SSC Annual Meeting, June 1999
Proceedings of the Survey Methods Section

PSEUDO HIERARCHICAL BAYES SMALL AREA ESTIMATION USING SAMPLING WEIGHTS

Yong You and J.N.K. Rao[1]

ABSTRACT

Unit-level random effects models are often used in small area estimation to obtain efficient model-based estimators of small area means. Such estimators typically do not make use of the survey weights. As a result, the estimators are not design consistent unless the sampling design is self-weighting within areas. The direct design-based estimators are design consistent but fail to borrow strength from the related small areas. Kott (1989) and Prasad and Rao (1999) proposed model-assisted estimators of small area means by utilizing the survey weights. These estimators are design-consistent and borrow strength across small areas. Prasad and Rao (1999) also derived an estimate of the MSE, assuming the number of small areas is large. In this paper, a two-step hierarchical Bayes approach is used to obtain design-consistent small area estimates. In the first step, the Bayes estimator (posterior mean) and a measure of its variability (posterior variance) are derived from the aggregated area-level model, assuming the variance parameters are known. In the second step, the posterior distribution of variance parameters from the unit-level model is employed to derive a pseudo posterior mean which is used as the estimator of a small area mean and a pseudo posterior variance which is used as a measure of variability of the estimator. The proposed method is implemented by using the Gibbs sampler to generate samples from the posterior distribution of the variance parameters. A limited simulation study shows that the pseudo posterior estimators are efficient and the proposed approach is promising and attractive.

KEY WORDS: Design consistent; Gibbs sampling; Posterior; Random effects model; Survey weights; Small area.

RÉSUMÉ

Les modèles à effets aléatoire au niveau des unités sont souvent utilisés dans l'estimation des petites régions afin d'obtenir des estimateurs efficaces basés sur un modèle de moyennes pour des petites régions. En général, ces estimateurs ne tiennent pas compte des poids de sondage, ce qui a pour effet de produire des estimateurs incohérents par rapport au plan de sondage à moins que le plan soit auto-représentatif dans chacune des régions. Les estimateurs directs basés sur le plan sont cohérents mais ces derniers n'empruntent pas de puissance des petites régions. Kott (1989) et Prasad et Rao (1999) ont proposé des estimateurs assistés par un modèle de moyennes pour les petits régions, qui tiennent compte des poids de sondage. Ces estimateurs sont cohérents et empruntent de la puissance aux petites régions. Prasad et Rao (1999) ont aussi obtenu un estimateur de l'EQM, sous l'hypothèse que le nombre de petites régions est grand. Dans cet article, une approche hiérarchique de Bayes en deux étapes est utilisée afin d'obtenir des estimateurs cohérents pour les petites régions. Dans un premier temps, l'estimateur de Bayes (moyenne a posteriori) et une mesure de sa variabilité (variance a posteriori) sont obtenus à partir du modèle agrégé au niveau de la région, en supposant les variances connues. Dans un deuxième temps, la distribution a posteriori des variances, à partir du modèle à niveau unique, est utilisée afin d'obtenir une moyenne pseudo a posteriori qui sert d'estimateur des moyennes des petites régions; de plus, une variance pseudo a posteriori est obtenue comme mesure de variabilité des estimateurs. La méthode proposée dans cet article utilise l'échantillonnage de Gibbs pour générer des échantillons à partir de la distribution a posteriori de variances. Une petite simulation montre que les estimateurs pseudo a posteriori sont efficaces et que l'approche proposée est attrayante et prometteuse.

MOTS-CLÉS : Cohérent par rapport au plan de sondage; Échantillonnage de Gibbs; a posteriori; modèles à effets aléatoires; poids de sondage; petites régions.

1. INTRODUCTION

Unit-level random effects models including nested error regression models and multi-level models are often used in small area estimation to obtain efficient model-based estimators of small area means. Such estimators typically do not make use of the survey weights. As a result, the estimators are not design consistent unless the sampling design is self-weighting within areas. On the other hand, direct design-based estimators are design consistent but fail to borrow strength from the related small areas. Kott (1989) advocated the use of design-consistent model-based estimators. He derived a design-consistent estimator of

[1] Yong You, Household Survey Methods Division, Statistics Canada, Ottawa, Canada, K1A 0T6, yongyou@statcan.ca;
J.N.K. Rao, School of Mathematics and Statistics, Carleton University, Ottawa, Canada, K1S 5B6, jrao@math.carleton.ca.

a small area mean under a simple random effects model. He also proposed an estimator of its mean squared error (MSE) which is model unbiased as well as design-consistent. But the MSE estimator is quite unstable and can even take negative values. Prasad and Rao (1999) proposed a two-step procedure to obtain a pseudo EBLUP estimator of a small area mean. In the first step, an area-level model is derived from the unit-level model using survey weights and optimal small area estimates are obtained assuming the variance parameters in the model are known. In the second step, the variance parameters are estimated from the unit-level model and then used in the optimal first-step estimator. The pseudo EBLUP estimators are design-consistent and borrow strength across small areas. Prasad and Rao (1999) also derived a stable model-based MSE estimator, assuming the number of small areas is large.

In this paper, a two-step hierarchical Bayes approach is used to obtain posterior estimators of the small area means. In the first step, the Bayes estimator (posterior mean) and a measure of its variability (posterior variance) are derived from the aggregated area-level model, assuming the variance parameters are known. In the second step, the posterior distribution of variance parameters from the unit-level model is employed to derive a pseudo posterior mean which is used as the estimate of a small area mean and a pseudo posterior variance which is used as a measure of variability of the estimate. A limited simulation study shows that the pseudo posterior estimators are efficient. Our results suggest that the proposed approach is very promising and attractive. The pseudo hierarchical Bayes approach can also be extended to general multi-level models.

2. SIMPLE RANDOM EFFECTS MODEL

A simple random effects unit level full model is given by

$$y_{ij} = \theta_i + e_{ij}, \quad j = 1,\ldots,N_i; i = 1,\ldots,m, \qquad (1)$$

where y_{ij} are the unit population values, e_{ij} are uncorrelated random errors with $E(e_{ij}) = 0$ and $V(e_{ij}) = \sigma_i^2$, and N_i is the population size in the i-th area ($i = 1,\ldots,m$). For simplicity, we take θ_i as the i-th small area mean $\bar{Y}_i = \sum_j y_{ij}/N_i$. Note that $\bar{Y}_i = \theta_i + \bar{E}_i$ and $\bar{E}_i = \sum_j e_{ij}/N_i \approx 0$ if N_i is large, so that $\bar{Y}_i \approx \theta_i$. It is customary to assume equal error variances $\sigma_i^2 = \sigma_e^2$, although the case of unequal and random error variances has also been studied (Kleffe and Rao, 1992; You and Rao, 1999). The linking model or population model assumes the small area mean θ_i is a random variable satisfying

$$\theta_i = \mu + v_i, \quad i = 1,\ldots,m, \qquad (2)$$

where the v_i's are uncorrelated random variables with $E(v_i) = 0$ and $V(v_i) = \sigma_v^2$. Further, v_i and e_i are assumed to be uncorrelated for all i and j. Assuming that the full model (1) also holds for the sample $\{y_{ij}, j = 1,\ldots,n_i; i = 1,\ldots,m\}$ and combining the sample model with the linking model, we get the unit-level simple random effects model

$$y_{ij} = \mu + v_i + e_{ij}, \quad j = 1,\ldots,n_i; i = 1,\ldots,m. \qquad (3)$$

This is the well-known two variance components model and is used in small area estimation to obtain model-based estimators of small area means. However, the model-based estimators ignore the sampling design, in particular, the design weights attached to the sample $\{y_{ij}, j = 1,\ldots,n_i; i = 1,\ldots,m\}$.

To bring the design weights into the estimation, suppose $\tilde{w}_{ij}$ denotes the basic design weight attached to the j-th sample unit ($j = 1,\ldots,n_i$) in the i-th small area ($i = 1,\ldots,m$). A direct design-based estimator of small area mean θ_i is given by the ratio estimator

$$\bar{y}_{iw} = \frac{\sum_{j=1}^{n_i} \tilde{w}_{ij} y_{ij}}{\sum_{j=1}^{n_i} \tilde{w}_{ij}} = \sum_{j=1}^{n_i} w_{ij} y_{ij}, \qquad (4)$$

where $w_{ij} = \tilde{w}_{ij} / \sum_{j=1}^{n_i} \tilde{w}_{ij}$. The direct estimator $\bar{y}_{iw}$ is design consistent but fails to borrow strength from other areas. By combining the direct estimator (4) with the sampling model (3), we get the aggregated area level model given by

$$\bar{y}_{iw} = \mu + v_i + \bar{e}_{iw}, \qquad (5)$$

where $\bar{e}_{iw} = \sum_{j=1}^{n_i} w_{ij} e_{ij}$ are uncorrelated random variables with $E(\bar{e}_{iw}) = 0$ and $V(\bar{e}_{iw}) = \sigma_e^2 \sum_{j=1}^{n_i} w_{ij}^2 \equiv \delta_i$. The aggregated model (5) is an area level model similar to the well known Fay-Herriot Model (Fay and Herriot, 1979). We will use the area level model (5) to obtain the posterior estimates of the small area means θ_i.

3. BAYES ESTIMATION

We are interested in obtaining the posterior estimates of the small area mean θ_i based on the design consistent estimator $\bar{y}_{iw}$. Under model (5) with normal assumption on v_i and e_{ij}, we first obtain the posterior distribution of θ_i given $\bar{Y}_w = \{\bar{y}_{iw}, i = 1,...,m\}$, σ_e^2 and σ_v^2, by assuming a flat prior distribution on μ. Calculations show that

$$\theta_i \mid \bar{Y}_w, \sigma_e^2, \sigma_v^2 \sim N(\tilde{\theta}_i, \tilde{V}(\theta_i)), \quad (6)$$

where

$$\tilde{\theta}_i = E(\theta_i \mid \bar{Y}_w, \sigma_e^2, \sigma_v^2) = r_{iw}\bar{y}_{iw} + (1 - r_{iw})\tilde{\mu} \quad (7)$$

and

$$\begin{aligned} \tilde{V}(\theta_i) &= V(\theta_i \mid \bar{Y}_w, \sigma_e^2, \sigma_v^2) \\ &= r_{iw}\delta_i + (1 - r_{iw})^2 \left(\sum_{i=1}^{m} \frac{1}{\sigma_v^2 + \delta_i}\right)^{-1} \quad (8) \\ &\equiv g_{1i}(\sigma_e^2, \sigma_v^2) + g_{2i}(\sigma_e^2, \sigma_v^2), \end{aligned}$$

where $r_{iw} = \sigma_v^2 / (\sigma_v^2 + \delta_i)$, $\tilde{\mu} = \sum_{i=1}^{m} r_{iw}\bar{y}_{iw} / \sum_{i=1}^{m} r_{iw}$. Note that the posterior mean $\tilde{\theta}_i$ given by (7) is the same as the BLUP estimator of θ_i in Prasad and Rao (1999) and the posterior variance $\tilde{V}(\theta_i)$ given by (8) is the same as the mean squared error of the BLUP estimator.

The posterior mean $\tilde{\theta}_i$ and posterior variance $\tilde{V}(\theta_i)$ depend on σ_e^2 and σ_v^2, whereas σ_e^2 and σ_v^2 are unknown in practice. By taking account of the uncertainty about σ_e^2 and σ_v^2, we obtain the posterior mean and posterior variance of θ_i given $\bar{Y}_w$ as

$$\begin{aligned} E(\theta_i \mid \bar{Y}_w) &= E(E(\theta_i \mid \bar{Y}_w, \sigma_e^2, \sigma_v^2)) \\ &= \int E(\theta_i \mid \bar{Y}_w, \sigma_e^2, \sigma_v^2)\pi(\sigma_e^2, \sigma_v^2 \mid \bar{Y}_w) d\sigma_e^2 d\sigma_v^2 \end{aligned} \quad (9)$$

and

$$\begin{aligned} V(\theta_i \mid \bar{Y}_w) &= E(V(\theta_i \mid \bar{Y}_w, \sigma_e^2, \sigma_v^2)) \\ &\quad + V(E(\theta_i \mid \bar{Y}_w, \sigma_e^2, \sigma_v^2)) \\ &= \int \tilde{V}(\theta_i)\pi(\sigma_e^2, \sigma_v^2 \mid \bar{Y}_w) d\sigma_e^2 d\sigma_v^2 \quad (10) \\ &\quad + \int \tilde{\theta}_i^2 \pi(\sigma_e^2, \sigma_v^2 \mid \bar{Y}_w) d\sigma_e^2 d\sigma_v^2 \\ &\quad - [\int \tilde{\theta}_i \pi(\sigma_e^2, \sigma_v^2 \mid \bar{Y}_w) d\sigma_e^2 d\sigma_v^2]^2, \end{aligned}$$

where $\pi(\sigma_e^2, \sigma_v^2 \mid \bar{Y}_w)$ is the posterior distribution of σ_e^2 and σ_v^2 given the aggregated data $\bar{Y}_w$. The marginal distribution of $\bar{y}_{iw}$ from the aggregated area-level model (5) is

$$\bar{y}_{iw} \sim N(\mu, \sigma_v^2 + \sigma_e^2 \sum_{j=1}^{n_i} w_{ij}^2), \quad i = 1,...,m. \quad (11)$$

Let $d_i = \sum_{j=1}^{n_i} w_{ij}^2$. Since both σ_e^2 and σ_v^2 are not known in the marginal model (11), σ_e^2 and σ_v^2 are not identifiable in the marginal area-level model (11) if all the d_i's are equal, say, $d_i = d_0$. In general, we cannot get reliable estimates of σ_e^2 and σ_v^2 using only $\{\bar{y}_{iw}\}$ based on (11). Using $\pi(\sigma_e^2, \sigma_v^2 \mid \bar{Y}_w)$ in the posterior estimators (9) and (10) may result in misleading results, particularly in Monte Carlo evaluation. Bell (1998) also indicated similar problems. Thus unit-level data and model should be used to get reliable posterior estimates of σ_e^2 and σ_v^2.

4. PSEUDO BAYES ESTIMATION

By formulating the unit level sampling model (3) in a hierarchical Bayes framework, we can obtain the posterior distribution of σ_e^2 and σ_v^2 given the original sample $Y = \{y_{ij}, j = 1,...,n_i; i = 1,...,m\}$ as $\pi(\sigma_e^2, \sigma_v^2 \mid Y)$. Replacing $\pi(\sigma_e^2, \sigma_v^2 \mid \bar{Y}_w)$ by $\pi(\sigma_e^2, \sigma_v^2 \mid Y)$ in equation (9) and (10), we obtain the pseudo posterior mean and pseudo posterior variance of θ_i denoted as

$$E^P(\theta_i \mid \bar{Y}_w, Y) = \int E(\theta_i \mid \bar{Y}_w, \sigma_e^2, \sigma_v^2)\pi(\sigma_e^2, \sigma_v^2 \mid Y) d\sigma_e^2 d\sigma_v^2 \quad (12)$$

and

$$\begin{aligned} V^P(\theta_i \mid \bar{Y}_w, Y) &= \int \tilde{V}(\theta_i)\pi(\sigma_e^2, \sigma_v^2 \mid Y) d\sigma_e^2 d\sigma_v^2 \\ &\quad + \int \tilde{\theta}_i^2 \pi(\sigma_e^2, \sigma_v^2 \mid Y) d\sigma_e^2 d\sigma_v^2 \\ &\quad - [\int \tilde{\theta}_i \pi(\sigma_e^2, \sigma_v^2 \mid Y) d\sigma_e^2 d\sigma_v^2]^2. \end{aligned} \quad (13)$$

Numerical evaluation of (12) and (13) involves two dimensional integration. Following Box and Tiao (1973), we can obtain the joint posterior distribution $\pi(\sigma_e^2, \sigma_v^2 \mid Y)$. The form of $\pi(\sigma_e^2, \sigma_v^2 \mid Y)$ is very complicated (see You, 1999). Direct integration of (12) and (13) involves combining (12) and (13) with $\pi(\sigma_e^2, \sigma_v^2 \mid Y)$ and is very difficult, even though we only have two dimensional integration. In general, high dimensional integration may be involved, for example, in the cases of unequal error variances

(σ_i^2 unequal) and multi-level models (You and Rao, 1999). Thus Monte Carlo integration should be used and is more feasible in practice. A sampling based approach, in particular, the Gibbs sampling method, is used to estimate the pseudo posterior mean and pseudo posterior variance of θ_i.

5. A GIBBS SAMPLING APPROACH

For Monte Carlo evaluation of (12) and (13), the Gibbs sampling method (Gelfand and Smith, 1990) is used to draw samples from the joint posterior distribution $\pi(\sigma_e^2, \sigma_v^2 \mid Y)$. To apply the Gibbs sampling method, we present the unit level sampling model (3) in a hierarchical Bayes framework as follows:

$$y_{ij} \mid \theta_i, \sigma_e^2 \sim N(\theta_i, \sigma_e^2), \quad \theta_i \mid \mu, \sigma_v^2 \sim N(\mu, \sigma_v^2),$$
$$\mu \sim N(0, \tau), \quad \sigma_e^2 \sim IG(a_1, b_1), \quad \sigma_v^2 \sim IG(a_2, b_2), \tag{14}$$

where μ has a normal prior, σ_e^2 and σ_v^2 have an inverse gamma (IG) prior distribution, respectively. τ is chosen to be very large and $a_i > 0, b_i > 0, (i = 1,2)$ are chosen to be very small to reflect our vague knowledge. Thus proper priors are used here to avoid improper posteriors (Hobert and Casella, 1996).

The full conditional distributions for the Gibbs sampler can be found in You (1999). Suppose a sample of size G, $\{\theta_i^{(k)}, \mu^{(k)}, \sigma_e^{2(k)}, \sigma_v^{2(k)}; k = 1, \ldots, G\}$, is generated from the Gibbs sampler based on the Bayesian unit level sampling models (14). Then it follows from (12) and (13) that the estimated pseudo posterior mean and pseudo posterior variance of θ_i are given respectively by

$$\hat{\theta}_i^p = \frac{1}{G}\sum_{k=1}^{G} E(\theta_i \mid \bar{Y}_w, \sigma_e^{2(k)}, \sigma_v^{2(k)}) = \frac{1}{G}\sum_{k=1}^{G} \tilde{\theta}_i(\sigma_e^{2(k)}, \sigma_v^{2(k)}) \tag{15}$$

and

$$\hat{V}^p(\theta_i) = \frac{1}{G}\sum_{k=1}^{G}[g_{1i}(\sigma_e^{2(k)}, \sigma_v^{2(k)}) + g_{2i}(\sigma_e^{2(k)}, \sigma_v^{2(k)})] + \frac{1}{G}\sum_{k=1}^{G}[\tilde{\theta}_i(\sigma_e^{2(k)}, \sigma_v^{2(k)})]^2 - [\frac{1}{G}\sum_{k=1}^{G}\tilde{\theta}_i(\sigma_e^{2(k)}, \sigma_v^{2(k)})]^2, \tag{16}$$

where $\tilde{\theta}_i$ and g_{1i}, g_{2i} are given by (7) and (8), respectively.

By using the output from the Gibbs sampler, we can also find $\hat{\theta}_i^m$ and $\hat{V}^m(\theta_i)$, the Rao-Blackwellized estimators of the completely model-based posterior mean $E(\theta_i \mid Y)$ and posterior variance $V(\theta_i \mid Y)$. Formulas for $\hat{\theta}_i^m$ and $\hat{V}^m(\theta_i)$ can be found in You (1999). We will compare the performance of the pseudo posterior mean and pseudo posterior variance with that of the model-based posterior mean and posterior variance in the simulation study.

6. A SIMULATION STUDY

To evaluate the finite sample performance of the proposed pseudo posterior estimators, we conducted a limited simulation study. We first constructed three fixed finite populations with m=30 areas in each population. Each area has $N_i = 200$ population units in it. The three synthetic populations were generated from the unit level full models (1) and (2) by taking $\mu = 50$, $\sigma_e^2 = 225$ and $\sigma_v^2 = 100, 64, 36$, respectively. From each population, PPS (probability proportional to size) samples within each area were drawn independently. To implement PPS sampling, size measures x_{ij} for each unit ij were generated from an exponential distribution with mean 200. The selection probability for each unit ij is computed as $p_{ij} = x_{ij} / \sum_j x_{ij}$. For the three populations, we chose equal sample size $n_i = n$ within each area, by taking n=5, 20 and 40, respectively. The basic design weights are given by $\tilde{w}_{ij} = n^{-1} p_{ij}^{-1}$ so that $w_{ij} = p_{ij}^{-1} / \sum_j p_{ij}^{-1}$. Using these weights and the associated sample values y_{ij}, we computed the direct ratio estimate $\bar{y}_{iw}$. To compute the pseudo posterior mean estimate $\hat{\theta}_i^p$ and the associated pseudo posterior variance estimates $\hat{V}^p(\theta_i)$, we need to first find the posterior estimates of σ_e^2 and σ_v^2. Samples $\{y_{ij}\}$ were used in the Gibbs sampler and after 3000 "burn-in" period of iterations, 1000 iterations were kept for analysis. By using the Gibbs sampling output, we calculated pseudo posterior estimators $\hat{\theta}_i^p$ and $\hat{V}^p(\theta_i)$ based on (15) and (16) and the model-based posterior estimators $\hat{\theta}_i^m$ and $\hat{V}^m(\theta_i)$, respectively. This whole process was repeated R=100 times, and for each run $r\,(r = 1, \ldots, R)$, $\bar{y}_{iw}(r)$, $\hat{\theta}_i^p(r)$ and $\hat{\theta}_i^m(r)$ were computed, as well as the associated posterior variances $\hat{V}^p(\theta_i)(r)$ and $\hat{V}^m(\theta_i)(r)$. This simulation approach is design-based in the sense that it refers to the design and the PPS samples were drawn from a fixed population.

Our purpose of this simulation study is to test the performance of the proposed pseudo posterior estimators. For this purpose, we calculated the absolute relative bias (ARB) of $\hat{\theta}_i^p$ with respect to the fixed small area mean θ_i as $ARB(\hat{\theta}_i^p) = | E^*(\hat{\theta}_i^p)/\theta_i - 1 |$, where E^* denotes expectation with respect to the design, that is, the average value over R=100 runs, $E^*(\hat{\theta}_i^p) = \sum_{r=1}^{R} \hat{\theta}_i^p(r)/R$. Similarly we calculated the absolute relative bias of $\hat{\theta}_i^m$ as $ARB(\hat{\theta}_i^m)$. The empirical coefficient of variation (CV) of $\hat{\theta}_i^p$ and $\hat{\theta}_i^m$ were calculated as $CV(\hat{\theta}_i^p) = E^*(\sqrt{\hat{V}^p(\theta_i)}/\hat{\theta}_i^p)$ and $CV(\hat{\theta}_i^m) = E^*(\sqrt{\hat{V}^m(\theta_i)}/\hat{\theta}_i^m)$. The relative efficiency (RE) of $\hat{\theta}_i^p$ and $\hat{\theta}_i^m$ over the direct survey estimator $\bar{y}_{iw}$ were computed as $RE(\hat{\theta}_i^p) = MSE^*(\bar{y}_{iw})/MSE^*(\hat{\theta}_i^p)$ and $RE(\hat{\theta}_i^m) = MSE^*(\bar{y}_{iw})/MSE^*(\hat{\theta}_i^m)$, where MSE^* denotes the MSE over R=100 runs. For example, $MSE^*(\hat{\theta}_i^p) = \sum_{r=1}^{R} (\hat{\theta}_i^p(r) - \theta_i)^2/R$.

Table 1 presents the average values of ARB, CV and RE for $\hat{\theta}_i^p$ and $\hat{\theta}_i^m$ over the small areas $i = 1,\ldots,30$ for the three populations. First of all, the ARB is very small for both $\hat{\theta}_i^p$ and $\hat{\theta}_i^m$, even when the sample size n is small. For example, when n=5, $\hat{\theta}_i^p$ has about 7% and $\hat{\theta}_i^m$ has about 6% ARB. As sample size n increases, the ARB decreases for both $\hat{\theta}_i^p$ and $\hat{\theta}_i^m$. The pseudo posterior mean $\hat{\theta}_i^p$ has slightly larger bias than the posterior mean $\hat{\theta}_i^m$ due to the bias induced by the direct survey estimate $\bar{y}_{iw}$ in $\hat{\theta}_i^p$. For the CV, $\hat{\theta}_i^p$ has about 2% larger CV than $\hat{\theta}_i^m$ due to the extra uncertainty associated with the $\bar{y}_{iw}$. So in terms of ARB and CV, $\hat{\theta}_i^p$ performs very much like the $\hat{\theta}_i^m$, except for the extra bias and uncertainty induced by $\bar{y}_{iw}$ in $\hat{\theta}_i^p$. Since the sample obeys the assumed model, $\hat{\theta}_i^m$ performs well.

For the RE, both $\hat{\theta}_i^p$ and $\hat{\theta}_i^m$ have large RE gain over $\bar{y}_{iw}$. As sample size n decreases, the RE gain increases. For example, in Table 1, $RE(\hat{\theta}_i^p) = 1.2$ if n=40; $RE(\hat{\theta}_i^p) = 1.3$ if n=20; $RE(\hat{\theta}_i^p) = 2.0$ if n=5. So the model-assisted estimators indeed borrowing strength from related areas and the gains in efficiency over $\bar{y}_{iw}$ are more significant when the sample size is small. This is because when n is small, the survey estimator $\bar{y}_{iw}$ is not reliable, thus the improvement of the model-assisted estimator is more evident. When the sample size is the same, $\hat{\theta}_i^m$ has larger RE gain than $\hat{\theta}_i^p$ as expected. Note that as sample size increases, the pseudo posterior mean $\hat{\theta}_i^p$ moves toward $\bar{y}_{iw}$ (design-consistent), whereas $\hat{\theta}_i^m$ moves toward the sample mean $\bar{y}_i = 1/n_i \sum_j y_{ij}$. Given that the model is true, we know that $\bar{y}_i$ is the BLUE of θ_i and is always better than $\bar{y}_{iw}$. Thus, as it is shown in the results, as n increases, $RE(\hat{\theta}_i^p)$ decreases and moves toward 1, whereas $RE(\hat{\theta}_i^m)$ is always larger than $RE(\hat{\theta}_i^p)$. Overall, the pseudo posterior mean $\hat{\theta}_i^p$ has less RE gain than $\hat{\theta}_i^m$. This is the trade-off for bringing the survey weights into the model and making $\hat{\theta}_i^p$ design-consistent. The RE gain also depends on the ratio σ_v^2/σ_e^2. When the sample size is fixed, as σ_v^2/σ_e^2 increases, the weight r_{iw} in $\hat{\theta}_i^p$ will increase and more weight will be put on $\bar{y}_{iw}$ so that the RE gain of $\hat{\theta}_i^p$ over $\bar{y}_{iw}$ decreases. For example, when n=20, under Pop 3 ($\sigma_v^2/\sigma_e^2 =$ 0.16), $RE(\hat{\theta}_i^p)$=2.0; under Pop 2 ($\sigma_v^2/\sigma_e^2 =$ 0.28), $RE(\hat{\theta}_i^p)$=1.6; under Pop 1 ($\sigma_v^2/\sigma_e^2 =$ 0.44), $RE(\hat{\theta}_i^p)$=1.3.

7. CONCLUSION

We have proposed a Bayesian approach to obtain a pseudo hierarchical Bayes estimator of a small area mean under simple random effects models. This pseudo Bayes estimator depends on the survey weights and is design-consistent as sample size increases. We also obtained the pseudo posterior variance, which is used as a measure of variability of the estimate. Our simulation study shows that the proposed pseudo posterior estimators are efficient and the proposed approach is very promising and attractive. The pseudo Bayes approach can also be extended to general multi-level models to make use of survey weights and borrow strength across areas (You, 1999; You and Rao, 1999).

Table 1: Comparison of $\hat{\theta}_i^p$ and $\hat{\theta}_i^m$

Measure	Population 1			Population 2			Population 3		
	n=5	n=20	n=40	n=5	n=20	n=40	n=5	n=20	n=40
$ARB(\hat{\theta}_i^p)\%$	6.7	3.5	2.3	7.0	3.6	2.4	7.2	4.1	2.6
$ARB(\hat{\theta}_i^m)\%$	5.9	3.1	2.1	6.1	3.1	2.0	6.3	3.2	1.9
$CV(\hat{\theta}_i^p)\%$	12.5	8.5	6.7	11.7	8.2	6.6	9.3	7.6	6.4
$CV(\hat{\theta}_i^m)\%$	11.0	6.3	4.6	10.6	6.2	4.6	8.5	5.9	4.4
$RE(\hat{\theta}_i^p)$	2.0	1.3	1.2	2.9	1.6	1.4	3.9	2.0	1.6
$RE(\hat{\theta}_i^m)$	2.5	2.2	2.0	3.2	2.6	2.3	4.5	2.9	2.5

ACKNOWLEDGEMENT

This work was supported by a research grant from NSERC of Canada.

REFERENCES

Bell, W. (1998). Models for county and state poverty estimates. In *Small-Area Estimation of School-Age Children in Poverty: Interim Report 2-Evaluation of Revised 1993 County Estimates for Title I Allocations*, Edited by C.F. Citro, M.L. Cohen, and G. Kalton, National Research Council, Washington D.C.: National Academy Press.

Box, G.E.P. and Tiao, G.C. (1973). *Bayesian Inference In Statistical Analysis*, Addison-Wesley.

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.

Gelfand, A.E. and Smith, A.F.M. (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85, 398-409.

Hobert, J.P. and Cassella, G. (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*, 91, 1461-1473.

Kleffe, J. and Rao, J.N.K. (1992). Estimation of mean square error of empirical best linear unbiased predictors under a random error variance linear model. *Journal of Multivariate Analysis*, 43, 1-15.

Kott, P. (1989). Robust small domain estimation using random effects modelling. *Survey Methodology*, 23, 81-89.

Prasad, N.G.N. and Rao, J.N.K. (1999). On robust small area estimation using a simple random effects model. *Survey Methodology*, in press.

You, Y. (1999). *Hierarchical Bayes and Related Methods for Model-Based Small Area Estimation.* Ph.D. Thesis, Carleton University, Ottawa, Canada.

You, Y and Rao, J.N.K. (1999). Hierarchical Bayes estimation of small area means using multi-level models. Invited paper, IASS Satellite Conference on Small Area Estimation, Riga, Latvia.