A Hierarchical Bayesian approach to small area estimation of health insurance coverage in Ethiopian administrative zones for better policies and programs

Yegnanew A. Shiferaw[1], Seyifemickael Amare Yilema[2]* and Yikeber Abebaw Moyehodie[2]

Abstract

Sample surveys are extensively used to provide reliable direct estimates for large areas or domains with enough sample sizes at national and regional levels. However, zones are unplanned domains by the Demographic and Health Survey (DHS) program and need more sample sizes to produce direct survey estimates with adequate precision. Conducting surveys in small areas (like zones) is too expensive and time-consuming, making it unfeasible for developing countries like Ethiopia. Therefore, this study aims to use the Hierarchical Bayes (HB) Small Area Estimation (SAE) model to estimate the Community-Based Health Insurance (CBHI) coverage at the zone levels in Ethiopia. To achieve this, we combined the 2019 Ethiopia Mini-Demographic and Health Survey (EMDHS) data with the 2007 population census data. SAE has addressed the challenge of producing reliable parameter estimates for small or even zero sample sizes across Ethiopian zones by utilizing auxiliary information from the population census. The results show that model-based estimates generated by the SAE approach are more accurate than direct survey estimates of CBHI. A map of CBHI scheme coverage was also used to visualize the spatial variation in the distribution of CBHI scheme coverage. From the CBHI scheme coverage map, we noticed notable variations in CBHI scheme coverage across Ethiopian zones. Additionally, this research identified areas with high and low CBHI scheme coverage to improve decision-making and increase coverage in Ethiopia. One of the novelties of this paper is estimating the non-sampled zones; therefore, the policymakers will give equal attention similar to the sampled zones.

Keywords Disaggregated level CBHI scheme, Hierarchical Bayes model, Small area estimation, Ethiopia

Introduction

Health insurance is vital to the realization of universal healthcare and the prevention of disease. Community-based health insurance (CBHI) is the most effective means to provide universal health coverage (UHC) with enough financial protection against healthcare expenses, promote equitable access to high-quality healthcare, raise financial security, and improve social cohesion and solidarity [1–3]. Out-of-pocket medical expenses, however, lead to lower healthcare service utilization and higher catastrophic medical expenses. Certain countries have implemented CBHI programs to reduce out-of-pocket expenditures and enhance access to healthcare services, particularly for people living in rural areas or working in the unorganized sector [4, 5].

Around 100 million people are forced into extreme poverty yearly due to out-of-pocket health spending, and

*Correspondence:
Seyifemickael Amare Yilema
samarey1981@gmail.com
[1] Department of Statistics, University of Johannesburg, PO Box 524, Johannesburg 2006, South Africa
[2] Department of Statistics, College of Natural and Computational Science, Debre Tabor University, PO Box 272, Debre Tabor, Ethiopia

Shiferaw *et al. Health Economics Review*   (2024) 14:29

Page 2 of 14

over 930 million individuals worldwide spend at least 10% of their household income on healthcare [6]. CBHI is becoming more popular in low- and middle-income countries to enhance healthcare utilization and shield households from financial catastrophes brought on by out-of-pocket medical charges [7]. A significant issue in Africa has been the catastrophic nature of healthcare financing arrangements for the poor, particularly the rural population [8]. Out-of-pocket charges can significantly hinder access to quality medical treatment in Sub-Saharan African nations[9]. Most sub-Saharan African nations, including Ethiopia, have adopted the CBHI program to increase the population's fair access to health services. However, enrollment rates and renewals for CBHI in sub-Saharan Africa are persistently low [10].

In 2019, it was estimated that 28% of the Ethiopian population nationwide had health insurance, primarily community-based health insurance [11, 12]. In other words, nearly 82 million Ethiopians, or about 72% of the country's total population, lacked health insurance enrollment. According to [13], a lack of health insurance is significantly associated with worse health outcomes and restricted access to medical care. People without insurance may experience health problems due to their lack of coverage. This could also affect the quality and organization of local healthcare systems [14].

The Sustainable Development Goal for Health (SDG3) aimed to ensure healthy lives and promote well-being for people of all ages [15]. SDG3 includes a sub-target (3.8) on achieving UHC, including financial risk protection and access to quality, essential health services. One of the SDGs for 2030 is UHC, which means everyone can access a full range of high-quality health services when and where they need them without financial hardship [16]. This goal still needs to be accomplished in six years. Although the UHC service coverage index (SDG indicator 3.8) has increased from 45 in 2000 to 67 in 2019, 30% of the world's population still lacks access to essential health services [17]. Ethiopia's health insurance coverage in 2019 was 28% [12], far below the SDG target of 80% by 2030. The inability to reach other SDG health-related targets results from this low coverage rate compared to other nations in the region. Additionally, 31% of all medical spending in 2016–17 was attributed to catastrophic out-of-pocket costs, which have remained high over time [18].

To achieve UHC in Ethiopia, it is essential to understand the impact of health insurance on healthcare utilization and financial protection. Studies have highlighted the potential factors related to maternal health-seeking behavior and the association between health insurance membership and antenatal care utilization in Ethiopia [19]. Additionally, the impact of CBHI on universal health coverage in Ethiopia has been evaluated, indicating positive impacts on health service issues by reducing catastrophic costs. However, the low universal health coverage for a limited population suggests a broader national scheme covering the entire population [20].

Aside from that, coverage varies according to administrative regions and other sociodemographic factors [12]. According to a recent report by [17], despite improvements in healthcare coverage nationally, intra-country disparities are obscured by aggregate data, making it challenging for UHC to address inequalities.

The studies above compiled national and regional health insurance coverage estimates through household surveys. These surveys are used to obtain accurate statistical information about a specific target population in broad regions, such as the nation, the state, the province, and so on [21, 22]. However, estimates obtained from local surveys are often unreliable due to inadequate sample sizes. In the case of Ethiopia, estimates for lower administrative levels, such as zones, are unavailable due to small sample sizes. To address this issue and produce more precise estimates, the small area estimation (SAE) model combines survey and census data.

Small areas refer to subpopulations or unplanned domains in large-scale sample surveys. In these cases, model-based methods, also known as the SAE approach, have consistently attracted attention for their ability to provide appropriate estimates for such a small area or domain. The 2030 Agenda for Sustainable Development Goals (SDGs) emphasizes the importance of decentralized statistics for micro-level planning, policy formation, and targeted social advancement. SAE is indispensable for reconciling the demand for reliable, representative, and disaggregated official statistics [21].

In Ethiopia, the demographic and health survey (DHS) program only includes subnational administrative levels like regions. In other words, the zones, Ethiopia's third administrative layer, were not the focus of the DHS program's design when providing accurate health-related information. For public health policymakers, research that only uses survey data at the lowest administrative levels, like zones, may not be reliable and practical since the estimates have large standard errors and coefficients of variation. As a result, additional sample data is needed to generate precise projections at the zone level [21]. We cannot increase the sample size due to financial and operational constraints.

There has been a growing demand for small area statistics in both the public and private sectors. Currently, SAE is experiencing high demand, with occasional nationwide spikes. In order to allocate resources and focus initiatives at lower-level administrations, SAE is a crucial technique for estimating health, education, and environmental

Shiferaw *et al. Health Economics Review*   (2024) 14:29

Page 3 of 14

indicators [23, 24]. The DHS program, the World Bank group, the International Monetary Fund (IMF), and the World Health Organization (WHO) were all very concerned about how to disaggregate DHS data and provide it to third-level government administrations [25–27]. The DHS host nations often request the DHS Program for subregional indicator estimates for development planning and policymaking [28]. In order to facilitate the mapping of health and economic inequality, the United Nations Population Fund (UNFPA) works to strengthen countries' capacities to produce and distribute high-quality, disaggregate data on population and development indicators, as well as demographic discrepancies, and to support humanitarian programming [29].

Estimating health-related parameters at unplanned geographic levels is crucial for successful health promotion and initiatives, improved healthcare delivery, and population-specific health planning and management. The lack of a readily available, adequate dataset for characteristics of interest in small areas has a detrimental effect on the ability of local and national organizations to manage significant health issues and associated risks in the community [12, 15, 18]. Developing a method for simulating reliable small-area statistics could solve this problem.

In the Ethiopian DHS, some zones have very few or no samples available. However, estimating the characteristics of interest, such as health insurance coverage, at the zone level for effective healthcare planning and resource allocation is crucial [30, 31]. To address this need, the study aims to estimate the level of health insurance coverage in Ethiopia at the zone level. The SAE approach's strength lies in using additional data sources, including census and administrative data [22, 32]. Therefore, this study used the HB approach to SAE, which combines data from the 2007 population census and the 2019 Ethiopian DHS.

Accurately estimating health insurance coverage at the zone level is vital in Ethiopia for several reasons. Firstly, it allows policymakers and healthcare providers to allocate resources more efficiently by providing precise information on the areas in need of healthcare services. Secondly, understanding the distribution of health insurance coverage ensures that healthcare services are accessible to all individuals, regardless of their location, thereby promoting equitable healthcare access. Thirdly, precise estimates of health insurance coverage at the zone level provide valuable insights for evaluating the effectiveness of existing policies and planning future interventions. Lastly, by estimating health insurance coverage at the zone level, healthcare providers can identify zones with high uninsured populations, which helps to design strategies to increase coverage, leading to a more efficient and equitable healthcare system. This knowledge can also help reveal geographic disparities and enable healthcare providers to address them more effectively [33, 34].

The Fay-Herriot model and its extensions have been used in various studies in the SAE framework. These include exploring spatiotemporal models using undernourished children in studies such as [34–36] and the Bayesian spatiotemporal approach to quantify SARS-CoV-2 [37]. Spatial SAE is also employed by [33, 38, 39]. In order to explore binary outcomes of response variables for SAE, a logit-normal model [40], logit-normal spatial model [41], HB logit-normal model [8, 42], HB logit-normal spatial model [43], zero-inflated Binomial [44], and zero-inflated negative Binomial model [45] are used under the Fay-Herriot model.

Although the SAE approach is popular in literature, it has yet to be widely practiced in Ethiopia. Therefore, this study aims to implement the HB SAE approach using the CBHI binary outcomes from the 2019 EMDHS and auxiliary variables from the 2007 population census datasets [46]. The main objective of this research is to estimate the community-based health insurance (CBHI) coverage at the zonal level in Ethiopia using the HB SAE approach. The study is expected to offer important insights into the spatial distribution of health insurance coverage, which is vital for making informed policy decisions and developing effective healthcare interventions at the zone level.

The paper is structured as follows: Section two explains the data sources, sampling design, and methods applied in the study, namely the Fay Herriot model under the different linking models and the HB SAE model. Section three presents the results and discussions, while Section four concludes.

Methods and Materials
Sampling design
The EMDHS sample for 2019 was stratified and chosen in two stages. At each lower administrative level, implicit stratification and proportional allocation were achieved by sorting the sampling frame inside each sampling stratum before sample selection, according to administrative units at various levels, and by employing a probability proportional to size selection at the first step of sampling. In the first step, 305 enumeration areas (EAs) were chosen proportional to EA size and with independent selection in each sampling stratum.

Each large EA chosen for the 2019 EMDHS was split to reduce the process of household listing. The survey was limited to one segment, with the likelihood proportional to segment size. Only households in the specified segment were listed; a 2019 EMDHS cluster is either an EA or a section of an EA. The second selection round involved selecting 30 households per cluster using an

Shiferaw *et al. Health Economics Review*   (2024) 14:29

Page 4 of 14

equal probability systematic selection from the freshly constructed household listing. All women aged 15–49 and men aged 15–59 who were either permanent inhabitants of the selected homes or guests who stayed in the household the night before the survey were eligible to be addressed [47].

Data sources

This section describes the data sources used in this research. The 2019 EMDHS was the country's second mini-demographic and health survey. The first EMDHS is conducted in 2014. The response variable, community-based health insurance coverage, is taken from the 2019 EMDHS, whereas the auxiliary variables are from the 2007 population and housing census data. We assume that the functional relationship between health insurance and auxiliary variables remains constant across the survey and census periods.

The auxiliary variables for this study were derived from Ethiopia's population and housing census in 2007. The study considered 41 auxiliary variables. They were generated from the population census data in two ways: through variables relating to parents and young children under five. For example, children under five have been chosen based on their gender (male and female) and age (below one year, 1–2 years, and 4–5 years). On the other hand, the auxiliary variables of the parents included sex (male and female), place of residence (urban and rural), age (15–24, 25–34, 35–44, and 45–49), source of drinking water (improved and unimproved), educational levels (non-educated, elementary, secondary, and above), literacy (literate and illiterate), marital status (married, never married, and others), type of toilet facility (has toilet facility and does not have toilet facility), the number of sons who passed away (none, one, or two or more), the number of daughters who passed away (none, one, or two or more), the size of the family (less than five, and five or more), the presence or absence of disabilities (disabled and not), and the type of employment (government employment, private employment, self-employment, employer, unemployed, or other employment) [2, 48, 49]. These variables are considered at the zonal level as auxiliary variables in SAEs.

Figure 1 shows the Ethiopian zones with the 2019 EMDHS enumeration areas. Ethiopia has nine regions and two administrative cities at the subnational administrative level. However, there are 83 zones, of which 77 were sampled, and the remaining 6 were non-sampled zones for community-based health insurance coverage in the 2019 EMDHS data.

The Fay-Herriot model

In 1979, Fay and Herriot introduced the Fay-Herriot model to estimate income in counties in the United States [23]. Since then, numerous researchers have explored the concepts of SAE in various ways. Several authors have

Fig. 1 The enumeration area within Ethiopian zones

written comprehensive textbooks on the SAE approach at both the unit and area levels [22, 50], with the most recent one discussing the concept of SAE in mixed-effect models [46]. Pfeffermann [32] conducted an excellent review of the new developments in SAE.

The most popular small-area model for area-level data is the Fay-Herriot model [23, 24], which includes an area-linking model. In order to compare binary data, we looked at three popular models. The Fay-Herriot model [23] is the first with known sampling variances and normal distributions for both the sampling and linking models. The second model is the log-normal model, and the third is the normal-logistic model. The normal-logistic model differs from the Fay-Herriot model only in the linking model by using a logit-normal distribution instead of a normal distribution. The log and logit link functions transform the health insurance coverage data in the linking model.. The logit linking model is used here to ensure that the estimates fall within the proper range of (0, 1).

There are two models in a general area level model. A sampling model for the sampling error of direct survey estimates and the linking model linked the population value for area-to-area specific auxiliary variables $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})$ through a logit model $\theta_i = x_i\boldsymbol{\beta} + v_i$, where v_i the random area effects are the true means of the target variable in the m areas $i = 1, 2, \cdots, m$.

Consider simple random sampling within areas $(i = 1, \cdots, m)$ and sample proportion $\bar{y}_i$ as direct estimators (i.e., estimators from the 2019 EMDHS data only) of the population mean. Let N_i be the population sizes of an area i of finite population N. And let y_{ik} be the binary response characteristics of interest (i.e., households with and without health insurance coverage) for a unit k in the area i. The small area proportion parameter that needs to be estimated is given by $P_i = \sum y_{ik}/N_i$, and the variance P_i is denoted by Ψ_i and given by $\Psi_i = Var(P_i) = \frac{P_i(1-P_i)}{N_i}$.

The response variables y_i for the ith area response variable are CBHI enrolment. The variables are binary, which are CBHI users and non-users. Thus, the appropriate binomial distribution is $y_i \sim bin(N_i, P_i)$, where P_i are the observed local zones CBHI prevalence in the zones $(i = 1, 2, \cdots, m)$, y_i the observed number of CBHI enrolments in the zone i, and finally N_i, the sample sizes at the zone i. The logistic regression model, which is the binomial regression model, is appropriate to predict θ_i. Thus, the predictive *logit* regression model is given by $\theta_i = logit(P_i) = \mathbf{x}\beta + \mathbf{v_i}$ [8, 41, 44]. The three commonly used models are summarized as.

Model 1 (Normal Fay-Herriot model)
Sampling model: $y_i/\theta_i \sim^{iid} N(\theta_i, \Psi_i)$.
Linking model: $\theta_i/\boldsymbol{\beta}, \sigma_v^2 \sim^{iid} N(x_i\beta, \sigma_v^2)$.

Model 2 (Normal-log Fay-Herriot model)
Sampling model: $y_i/\theta_i \sim^{iid} N(\theta_i, \Psi_i)$.
Linking model: $log(\theta_i)/\boldsymbol{\beta}, \sigma_v^2 \sim^{iid} N(x_i\beta, \sigma_v^2)$.
Model 3 (Normal-logistic Fay-Herriot model):
Sampling model: $y_i/\theta_i \sim^{iid} N(\theta_i, \Psi_i)$.
Linking model: $logit(\theta_i)/\boldsymbol{\beta}, \sigma_v^2 \sim^{iid} N(x_i\beta, \sigma_v^2)$.

The sampling variance Ψ_i is the direct estimator estimated from the survey data. Model 1 is a matched model since the sampling and linking models can be combined to produce a simple linear mixed model. However, a nonlinear linkage model also unmatched model, such as model 2 or model 3, is generally recommended for modeling proportions [51, 52].

Hierarchical Bayes SAE model

We employed the HB SAE approach by assigning prior distributions on unknown parameters β and σ_v^2 to compute a posterior distribution of μ, which produces the point estimator as well as credible intervals [46]. The parameter of interest y_i and other model parameters are represented as random variables rather than fixed, unknown values under the Bayesian paradigm [51, 52]. Using Bayes' Theorem, this approach combines data-derived information from a likelihood function with previous knowledge about the parameter of interest and model parameters to produce a posterior distribution for the parameters.

Let $\mu = logit(\theta)$, and then the posterior distribution is

$$P\left(\mu, \beta, \sigma_v^2 | data\right) \propto P\left(data | \mu, \beta, \sigma_v^2\right) P\left(\mu, \beta, \sigma_v^2\right) \quad (2.1)$$

The HB approach also requires specifying prior distribution for population parameters $\boldsymbol{\beta}$ model coefficients as

$$f(\boldsymbol{\beta}) \propto 1$$

For the between area variations parameters or model component parameters, the prior is considered the non-informative uniform distribution given as

$$f(\sigma_v^2) \propto Unif(0, 1)$$

We use Markov Chain Monte Carlo (MCMC) techniques to generate the model parameters. MCMC algorithms approximate posterior distributions by sampling from a probability distribution. In order to obtain the results of each HB model we fit, we run 50,000 sampling iterations (the length of each Markov chain) [53, 54]. One measure of model compatibility that can be used in evaluating the compatibility of Bayes statistics is the Deviance Information Criterion (DIC) [55]. The optimal candidate model is chosen using the DIC. It is interpreted similarly to the AIC, with models with smaller DIC being

Shiferaw *et al. Health Economics Review*   (2024) 14:29

Page 6 of 14

preferred. It is based on the model's deviation and penalized for model complexity.

The coefficient of variation (CV), which is displayed as a percentage of the estimate, shows the sampling variability. Large CVs are considered untrustworthy for small area estimates. If the CV for a small area estimate is unacceptably high, it will be deemed untrustworthy [56]. Although there are no globally recognized tables to determine what is "too large" [57], the smallest CV is preferred. The study uses the BayesSAE R software package [58].

Bias diagnostic measures are crucial in SAE to assess the accuracy of model-based estimates. One such measure is to plot the direct estimates on the y-axis and the model-based estimates on the x-axis. By examining the deviation of the small area estimates from the regression line fitted values, we can determine the level of bias.

The Generalized Variance Function (GVF)

When dealing with a complex situation, the Generalized Variance Function (GVF) can be used to reduce the uncertainty in design-based variance estimates [20, 58, 59]. Generally, the standard design-based estimators tend to have high variability when the sample sizes are small. Therefore, having stable estimators is crucial for smaller domains or zones. The GVF provides more reliable estimators for domain-level variance than direct design-based variance estimators [60]. The GVF was fitted by using simple linear regression as

$$y_i = \beta_0 + \beta_1 x_i + e_i, i = 1, 2, \cdots, m$$

Where the dependent variable y_i is $log(Var(P_i))$, and the independent variable x_i is, the $log(P_i)$ success probability at the zone level. β_0 and β_1 are the constant and slope parameters estimates, and $e_i \sim^{iid} N(0, \sigma^2)$ therefore, the GVF function is $\widehat{GVF} = exp(\hat{\sigma}^2/2)exp(\hat{\beta}_0 + \hat{\beta}_1 x_i)$ [30, 59–61].

Results

Model selection and fitting

We used Ethiopia's population and housing census 2007 data to select suitable auxiliary variables for Fay-Herriot modeling [23, 62]. The auxiliary variables are used for linking to the response variable using the linking small area model. The response variable, community-based health insurance coverage in Ethiopia, is binary; therefore, data transformation is conducted. The log and logit transformations are used to transform the binary outcome variable.

There are almost 41 auxiliary variables available from the census data, and we have done exploratory analysis

Table 1 Model comparison

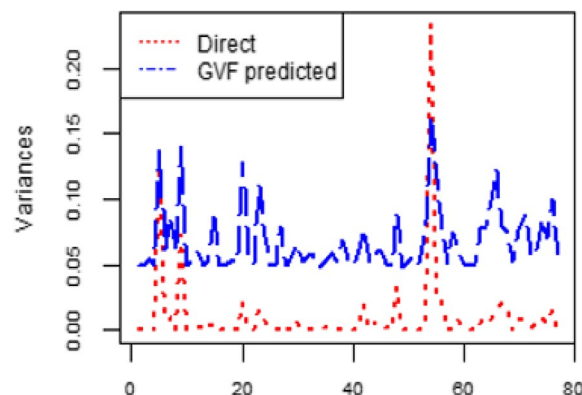| Model	DIC
Fay-Herriot	1133.3222
Log Fay-Herriot	1133.3750
Logit Fay-Herriot	1130.0822

Fig. 2 Dispersion plots of GVF fit for health insurance coverage

before selecting appropriate variables. Eight auxiliary variables were chosen to link the survey data using the appropriate variable selection method. We applied the DIC to select the most suitable candidate model for estimating health insurance coverage. The smaller the deviance is, the better the model. The logit transformations of the Fay-Herriot model have small DIC. Thus, the logit Fay-Herriot model is the best for analyzing health insurance coverage Table 1.

The generalized variance function

Standard design-based estimators are not precise enough when sample sizes are too small. In this study, the GVF was used to produce area-level variance estimators that would be more stable [63]. The GVF aids in smoothing out unreliable and noisy estimated variances [59, 60]. Figure 2 shows the direct survey estimated variance (red color) and the GVF predicted variance (blue color). The GVF predicted variance is relatively smoother than the direct survey variances.

The table (Table 2) displays the posterior mean, standard deviation (SD), and 95% credible interval (CI) for the regression coefficients of the normal-logistic Fay-Herriot model. It is evident from the table that marital status (married) and having more than five children are significant predictors of health insurance coverage since the CIs of both variables do not include zero.

Shiferaw *et al. Health Economics Review* (2024) 14:29

Page 7 of 14

Table 2 The 95% CIs for the regression coefficients under the normal-logistic Fay-Herriot model

Coefficient	Auxiliary variables	Posterior mean	SD	95% credible interval	
				Lower limit	Upper limit
β_0	Constant	-2.75	22.53	-46.43	41.00
β_1	Age category more than 44 years	-24.43	13.41	-50.51	2.73
β_2	Urban	-0.40	1.49	-3.29	2.52
β_3	Female	-5.13	3.50	-11.99	1.80
β_4	Not disabled	14.29	22.93	-31.21	58.92
β_5	Married	-4.56	2.28	-9.07	-0.095
β_6	Nunber of children more than five	4.19	1.73	0.75	7.65
β_7	Government employement	-5.40	3.59	-12.44	1.64
β_8	Primary level of education	-0.68	2.09	-3.39	4.75
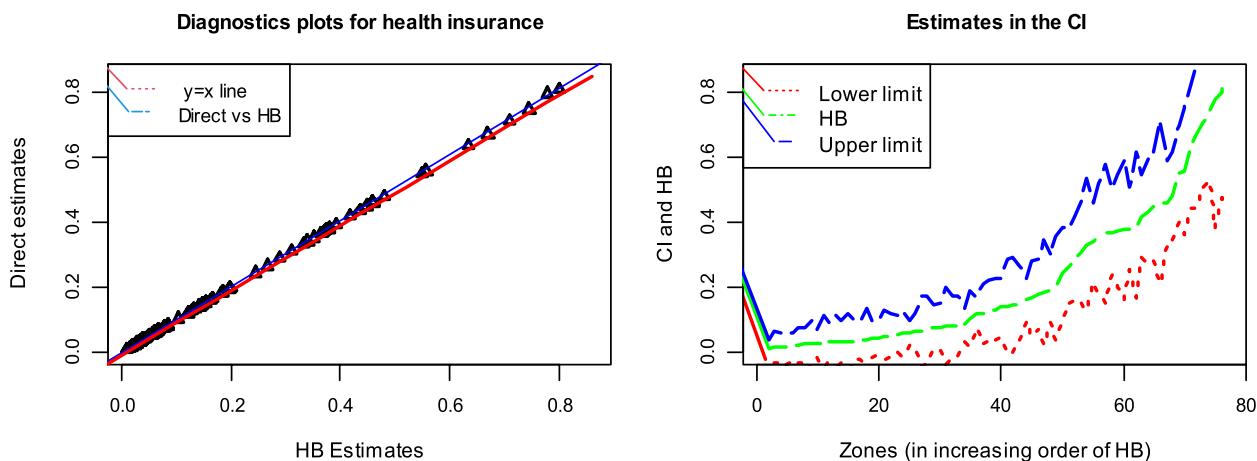
Fig. 3 Bias diagnostics plots (left) and zonal-wise 95% CIs for HB estimates (right)

Bias diagnostic

The bias diagnostic measure assesses the validity, while the 95% CI and coefficient of variation (CV) assess the precision of the model-based estimates [57]. The regression of the direct estimates on the true population values is projected to be linear with the identity line since they are unbiased estimators of the population parameters. Suppose the model-based small area health insurance estimates are close to the population's true values. In that case, the regression of direct estimates on model-based small area estimates should be comparable. As a result, we plotted the direct and model-based estimates on the y and x axes, respectively, and looked at how much the small area estimates deviated from the regression line fitted values. Figure 3 shows the estimates of HB bias diagnostics plots against the baseline $y = x$.

Figure 4 displays the direct survey and model-based HB estimates of the prevalence of CBHI in Ethiopian zones. In principle, the values of HB and direct survey estimates are equivalent [21]. Therefore, HB and direct survey estimates are comparable and consistent for all zones (Fig. 3).

The Bayesian model assumes that the random area-specific effects are normally distributed with a zero mean and constant variance σ_v^2. If the underlying model assumptions hold, the zonal level residuals should be distributed randomly around zero. To test the normality assumption of the residuals of the fitted model, we used the normal probability (Q-Q) plot. Figure 5 displays the Q-Q plots of the zonal-level residuals for health insurance coverage. The Shapiro–Wilk test was used to assess the normality assumption of the zonal-specific random effects, and it was found to meet the criteria for the assumptions.

Improvements of model based HB estimates of health insurance coverage

The root MSE and CV of the HB estimates for the prevalence of health insurance coverage are shown in Fig. 6. The CVs and root mean square error (root MSE) of the
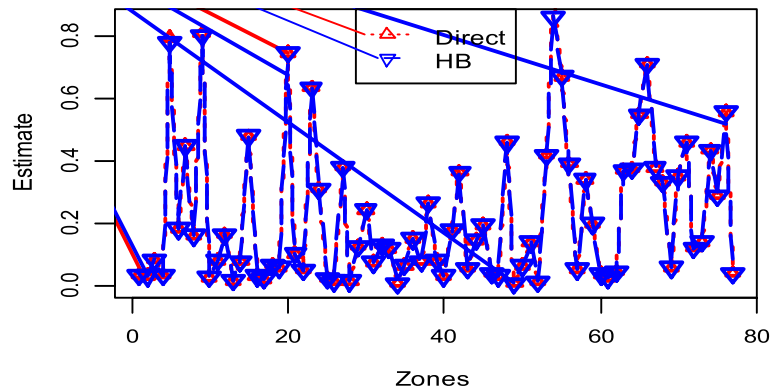
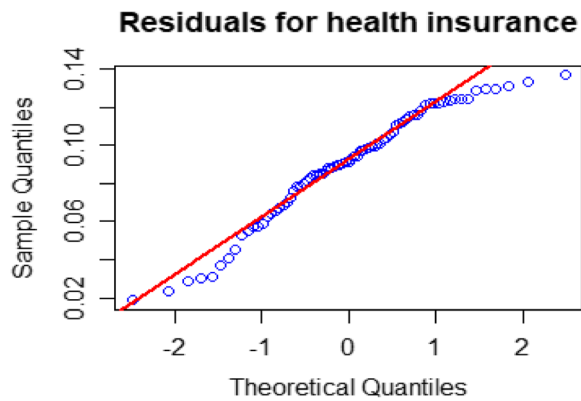**Fig. 4** The consistency of HB and the direct survey estimates

Fig. 5 Q-Q plot of the Fay-Herriot model residuals

HB estimates were smaller than the corresponding direct survey estimates. These findings demonstrate that HB estimates, which borrow strength from auxiliary variables in the census data, are more accurate than direct survey estimates [32, 64]. Model based HB estimates are empirical best linear unbiased predictors (EBLUP).

Table 3 shows the efficiency improvements in the CV of HB estimates over the direct survey estimates. The minimum, first quartile, mean, third quartile, and maximum values for the CV for HB estimates are found to be lower than those for the equivalent direct survey estimates. As a result, compared to estimates from the direct survey, those from the HB Fay-Herriot model are more accurate and reliable. Compared to direct survey estimates, CV efficiency has increased by a mean of 46.75 and a
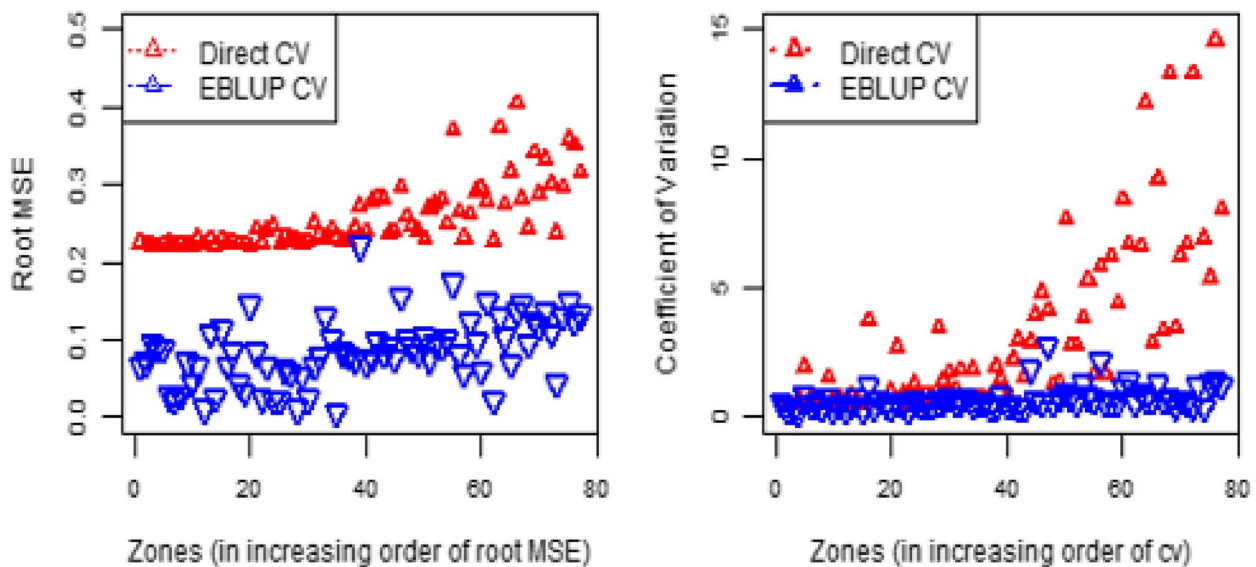
Fig. 6 Zone-specific root MSE (left) and percentage CV (right) of direct and HB health insurance coverage estimates

Shiferaw *et al. Health Economics Review*   (2024) 14:29

Page 9 of 14

Table 3 Comparatives of % CV for the direct and HB estimates of health insurance coverage

Quantities	Direct CV	EBLUP CV	Efficiency gain
Minimum	0.2337	0.2338	-6.03
1st quartile	1.8876	1.4892	21.11
Median	5.9076	3.3748	42.87
mean	6.8314	3.6376	46.75
3rd quartile	8.9087	4.3265	51.43
Maximum	15.6256	4.5400	70.95

Table 4 The CV of health insurance coverage for non-sampled zones

Value	Direct CV	EBLUP CV
Minimum	-	0.5600
1st quartile	-	2.1563
Median	-	3.2135
Mean	-	3.7568
3rd quartile	-	4.9320
Maximum	-	5.1223

data from census data, we provided an estimate for health insurance coverage. Table 4 displays the CV of HB estimates for small area health insurance for non-sampled zones.

Figure 7 depicts the prevalence of health insurance coverage across Ethiopian administrative zones. On the Ethiopian maps, the prevalence indicates that green indicates better and red indicates inadequate health insurance coverage. Zones such as North Wollo, Awi, East Gojam, South Wollo, and East Tigray, located in the country's north and northwest, have better estimates of the prevalence of health insurance coverage. However, coverage for health insurance needs to be improved in the Metekel, all Afar and Somalia regional zones, Kelem Wollega, Gedio, Shaka, Mezhenger, Sidama, Horo Guduru Wollega, and Dawro zones.

Discussion

This research focused on estimating the zone-level CBHI in Ethiopia by utilizing the 2019 EMDHS and the 2007 population census data. The CBHI enrollment status is represented using binary outcomes, where zero indicates non-insured individuals and one represents those enrolled in CBHI [65]. For binary outcomes, we employed
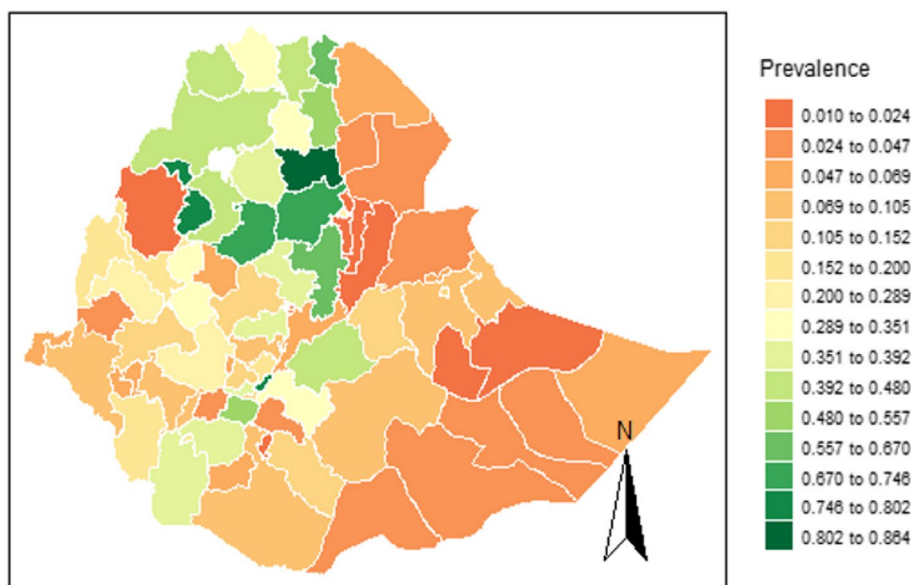
Fig. 7 The distribution of the HB estimates for the prevalence of health insurance coverage across Ethiopian zones

maximum of 70.95. The HB estimates had an advantage over direct survey estimates because they borrowed strength from the census data source.

Due to the zero sample size, no direct survey estimates of health insurance coverage for non-sampled zones exist. However, based on the HB approach and auxiliary

the Logit Normal model with the HB SAE method [8, 22, 41, 50]. This method is advantageous because it simplifies model specifications and is relatively easy to compute using the MCMC technique [44]. This study utilizes the GVF to effectively smooth unreliable and noisy estimates of sampling error variances for SAE obtained through

direct surveys. The results indicate that GVF helps to smooth out unreliable and noisy estimated variances, which is consistent with previous SAE literature [59, 60, 65–67]. Figure 2 shows the direct survey estimated variance (red color) and the GVF predicted variance (blue color). The GVF predicted variance is smoother compared to the variances obtained through direct survey estimates. In HB SAE applications, model diagnostics for small area estimates are typically employed. These techniques are used to verify the assumptions made by the model [59, 68]. The Shapiro–Wilk normality test and Q-Q plot are the most commonly used methods to establish normality [53, 69]. The model assumptions of the random area effect v are assumed to have identically and independently distributed with mean zero and variance ($\sigma_v^2 = 0.3243$). The results of the Shapiro–Wilk test indicate that the residuals at the zonal level (W $= 0.950$ and p-value $= 0.085$) support the normality assumption for the data on health insurance coverage. This suggests that the normality assumption for the health insurance coverage data is adequately met. Similar studies in Ethiopia about children under the age of five have agreed with the normality assumptions [31, 34].

The bias diagnostics plot is the second diagnostic measure used to evaluate the reliability of HB estimates [68]. It helps identify potential sources of bias and assesses the model's validity. When estimating health insurance coverage through HB, it should be comparable to direct survey estimates. Figure 3 shows the direct survey estimates on the y-axis and the HB estimates on the x-axis. The graph depicts a strong relationship between both estimates, which is consistent with various literature sources [33, 53]. The direct survey-based and HB estimates agree, indicating that they are consistent and aligned. To put it differently, the estimates obtained from the HB method and the direct survey method are consistent with each other.

Additionally, three widely used diagnostic measurements, such as percent CV, root MSE, and 95% credible intervals, are also used to assess the accuracy of the HB estimates [69]. Even though there are no precise cut-off values for CV to determine precision, the lower the CV, the more precise and consistent the estimates [70]. According to the condensed tables and figures (Fig. 6 and Table 3), the HB estimates have a lower CV in each zone. The HB estimates are relatively more accurate than the corresponding direct estimates in all zones, according to the CV, root MSE, and 95% CI, which are consistent with the different study results [57, 71, 72].

The results of this study found that auxiliary variables such as marriage and the number of children of more than five are statistically significant predictors of CBHI prevalence. This finding explicitly revealed that

respondents had more than five children in the household, which is a significant effect of the CBHI scheme. Other studies are consistent with our results [3, 73, 74]. The number of children over five significantly impacts the CBHI prevalence, which is consistent with the studies [2, 3]. Similarly, the studies conducted in South Gondar zones of Ethiopia show that those having more than five children in the household were two times more likely to use the CBHI scheme when compared with those who have fewer than five children [2, 3].

This could be because families with more children require more medical care, such as immunizations, which could explain why they seek enrollment in CBHI programs. However, recent research contradicts earlier studies in Kenya, which suggested that the number of children did impact CBHI enrollment [75]. Under UHC, healthcare should be accessible to all without causing financial hardship [17]. Achieving UHC is essential for long-term development and reducing poverty. It is also critical in any effort to combat social inequality and improve access to healthcare [5, 16].

The study also demonstrates that auxiliary variables, like marital status, can significantly influence the possibility of CBHI adoption in Ethiopia. This suggests that evaluating and enhancing the uptake of CBHI schemes requires an understanding of the demographic and social factors of the population. This finding aligns with a study in the North Shoa Zone of Ethiopia by [76], which also found a significant association between marital status and CBHI utilization.

Using the estimation from the disaggregated level HB estimates, we were able to identify disparities in CBHI coverage between Ethiopian administrative zones. It was found that the Amhara and Tigray regional zones have better CBHI coverage than other regions of the nation, as shown in Fig. 7. This is because the Ethiopian government approved and initiated the CBHI practice in the Tigray and Amhara regions before rolling it out nationwide to improve healthcare quality [49]. As a result, the prevalence of CBHI in Ethiopia differs from region to region, with higher enrollment rates in the Amhara and Tigray regions than in others [77]. This finding agrees with a study by [2], which also reported higher enrollment rates in these regions. Similarly, another study by [48] found that the overall pooled prevalence of CBHI enrollment in Ethiopia was 45.5%, with better coverage in the Amhara and Tigray regions. Our research also confirms that there are variations in health insurance coverage across different administrative zones [77, 78]. It is evident that health insurance coverage varies geographically.

Our study is unique in that it provides synthetic estimates for small areas with zero sample sizes. These estimates are generated by either combining data from

different sources or using statistical models to infer values for smaller areas based on larger, more reliable datasets. Synthetic estimates help make informed decisions when direct measurements are unavailable [22, 50]. In Table 4, the HB model-based SAE produced CBHI estimates for non-sampled zones. We used the HB model-based SAE to estimate CBHI coverage in non-sampled zones, which borrowed auxiliary variables from census datasets. This approach is consistent with most studies in SAE contexts [39, 78, 79].

Conclusions and policy implications

The SAE approach is a statistical method used to estimate the proportion of socioeconomic indicators at the local or domain level. This method has been extensively studied and adopted as a target by national statistics offices in developed countries. However, there is a need for more research on how to apply SAE in developing countries such as Ethiopia. For example, in order to track progress towards achieving the SDGs, it is essential to have disaggregated estimates at lower administrative levels or specific domain levels. This initiative aligns with the global framework of indicators for monitoring the SDGs, which suggests that information should be disaggregated not only geographically (in subregions such as provinces, municipalities, or districts) but also by income group, gender, age, race, ethnicity, immigration status, and disability status [80].

This study found that by combining data from the 2019 EMDHS with additional information from the 2007 population census, more accurate estimates of health insurance coverage at the zone level were obtained using the SAE approach. The research also suggests that relying solely on survey data for estimating the CBHI coverage in Ethiopia's administrative zones is less precise than when auxiliary variables are also taken into account.

Ethiopia is strongly committed to achieving UHC, which means providing high-quality healthcare services that are accessible and equitable to all. Our study has highlighted the differences across Ethiopian zones regarding their coverage of CBHI schemes. We have also identified some areas with high coverage and some with low coverage. This information can be very useful in improving local planning and policy decision-making, which can ultimately lead to an increase in the coverage of CBHI schemes across Ethiopia.

Overall, this study stands out due to its innovative approach of integrating data from two sources, namely the 2007 population census and the 2019 EMDHS, to improve the accuracy of health insurance coverage estimation at the zone level. The study employed the HB modeling framework in SAE to estimate health insurance coverage at the lower administrative level (zones), which

can help policymakers gain a more nuanced understanding of coverage disparities and target resources and interventions more effectively. Furthermore, this paper's novelty lies in estimating the non-sampled zones, which will receive equal attention from policymakers as the sampled zones. This study contributes to evidence-based decision-making in healthcare planning and resource allocation, making it valuable for researchers and policymakers.

Summary of key policy implications

The study's key policy implications are summarized below:

(i) Policymakers can gain valuable insights into the spatial distribution of health insurance coverage at the zone level. The map in Fig. 7 can identify zones with inadequate coverage, which can help policymakers prioritize interventions, allocate additional funding, and implement specific outreach programs to increase health insurance enrollment in those zones.

(ii) The study highlights disparities in health insurance coverage across Ethiopia. Policymakers can use this information to improve coverage in underserved areas with targeted interventions such as financial incentives, awareness campaigns, and improved access. Ensuring equity in access and benefits for all populations is key, which aligns with the UHC.

(iii) The study offers a robust methodology to estimate health insurance coverage at lower administrative levels, which are zones in the Ethiopian case. The methodology can be replicated to monitor coverage changes and evaluate policy interventions. Policymakers can use the findings to set targets, monitor progress, and make informed decisions about health insurance expansion.

(iv) The study emphasizes the significance of having trustworthy and current data sources for precisely estimating health insurance coverage. Policymakers can use this research to advocate for better data collection efforts, including regular surveys and updates to census data. By investing in data infrastructure and ensuring data quality, policymakers can monitor and evaluate programs more effectively, identify areas for improvement, and make data-driven policy decisions.

Abbreviations

CBHI	Community-based health insurance
CI	Credible interval
CV	Coefficient of variation
DHS	Demographic and health survey
EA	Enumeration areas
EBLUP	Empirically best lieaner unbiased predictors

Shiferaw *et al. Health Economics Review*   (2024) 14:29

Page 12 of 14

EMDHS   Ethiopian mini DHS
GVF   Generalized variance function
HB   Hierarchical Bayes
MCMC   Morkov Chain Monte Carlo
MSE   Mean square error
SAE   Small area estimation
SDG   Sustainable development goal for health
SD   Standard deviation
UHC   Universal health coverage

Availability of data and materials
The 2019 DHS data for this study was received through a registration request from the DHS program website, https://www.dhsprogram.com. The GPS enumeration area shapefiles were received similarly from https://www.dhsprogram.com. Shapefiles for Ethiopian zonal administrative borders are also accessible on the website https://africaopendata.org. For research purposes, the Ethiopian Central Statistical Agency gives access to a 10% sample of the 2007 census.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
Not applicable.

References
1. Tetzlaff J, Luy M, Epping J, Geyer S, Beller J, Stahmeyer JT, et al. Estimating trends in working life expectancy based on health insurance data from Germany – Challenges and advantages. SSM - Popul Heal. 2022;19:101215. https://doi.org/10.1016/J.SSMPH.2022.101215.
2. Moyehodie YA, Mulugeta SS, Amare Yilema S. The effects of individual and community-level factors on community-based health insurance enrollment of households in Ethiopia. PLoS One 2022;17:e0275896. https://doi.org/10.1371/journal.pone.0275896.
3. Moyehodie YA, Fenta SM, Mulugeta SS, Agegn SB, Yismaw E, Biresaw HB, et al. Factors Associated With Community Based Health Insurance Healthcare Service Utilization of Households in South Gondar Zone, Amhara, Ethiopia. A Community-Based Cross-Sectional Study. Heal Serv Insights 2022;15. https://doi.org/10.1177/11786329221096065.
4. Osei Afriyie D, Krasniq B, Hooley B, Tediosi F, Fink G. Equity in health insurance schemes enrollment in low and middle-income countries: A systematic review and meta-analysis. Int J Equity Health. 2022;21:1–12. https://doi.org/10.1186/s12939-021-01608-x.
5. Acharya A, Vellakkal S, Taylor F, Masset E, Satija A, Burke M, et al. The impact of health insurance schemes for the informal sector in low-and middle-income countries: A systematic review. World Bank Res Obs. 2013;28:236–66. https://doi.org/10.1093/wbro/lks009.
6. WHO. Selection of essential in vitro diagnostics at country level: using the WHO Model List of Essential In Vitro Diagnostics to develop and update a national list of essential in vitro diagnostics. 2021. https://iris.who.int/handle/10665/343385.
7. Ranabhat CL, Kim CB, Singh DR, Park MB. A Comparative Study on Outcome of Government and Co-Operative Community-Based Health Insurance in Nepal. Front Public Heal 2017;5. https://doi.org/10.3389/fpubh.2017.00250.
8. Sari AD, Yanuar F. Hierarchical Bayesian Modelling in Small Area for Estimating Binary Data. J Phys Conf Ser 2020;1554. https://doi.org/10.1088/1742-6596/1554/1/012049.
9. Artignan J, Bellanger M. Does community-based health insurance improve access to care in sub-Saharan Africa? A rapid review. Health Policy Plan 2021;36. https://doi.org/10.1093/heapol/czaa174.
10. Shewamene Z, Tiruneh G, Abraha A, Reshad A, Terefe MM, Shimels T, et al. Barriers to uptake of community-based health insurance in sub-Saharan Africa: a systematic review. Health Policy Plan 2021;36. https://doi.org/10.1093/heapol/czab080.
11. Merga BT, Balis B, Bekele H, Fekadu G. Health insurance coverage in Ethiopia: financial protection in the Era of sustainable cevelopment goals (SDGs). Health Econ Rev. 2022;12:1–7. https://doi.org/10.1186/s13561-022-00389-5.
12. CSA, ICF. Ethiopia mini demographic and health survey: Key indicators. Addis Ababa, Ethiopia; 2019. http://www.ephi.gov.et.
13. Michael McWilliams J. Health Consequences of Uninsurance among Adults in the United States: Recent Evidence and Implications. Milbank Q 2009;87. https://doi.org/10.1111/j.1468-0009.2009.00564.x.
14. Pauly M V., Pagán JA. Spillovers and vulnerability: The case of community uninsurance. Health Aff 2007;26. https://doi.org/10.1377/hlthaff.26.5.1304.
15. Assembly, G. S development goals. SDGs Transform Our World, 2030. 2015. Retrieved from https://sdgs.un.org/2030agenda on 13 January 2023.
16. United Nations General Assembly. Transforming our world: The 2030 agenda for sustainable development. Technical report. United Nations. 2015. https://sdgs.un.org/2030agenda.
17. WHO. Universal health coverage (UHC). 2022. Retrieved from https://www.who.int/news-room/fact-sheets/detail/universal-health-coverage-(uhc) on 13 January 2023.
18. FDRE M o. H. Health Sector Transformation Plan. Addis Ababa, Ethiopia; 2015-2019. https://ephi.gov.et/wp-content/uploads/2023/01/2.-Ethiopia-Health-Sector-Transformation-Plan-I-2015-2020-Endline-Review.pdf.
19. Merga BT, Raru TB, Deressa A, Regassa LD, Gamachu M, Negash B, et al. The effect of health insurance coverage on antenatal care utilizations in Ethiopia: evidence from national survey. Front Heal Serv. 2023;3:1–7. https://doi.org/10.3389/frhs.2023.1101164.
20. Bayked EM, Toleha HN, Kebede SZ, Workneh BD, Kahissay MH. The impact of community-based health insurance on universal health coverage in Ethiopia: a systematic review and meta-analysis. Glob Health Action 2023;16. https://doi.org/10.1080/16549716.2023.2189764.
21. Yunhan Wu, Richard Li Z, Mayala BK, Wang H, Gao PA, Paige J, et al. Spatial Modeling for Subnational Administrative Level 2 Small-Area Estimation. vol. 21. 2021. https://dhsprogram.com/pubs/pdf/SAR21/SAR21.pdf.
22. Rao JNK, Molina I. Small area estimation. 2nd ed. Hoboken: John Wiley & Sons, Inc., Hoboken, New Jersey; 2015. https://doi.org/10.1002/9781118735855.
23. Fay R, Herriot R. Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data. J Am Stat Assoc. 1979;74:269–77. https://doi.org/10.1080/01621459.1979.10482505.
24. Rao JNK, Molina I. Small area estimation. 2nd ed. Hoboken, New Jersey: John Wiley & Sons Inc; 2015.
25. Corral P, Molina I, Cojocaru A, Segovia S. Guidelines to Small Area Estimation for Poverty Mapping. 2022. https://doi.org/10.1596/37728.
26. Asian Development Bank. Introduction to Small Area Estimation Techniques: A Practical Guide for National Statistics Offices. 2020. https://www.adb.org/sites/default/files/publication/609476/small-area-estimation-guidensos.pdf.
27. Newhouse DL, Merfeld JD, Ramakrishnan A, Swartz T, Lahiri P. Small Area Estimation of Monetary Poverty in Mexico Using Satellite Imagery and

Machine Learning. SSRN Electron J. 2022. https://doi.org/10.2139/ssrn.4235976.

28. Ren R. Design-based Small Area Estimation : An Application to the DHS Surveys DHS Work Pap No 180. Rockville, Maryland: USA ICF; 2021.

29. UNFPA. Better Data for More Effective Policies and Programmes: Small area estimation 2022. https://www.398unfpa.org/publications/small-area-estimation.

30. Yilema SA, Shiferaw YA, Zewotir T, Muluneh EK. Improving survey based estimates of malnutrition using small area estimation. Stat J IAOS. 2022;38:1–11. https://doi.org/10.3233/SJI-210892.

31. Yilema SA, Shiferaw YA, Zewotir T, Muluneh EK. Multivariate small area estimation of undernutrition for children under five using official statistics. Stat J IAOS. 2022;38:1–12. https://doi.org/10.3233/SJI-220935.

32. Pfeffermann D. New important developments in small area estimation. Stat Sci. 2013;28:40–68. https://doi.org/10.1214/12-STS395.

33. Yilema SA, Shiferaw YA, Zewotir TT, Muluneh EK. Spatial small area estimates of undernutrition for under five children in Ethiopia via combining survey and census data. Spat Spatiotemporal Epidemiol. 2022;42:100509. https://doi.org/10.1016/J.SSTE.2022.100509.

34. Yilema SA, Shiferaw YA, Zewotir T, Muluneh EK. Small Area Estimation of Undernutrition under Age Five Children Based on Spatial-Temporal Models. Sci World J 2022;2022. https://doi.org/10.1155/2022/6882047.

35. Marhuenda Y, Molina I, Morales D. Small area estimation with spatiotemporal Fay-Herriot models. Comput Stat Data Anal. 2013;58:308–25.

36. Singh BB, Shukla GK, Kundu D. Spatio-Temporal Models in Small Area Estimation Spatio-Temporal Models in Small Area Estimation 2005. Vol. 31, No. 2. https://www150.statcan.gc.ca/n1/pub/12-001-x/2005002/article/9053-eng.pdf.

37. Bauer C, Li X, Zhang K, Lee M, Guajardo E, Fisher-Hoch S, et al. A Novel Bayesian Spatial-Temporal Approach to Quantify SARS-CoV-2 Testing Disparities for Small Area Estimation. Am J Public Health. 2023;113:40–8. https://doi.org/10.2105/AJPH.2022.307127.

38. Petrucci A, Salvati N, Petrucci A, Salvati N. Small Area Estimation for Spatial Correlation in Watershed Erosion Assessment. Agric Biol Environ Stat. 2006;11:169–82. https://doi.org/10.1198/108571106X110531.

39. Wawrowski Ł. The Spatial Fay-Herriot Model in Poverty Estimation. Folia Oeconomica Stetin. 2016;16:191–202. https://doi.org/10.1515/foli-2016-0034.

40. Hwang H-J, Shin K-I. Logistic Regression Type Small Area Estimations. Korean J Appl Stat. 2011;24:445–53. https://doi.org/10.5351/KJAS.2011.24.3.445.

41. E Sunandi 2018 Logit-Normal Spatial Model for Small Area Estimation : Case Study of Poverty in Bengkulu Logit-Normal Spatial Model for Small Area Estimation : IOP Conf Ser https://doi.org/10.1088/1755-1315/187/1/012051

42. Purwa T, Rumiati AT, Zain I. Small Area Estimation with Bivariate Hierarchical Bayes (HB) Approach to Estimate Monthly Average per Capita Expenditure of Food and Non-Food Commodities in Province of Bali. IOP Conf. Ser. Mater. Sci. Eng., vol. 546, 2019. https://doi.org/10.1088/1757-899X/546/5/052054.

43. Anjoy P, Chandra H. Hierarchy Bayes estimation of small area means under a spatial nonstationary Fay-Herriot model. Commun Stat Simul Comput. 2021. https://doi.org/10.1080/03610918.2021.1926501.

44. Franco C, Bell WR. Applying Bivariate Binomial/Logit Normal Models to Small Area Estimation. JSM Proceedings, Surv Res Methods Sect. 2013:690–702. https://www.census.gov/content/dam/Census/library/workingpapers/2013/demo/francobell2013asaproceedings.pdf.

45. Nadhiroh IM, Notodiputro KA. Zero Inflated Negative Binomial Models in Small Area Estimation. Semin Mat Dan Pendidik Mat. 2008:183–93. https://eprints.uny.ac.id/6861/1/M-18%20Statistika%28Irene%29.pdf.

46. Sugasawa S, Kubokawa T. Mixed-Effects Models and Small Area Estimation. Tokyo: JSS Research Series in Statistics; 2023. https://doi.org/10.1007/978-981-19-9486-9.

47. Ethiopian Public Health Institute (EPHI), ICF. Ethiopia Mini Demographic and Health Survey 2019: Final Report. 2019. https://dhsprogram.com/pubs/pdf/FR363/FR363.pdf.

48. Habte A, Tamene A, Ejajo T, Dessu S, Endale F, Gizachew A, et al. Towards universal health coverage: The level and determinants of enrollment in the Community-Based Health Insurance (CBHI) scheme in Ethiopia: A systematic review and meta-analysis. PLoS ONE. 2022;17:1–21. https://doi.org/10.1371/journal.pone.0272959.

49. Dagnaw FT, Azanaw MM, Adamu A, Ashagrie T, Mohammed AA, Dawid HY, et al. Community-based health insurance, healthcare service utilization and associated factors in South Gondar Zone Northwest, Ethiopia, 2021: A comparative cross-sectional study. PLoS ONE. 2022;17:1–11. https://doi.org/10.1371/journal.pone.0270758.

50. Rao JNK. Small Area Estimation. John Wiley & Sons, Inc.; 2003. https://doi.org/10.1002/0471722189.

51. Trevisani M, Torelli N. Hierarchical Bayesian models for small area estimation with count data. 2007. https://doi.org/10.4236/ojs.2017.73036.

52. Trevisani M, Torelli N. A Comparison of Hierarchical Bayesian Models for Small Area Estimation of Counts. Open J Stat. 2017;07:521–50. https://doi.org/10.4236/ojs.2017.73036.

53. White GW, McConville KS, Moisen GG, Frescino TS. Hierarchical Bayesian Small Area Estimation Using Weakly Informative Priors in Ecologically Homogeneous Areas of the Interior Western Forests. Front For Glob Chang. 2021;4:1–15. https://doi.org/10.3389/ffgc.2021.752911.

54. Anjoy P, Chandra H. Hierarchical Bayes estimation of small area means under a spatial nonstationary Fay–Herriot model. Commun Stat Simul Comput 2021:1–19. https://doi.org/10.1080/03610918.2021.1926501.

55. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit (with discussion). J R Stat Soc. 2002:583–639. https://doi.org/10.1111/1467-9868.00353.

56. Datta G, Kubokawa T, Molina I, Rao JNK. Estimation of mean squared error of model-based small area estimators. TEST. 2011;20:367–88. https://doi.org/10.1007/s11749-010-0206-2.

57. Chandra H, Salvati N, Sud UC. Disaggregate-level estimates of indebtedness in the state of Uttar Pradesh in India: An application of small-area estimation technique. J Appl Stat 2011;38. https://doi.org/10.1080/02664763.2011.559202.

58. Ridson P, Ubaidillah A. Package 'saeHB'. 2023. https://github.com/zazaperwira/saeHB.

59. Shiferaw YA. Analysis of the spatial distribution of under-5 mortality rate in local areas of South Africa. Stat J IAOS. 2020;36:1161–73. https://doi.org/10.3233/SJI-200650.

60. Shiferaw YA. Model-based estimation of small area food insecurity measures in Ethiopia using the Fay-Herriot EBLUP estimator. Stat J IAOS. 2020;36:S177–87. https://doi.org/10.3233/SJI-200673.

61. Wolter L. Introduction. 2007. https://doi.org/10.1007/978-3-658-04536-4_1.

62. CSA. Population and Housing Census of Ethiopia. Adds Ababa; 2007. https://www.statsethiopia.gov.et/wpcontent/uploads/2019/06/Population-and-Housing-Census-2007-National_Statistical.pdf.

63. Esteban MD, Morales D, Pérez A, Santamaría L. Small area estimation of poverty proportions under area-level time models. Comput Stat Data Anal. 2012;56:2840–55. https://doi.org/10.1016/j.csda.2011.10.015.

64. Pfeffermann D, Ben-Hur D. Estimation of Randomisation Mean Square Error in Small Area Estimation. Int Stat Rev 2018;87:1–19. https://doi.org/10.1111/insr.12289.

65. Chandra H, Salvati N, Chambers R. Small area prediction of counts under a non-stationary spatial model. Spat Stat 2017;20. https://doi.org/10.1016/j.spasta.2017.01.004.

66. Islam S, Chandra H, Aditya K, Lal SB. Small area estimation under a spatial model using data from two surveys. Int J Agric Stat Sci. 2018;14:231–7.

67. Wolter KM. Introduction to Variance Estimation. 2nd ed. Chicago; 2007. https://doi.org/10.1007/978-3-658-04536-4_1.

68. Brown G, Chambers R, Heady P, Heasman D. Evaluation of Small Area Estimation Methods - An Application to unemployement Estimats from the UK LFS. Proc. Stat. Canada Symp., 2001, p. 1–10. https://doi.org/10.1201/9780203166314.ch1.

69. Anjoy P, Chandra H, Basak P. Estimation of Disaggregate-Level Poverty Incidence in Odisha Under Area-Level Hierarchical Bayes Small Area Model. Soc Indic Res. 2018;144:251–73. https://doi.org/10.1007/s11205-018-2050-9.

70. Chandra H. Exploring spatial dependence in area-level random effect model for disaggregate-level crop yield estimation. J Appl Stat. 2013;40:823–42. https://doi.org/10.1080/02664763.2012.756858.

71. Chandra H, Salvati N, Chambers R. Small area estimation for spatially correlated populations-a comparison of direct and indirect model-based methods; 2007. https://ro.uow.edu.au/infopapers/3022.

72. Guha S, Chandra H. Measuring disaggregate level food insecurity via multivariate small area modelling: evidence from rural districts of Uttar

Pradesh. India Food Secur. 2021;13:597–615. https://doi.org/10.1007/
S12571-021-01143-1.

73. Bantie GM, Woya AA, Zewdie BM. Community-based health insurance and
associated factors in north-western Ethiopia. The case of Bahir Dar city. Int
J Gen Med 2020;13:1207–17. https://doi.org/10.2147/IJGM.S264337.

74. Atnafu DD, Tilahun H, Alemu YM. Community-based health insurance
and healthcare service utilisation, North-West, Ethiopia: A comparative,
cross-sectional study. BMJ Open. 2018;8:1–6. https://doi.org/10.1136/
bmjopen-2017-019613.

75. Maina JM, Kithuka P, Tororei S. Perceptions and uptake of health insurance
for maternal care in rural Kenya: A cross sectional study. Pan Afr Med J
2016;23:1–10. https://doi.org/10.11604/pamj.2016.23.125.8936.

76. Getahun T, Teklesilassie L, Habtemichael M, Abebe Y, Getahun H. Mag-
nitude of community-based health insurance utilization and associated
factors in Bassona Worena District, North Shoa Zone, Ethiopia: a com-
munity-based cross-sectional study. BMC Health Serv Res. 2022;22:1–8.
https://doi.org/10.1186/s12913-022-08794-6.

77. Terefe B, Alemu TG, Techane MA, Wubneh CA, Assimamaw NT, Belay GM,
et al. Spatial distribution and associated factors of community based
health insurance coverage in Ethiopia: further analysis of Ethiopian
demography and health survey, 2019. BMC Public Health. 2022;22:1–13.
https://doi.org/10.1186/s12889-022-13950-y.

78. Guha S, Chandra H. Measuring and Mapping Disaggregate Level Dispari-
ties in Food Consumption and Nutritional Status via Multivariate Small
Area Modelling. Soc Indic Res. 2021;154:623–46. https://doi.org/10.1007/
s11205-020-02573-8.

79. Hossain MJ, Das S, Chandra H, Islam MA. Disaggregate level estimates
and spatial mapping of food insecurity in Bangladesh by linking survey
and census data. PLoS ONE. 2020;15:1–16. https://doi.org/10.1371/journ
al.pone.0230906.

80. Molina I. Disaggregating data in household surveys: using small area
estimation methodologies. Santiago: Statistics series; 2022. https://repos
itorio.cepal.org/bitstreams/e4e38563-d196-4f01-b903-5f64faa6866b/
download.

Publisher's Note