Journal of the Korean Data & Information Science Society
2013, **24**(4), 885–893

http://dx.doi.org/10.7465/jkdi.2013.24.4.885
한국데이터정보과학회지

Bayesian small area estimations with measurement errors

You Mee Goo[1] · Dal Ho Kim[2]

[12]Department of Statistics, Kyungpook National University

Received 6 May 2013, revised 16 May 2013, accepted 21 May 2013

Abstract

This paper considers Bayes estimations of the small area means under Fay-Herriot model with measurement errors. We provide empirical Bayes predictors of small area means with the corresponding jackknifed mean squared prediction errors. Also we obtain hierarchical Bayes predictors and the corresponding posterior standard deviations using Gibbs sampling. Numerical studies are provided to illustrate our methods and compare their efficiencies.

Keywords: Empirical Bayes, Fay-Herriot model, Gibbs sampler, hierarchical Bayes, jackknife method, mean squared prediction error, measurement error, small areas.

1. Introduction

Sample surveys are generally designed to provide estimates of totals or means for large subpopulation (or domain). Such estimates are "direct" in the sense of using only the domain-specific sample data, and the domain sample sizes are large enough to support reliable direct estimates.

In recent years, demand for reliable estimates for small domains (small areas) has greatly increased due to their growing use in formulating policies and programs, allocation of government funds, regional planning, marketing decisions at local level, income for small places, and others. However, due to cost and operational considerations, it is seldom possible to obtain a large enough overall sample size to support direct estimates for all domains of interest. We use the term "small area" to denote any domain for which direct estimates of adequate precision cannot be produced due to small domain-specific sample size.

It is often necessary to employ "indirect" estimates for small areas that can increase the "effective" domain sample size by "borrowing strength" from related area through linking models, using auxiliary data associated with the small areas. A compressive account of model-based small area estimation is given by Rao (2003).

However, it may happen the situation in which the auxiliary data for use in small area estimation may be measured with errors. For example, one might use auxiliary data from another survey. Bolfarine *et al.* (1996) and Goo and Kim (2012) studied measurement error

[1] Ph.D. candidate, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea.
[2] Corresponding author: Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea. E-mail: dalkim@knu.ac.kr

regression models in finite population sampling. Ghosh *et al.* (2006), Ghosh and Sinha (2007), Torabi *et al.* (2010), Datta *et al.* (2010), and Ybarra and Lohr (2008) considered small area models with covariates subject to measurement error.

In this paper, we focus on empirical Bayes (EB) and hierarchical Bayes (HB) estimation for small areas under area-level Fay-Herriot (1979) models with measurement errors in the area-level covariate values. The outline of the remaining section is as follows. In Section 2, we provide EB estimators of small area means with the corresponding mean squared prediction errors (MSPE) under Fay-Herriot measurement error model. Also we provide HB estimators of small area means with measurement errors and the corresponding posterior standard deviations (PSD) using Gibbs sampling. In Section 3, a numerical studies are provided to illustrate the results of the preceding section. A simualtion study is conducted to compare the performances of the EB and HB estimators.

2. Bayes estimation of small area means with measurement errors

We consider a basic area-level model, well known as the Fay and Herriot (1879) model, when the area-level covariates in the model are subject to measurement errors. If the true covariate vector $\boldsymbol{X}_i$ for each area i is known, then the Fay-Herriot model is given by

$$y_i \; = \; \boldsymbol{X}_i^T \boldsymbol{b} + \nu_i + e_i, \;\; i = 1, ..., m, \tag{2.1}$$

where y_i is a direct survey estimator of the area mean θ_i with sampling error $e_i \overset{ind}{\sim} N(0, \psi_i)$ and known sampling variance ψ_i, and θ_i is modeled as $\theta_i = \boldsymbol{X}_i^T \boldsymbol{b} + \nu_i$ with model errors $\nu_i \overset{ind}{\sim} N(0, \sigma_\nu^2)$ independent of e_i for all i. If the model parameters $\boldsymbol{b}$ and σ_ν^2 are known, then the best (or Bayes) predictor of θ_i is given by

$$\tilde{\theta}_{iFH} \; = \; \gamma_{i\nu} y_i + (1 - \gamma_{i\nu}) \boldsymbol{X}_i^T \boldsymbol{b}, \tag{2.2}$$

where $\gamma_{i\nu} = \sigma_\nu^2 / (\sigma_\nu^2 + \psi_i)$.

Suppose now that $\boldsymbol{X}_i$ is not known and that $\boldsymbol{X}_i$ in (2.1) is replaced by an estimator $\hat{\boldsymbol{X}}_i$ from an independent survey, where $\hat{\boldsymbol{X}}_i \overset{ind}{\sim} N(\boldsymbol{X}_i, \boldsymbol{C}_i)$ and $\boldsymbol{C}_i$ is assumed to be known.

2.1. Empirical Bayes estimation

We consider empirical Bayes predictors of small area means with measurement errors. Now, if we substitute $\hat{\boldsymbol{X}}_i$ for $\boldsymbol{X}_i$ in (2.2), then the resulting substitution predictor, $\tilde{\theta}_{iS}$, is worse than $\tilde{\theta}_{iFH}$ in the sense of the mean squared prediction error (MSPE). Moreover, if $\boldsymbol{b}^T \boldsymbol{C}_i \boldsymbol{b} > \sigma_\nu^2 + \psi_i$, then using $\tilde{\theta}_{iS}$ is worse than using the direct estimator (Ybarra and Lohr, 2008).

Now following the idea given in Datta *et al.* (2010) we obtain a pseudo-Bayes predictor of θ_i by estimating $\boldsymbol{X}_i$ efficiently. We observe that both y_i and $\hat{\boldsymbol{X}}_i$ provide information on $\boldsymbol{X}_i$, noting that $y_i | \boldsymbol{X}_i \sim N(\boldsymbol{X}_i^T \boldsymbol{b}, \sigma_\nu^2 + \psi_i)$ and $\hat{\boldsymbol{X}}_i | \boldsymbol{X}_i \sim N(\boldsymbol{X}_i, \boldsymbol{C}_i)$. Expressing the joint density function $f(y_i, \hat{\boldsymbol{X}}_i | \boldsymbol{X}_i) = f(y_i | \boldsymbol{X}_i) f(\hat{\boldsymbol{X}}_i | \boldsymbol{X}_i))$ as a likelihood function $L(\boldsymbol{X}_i)$ and then maximizing $L(\boldsymbol{X}_i)$ with respect to $\boldsymbol{X}_i$. The "score" equation is

$$l(\boldsymbol{X}_i) = \frac{\partial}{\partial \boldsymbol{X}_i} \log L(\boldsymbol{X}_i) \propto \frac{(y_i - \boldsymbol{X}_i^T \boldsymbol{b}) \boldsymbol{b}}{\sigma_\nu^2 + \psi_i} + \boldsymbol{C}_i^{-1} (\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i) = \boldsymbol{0}. \tag{2.3}$$

Solving (2.3) for $\boldsymbol{X}_i$ we get the maximum likelihood estimator of $\boldsymbol{X}_i^T\boldsymbol{b}$ as

$$\tilde{\boldsymbol{X}}_i^T\boldsymbol{b} = \delta_i\hat{\boldsymbol{X}}_i^T\boldsymbol{b} + (1-\delta_i)y_i, \tag{2.4}$$

where $\delta_i = (\sigma_\nu^2+\psi_i)/(\sigma_\nu^2+\psi_i+\boldsymbol{b}^T\boldsymbol{C}_i\boldsymbol{b})$. Now, substituting $\tilde{\boldsymbol{X}}_i^T\boldsymbol{b}$, given by (2.4), for $\boldsymbol{X}_i^T\boldsymbol{b}$ in (2.2), we get the pseudo-Bayes predictor of θ_i,

$$\tilde{\theta}_{iPB} \;=\; \gamma_i y_i \;+\; (1-\gamma_i)\hat{\mathbf{X}}_i^T\mathbf{b}, \tag{2.5}$$

where $\gamma_i = (\sigma_\nu^2+\boldsymbol{b}^T\boldsymbol{C}_i\boldsymbol{b})/(\sigma_\nu^2+\boldsymbol{b}^T\boldsymbol{C}_i\boldsymbol{b}+\psi_i)$.

Note that Ybarra and Lohr (2008) obtained the minimum MSPE predictor, $\tilde{\theta}_{iME}$, among all linear combination of y_i and $\hat{\boldsymbol{X}}_i^T\boldsymbol{b}$ of the form of $a_iy_i+(1-a_i)\hat{\boldsymbol{X}}_i^T\boldsymbol{b}$. It turns out that the Ybarra-Lohr predictor $\tilde{\theta}_{iME}$ is the same as the pseudo-Bayes predictor $\tilde{\theta}_{iPB}$.

A pseudo-empirical Bayes predictor is obtained by replacing $\boldsymbol{b}$ and σ_ν^2 by consistent estimators. To do this, we use modified least squares to estimate the parameters (Cheng and Van Ness, 1999, pp. 85, 146). Let $w_1,\cdots,w_m$ be a set of finite weights bounded away from 0. The estimated regression parameters $\hat{\boldsymbol{b}}_w$ satisfy the equation $\sum_{i=1}^m w_i(\hat{\boldsymbol{X}}_i\hat{\boldsymbol{X}}_i^T-\boldsymbol{C}_i)\hat{\boldsymbol{b}} = \sum_{i=1}^m w_i\hat{\boldsymbol{X}}_iy_i$, when the solution exists. Thus,

$$\hat{\boldsymbol{b}}_w = \{\sum_{i=1}^m w_i(\hat{\boldsymbol{X}}_i\hat{\boldsymbol{X}}_i^T-\boldsymbol{C}_i)\}^{-1}\sum_{i=1}^m w_i\hat{\boldsymbol{X}}_iy_i \tag{2.6}$$

estimates $\boldsymbol{b}$ if the inverse exists. Ybarra and Lohr (2008) showed that $\hat{\boldsymbol{b}}_w$ is a consistent estimator of $\boldsymbol{b}$ as $m\to\infty$. Furthermore,

$$\hat{\sigma}_\nu^2(w) \;=\; (m-p)^{-1}\sum_{i=1}^m\{(y_i-\hat{\boldsymbol{X}}_i^T\hat{\boldsymbol{b}}_w)^2-\psi_i-\hat{\boldsymbol{b}}_w^T\boldsymbol{C}_i\hat{\boldsymbol{b}}_w\} \tag{2.7}$$

is a consistent estimator of σ_ν^2.

The method of weighted least squares is typically used to estimate $\boldsymbol{b}$ in the Fay-Herriot model, with $\hat{w}_i = 1/(\hat{\sigma}_\nu^2+\psi_i)$. In our situation, we would like to use weights $w_i = 1/(\sigma_\nu^2+\psi_i+\boldsymbol{b}^T\boldsymbol{C}_i\boldsymbol{b})$. In practice, we initially set $w_i=1$ and estimate $\boldsymbol{b}$ and σ_ν^2 using (2.6) and (2.7). We then substitute these estimates into the expression for the desired weights to obtain $\hat{w}_i$, which is consistent for w_i. If desired, this process can be iterated.

Then, let $\hat{\boldsymbol{\phi}} = (\hat{\sigma}_\nu^2,\ \hat{\mathbf{b}}^T)^T$ be an estiamtor of $\boldsymbol{\phi} = (\sigma_\nu^2,\ \mathbf{b}^T)^T$ with

$$\mathrm{cov}(\hat{\boldsymbol{\phi}}) \;=\; \begin{pmatrix} \mathrm{var}(\hat{\sigma}_\nu^2) & \boldsymbol{d}_m^T \\ \boldsymbol{d}_m & \boldsymbol{B}_m \end{pmatrix} + o(m^{-1}).$$

Assume that $\hat{\boldsymbol{\phi}}$ is independent of $(\hat{\boldsymbol{X}}_i,\ y_i)$ and that the sixth central moments of $\hat{\boldsymbol{\phi}}$ are $o(m^{-1})$. Then

$$\begin{aligned} MSE(\hat{\theta}_{iPB}) \;=\;& \gamma_i\psi_i \;+\; (1-\gamma_i)^2\; tr\{(\boldsymbol{C}_i+\boldsymbol{X}_i\boldsymbol{X}_i^T)\boldsymbol{B}_m\} \\ &+ \frac{\psi_i^2}{(\boldsymbol{b}^T\boldsymbol{C}_i\mathbf{b}+\sigma_\nu^2+\psi_i)^3}\; E(\hat{\sigma}_\nu^2+\hat{\boldsymbol{b}}^T\boldsymbol{C}_i\hat{\mathbf{b}}-\sigma_\nu^2-\boldsymbol{b}^T\boldsymbol{C}_i\boldsymbol{b})^2 \\ &+ 2E\{(1-\hat{\gamma}_i)^2(\hat{\boldsymbol{b}}-\boldsymbol{b})^T\}\boldsymbol{C}_i\boldsymbol{b}+o(m^{-1}). \end{aligned} \tag{2.8}$$

The mean squared error may be estimated by analytically obtaining estimators of the terms in (2.8), as done in Prasad and Rao (1990), Datta and Lahiri (2000) and Datta *et al.* (2005) for the Fay-Herriot estimator. In this approach, estimators are substituted for $\boldsymbol{B}_m$ and the expected values in (2.8).

An alternative approach is to use the jackknife derived in Jiang *et al.* (2002) to estimate the mean squared error. Under the conditions in the above equation and assuming that y_i and $\hat{\mathbf{X}}_i$ are normally distributed, we can write $MSE(\hat{\theta}_{iPB}) = M_{1i} + M_{2i}$, where $M_{1i} = \gamma_i \psi_i$ and $M_{2i} = E(\hat{\theta}_i - \tilde{\theta}_i)^2$.

A jackknife bias correction is used to estimate M_{1i} by

$$\hat{M}_{1i} = \hat{\gamma}_i \psi_i + \frac{m-1}{m} \sum_{j=1}^{m} (\hat{\gamma}_i \psi_i - \hat{\gamma}_{i(-j)} \psi_i), \tag{2.9}$$

where the notation $(-j)$ indicates an estimator of the same form but based on the dataset without area j. We estimate M_{2i} by

$$\hat{M}_{2i} = \frac{m-1}{m} \sum_{j=1}^{m} (\hat{\theta}_{i(-j)} - \hat{\theta}_i)^2, \tag{2.10}$$

where $\hat{\theta}_{i(-j)} = \hat{\gamma}_{i(-j)} y_i + (1 - \hat{\gamma}_{i(-j)}) \hat{\boldsymbol{X}}_i^T \hat{\boldsymbol{b}}_{(-j)}$. Then the jackknife estimator of $MSE(\hat{\theta}_{iPB})$ is given by

$$\begin{aligned} mse(\hat{\theta}_{iPB}) &= \hat{M}_{1i} + \hat{M}_{2i} \\ &= \hat{\gamma}_i \psi_i + \frac{m-1}{m} \sum_{j=1}^{m} (\hat{\gamma}_i \psi_i - \hat{\gamma}_{i(-j)} \psi_i) \\ &\quad + \frac{m-1}{m} \sum_{j=1}^{m} [\{\hat{\gamma}_{i(-j)} y_i + (1 - \hat{\gamma}_{i(-j)}) \hat{\boldsymbol{X}}_i^T \hat{\boldsymbol{b}}_{(-j)}\} - \hat{\theta}_i]^2. \end{aligned} \tag{2.11}$$

2.2. Hierarchical Bayes estimation

We consider a hierarchical Bayesian framework to predict the small area means θ_i. To this end, we begin with the following simple Fay-Herriot model with measurement errors:

I. $y_i | \theta_i \overset{ind}{\sim} N(\theta_i,\ \psi_i),\ i = 1, ..., m$, where ψ_i is known.

II. $\theta_i | \boldsymbol{b}, \sigma_\nu^2 \overset{ind}{\sim} N(\boldsymbol{X}_i^T \boldsymbol{b}, \sigma_\nu^2)$ and $\hat{\boldsymbol{X}}_i | \boldsymbol{X}_i \overset{iid}{\sim} N(\boldsymbol{X}_i, \boldsymbol{C}_i),\ i = 1, ..., m$ where $\boldsymbol{C}_i$ is known.

III. $\boldsymbol{X}_i \overset{iid}{\sim} \pi(\boldsymbol{X}_i) \propto 1$.

IV. $\boldsymbol{b}, \sigma_\nu^2$ are mutually independent with $\pi(\boldsymbol{b}) \propto 1$ and $\sigma_\nu^2 \sim IG(a_\nu/2,\ b_\nu/2)$. Here $IG(a, b)$ denotes an inverse Gamma distribution with pdf $f(z) \propto exp(-a/z) z^{(-b-1)} I_{[z>0]}$.

The implementation of the Bayesian procedure is greatly facilitated by the MCMC numerical integration technique, in particular the Gibbs sampler. This requires generating samples from the full conditionals of each of $\theta_i,\ \boldsymbol{b},\ \sigma_\nu^2$, and $\boldsymbol{X}_i$ given the remaining parameters and the data. The details are given below.

Under the HB model, the joint posterior distribution is given by

$$\pi(\theta_1, ..., \theta_m, \boldsymbol{b},\ \sigma_\nu^2, \boldsymbol{X}_1, ..., \boldsymbol{X}_m | \boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m)$$
$$\propto \exp[-\frac{1}{2}\sum_{i=1}^{m}\frac{(y_i-\theta_i)^2}{\psi_i}] \times (\sigma_\nu^2)^{-\frac{m}{2}} \exp[-\frac{1}{2\sigma_\nu^2}\sum_{i=1}^{m}(\theta_i - \boldsymbol{X}_i^T\boldsymbol{b})^2]$$
$$\times \exp[-\frac{1}{2}\sum_{i=1}^{m}(\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i)^T\boldsymbol{C}_i^{-1}(\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i)] \times \ \exp[-\frac{a_\nu}{2\sigma_\nu^2}](\sigma_\nu^2)^{-(b_\nu/2+1)}.$$

Then the full conditionals are obtained as follows:

(i) $[\theta_i|\boldsymbol{b}, \sigma_\nu^2, \boldsymbol{X}_1, ..., \boldsymbol{X}_m, \boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m]$
$\overset{ind}{\sim} N[\ (\psi_i^{-1} + \sigma_\nu{}^{-2})^{-1}\{\psi_i^{-1}y_i + \sigma_\nu^{-2}\boldsymbol{X}_i^T\boldsymbol{b}\},\ (\psi_i^{-1} + \sigma_\nu{}^{-2})^{-1}], i = 1, \cdots, m;$
(ii) $[\boldsymbol{b}|\sigma_\nu^2, \boldsymbol{X}_1, ..., \boldsymbol{X}_m, \theta_1, ..., \theta_m, \boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m]$
$\sim N((\sum_{i=1}^m \boldsymbol{X}_i\boldsymbol{X}_i^T)^{-1}(\sum_{i=1}^m \theta_i\boldsymbol{X}_i),\ \sigma_\nu^2(\sum_{i=1}^m \boldsymbol{X}_i\boldsymbol{X}_i^T)^{-1});$
(iii) $[\sigma_\nu^2|\boldsymbol{X}_1, ..., \boldsymbol{X}_m, \theta_1, ..., \theta_m, \boldsymbol{b}, \boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m]$
$\sim IG(\frac{1}{2}\{\sum_{i=1}^m(\theta_i - \boldsymbol{X}_i^T\boldsymbol{b})^2 + a_\nu\},\ \frac{m+b_\nu}{2});$
(iv) $[\boldsymbol{X}_i|\theta_1, ..., \theta_m, \boldsymbol{b}, \sigma_\nu^2, \boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m]$
$\overset{ind}{\sim} N[(\boldsymbol{C}_i^{-1} + \sigma_\nu^{-2}\boldsymbol{b}\boldsymbol{b}^T)^{-1}(\boldsymbol{C}_i^{-1}\hat{\boldsymbol{X}}_i + \theta_i\boldsymbol{b}),\ (\boldsymbol{C}_i^{-1} + \sigma_\nu^{-2}\boldsymbol{b}\boldsymbol{b}^T)^{-1}], i = 1, \cdots, m.$

Using the Gibbs sampler, we obtain the HB estimators for small area means

$$\begin{aligned} E(\theta_i|\boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m) &= E[E[\theta_i|\boldsymbol{b}, \sigma_\nu^2, \boldsymbol{X}_1, ..., \boldsymbol{X}_m, \boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m]|\boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m] \\ &\approx (Ld)^{-1}\sum_{l=1}^{L}\sum_{k=d+1}^{2d}(\frac{1}{\psi_i} + \frac{1}{\sigma_\nu^{2(lk)}})^{-1}(\frac{y_i}{\psi_i} + \frac{\boldsymbol{X}_i^{(lk)T}\boldsymbol{b}^{(lk)}}{\sigma_\nu^{2(lk)}}), \end{aligned}$$

where $L \geq 2$ independent sequences are generated, each of length 2d, the first d iterations of each sequence are discarded. We have $L\times$d simulated values for each parameter. And the corresponding posterior variance is given by

$$\begin{aligned} &V(\theta_i|\boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m) \\ &= E[V[\theta_i|\boldsymbol{b}, \sigma_\nu^2, \boldsymbol{X}_1, ..., \boldsymbol{X}_m, \boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m]|\boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m] \\ &\quad +V[E[\theta_i|\boldsymbol{b}, \sigma_\nu^2, \boldsymbol{X}_1, ..., \boldsymbol{X}_m, \boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m]|\boldsymbol{y}, \hat{\boldsymbol{X}}_1, ..., \hat{\boldsymbol{X}}_m] \\ &\approx (Ld)^{-1}\sum_{l=1}^{L}\sum_{k=d+1}^{2d}(\frac{1}{\psi_i} + \frac{1}{\sigma_\nu^{2(lk)}})^{-1} \\ &\quad +(Ld)^{-1}\sum_{l=1}^{L}\sum_{k=d+1}^{2d}(\frac{1}{\psi_i} + \frac{1}{\sigma_\nu^{2(lk)}})^{-2}(\frac{y_i}{\psi_i} + \frac{\boldsymbol{X}_i^{(lk)T}\boldsymbol{b}^{(lk)}}{\sigma_\nu^{2(lk)}})^2 \\ &\quad -[(Ld)^{-1}\sum_{l=1}^{L}\sum_{k=d+1}^{2d}(\frac{1}{\psi_i} + \frac{1}{\sigma_\nu^{2(lk)}})^{-1}(\frac{y_i}{\psi_i} + \frac{\boldsymbol{X}_i^{(lk)T}\boldsymbol{b}^{(lk)}}{\sigma_\nu^{2(lk)}})]^2. \end{aligned}$$

3. Numerical studies

In this section we conduct the analysis of data set to illustrate our methods obtained in previous section. We also perform a small simulation experiment to investigate the performance of proposed estimators.

3.1. Data analysis

The U.S. Department of Health and Human Services (HHS) has a direct need for the income data at the state level (the 50 states and the District of Columbia) for formulation its energy assistance program to low income familes. Such estimates are provided to the HHS annually by the U.S. Census Bureau.

Starting with income year 1974, the U.S. Census Bureau has computed model-based estimates of median annual income for 4-person families by state using data from the decennial censuses, March sample of the Current Population Survey (CPS), and estimates of per capita income (PCI) from the Bureau of Economic Analysis (BEA).

In original data we have the following variables.

- y_i = CPS estimates of 1989 4-person family median income
- ψ_i = (SE of y_i)^2
- x_i = census estimates of 1979 4-person family median income

Now we synthesize the data for illustrative purpose. We assume the the x_i are unknown and their estimates are available.

- $\hat{x}_i$ = estimates of the x_i from an independent survey

That is, $\hat{x}_i = x_i + \eta_i$, where $\eta_i \overset{iid}{\sim} N(0, 60),\ i\ =\ 1, ...,\ 51.$

Now we provide the Bayes estimates of median income for 4-person families by state using EB and HB procedures with measurement errors. Using the given data set $(y_i, \hat{x}_i)$, $i = 1, ...m$, we calculate the pseudo-EB estimates and the corresponding jackknifed root mean squared errors (RMSE). In the iterative procedures for estimating $\hat{\boldsymbol{b}}$ and $\hat{\sigma}^2_\nu$, we set to zero for the value of $\hat{\sigma}^2_\nu$ if $\hat{\sigma}^2_\nu$ is estimated to be negative. To obtain the HB estimates, we run 5 Gibbs chains of size 10,000 with a burn-in of the first 5,000. After burning out the first half (to eliminate any possible instability in the initial generated samples), we use the average principle and take the average of the HB estimates over the remaining sets to obtain the final HB estimate. The same method is applied to calculate the corresponding posterior standard deviations (PSD). The results are given in Table 3.1.

From Table 3.1, we can see that two Bayes estimates of the small area means are quite close to each other. But from Table 3.1, it appears that the estimated standard errors given by jackknife method are slightly unstable compared to those given by Gibbs sampling.

To check the performance of our estimates, we use the following four criteria to compare the different estimates.

- average relative bias (ARB) = $(51)^{-1} \sum_{i=1}^{51} \frac{|c_i - e_i|}{c_i}$
- average squared relative bias (ASRB) = $(51)^{-1} \sum_{i=1}^{51} \frac{|c_i - e_i|^2}{c_i^2}$
- average absolute bias (AAB) = $(51)^{-1} \sum_{i=1}^{51} |c_i - e_i|$
- average squared deviation (ASD) = $(51)^{-1} \sum_{i=1}^{51} (c_i - e_i)^2$

Here c_i and e_i respectively denote the census and estimates for the i^{th} state ($i = 1, ...51$). The lower values of these measures would imply a better procedure. The values of the four criteria are provided in Table 3.2.

Table 3.2 indicates that the HB estimates are relatively better than the EB estimates under all four criteria for this data set.

Table 3.1 EB and HB estimates of median 4-person family income

state	x	EB	RMSE	HB	PSD
ME	37120	33256	0.588357	33439	0.181587
NH	46613	39332	0.079283	40131	0.191338
VT	40477	35024	0.404817	35451	0.178510
MA	48302	42061	0.336552	43409	0.189269
RI	40743	39389	0.079507	39860	0.189809
CT	62462	45148	0.657888	45947	0.192483
NY	39397	40322	0.164851	40776	0.160282
NJ	48849	45698	0.707198	46721	0.181676
PA	37567	39711	0.110156	39882	0.160571
OH	37933	41284	0.260199	41735	0.167421
IN	31185	40904	0.223823	40942	0.179429
IL	38966	44706	0.615783	44631	0.164885
MI	39606	44246	0.567056	44276	0.162013
WI	38107	41338	0.266217	41492	0.173030
MN	36779	42535	0.388040	42841	0.182903
IO	32923	39797	0.114904	30701	0.171984
MO	33700	39000	0.060069	39014	0.184392
ND	32972	36811	0.219594	36771	0.172268
SD	31538	33855	0.521652	33591	0.173608
NE	36424	38305	0.080305	38542	0.174971
KS	36011	39385	0.079393	39633	0.182735
DE	41499	41885	0.322510	41856	0.184519
MD	45329	46074	0.754058	46430	0.190758
DC	31116	39049	0.065240	38928	0.189243
VA	41226	40433	0.172625	41001	0.190302
WV	29640	36314	0.268916	35801	0.176279
NC	34799	35580	0.347085	35754	0.159257
SN	33477	35935	0.306579	35544	0.188925
GA	32923	37045	0.197431	37489	0.188925
FL	31573	37711	0.131591	37612	0.156486
KY	30941	35468	0.353176	35122	0.181168
TN	30172	35502	0.349204	35055	0.178198
AL	30368	35805	0.319034	35438	0.179574
MA	28617	32990	0.608796	32702	0.177689
AR	25842	32544	0.657919	32435	0.176412
LA	29181	38340	0.082710	38179	0.182470
OK	29374	37011	0.199861	37185	0.185778
TX	31571	40092	0.143445	39141	0.164732
MT	31369	37208	0.175106	36491	0.180914
ID	30815	35911	0.307427	35480	0.175560
WY	35808	43472	0.485885	43308	0.183417
CO	38321	42021	0.336509	41772	0.181983
NM	25743	34820	0.419200	34315	0.182573
AZ	33239	39047	0.061864	39032	0.190382
UT	33317	38455	0.073142	38428	0.181401
NE	35231	43168	0.455658	42579	0.179635
WA	41310	43068	0.442200	43146	0.175561
OR	36071	40355	0.166401	40477	0.180255
CA	39775	43635	0.501622	43429	0.160654
AK	43115	53798	1.542975	53420	0.184254
HI	44416	44012	0.546514	44221	0.190274

Table 3.2 Comparative measures

method	ARB	ASRB	AAB	ASD
EB	0.227530	0.069202	0.503357	0.349764
HB	0.218647	0.063862	0.482844	0.322999

3.2. Simulation study

We conduct a small simulation experiment with $m = 10$ to investigate the performance of the proposed estimators. First the x_i are generated from a $N(5, 3^2)$. Then we generate $\theta_i = 1 + 3x_i + \nu_i$ where $\nu_i \sim N(0, \sigma_\nu^2)$ with $\sigma_\nu^2 = 1, 2, 4$ and $y_i = \theta_i + e_i$ where $e_i \sim N(0, \psi_i)$ with $\psi_i = 0.5, 1, 2$. The $\hat{x}_i$ are generated by $\hat{x}_i = x_i + \eta_i$ where $\eta_i \sim N(0, c_i)$ with $c_i = 1, 3$. So we have 18 cases of parameter values.

For each case we take $R = 5,000$ iterations for each area i after deleting a burn-in of the first 2,000 samples. Then we obtain the true small area means $\theta_i^{(r)}$, the pseudo-EB estimates $\hat{\theta}_i^{EB(r)}$ and HB estimates $\hat{\theta}_i^{HB(r)}$, $i = 1, ..., m$, $r = 1, ..., R$. In calculating HB estimates, we use the small hyperparameter values of $a_\nu = b_\nu = 0.005$ for the diffused Gamma prior.

The empirical MSPE of $\hat{\theta}_i^{EB}$ and $\hat{\theta}_i^{HB}$ are then calculated as

$$EMSPE(\hat{\theta}_i^{EB}) = (R)^{-1} \sum_{r=1}^{R} \ (\ \hat{\theta}_i^{EB(r)} - \theta_i^{(r)} \ )^2,$$

and

$$EMSPE(\hat{\theta}_i^{HB}) = (R)^{-1} \sum_{r=1}^{R} \ (\ \hat{\theta}_{ir}^{HB(r)} - \theta_i^{(r)} \ )^2.$$

Table 3.3(a)-(c) shows that in terms of empirical MSPE, $\hat{\theta}_i^{HB}$ is slightly more efficient than $\hat{\theta}_i^{EB}$. The different values of m was also tried in simulation study, but the results were very similar.

Table 3.3(a) Empirical MSPE of $\hat{\theta}_i^{EB}$ and $\hat{\theta}_i^{HB}$ with $\sigma_\nu^2 = 1$

m	$c_i = 1$						$c_i = 3$					
	ψ_i=0.5		ψ_i=1		ψ_i=2		ψ_i=0.5		ψ_i=1		ψ_i=2	
	EB	HB	EB	HB	EB	HB	EB	HB	EB	HB	EB	HB
1	0.4909	0.4899	0.9004	0.8949	1.7970	1.7996	0.4778	0.4744	1.0937	1.0827	1.9077	1.8741
2	0.5126	0.5068	0.9170	0.9128	1.9126	1.8463	0.4984	0.5011	0.9919	1.0017	1.7511	1.7078
3	0.4672	0.4610	0.9455	0.9263	1.7100	1.6562	0.5020	0.5003	1.0083	0.9864	1.8389	1.7709
4	0.5049	0.5007	0.9746	0.9538	1.7958	1.7781	0.4819	0.4825	0.9786	0.9870	1.8707	1.8012
5	0.5088	0.5055	0.8708	0.8696	2.0249	1.9537	0.4999	0.4962	0.9864	0.9793	1.9687	1.9622
6	0.4725	0.4692	0.9629	0.9609	1.7166	1.6802	0.5107	0.5072	0.9653	0.9447	1.8931	1.9155
7	0.5082	0.5040	0.9602	0.9687	1.7669	1.7295	0.4981	0.4983	0.9009	0.9017	1.9885	1.9678
8	0.4945	0.4927	0.9275	0.9257	1.8181	1.7669	0.4846	0.4819	0.9723	0.9776	1.8313	1.7975
9	0.4481	0.4462	0.9337	0.9254	1.6254	1.5876	0.5110	0.5089	0.9798	0.9742	1.8540	1.8167
10	0.4904	0.4850	0.9883	0.9757	1.7375	1.7255	0.4806	0.4732	1.0096	1.0152	1.9218	1.8687

Table 3.3(b) Empirical MSPE of $\hat{\theta}_i^{EB}$ and $\hat{\theta}_i^{HB}$ with $\sigma_\nu^2 = 2$

m	$c_i = 1$						$c_i = 3$					
	ψ_i=0.5		ψ_i=1		ψ_i=2		ψ_i=0.5		ψ_i=1		ψ_i=2	
	EB	HB	EB	HB	EB	HB	EB	HB	EB	HB	EB	HB
1	0.4674	0.4654	1.0494	1.0343	1.7775	1.7555	0.4784	0.4755	1.0930	1.0819	1.9092	1.8776
2	0.4885	0.4914	0.9673	0.9747	1.6511	1.6114	0.4986	0.5017	0.9934	1.0052	1.7544	1.7155
3	0.4934	0.4913	0.9722	0.9530	1.7064	1.6440	0.5023	0.5009	1.0099	0.9898	1.8415	1.7750
4	0.4766	0.4758	0.9521	0.9579	1.7567	1.7010	0.4820	0.4826	0.9806	0.9893	1.8736	1.8090
5	0.4946	0.4890	0.9472	0.9431	1.8354	1.8184	0.5003	0.4968	0.9865	0.9808	1.9757	1.9681
6	0.5005	0.4984	0.9372	0.9220	1.7828	1.8023	0.5112	0.5084	0.9661	0.9477	1.8986	1.9251
7	0.4865	0.4870	0.8750	0.8785	1.8414	1.8446	0.4983	0.4990	0.9031	0.9035	1.9954	1.9778
8	0.4787	0.4752	0.9607	0.9638	1.7153	1.6912	0.4848	0.4825	0.9735	0.9797	1.8387	1.8104
9	0.5058	0.5043	0.9494	0.9436	1.7079	1.6755	0.5116	0.5105	0.9809	0.9770	1.8566	1.8222
10	0.4704	0.4638	0.9720	0.9769	1.7936	1.7588	0.4805	0.4731	1.0105	1.0158	1.9249	1.8756

Table 3.3(c) Empirical MSPE of $\hat{\theta}_i^{EB}$ and $\hat{\theta}_i^{HB}$ with $\sigma_\nu^2 = 4$

m	$c_i = 1$						$c_i = 3$					
	ψ_i=0.5		ψ_i=1		ψ_i=2		ψ_i=0.5		ψ_i=1		ψ_i=2	
	EB	HB	EB	HB	EB	HB	EB	HB	EB	HB	EB	HB
1	0.4711	0.4697	1.0562	1.0409	1.8069	1.7906	0.4792	0.4773	1.0923	1.0817	1.9129	1.8851
2	0.4917	0.4952	0.9852	0.9922	1.6872	1.6583	0.4988	0.5025	0.9963	1.0097	1.7602	1.7276
3	0.4969	0.4961	0.9839	0.9699	1.7375	1.6815	0.5030	0.5020	1.0125	0.9947	1.8456	1.7830
4	0.4787	0.4785	0.9633	0.9717	1.8014	1.7502	0.4823	0.4828	0.9835	0.9931	1.8801	1.8249
5	0.4966	0.4915	0.9561	0.9527	1.8731	1.8612	0.5008	0.4979	0.9874	0.9836	1.9872	1.9786
6	0.5044	0.5040	0.9445	0.9361	1.8275	1.8511	0.5120	0.5102	0.9670	0.9524	1.9088	1.9398
7	0.4905	0.4915	0.8820	0.8883	1.8810	1.8949	0.4988	0.4997	0.9059	0.9061	2.0047	1.9940
8	0.4794	0.4782	0.9687	0.9741	1.7573	1.7449	0.4852	0.4834	0.9751	0.9816	1.8492	1.8281
9	0.5103	0.5097	0.9597	0.9584	1.7526	1.7258	0.5128	0.5127	0.9826	0.9811	1.8637	1.8351
10	0.4712	0.4664	0.9790	0.9873	1.8298	1.8045	0.4805	0.4736	1.0118	1.0177	1.9301	1.8868

4. Summary and conclusion

We have derived EB and HB predictors of a small area means under Fay-Herriot medel with measurement errors. Our numerical studies show that HB predictors are slightly better than the EB predictors in the closeness of census values. Also in terms of empirical MSPE, HB predictors are are slightly efficient than the EB predictors.

References

Bolfarine, H., Zacks, S. and Sandoval, M. (1996). On predicting the population total under regression models with measurement errors. *Journal of Statistical Planning and Inference*, **55**, 63-76.

Cheng, C. L. and Van Ness, J. W. (1999). *Statistical regression with measurement error*, Arnold, London.

Datta, G. S. and Lahiri, P. (2000). A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Statistica Sinica*, **10**, 613-627.

Datta, G. S., Rao, J. N. K. and Smith, D. D. (2005). On measuring the validity of small area estimators under a basic area level model. *Biometrika*, **92**, 183-196.

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

Fay, R. E. and 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.

Ghosh, M. and Sinha, K. (2007). Empirical Bayesian estimation in finite population sampling under functional measurement error model. *Journal of Statistical Planning and Inference*, **137**, 2759-2773.

Ghosh, M., Sinha, K. and Kim, D. (2006). Empirical and hierarchical Bayesian estimation in finite population sampling under structural measurement error model. *The Scandinavian Journal of Statistics*, **33**, 591-608.

Goo, Y. M. and Kim, D. H. (2012). Bayesian inference in finite population sampling under measurement error model. *Journal of the Korean Data & Information Science Society*, **23**, 1241-1247.

Jiang, J., Lahiri, P. and Wan, S.M. (2002) A unified Jackknife theory for empirical best prediction with M-estimation. *Annals of Statistics*, **30**, 1782-1810.

Prasad, N. G. N. and Rao, J. N. K. (1990). The estimation of mean squared error of small area estimators. *Journal of the American Statistical Association*, **85**, 163-171.

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

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

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