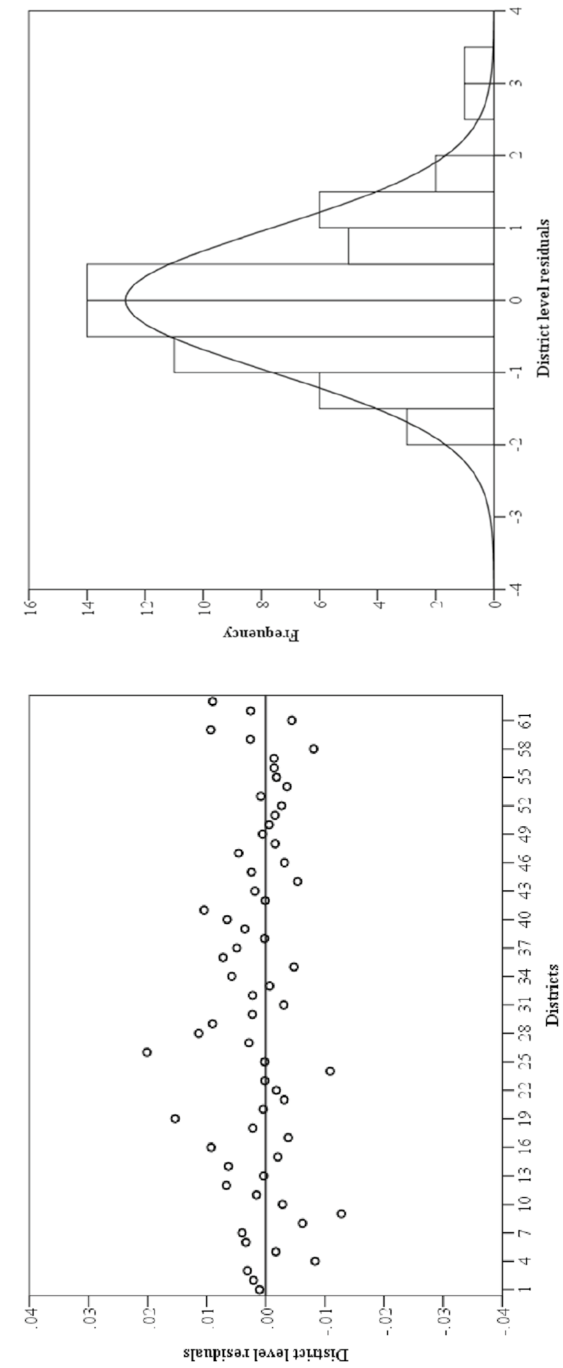


Fig. 4 Plot comparing the CV% for direct and model based SAE estimates across districts in Uttar Pradesh

Madhya Pradesh border, hence rural people tend to commute to cities of adjacent state. Development of large cities, urbanization and improved transport facilities with reduced cost are some of the reasons attributed to higher number of commuters out-commuting from rural places in these districts. Such spatial mapping of commuting pattern may be useful to the policy makers to plan urbanization accordingly to attract workforce or in contrast, generate employment in rural areas itself to retain workers at rural places. This attempt of policy making is deliberated by the administration and their strategy to shape a state may be guided by the disaggregated level figures eying stable mobility behaviour of workers.

Further, there are certain issues which warrants additional investigations. For example, majority of the rural workers are commuting for regular salaried job (ref. Table 2) which demands relatively higher educational as well as skill status by the employed workers. This also may imply that in rural settlements, their job is being either underpaid or temporary, hence they commute to urban places having diverse opportunities to secure their job as stable or permanent. Again, across districts of Uttar Pradesh spatial mapping of the certain technical skills of the workers may be sketched to understand any possible association with the definite mobility pattern. Mode of transport of the workers is another important factor which governs distance of commuting and currently the State administration is emphasizing a lot on public transportation facilities. Data on mode of transport is available from Census database which may be utilized to correlate the mobility behaviour.

Fig. 5 Distribution of the district level residuals (left) and histogram of the residuals (right) for the model based small area estimates



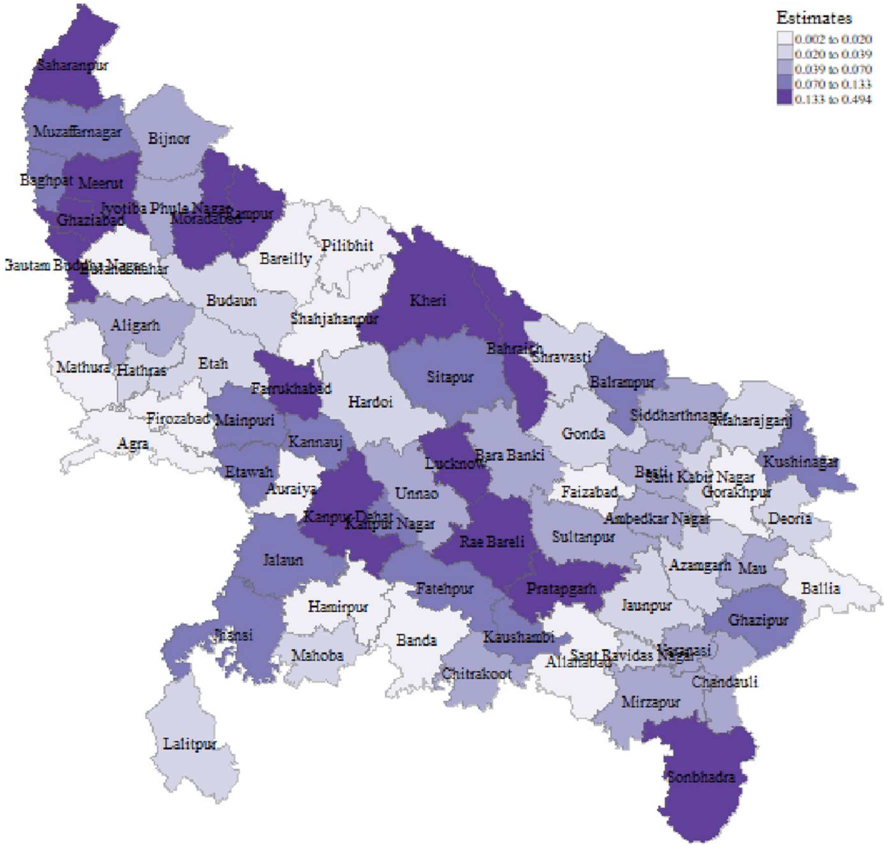


Fig. 6 Spatial map of rural to urban workers mobility pattern across districts in Uttar Pradesh generated using SAE approach

Concluding Remarks

The article on theoretical front focusses on using HBME SAE approach for binary data with an innovative application to generate the number of commuters at small area level. Disaggregated level pattern of commuting or mobility data of workforce is also indicative of the real situation of overall employment pattern. What we actually see at the state or the national level in context of commuting patterns of workers is quite different than this distribution at disaggregated level. State level or central level allocation of fund on need basis can be guided by these micro level statistics. The purpose of this work is well motivated from the background of microscopic scrutinization of commuting pattern in rural and urban India. Mobility is an issue caused by several factors. The reasons can be traced differently in rural areas as contrast to urban areas. But the main trend of this is driven by unemployment or under-employment. Again number of rural migrators are much more than urban migrators, because the employment opportunities are created due to urbanization. In this article, based on PLFS 2017–18 and 2018–19 data the commuting pattern across the states, sectors, genders and work status has been analysed. Then, the mobility behaviour has been examined at disaggregated level using HBME SAE model. The performance of HBME SAE method has been found to be promising in the social science application presented in this article. Implementation of direct estimation method to the unplanned regions having negligible or zero sample sizes may not provide the proficient estimates. Again in absence of good auxiliaries from census or administrative records, use of ME small area model is crucial linking two surveys. Small area estimation method is officially used in many countries to produce several official estimates and even more spread of such approach is need of the day with emerging necessities for micro level data.

Appendix

See Figs. 7, 8 and 9.

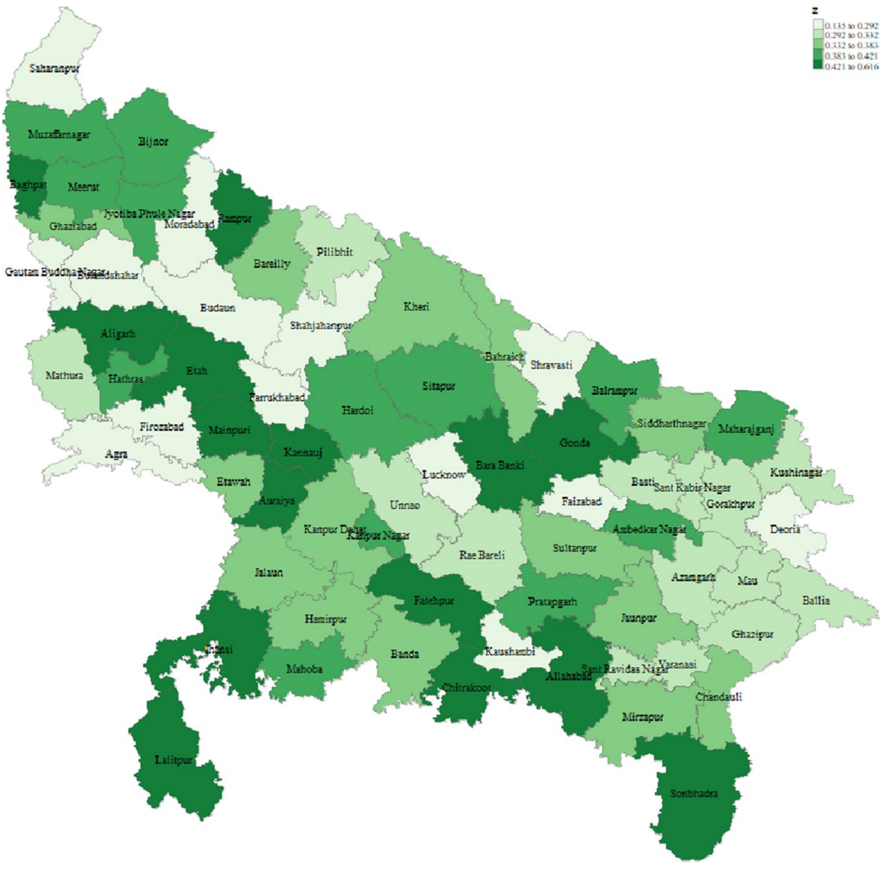


Fig. 7 Spatial map showing district-wise rural workforce participation rate in Uttar Pradesh from population census 2011

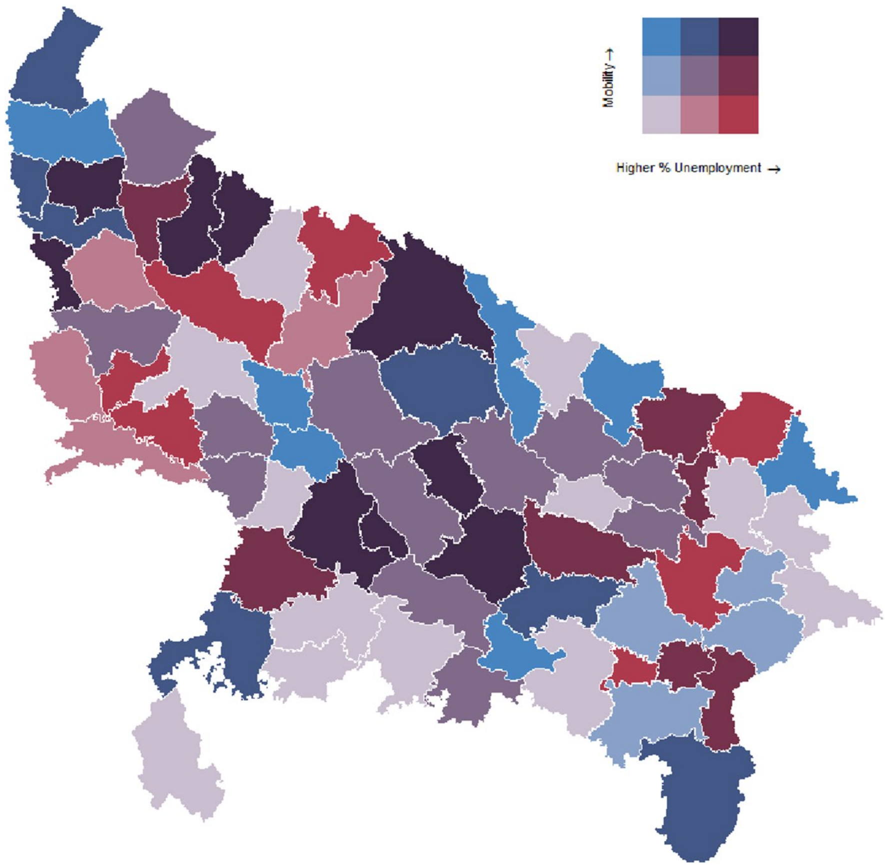


Fig. 8 Bivariate Choropleth map showing unemployment and mobility pattern in Uttar Pradesh

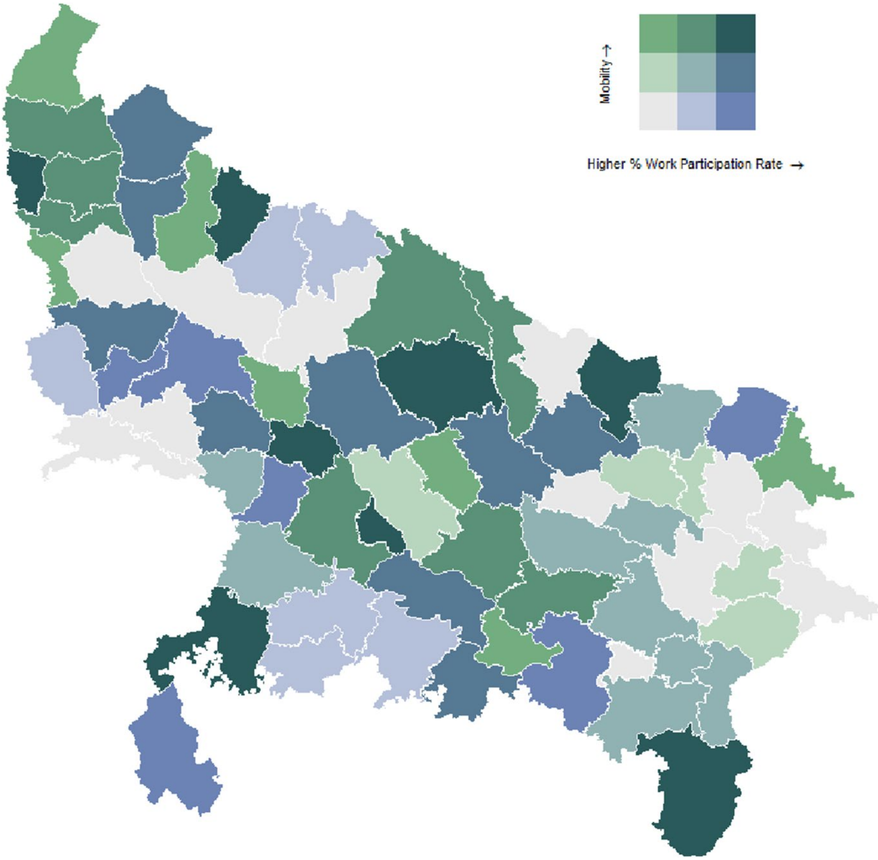


Fig. 9 Bivariate Choropleth map showing workforce participation rate and mobility pattern in Uttar Pradesh

Acknowledgements The author would like to acknowledge Ministry of Statistics and Programme Implementation, Govt. of India for the database available in its official website (http://www.mospi.gov.in), the required PLFS data for the analysis were obtained from here. The views expressed in the manuscript are personal (author's) and not that of the Govt. of India.

Author Contributions The first and single author has contributed to the concept of paper, analysis and writing of the manuscript.

Funding No funding agency was involved in this work.

Data availability The data on Periodic Labour Force Survey (PLFS) used in this article has been obtained from official website of Ministry of Statistics and Programme Implementation (http://www.mospi.gov.in).

Declarations

Conflict of interest The author declares that there is no potential conflict of interest.

References

Adamson, F.B. 2006. Crossing borders: International migration and national security. *International Security* 31 (1): 165–199.

Anjoy, P., and H. Chandra. 2020. District and social group-wise estimation and spatial mapping of food insecurity in the state of Odisha in India. *Journal of the Indian Society of Agricultural Statistics* 74 (2): 107–120.

Anjoy, P., and H. Chandra. 2021. Hierarchical Bayes estimation of small area means under a spatial nonstationary Fay–Herriot model. *Communications in Statistics Simulation and Computation*. https://doi.org/10.1080/03610918.2021.1926501.

Anjoy, P., H. Chandra, and P. Basak. 2019. 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.

Annual Report of Periodic Labour Force Survey July 2017-June 18 (2019) National Statistical Office, Ministry of Statistics and Programme Implementation. Government of India

Arima, S., G.S. Datta, and B. Liseo. 2012. Objective Bayesian analysis of a measurement error small area model. *Bayesian Analysis* 7: 363–384.

Arima, S., G.S. Datta, and B. Liseo. 2015. Bayesian estimators for small area models when auxiliary information is measured with error. *Scandinavian Journal of Statistics* 42: 518–529.

Bhatt, V., S. Chandrasekhar, and A. Sharma. 2020. Regional patterns and determinants of commuting between rural and urban India. *The Indian Journal of Labour Economics* 63: 1041–1063.

Brown, D., T. Champion, M. Coombes, and C. Wymer. 2015. The migration-commuting nexus in rural England: A longitudinal analysis. *Journal of Rural Studies* 41: 118–128.

Burgard, J.P., M.D. Esteban, D. Morales, and A. Pérez. 2019. A Fay–Herriot model when auxiliary variables are measured with error. *TEST*. https://doi.org/10.1007/s11749-019-00649-3.

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. 2018. 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: 29–44.

Chandra, H., and G. Chandra. 2015. *An overview of small area estimation techniques. Statistics in forestry: Methods and applications*. Coimbatoor: Bonfring.

Chandra, H., and G. Chandra. 2020. Small area estimation for total basal cover in The State of Maharashtra In India. In *Statistical methods and applications in forestry and environmental sciences*, ed. V. Ddd. Singapore: Springer.

Chandra, H., N. Salvati, and R. Chambers. 2017. Small area prediction of counts under a non-stationary spatial model. *Spatial Statistics* 20: 30–56.

Chandra, H., N. Salvati, and R. Chambers. 2018. Small area estimation under a spatially non-linear model. *Computational Statistics and Data Analysis* 126: 19–38.

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

Chandrasekhar, S., M. Naik, and S.N. Roy. 2017. On the importance of triangulating data sets to examine Indians on the move. *Economic and Political Weekly* 52: 60–68.

Datta, G.S., J.N.K. Rao, and M. Torabi. 2010. Pseudo-empirical Bayes estimation of small area means under a nested error linear regression model with functional measurement errors. *Journal of Statistical Planning and Inference* 140: 2952–2962.

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.

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 1: 515–533.

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.

Ghosh, M., and K. Sinha. 2007. Empirical Bayes estimation in finite population sampling under functional measurement error models. *Journal of Statistical Planning and Inference* 137: 2759–2773.

Ghosh, M., K. Sinha, and D. Kim. 2006. Empirical and hierarchical Bayesian estimation in finite population sampling under structural measurement error model. *Scandinavian Journal of Statistics* 33: 591–568.

Green, A. 2004. Is relocation redundant? Observations on the changing nature and impacts of employment-related geographical mobility in the UK. *Regional Studies* 38 (6): 629–641.

Guha, S., and H. Chandra. 2022. Multivariate small area modelling for measuring micro level earning inequality: Evidence from periodic labour force survey of India. *Social Indicators Research*. https://doi.org/10.1007/s11205-021-02857-7.

Hidiroglou, M.A., and Y. You. 2016. Comparison of unit level and area level small area estimators. *Survey Methodology* 42: 41–61.

Jiang, J., and P. Lahiri. 2006. Mixed model prediction and small area estimation. *Test* 15: 111–999.

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.

Liu, B., P. Lahiri, and G. Kalton. 2014. Hierarchical Bayes modeling of survey-weighted small area proportions. *Survey Methodology* 40: 1–13.

McAuliffe, M., and B. Khadria. 2020. *World migration report 2020*. Geneva: International Organization for Migration.

National Industrial Classification [All Economic Activities] (2008) Central Statistical Organization, Ministry of Statistics and Programme Implementation. Government of India

Nayka, S., and K.S. Sridhar. 2019. Urban commuters in Indian states and cities: Modes of transport and distances. *Urbanisation* 3 (2): 1–39.

Rao, J.N.K. 2003. *Small area estimation*. New York: Wiley.

Rao, J.N.K., and I. Molina. 2015. *Small area estimation*, 2nd ed. New York: Wiley.

Souza, D.F., F.A.S. Moura, and H.S. Migon. 2009. Small area population prediction via hierarchical models. *Survey Methodology* 35: 203–214.

Torabi, M., G.S. Datta, and J.N.K. Rao. 2009. Empirical Bayes estimation of small area means under nested error linear regression model with measurement errors in the covariates. *Scandinavian Journal of Statistics* 36: 355–368.

United Nations (2019) World urbanization prospects: The 2018 revision. Department of Economic and Social Affairs, Population Division (ST/ESA/SER.A/420). New York

Ybarra, L.M.R., and S.L. Lohr. 2008. Small area estimation when auxiliary information is measured with error. *Biometrika* 95 (4): 919–931.

You, Y. 2008. An integrated modeling approach to unemployment rate estimation for sub-provincial areas of Canada. *Survey Methodology* 34: 19–27.

You, Y., and J.N.K. Rao. 2002. Small area estimation using unmatched sampling and linking models. *The Canadian Journal of Statistics* 30: 3–15.

Zhu, Z., Z. Li, Y. Liu, and H. Chen. 2017. The impact of urban characteristics and residents' income on commuting in China. *Transportation Research Part D Transport and Environment* 57: 474–483.

Authors and Affiliations

Priyanka Anjoy[1]

[1] National Accounts Division, Ministry of Statistics and Programme Implementation, Khurshid Lal Bhawan, Janpath, New Delhi, Delhi 110001, India