

DISTRICT LEVEL CHILD NUTRITION STATUS IN BANGLADESH: AN APPLICATION OF SMALL AREA ESTIMATION METHOD

MOSSAMET KAMRUN NESA

*School of Mathematics and Applied Statistics, University of Wollongong
Wollongong 2500, NSW, Australia and*

*Department of Statistics, Shahjalal University of Science and Technology
Sylhet 3100, Bangladesh*

Email: mkn663@uowmail.edu.au

SUMMARY

National level indicators of child undernutrition often hide the real scenario across a country. In order to construct a child nutrition map, accurate estimates of undernutrition are required at very small spatial scales, typically the administrative units of a country or a region within a country. Although comprehensive data on child nutrition are collected in national surveys, the small scale estimates cannot be calculated using the standard estimation methods employed in national surveys, since such methods are designed to produce national or regional level estimates, and assume large samples. Small area estimation method has been widely used to find such micro-level estimates. Due to lack of unit level data, area level small area estimation methods (e.g., Fay-Herriot method) are widely used to calculate small-scale estimates. In Bangladesh, a few works have been done to estimate district level child nutrition status. The Bangladesh Demographic Health Survey covers all districts but district wise sample sizes are very small to get consistent estimates. In this paper, Fay-Herriot Model has been developed to calculate district wise estimates with efficient mean squared error. The Bangladesh Demographic Health Survey 2011 and Population Census 2011 are utilized for this study.

Keywords and phrases: Small Area Estimation; Bivariate Fay-Herriot Model; Univariate Fay-Herriot Model; Mean Squared Error.

1 Introduction

Small Area Estimation (SAE) techniques have received much attention in recent times due to increasing demand for micro-level official statistics. Because of small domain-specific sample sizes (even zero size), direct estimation may lead to estimates with large sampling variability (Rao and Molina, 2015). The basic idea of SAE method is to link the variable of interest with auxiliary information (e.g., Census and Administrative data) in a random

effects model. SAE method is broadly classified into two methods - unit level SAE and area level SAE. When unit level survey data is not available, area level SAE is utilized for small area estimates. One of the basic area level models is the Fay-Herriot (FH) model (Fay and Herriot, 1979) which relates small area direct survey estimates to area specific covariates. When multiple dependent variables are considered correlated, multivariate Fay-Herriot model may produce better results than univariate FH model (Rao and Molina, 2015), but these models have received relatively little attention. In practice, however, univariate FH (UFH) models are most common. They are either applied to a single variable, or separately to each variable of interest. Arima et al. (2017) points out two benefits of multivariate FH (MFH) models: they may result in lower prediction MSEs, and they are more suitable when functions of different parameters are estimated for each small area (for example annual change in an indicator).

The MFH estimator has seen increasing attention in recent years. Datta et al. (1998) compare the precision of small area estimators obtained from UFH and MFH models, showing that more precise small area estimators are obtainable using MFH models rather than separate UFH models. González-Manteiga et al. (2008) study MFH models with a common scalar random effect for the two variables. MFH estimators have lower MSEs than UFH estimators for their simulated data. González-Manteiga et al. (2008) also find that a bootstrap MSE estimator does better than an analytic estimator of the MSE. Benavent and Morales (2016) relax the assumption of a common scalar random effect, and calculate MFH estimators of provincial poverty rates using data from the 2005 and 2006 Spanish living condition surveys. Monte-Carlo simulations show that MFH estimators have lower MSEs than UFH estimators. Porter et al. (2015) develop models allowing for both spatial dependence and multivariate outcomes, finding that a generalized conditional auto-regressive model gives excellent results. A simulation study based on the Missouri county lattice structure finds gains in MSE of around 10% when cross-correlations are weaker, and around 30-70% when they are stronger, compared to a non-spatial multivariate model based on Datta et al. (1998).

The child nutrition status in Bangladesh are based only on surveys which produce national or regional level estimates. However, national level indicators of child undernutrition often hide the real spatial distribution across the country. The standard direct estimation methods cannot be used due to small sample size for a significant number of administrative units such as districts. As a result, district level child nutrition indicators have not previously been calculated. The Bangladesh Demographic and Health Survey (BDHS) 2011 covers all districts but only includes a small sample of children in each district thereby making consistent estimates of malnourished children are difficult to obtain. The purpose of this study is to develop an area level Fay-Herriot models to calculate district level estimates with their efficiency and to identify the conditions under which bivariate Fay-Herriot (BFH) improves prediction. Comparison between univariate and bivariate FH models has also been done. BFH and UFH estimators and the MSE estimator are applied to data from the Bangladesh Demographic and Health Survey 2011 and Bangladesh Population and Housing Census 2011.

The paper is organised as follows. Section 2 outlines the model structure and the MSE of the UFH and BFH estimators. The relative performance of the UFH and BFH estimators is measured by the relative efficiency of these MSEs. In Section 3, some special cases regarding the relative efficiencies are discussed. Section 4 describes a numerical study including a traditional descriptive summary of the results as well analysis of variance approach. Section 5 describes the empirical study and Section 6 summarises the findings.

2 Approximate Relative Efficiency of Univariate and Bivariate Fay-Herriot Estimators

The purpose of this section is to make simple approximate comparisons of the MSEs of the BFH and UFH estimators assuming known model parameters including variance components and regression coefficients. Of course in practice these parameters are unknown and this must be reflected in MSE estimation when FH estimators are calculated in practice.

2.1 Univariate Fay-Herriot Estimation

Let $\theta_d = g(\bar{Y}_d)$, where \bar{Y}_d is the small area mean for the d^{th} area. Suppose θ_d is related to area-specific auxiliary data $\mathbf{Z}_d = (Z_{d1}, Z_{d2}, \dots, Z_{dp})'$ through the model:

$$\theta_d = \mathbf{Z}'_d \boldsymbol{\beta} + v_d, \quad d = 1, 2, \dots, D$$

and the direct estimators of θ_d follow a sampling model,

$$\hat{\theta}_d = \theta_d + e_d, \quad d = 1, 2, \dots, D.$$

Combining the above two assumptions, the UFH model is

$$\hat{\theta}_d = \mathbf{Z}'_d \boldsymbol{\beta} + v_d + e_d, \quad d = 1, 2, \dots, D, \quad (2.1)$$

where D is the number of areas, \mathbf{Z}_d is a $p \times 1$ vector of auxiliary variables, and $\boldsymbol{\beta}$ is a $p \times 1$ vector of regression coefficients. Furthermore, v_d and e_d are area specific random effects and sampling errors respectively, assumed to be independent over $d = 1, \dots, D$, with $v_d \sim N(0, \sigma_v^2)$ and $e_d \sim N(0, \psi_d)$. The disturbance terms v_d and e_d are assumed to be independent from each other. Their respective variances, σ_v^2 and ψ_d , are the area-specific random effect variance and the design-based sampling variance. The BLUP estimator of $\hat{\theta}_d$, denoted by $\hat{\theta}_d^*$, for known $\boldsymbol{\beta}$, is

$$\hat{\theta}_d^* = \frac{\sigma_v^2}{\sigma_v^2 + \psi_d} \hat{\theta}_d + \frac{\psi_d}{\sigma_v^2 + \psi_d} \mathbf{Z}'_d \boldsymbol{\beta}$$

and its MSE is

$$\text{MSE}(\hat{\theta}_d^*) = \frac{\sigma_v^2 \psi_d}{\sigma_v^2 + \psi_d} \quad (2.2)$$

The log-likelihood function of the data $\{\hat{\theta}_d\}$ is

$$\log L = -\frac{D}{2} \log(2\pi) - \frac{1}{2} \sum_{d=1}^D \log(\sigma_v^2 + \psi_d) - \frac{1}{2} \sum_{d=1}^D (\hat{\theta}_d - \mathbf{Z}'_d \boldsymbol{\beta})^2 (\sigma_v^2 + \psi_d)^{-1}$$

2.2 Bivariate Fay-Herriot Estimation

Let $\boldsymbol{\theta}_d$ be the 2×1 vector of the small area statistics in area d for the two target variables. The BFH model assumes

$$\boldsymbol{\theta}_d = \boldsymbol{\beta}' \mathbf{Z}_d + \mathbf{v}_d, \quad d = 1, 2, \dots, D, \quad (2.3)$$

where \mathbf{Z}_d is a $p \times 1$ vector of auxiliary variables, $\boldsymbol{\beta}$ is a $(p \times 2)$ matrix of regression coefficients and \mathbf{v}_d are the area specific random effects with mean zero and variance-covariance matrix $\boldsymbol{\Sigma}_v$ given by:

$$\boldsymbol{\Sigma}_v = \begin{bmatrix} \Sigma_{v11} & \Sigma_{v12} \\ \Sigma_{v12} & \Sigma_{v22} \end{bmatrix}$$

The direct estimator of $\boldsymbol{\theta}_d$ follows a sampling model,

$$\hat{\boldsymbol{\theta}}_d = \boldsymbol{\theta}_d + \mathbf{e}_d, \quad d = 1, 2, \dots, D, \quad (2.4)$$

where $\hat{\boldsymbol{\theta}}_d = (\hat{\theta}_{d1}, \hat{\theta}_{d2})'$ are the direct estimates of $\boldsymbol{\theta}_d = (\theta_{d1}, \theta_{d2})'$. Let $\mathbf{e}_d = (e_{d1}, e_{d2})'$ be the sampling error with mean zero and variance-covariance matrix:

$$\boldsymbol{\Psi}_d = \begin{bmatrix} \Psi_{d11} & \Psi_{d12} \\ \Psi_{d12} & \Psi_{d22} \end{bmatrix}$$

Combining (2.3) and (2.4), the BFH model is

$$\hat{\boldsymbol{\theta}}_d = \boldsymbol{\beta}' \mathbf{Z}_d + \mathbf{v}_d + \mathbf{e}_d, \quad d = 1, 2, \dots, D \quad (2.5)$$

The BLUP estimator of $\hat{\boldsymbol{\theta}}_d$ for known $\boldsymbol{\beta}$ is

$$\left. \begin{aligned} \hat{\theta}_{d1}^* &= \boldsymbol{\beta}'_1 \mathbf{Z}_d + (\boldsymbol{\Sigma}_v + \boldsymbol{\Psi}_d)^{-1} \boldsymbol{\Sigma}_{v1} (\hat{\boldsymbol{\theta}}_d - \boldsymbol{\beta}' \mathbf{Z}_d) \\ \hat{\theta}_{d2}^* &= \boldsymbol{\beta}'_2 \mathbf{Z}_d + (\boldsymbol{\Sigma}_v + \boldsymbol{\Psi}_d)^{-1} \boldsymbol{\Sigma}_{v2} (\hat{\boldsymbol{\theta}}_d - \boldsymbol{\beta}' \mathbf{Z}_d) \end{aligned} \right\}, \quad (2.6)$$

where $\boldsymbol{\Sigma}_{v1} = (\Sigma_{v11}, \Sigma_{v12})'$ and $\boldsymbol{\Sigma}_{v2} = (\Sigma_{v12}, \Sigma_{v22})'$. The BLUP estimator (2.6) can be written as

$$\hat{\boldsymbol{\theta}}_d^* = \boldsymbol{\beta}' \mathbf{Z}_d + \boldsymbol{\Sigma}'_v (\boldsymbol{\Sigma}_v + \boldsymbol{\Psi}_d)^{-1} (\hat{\boldsymbol{\theta}}_d - \boldsymbol{\beta}' \mathbf{Z}_d) \quad (2.7)$$

See Rao and Molina (2015, p.236) who derive the BLUP when $\boldsymbol{\beta}$ is unknown.

The MSE of the BLUP estimator $\hat{\theta}_d^*$ is

$$\left. \begin{aligned} \text{MSE}(\hat{\theta}_{d1}^*) &= \Sigma_{v11} - \Sigma'_{v1}(\Sigma_v + \Psi_d)^{-1}\Sigma_{v1} \\ \text{MSE}(\hat{\theta}_{d2}^*) &= \Sigma_{v22} - \Sigma'_{v2}(\Sigma_v + \Psi_d)^{-1}\Sigma_{v2} \end{aligned} \right\} \quad (2.8)$$

which can be expressed as

$$\text{MSE}(\hat{\theta}_d^*) = \Sigma_v - \Sigma_v(\Sigma_v + \Psi_d)^{-1}\Sigma_v$$

The log likelihood function is

$$\begin{aligned} \log L &= -\frac{Dr}{2}\log(2\pi) - \frac{1}{2}\sum_{d=1}^D \log|\Sigma_v + \psi_d| \\ &\quad - \frac{1}{2}\sum_{d=1}^D (\hat{\theta}_d - \beta'Z_d)'(\Sigma_v + \psi_d)^{-1}(\hat{\theta}_d - \beta'Z_d) \end{aligned}$$

2.3 Relative Efficiency of UFH and BFH estimators

From (2.2) and (2.8), the approximate relative efficiencies of the bivariate BLUP estimator compared to the univariate BLUP estimator are

$$\left. \begin{aligned} \text{RE}_{d1} &= \frac{\Sigma_{v11} - \Sigma'_{v1}(\Sigma_v + \Psi_d)^{-1}\Sigma_{v1}}{\Sigma_{v11} - \Sigma_{v11}^2(\Sigma_{v11} + \Psi_{d11})^{-1}} \\ \text{RE}_{d2} &= \frac{\Sigma_{v22} - \Sigma'_{v2}(\Sigma_v + \Psi_d)^{-1}\Sigma_{v2}}{\Sigma_{v22} - \Sigma_{v22}^2(\Sigma_{v22} + \Psi_{d22})^{-1}} \end{aligned} \right\} \quad (2.9)$$

3 Special Cases

The approximate relative efficiency of BFH to UFH estimators will now be examined for some special cases where simple analytic results are possible. These special scenarios will provide insights into when BFH would be worthwhile in practice.

Case 1: No Sampling Error

When sampling errors are zero i.e. when $\Psi_d = \mathbf{0}$, the MSEs of both the UFH and BFH estimators in (2.2) and (2.8) become 0. So if sampling errors are small, then we can expect the UFH and BFH to have similar MSEs.

Case 2: Uncorrelated Sampling Errors and Uncorrelated Random Effects

The relative efficiencies in (2.9) become one when the correlation between the sampling errors (ρ_e) for each variable and the correlation between the random effects (ρ_v) for each

variable are both zero. This special case is an obvious one to consider and discussed in Rao and Molina (2015, p.236) where the authors state that the multivariate model leads to more efficient estimators compared to those based on univariate models when correlation exists between the response variables.

Case 3: Proportional Variance-Covariance Matrices of Sampling Errors and Random Effects

It is also straightforward to show that there is no gain when the variance-covariance matrices of the sampling errors and the random effects are proportional. That is, when $\Psi_d = k\Sigma_v$, where k is any scalar constant. In this case, $RE_{d1} = RE_{d2} = 1$. So even if the correlation coefficients ρ_e and ρ_v are non-zero, if the variance-covariance matrices of the sampling errors and the random effects are proportional, then no gains are obtained. Birrell et al. (2011) proved a similar result in the different context of multivariate time series models for prediction.

Case 1 and 2 are unsurprising, but Case 3 gives a new insight into why the gains from bivariate modelling may often be small. The result means that even if the cross-variable correlations are high, there is still no gain from bivariate modelling, unless either the correlation between the sampling errors and the correlation between the random effects differ, or the ratio of the sampling variance and the random effects variance differs from variable to variable.

Of course in practice none of the three cases will hold exactly, so a numerical study will be conducted in the next section to evaluate the relative efficiencies of UFH and BFH estimators over a wide range of situations.

4 Numerical Study

Birrell et al. (2011) show that the gains are attainable for a multivariate time series model when the ratio of variances of one sub-series is very different from the ratio of variances of another sub-series. Although the context is different, their results suggest a focus on the differences between the variance-covariance matrices of sampling errors and the random area effects, and also on the variances of the first variable relative to the second variable.

In this section, a numerical study is performed to investigate the relative efficiency of the bivariate FH estimator to the univariate FH estimator given known parameters. The settings of the parameters are discussed in the following section.

4.1 Setting the Parameters

We firstly rewrite the variance-covariance matrices of sampling errors and random effects to focus on the correlation coefficients.

Let ρ_v and ρ_e be the correlation of the random effects and the sampling errors respectively, which are defined as:

$$\rho_v = \frac{\Sigma_{v12}}{\sqrt{\Sigma_{v11}\Sigma_{v22}}} \text{ and } \rho_e = \frac{\Psi_{d12}}{\sqrt{\Psi_{d11}\Psi_{d22}}}.$$

The variance-covariance matrices Σ_v and Ψ_d can be expressed in terms of ρ_v and ρ_e :

$$\Sigma_v = \begin{bmatrix} \Sigma_{v11} & \rho_v \sqrt{\Sigma_{v11}\Sigma_{v22}} \\ \rho_v \sqrt{\Sigma_{v11}\Sigma_{v22}} & \Sigma_{v22} \end{bmatrix}, \quad \Psi_d = \begin{bmatrix} \Psi_{d11} & \rho_e \sqrt{\Psi_{d11}\Psi_{d22}} \\ \rho_e \sqrt{\Psi_{d11}\Psi_{d22}} & \Psi_{d22} \end{bmatrix}.$$

Let r_v be the ratio of the variances of random effects Σ_v and let r_s be the ratio of the variances of sampling errors Ψ_d :

$$r_v = \frac{\Sigma_{v22}}{\Sigma_{v11}}, \quad r_s = \frac{\Psi_{d22}}{\Psi_{d11}}.$$

The ratio of these ratios is defined by

$$R = \frac{r_s}{r_v} = \frac{\Psi_{d22}\Sigma_{v11}}{\Psi_{d11}\Sigma_{v22}}. \quad (4.1)$$

For this study, the parameters to be varied are Σ_{v11} , ρ_v , r_v , ρ_e and r_s . Without loss of generality, Ψ_{d11} is set to 1 and the first variable has sampling variance less than or equal to that of second variable. The values of the parameters considered in the numerical study are shown in Table 1.

The total number of combinations for these parameter values is $9 \times 6 \times 9 \times 6 \times 5 = 14580$. Given the range of values set for r_v and r_s , the value for R ranges from 0.1 to 100. For each combination of the parameters, the approximate MSEs of the UFH and BFH estimators are calculated for each variable using (2.2) and (2.8) respectively. The approximate relative efficiencies of BFH to UFH estimators are calculated using (2.9). The values of the relative efficiencies are calculated by substituting the parameter settings directly into their algebraic expressions, since these approximate expressions are a function of the model parameters only. The numerical study includes nearly 15,000 combinations, so the results need to be analysed in a structured way. Two ways of summarizing the study are used in the next two subsections: a descriptive analysis and a analysis of variance (ANOVA) approach.

4.2 Descriptive Summary of Numerical Study

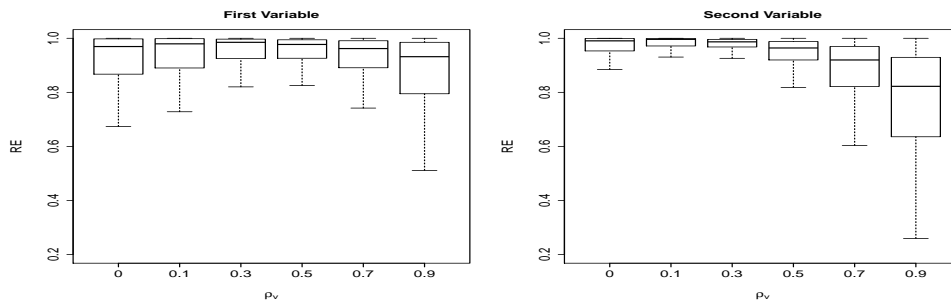
Firstly, we consider the effects of varying the correlation parameters, and then the effects of varying the values of the ratio R.

4.2.1 The Effects of the Correlations ρ_v and ρ_e on Relative Efficiency

The boxplots in Figure 1 summarise the results for both variables for each value of ρ_v . Each observation represents the relative efficiency of BFH to UFH from one of the scenarios in

Table 1: Values of parameters

Parameter	Possible values	No. values
Σ_{v11}	{0.1, 0.2, 0.3, 0.6, 1.0, 1.5, 3.0, 5.0, 10}	9
ρ_v	{0, 0.1, 0.3, 0.5, 0.7, 0.9}	6
r_v	{0.1, 0.2, 0.3, 0.6, 1.0, 1.5, 3.0, 5.0, 10}	9
ρ_e	{0, 0.1, 0.3, 0.5, 0.7, 0.9}	6
r_s	{1.0, 1.5, 3.0, 5.0, 10}	5

Figure 1: Results of relative efficiency, calculated using (2.9), for the first and second variables with different values of ρ_v .

the numerical study. Note that a low relative efficiency means a large gain from using BFH estimators. For the first variable, the potential gains from bivariate modelling are greater (lower relative efficiency) when ρ_v is large or small. For the second variable, potential gains are greater when ρ_v is large. For example, when $\rho_v = 0.9$, the relative efficiencies lie between 0.5 and 1 for the first variable and 0.25 and 1 for the second variable. However, as the range still reaches to 1 (or very nearly so), it is clear that a high correlation does not guarantee lower relative efficiency, as also found for Case 3 in Section 3. This is presumably because other parameters are also influencing the relative efficiencies.

The boxplots shown in Figure 2 summarise the relative efficiencies for different values of ρ_e . Figure 2 reveals a similar pattern over ρ_e to the one in Figure 1, except that the patterns of variable 1 and 2 are swapped. Here for $\rho_e = 0.9$, the relative efficiencies range from 0.25 to 1 for the first variable and from 0.5 to 1 for the second variable. This means that the relative efficiencies are not fully determined by ρ_e and that other parameters also play a role.

Overall, it can be said that the relative efficiencies of BFH estimators are affected by the correlations of the sampling errors and the random effects of the two variables. It was shown in Section 3 that when the correlations of sampling errors and random effects are both zero, no gain is achievable for both variables. At first thought, it might be expected

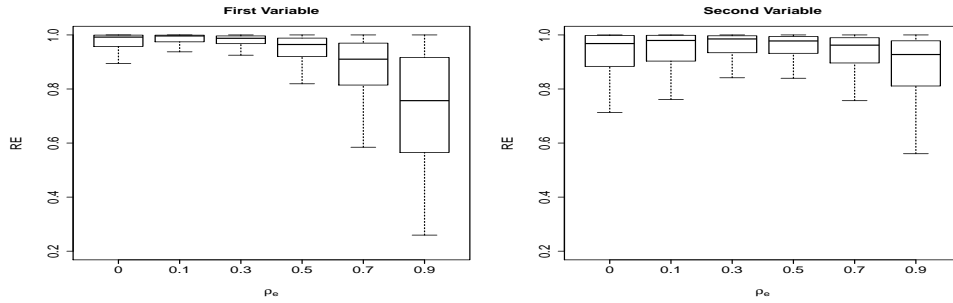


Figure 2: Results of relative efficiency, calculated using (2.9), for the first and second variables with different values of ρ_e .

that as either of these correlations increase, the relative efficiency would also improve. It has been shown here, that while that is true for most cases, it is definitely not true for all cases, as many of the boxes in Figure 1 and 2 are wide, showing that other parameters also affect the relative efficiency.

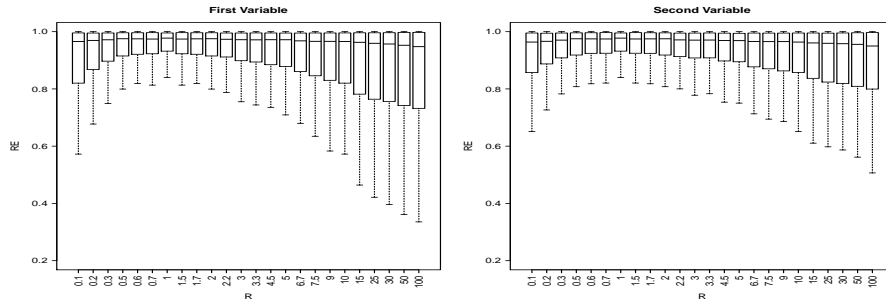


Figure 3: Results of relative efficiency, calculated using (2.9), against ratio, $R = r_s/r_v$.

4.2.2 The Effects of the Ratio of Ratio of Variances (R) on Relative Efficiency

The relative efficiency is plotted against the value of $R = (r_s/r_v)$ in Figure 3. The two panels clearly indicate that when $R=1$, BFH and UFH perform very similarly, while the relative efficiency increases for both variables when R moves away from one in either direction. For example, low relative efficiencies are possible when R is very large say 100 and R is small say 0.1, for both variables. However, the maximum relative efficiency is very close to 1 for all values of R , so it is clear that R alone cannot guarantee a gain from bivariate modelling.

Since the relative efficiency is also affected by both ρ_v and ρ_e (as shown 4.2.1) an investigation is now conducted to vary R for fixed ρ_v and ρ_e . Figure 4 plots the relative efficiency against R , fixing $\rho_v = \rho_e = 0.7$. The first noticeable feature is that when $R=1$ the relative efficiency becomes exactly one. Figure 4 also shows that the relative efficiency decreases

when R becomes further away from one in both direction. For example, better gains from BFH are obtainable when $R=100$ or $R=0.1$. Overall, it can be said that the designs which give the lowest relative efficiency have high ρ_v , high ρ_e and high R .

Figures 4-6 show the relative efficiencies vs R for different values of ρ_v and ρ_e . Figure 4 presents the results when $\rho_v = \rho_e = 0.7$. Figure 5 shows the case that greater gains of BFH over UFH are possible when $\rho_v = 0.7$ and $\rho_e = 0.1$, provided R is also large. The plots in Figure 5 show the opposite pattern for the two variables. The lower the R the higher the relative efficiency for the first variable while the reverse is true for the second variable. Figure 6 gives the result for $\rho_v = 0.1$ and $\rho_e = 0.7$. The picture flips around the two variables compared to Figure 5. Now useful gains are possible for the first variable as R increases and for the second variable as R decreases.

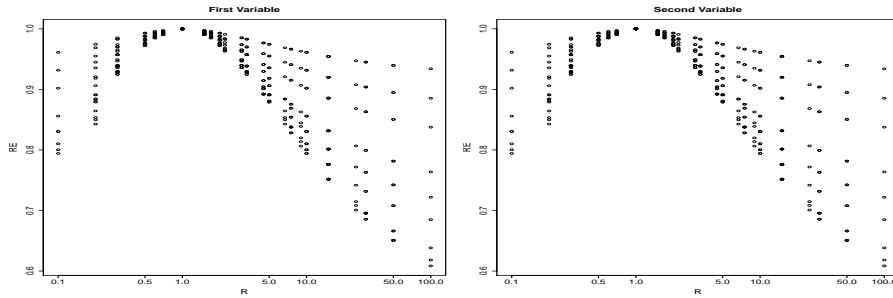


Figure 4: Results of relative efficiency, calculated using (2.9), against ratio, $R = r_s/r_v$, with $\rho_v = \rho_e = 0.7$.

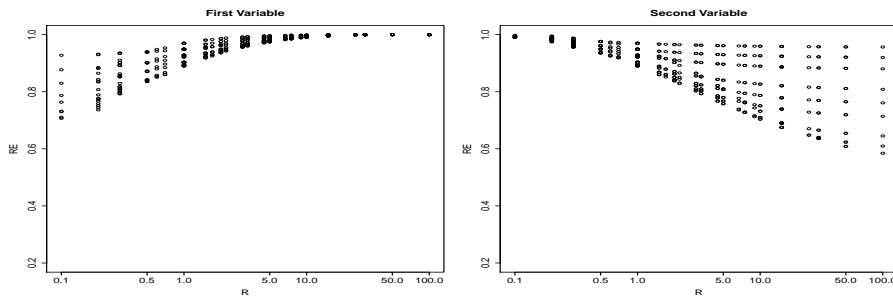


Figure 5: Results of relative efficiency, calculated using (2.9), against ratio, $R = r_s/r_v$, with $\rho_v = 0.7$, $\rho_e = 0.1$.

In summary, gains are achievable when the relative magnitude of the sampling variances differ from the relative magnitude of the random effects (i.e. R is not close to 1) and gains are also affected by the values of ρ_v and ρ_e .

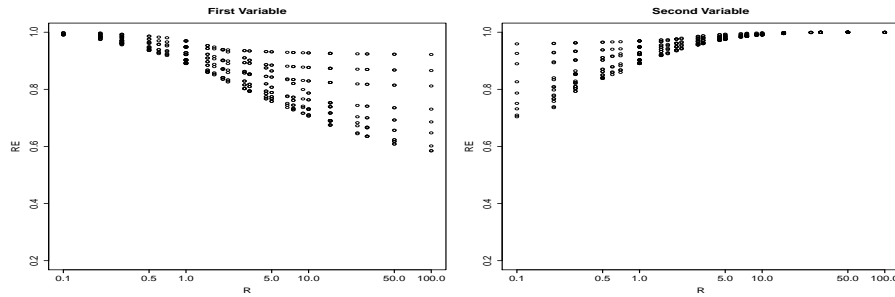


Figure 6: Results of relative efficiency, calculated using (2.9), against ratio, $R = r_s/r_v$, with $\rho_v = 0.1, \rho_e = 0.7$.

4.3 Formal Analysis of Numerical Study: Analysis of Variance

In Section 4.2, it was shown that the RE of BFH estimators depends on the different combinations of parameters. The numerical study can be regarded as a factorial experiment with RE as the dependent variable and the parameters with the different values shown in Table 1 as factors. The results can be formally analysed by a linear regression model of RE given the factors. See Bradley et al. (2015) for an example of analyzing the results of a simulation study using a formal ANOVA. Two separate regression models have been fitted for the RE of the first and second variables, in order to assess the main effect of each factor as well as two-way interaction effects, on the RE.

Figure 7 shows that the residuals of the regression model are approximately normal which makes the response variables suitable for a linear regression, unlike Bradley et al. (2015) where a fourth-root transformation was needed to achieve approximate normality.

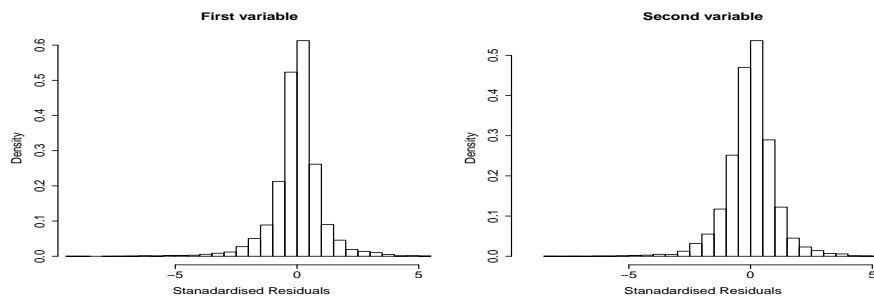


Figure 7: Histogram of standardized residuals.

Tables 2 and 3 show the results of ANOVA for both variables respectively where ‘Source’ indicates the source of variability used in the ANOVA, ‘DF’ denotes the degrees of freedom, ‘SS’ denotes the sum of squares, ‘MS’ denotes mean squared error, and F denotes the F-statistic associated with the source.

Table 2: Analysis of variance (ANOVA) on relative efficiency of the first variable, for up to two-way interaction

Source	DF	SS	MS	F value	R-squared
ρ_e	5	105.3	21.06	11511.2	0.9060
ρ_v	5	5.480	1.095	598.731	
R	5	16.05	3.209	1754.14	
Σ_{v11}	8	7.240	0.904	494.409	
$\rho_e \times \rho_v$	25	12.10	0.484	264.460	
$\rho_e \times R$	25	77.19	3.088	1687.70	
$\rho_e \times \Sigma_{v11}$	40	4.750	0.119	64.8710	
$\rho_v \times R$	25	21.53	0.861	470.716	
$\rho_v \times \Sigma_{v11}$	40	0.570	0.014	7.79800	
$R \times \Sigma_{v11}$	40	5.030	0.126	68.7200	
Residuals	14361	26.27	0.002		
Total	14579	281.5			

Table 2 presents the results of an ANOVA for the first variable with up to two-way interaction between the parameters. The ANOVA indicates that the explanatory variables including interactions explain 91% of the total variation of the RE for the first variable. In this case, the most variation (37%) is explained by the correlation of random effects, ρ_e , followed by the interaction of ρ_e with R (27%). The interaction between ρ_e and ρ_v explains 4% of the total variation.

The ANOVA for the second variable with up to two-way interactions is presented in Table 3. Here, all main effects and interactions explain 88% of the total variation. The highest variation (32%) is explained by the correlation of random effects, ρ_v . The interaction between ρ_v and R and the interaction between ρ_e and ρ_v explain 21% and 4% variation respectively.

Tables 2 and 3 show that most of the variation of the RE of the first variable is explained by ρ_e and of the second variable is explained by ρ_v . It can be concluded that ρ_e might be the most important parameter for the first variable and ρ_v might be the most important parameter for the second variable. The exploratory analysis described in Subsection 4.2 also revealed the importance of ρ_e and ρ_v for getting lower REs of the first and second variable respectively.

Table 3: Analysis of variance (ANOVA) on relative efficiency of the second variable, for up to two-way interactions

Source	DF	SS	MS	F value	R-squared
ρ_e	5	7.060	1.412	777.45	0.8836
ρ_v	5	72.55	14.51	7989.9	
R	5	9.720	1.945	1070.8	
Σ_{v11}	8	7.430	0.929	511.33	
$\rho_e \times \rho_v$	25	7.680	0.307	169.19	
$\rho_e \times R$	25	13.02	0.521	286.72	
$\rho_e \times \Sigma_{v11}$	40	3.650	0.091	50.210	
$\rho_v \times R$	25	47.04	1.882	1036.1	
$\rho_v \times \Sigma_{v11}$	40	18.32	0.458	252.18	
$R \times \Sigma_{v11}$	40	11.57	0.289	159.30	
Residuals	14361	26.08	0.002		
Total	14579	224.1			

5 Empirical Study

5.1 Materials and Methods

Small area models are developed using the data set of BDHS 2011 and Bangladesh Population and Housing Census 2011. The BDHS covers 600 communities (Primary Sampling Unit) across 396 sub-districts, comprising 64 districts in 7 divisions. Two anthropometric standard indices *Height-for-age* and *Weight-for-age* (Z-score) are used to calculate the proportion of stunted and underweight children at district level. Three district level statistics - *Proportion of children under 5 years*, *Proportion of household size with ≤ 4 members* and *Average household size* are considered as explanatory variables in the FH models. The BFH model (2.5) as well as the UFH model (2.1) will be used to estimate area prevalences of the indicators.

5.2 Results of Empirical study

Table 4 presents summary statistics of *Stunting* prevalence and *Underweight* prevalence across the districts. The mean value of *Stunting* and prevalence of *Underweight* are 40.16 and 35.73 respectively. The interquartile range is narrow for both variables.

The maximum likelihood (ML) estimate of Σ_v in the BFH model (2.5) for *Stunting* and

Table 4: Five number summary and mean of direct estimates of the prevalence of *Stunting* and prevalence of *Underweight* in percentage across the district

	Minimum	Lower Quartile	Median	Mean	Upper Quartile	Maximum
<i>Stunting</i>	21.21	34.14	38.80	40.16	45.53	68.75
<i>Underweight</i>	15.79	29.68	34.23	35.73	41.83	62.50

Underweight is:

$$\hat{\Sigma}_{\mathbf{v}} = \begin{bmatrix} 0.00001 & 0.0000099 \\ 0.0000099 & 0.00001 \end{bmatrix} \quad (5.1)$$

The estimated variances for *Stunting* and *Underweight* obtained from the corresponding fitted BFH model are same which is 0.00001. The variance component of *Stunting* is similar to *Underweight* which indicates that there is no much more between area variation in both variables.

Figure 8 presents the comparison of the direct, univariate and bivariate models against sample sizes. In Bangladesh, the national level proportion of stunted and underweight children aged under 5 are estimated to be 41.0 percent and 36.0 percent respectively (NIPORT et al., 2013), while district level estimates obtained from SAE models vary across the districts (Panels (a) and (b)). Direct estimates are highly variable for those districts with small (≤ 100) sample size. However, for districts with large (> 100) sample size the direct estimates are found stable and very close to the FH estimates (Panels (a) and (b)). The above statements are supported by the pattern of root mean squared errors (RMSE) or coefficient of variations (CV) of the estimates. The CVs of the direct estimates are almost double those of FH estimates when the samples are small, however the differences reduce with the increasing sample size (see Panels (c) and (d)). For FH, CV's are found below 15 percent for all sample sizes. Panels (e) and (f) show that the direct estimates are randomly distributed around the FH estimates and they are close to the diagonal line indicating no evidence of bias.

Although the differences in the UFH and BFH estimates are negligible (Panels (a) and (b)), the efficiency significantly varies with sample size. Panels (g) and (h) exhibit smaller RMSE of BFH estimates than those of UFH estimates for both stunting and underweight. The gain in efficiency (GE=ratio of difference between the RMSEs of UFH and BFH to RMSE of UFH) of BFH estimates decreases with sample size (Panels (i) and (j)).

District level maps of Bangladesh for stunting and underweight are shown in Figure 9. The first map (Panel a) shows that the districts of north-western (Rajshahi region) and central (Dhaka region) had lower prevalence of stunting while the districts of northern (Mymensingh region) and hill tracts area in south-eastern part had comparatively higher prevalence of stunting. The prevalence of underweight (Panel b) also followed the pattern of stunting.

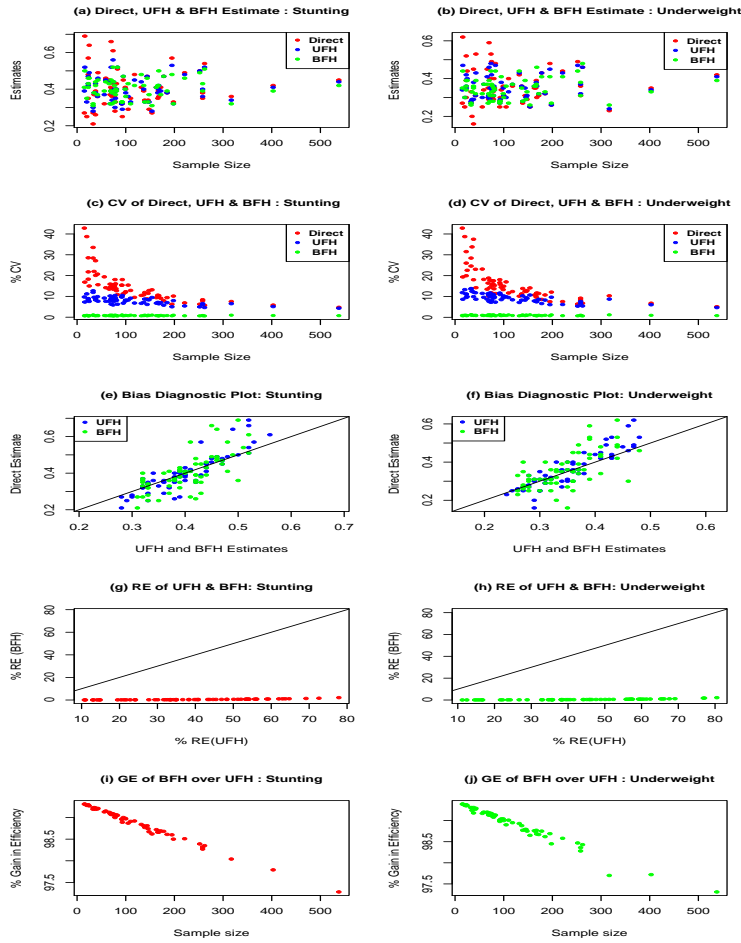


Figure 8: Comparison of Direct, UFH and BFH estimators

6 Summary

The efficiencies of BFH estimators and UFH estimators have been compared in a numerical study and an empirical study in this paper. The exploratory results of the numerical study reveal that the REs of both variables are affected by random effects correlation, sampling error correlation, and the ratio of the variances of sampling errors and random effects. Lower RE (good gains) were obtainable with higher values of random effects correlation and sampling error correlation. The results of numerical study also showed the influence of R on the RE. There was no gain when the R is equal to one.

BFH and UFH estimators were also applied to small area prevalences of Bangladesh Demographic Health Survey 2011. Small differences were observed between UFH and BFH

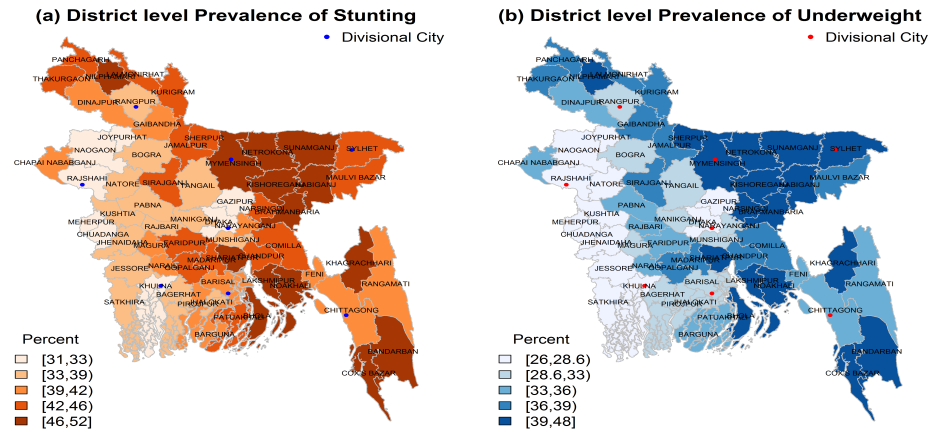


Figure 9: Spatial distribution of the estimated prevalence of stunting and underweight among under-5 children at district level using SAE (BFH) estimators

estimates for the considered two indicator variables *Stunting* and *Underweight*. Both the UFH and BFH estimators perform better showing lower RMSEs than the direct estimates for small sample size. However, BFH performs better than the UFH when the samples are small.

The generated maps show clear inequality among the districts. The map of both child malnutrition indicators might help the policy makers and the national and international organizations in aid distribution for the improvement of child vulnerability.

There is still much scope left for further numerical and empirical research on small area estimation. In small area estimation, generally, the random effects and sampling errors are considered to be independently and identically normally distributed with mean zero and constant variance. However, in practice it is not often true. The distribution of random area effects can be non-normal, temporally autocorrelated, and spatially-temporally autocorrelated. A recent review of small area studies Marhuenda et al. (2013) suggests a univariate spatio-temporal small area approach which accounts for the spatial autocorrelation pattern of the random effects. A multivariate Fay-Herriot model incorporating a multivariate version of spatio-temporal small area approach could also be considered in future research.

References

- Arima, S., Bell, W. R., Datta, G. S., Franco, C., and Liseo, B. (2017), “Multivariate Fay-Herriot Bayesian estimation of small area means under functional measurement error,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180, 1191–1209.

- Benavent, R. and Morales, D. (2016), “Multivariate Fay-Herriot models for small area estimation,” *Computational Statistics and Data Analysis*, 94, 372–390.
- Birrell, C. L., Steel, D. G., and Lin, Y.-X. (2011), “Seasonal adjustment of an aggregate series using univariate and multivariate basic structural models,” *Journal of Statistical Theory and Practice*, 5, 179–205.
- Bradley, J. R., Cressie, N., and Shi, T. (2015), “Comparing and selecting spatial predictors using local criteria,” *Test*, 24, 1–28.
- Datta, G. S., Day, B., and Maiti, T. (1998), “Multivariate Bayesian small area estimation: an application to survey and satellite data,” *Sankhyā: The Indian Journal of Statistics, Series A*, 60, 344–362.
- Fay, R. E. and Herriot, R. A. (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.
- González-Manteiga, W., Lombardía, M. J., Molina, I., Morales, D., and Santamaría, L. (2008), “Analytic and bootstrap approximations of prediction errors under a multivariate Fay–Herriot model,” *Computational Statistics and Data Analysis*, 52, 5242–5252.
- Marhuenda, Y., Molina, I., and Morales, D. (2013), “Small area estimation with spatio-temporal Fay–Herriot models,” *Computational Statistics and Data Analysis*, 58, 308–325.
- NIPORT, Mitra and Associates, and International, I. (2013), *Bangladesh Demographic and Health Survey 2011*, Dhaka, Bangladesh and Rockville, Maryland, U.S.A.: National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ICF International.
- Porter, A. T., Wikle, C. K., and Holan, S. H. (2015), “Small area estimation via multivariate Fay-Herriot models with latent spatial dependence,” *Australian and New Zealand Journal of Statistics*, 57, 15–29.
- Rao, J. N. and Molina, I. (2015), *Small Area Estimation*, Hoboken, New Jersey: John Wiley & Sons.

Received: August 30, 2018

Accepted: May 19, 2019