

A comparison of small area estimators of counts aligned with direct higher level estimates

Giorgio E. Montanari, M. Giovanna Ranalli and Cecilia Vicarelli

Abstract Indirect estimators for small areas use auxiliary variables to borrow strength from related areas through a linking model. Precision of indirect estimators depends on the validity of such a model. To protect against possible model failures, benchmarking procedures make the total of small area estimates match a design consistent estimate for a larger area. This is also particularly important for National Institutes of Statistics to ensure coherence between small area estimates and direct estimates produced at higher level planned domains. We investigate a self-benchmarked estimator in the case of a unit level logistic mixed model for a binary response, propose an estimator for its mean squared error and compare its performance with competing estimators through a simulation study.

Key words: Small area estimation, Logistic mixed model, Penalized Quasi Likelihood, Unit level model, MSE estimation.

1 Introduction

Sample surveys are designed to provide reliable estimates of finite population parameters for large domains. In fact, design-unbiased or approximately design-unbiased direct estimators are used for those domains whose sample size is sufficiently large. However, local and central governments increase their demand of detailed statistical information on small domains, either geographical areas that are smaller than those planned, or specific subpopulations, as those given by a fine classification on socio-demographic variables. Since such domains are not planned at the design stage, their sample size could be very small and even zero. Consequently, direct estimators of small area parameters may have very large variance and some-

Giorgio E. Montanari, M. Giovanna Ranalli and Cecilia Vicarelli
Dipartimento di Economia, Finanza e Statistica, Università degli Studi di Perugia,
e-mail: giovanna.ranalli@stat.unipg.it

times cannot be calculated due to lack of observations. This makes it necessary to find indirect estimators that connect related small areas through a linking model that is based on auxiliary information to borrow strength and increase the effective sample size. Precision of indirect estimators depends on the validity of such a model.

To protect against possible model failures, benchmarking procedures make the total of small area estimates match a design consistent estimate for a larger area. This is also particularly important for National Statistical Institutes to ensure coherence between small area estimates and direct estimates produced at higher level planned domains. There are two kinds of benchmarked estimators: estimators that are internally benchmarked (or self-benchmarked) and those that are externally benchmarked. For a recent review see Wang *et al.* (2008). Self benchmarked predictors are, for example, the pseudo-EBLUP introduced by You and Rao (2002) and the augmented estimator proposed by Wang *et al.* (2008). A drawback of this type of self benchmarked estimators is that they force the use of the same auxiliary information used for the direct – usually GREG-type – estimator also for the model-based small area predictors, whereas it could be very profitable to allow for different auxiliary variables at the small area level. Externally benchmarked predictors are obtained through an a-posteriori adjustment of model-based predictors. Among the others, Pfeiffermann and Barnard (1991) propose an externally restricted benchmarked estimator of small area means. This is constructed under an area linear mixed model for a continuous response variable. In this work, we investigate an extension of this approach to the case of a unit level logistic mixed model for a binary response. This has particular relevance any time the small area estimate takes the form of a count. You *et al.* (2004) address this issue by using a ratio type adjustment of Hierarchal Bayes estimators.

The paper is organized as follows. Section 2 describes the small area estimator proposed by Pfeiffermann and Barnard (1991) and then Section 3 illustrates the proposed methodology to tackle situations in which the variable of interest is binary. An algorithm to actually compute parameter estimates and random variables predictions is proposed together with an analytic estimator for its mean squared error. The results of a limited simulation study that compares the performance of different estimators are then presented in Section 4.

2 Some small area estimators with the benchmarking property

Consider a finite population $U = \{1, 2, \dots, N\}$ and a partition of U in d domains (small areas) U_i made of N_i units, with $i = 1, 2, \dots, d$, such that $\bigcup_{i=1}^d U_i = U$ and $\sum_{i=1}^d N_i = N$. We are interested in estimating small area totals $Y_i = \sum_{j \in U_i} y_{ij}$ of a binary y variable. To this purpose, we assume that a $(K + 1)$ -dimensional row-vector $\mathbf{x}_{ij} = [1, x_{ij1}, \dots, x_{ijK}]$ of an auxiliary variable $\mathbf{x}$ is known for each element $j \in U_i$, for $i = 1, 2, \dots, d$. Let $\mathbf{X}_N = \{\mathbf{x}_{ij}\}_{j \in U_i; i=1,2,\dots,d}$ be the matrix of population values $\mathbf{x}_{ij}$. Suppose we also know the d -dimensional row-vector $\mathbf{z}_{ij}$ that takes value 1 in its i -th position and 0 otherwise, for all units $j \in U_i$, for $i = 1, 2, \dots, d$. Let

$\mathbf{Z}_N = \{\mathbf{z}_{ij}\}_{j \in U_i; i=1,2,\dots,d}$ be the matrix of population values $\mathbf{z}_{ij}$. To estimate the population and domains total, a sample s of size n is drawn from the finite population U using a probabilistic sampling design $p(s)$. Let $s_i = s \cap U_i$ denote the sample with n_i elements realized in the i -th domain, for $i = 1, 2, \dots, d$. Estimators $\hat{Y}_i^{bench}$ have the benchmarking property if $\sum_{i=1}^d \hat{Y}_i^{bench} = \hat{Y}$, where $\hat{Y}$ is a consistent estimator, usually a GREG or Calibration type estimator of the population total $Y = \sum_{i=1}^d \sum_{j \in U_i} y_{ij}$. Model based estimators – the most widely used for small area parameters – usually do not have this property when $p(s)$ is a complex sampling design. Given a small area estimator $\hat{Y}_i$, that doesn't show the benchmarking property, a first simple way of achieving benchmarking is by a ratio type adjustment, i.e.

$$\hat{Y}_i^{bench} = \hat{Y}_i \frac{\hat{Y}}{\sum_{i=1}^d \hat{Y}_i}. \quad (1)$$

Pfeffermann and Barnard (1991) consider an area level model for farmland values and modify the optimal predictor under a linear mixed model to have the benchmarking property on the sample mean. To this end, they compute the model parameters and random effects as the penalized generalized least squares solution under the benchmark constraint. This procedure gives adjusted estimates that are equivalent to the (un-benchmarked) small area estimates under the linear mixed model plus a correction term. This correction term is given for each small domain estimate by a proportion, a_i , of the difference between the design consistent estimate and the sum of the model-based small area estimates. In particular,

$$\hat{Y}_i^{bench} = \hat{Y}_i + a_i(\hat{Y} - \sum_{i=1}^d \hat{Y}_i), \quad (2)$$

with $\sum_{i=1}^d a_i = 1$. These proportions take the form $a_i = \text{cov}(\hat{Y}_i, \sum_{i=1}^d \hat{Y}_i) / \text{var}(\sum_{i=1}^d \hat{Y}_i)$. In this work we consider a similar approach to deal with variables of interest that are binary and small area totals that are, therefore, counts. This is illustrated in the next section.

3 Benchmarked small area estimates of counts

We assume that the y_{ij} 's are independent bernoulli random variables with

$$\begin{cases} P(y_{ij} = 1 | \mathbf{x}_{ij}) = p_{ij} \\ \text{logit}(p_{ij}) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \eta_{ij} = \mathbf{x}_{ij}\mathbf{b} + \mathbf{z}_{ij}\mathbf{u} , \\ \mathbf{u} \sim N(\mathbf{0}, \sigma_u^2 \mathbf{I}) \end{cases} \quad (3)$$

for $j \in U_i$, $i = 1, 2, \dots, d$. We want to derive a benchmarked estimator $\hat{Y}_i^{bench}$ of Y_i for $i = 1, 2, \dots, d$ under model (3) that takes the following form

$$\hat{Y}_i^{bench} = \sum_{j \in s_i} y_{ij} + \sum_{j \in U_i \setminus s_i} \hat{p}_{ij} = \sum_{j \in s_i} y_{ij} + \sum_{j \in U_i \setminus s_i} \frac{\exp(\mathbf{x}_{ij}\hat{\mathbf{b}} + \mathbf{z}_{ij}\hat{\mathbf{u}})}{1 + \exp(\mathbf{x}_{ij}\hat{\mathbf{b}} + \mathbf{z}_{ij}\hat{\mathbf{u}})} \quad (4)$$

(see e.g. Saei and Chambers, 2003, or Rao, 2003), but for which the benchmarking property holds, i.e. such that

$$\sum_{i=1}^d \hat{Y}_i^{bench} = \hat{Y}. \quad (5)$$

To this end, we consider the Penalized Quasi log-Likelihood (PQL) of model (3) under the benchmark constraint in equation (5). The PQL approach considers a log-likelihood function of the sample values y_{ij} for $i \in s_i$ and $i = 1, \dots, d$ conditional on $\mathbf{u}$ with the addition of a penalty function due to the presence of the random effect, given by the logarithm of the probability density function of $\mathbf{u}$ (for a more detailed description of PQL methods see e.g. McCulloch and Searle, 2001). We modify such PQL method by adding the benchmark constraint through the method of Lagrange multipliers. As a result, the loglikelihood is made of three parts. In particular, let l_1 be the loglikelihood function of the sample values y_{ij} conditional on $\mathbf{u}$, let l_2 be the penalization coming from the marginal distribution of the random effects $\mathbf{u}$, and let l_3 be the penalization due to the benchmarking constraint. Therefore, the benchmarked PQL is given by $l = l_1 + l_2 + l_3$, with

$$\begin{aligned} l_1 &= \mathbf{y}_s' \boldsymbol{\eta}_s - \mathbf{1}_n' \log[1 + \exp(\boldsymbol{\eta}_s)], \\ l_2 &= -\frac{1}{2} [\boldsymbol{\sigma}_u^{-2} \mathbf{u}' \mathbf{u} + d \log(\boldsymbol{\sigma}_u^2)], \\ l_3 &= \lambda \left(\hat{Y} - \sum_{i=1}^d \hat{Y}_i^{bench} \right) = \lambda \left(\hat{Y} - \mathbf{1}_n' \mathbf{y}_s - \mathbf{1}_{N-n}' \frac{\exp(\boldsymbol{\eta}_r)}{1 + \exp(\boldsymbol{\eta}_r)} \right), \end{aligned}$$

where $\mathbf{y}_s$ denotes the vector of observed values y_{ij} for $j \in s_i$ and $i = 1, 2, \dots, d$; $\boldsymbol{\eta}_s$ is the linear predictor vector for $j \in s_i$, $i = 1, 2, \dots, d$, and $\boldsymbol{\eta}_r$ is the linear predictor vector for $j \in U_i \setminus s_i$, $i = 1, 2, \dots, d$.

To obtain maximum benchmarked PQL estimates of $\mathbf{b}$, $\mathbf{u}$, λ and the ML estimate of the variance component $\boldsymbol{\sigma}_u^2$ we use an algorithm based on Fisher Scoring. Saei and Chambers (2003) describe the procedure to obtain small area estimates from a logistic mixed model. Such procedure is adapted to accommodate l_3 and described below.

1. Assign initial values to $\mathbf{b}$, $\mathbf{u}$, λ , and $\boldsymbol{\sigma}_u^2$.
2. Update $\mathbf{b}$, $\mathbf{u}$ and λ using Fisher scoring, i.e. by

$$\begin{pmatrix} \mathbf{b}^{(r+1)} \\ \mathbf{u}^{(r+1)} \\ \lambda^{(r+1)} \end{pmatrix} = \begin{pmatrix} \mathbf{b}^{(r)} \\ \mathbf{u}^{(r)} \\ \lambda^{(r)} \end{pmatrix} + \mathbb{I}(\mathbf{b}^{(r)}, \mathbf{u}^{(r)}, \lambda^{(r)})^{-1} \mathbb{S}(\mathbf{b}^{(r)}, \mathbf{u}^{(r)}, \lambda^{(r)})$$

where $\mathbb{I}(\cdot)$ and $\mathbb{S}(\cdot)$ are the Information matrix and the score of l , respectively.

3. Calculate $\boldsymbol{\eta}_s$ and $\boldsymbol{\eta}_r$.
4. Calculate $\mathbf{B}_s = -(\partial^2 l_1 / \partial \boldsymbol{\eta}_s \partial \boldsymbol{\eta}_s')$ and $\mathbf{C}_r = -(\partial^2 l_3 / \partial \boldsymbol{\eta}_r \partial \boldsymbol{\eta}_r')$.

5. Calculate $\mathbf{T} = (\sigma_u^{2(r)} \mathbf{I} + \mathbf{Z}_s^t \mathbf{B}_s \mathbf{Z}_s)^{-1}$, where $\mathbf{Z}_s = \{\mathbf{z}_{ij}\}_{j \in s; i=1,2,\dots,d}$ and $\mathbf{Z}_r$ is its analogous on non-sampled units.
6. Update $\sigma_u^{2(r+1)} = d^{-1}(\text{trace}(\mathbf{T}) + \mathbf{u}^{t(r+1)} \mathbf{u}^{(r+1)})$.
7. Return to step 2 and repeat the procedure until convergence of the loglikelihood.

It is recommended that different starting values are considered to check whether the procedure got trapped in some local maximum. Good starting values for $\mathbf{b}$ and $\mathbf{u}$ are those coming from a fixed logistic model, for σ_u^2 is the variance of these $\mathbf{u}$ estimates. The choice for λ , on the other side, seems less critical; in the simulation studies a value of 1 has been used. Using the results of the algorithm we obtain estimates of parameters to be used in (4).

To calculate the mean squared error (MSE) of the proposed estimator we use an approximation that is an application of Prasad and Rao (1990) analytic form of MSE after a Taylor linearization of the model. In particular, let $\hat{\mathbf{y}}^{bench} = (\hat{Y}_1^{bench}, \dots, \hat{Y}_d^{bench})^t$,

$$MSE(\hat{\mathbf{y}}^{bench}) \cong \mathbf{G}_1(\sigma_u^2) + \mathbf{G}_2(\sigma_u^2) + \mathbf{G}_3(\sigma_u^2) + \mathbf{G}_4(\sigma_u^2) \quad (6)$$

where, if matrices $\mathbf{X}_r$ and $\mathbf{B}_r$ are the analogous of $\mathbf{X}_s$ and $\mathbf{B}_s$ for nonsampled units, and $\mathbf{A}_r$ is a $d \times (N - n)$ matrix such that $\mathbf{Y} = (Y_1, \dots, Y_d)^t = \mathbf{A}_s \mathbf{y}_s + \mathbf{A}_r \mathbf{y}_r$ and $\mathbf{y}_r$ is the analogous of $\mathbf{y}_s$ for nonsampled units,

$$\begin{aligned} \mathbf{G}_1(\sigma_u^2) &= \mathbf{A}_r \mathbf{H}_r \mathbf{Z}_r \mathbf{T} \mathbf{Z}_r^t \mathbf{H}_r^t \mathbf{A}_r^t, \\ \mathbf{G}_2(\sigma_u^2) &= \mathbf{A}_r \mathbf{H}_r (\mathbf{X}_r - \mathbf{Z}_r \mathbf{T} \mathbf{Z}_s^t \mathbf{B}_s \mathbf{X}_s) (\mathbf{X}_s^t \mathbf{V}_s^{-1} \mathbf{X}_s)^{-1} (\mathbf{X}_r^t - \mathbf{X}_s^t \mathbf{B}_s \mathbf{Z}_s \mathbf{T} \mathbf{Z}_r^t) \mathbf{H}_r^t \mathbf{A}_r^t, \\ \mathbf{G}_3(\sigma_u^2) &= \text{var}(\hat{\sigma}_u^2) [\mathbf{h}_k^{t'} \mathbf{Z}_s^t \mathbf{B}_s \mathbf{V}_s \mathbf{B}_s \mathbf{Z}_s \mathbf{h}_l^t]_{k,l=1,\dots,d}, \text{ and} \\ \mathbf{G}_4(\sigma_u^2) &= \mathbf{A}_r \mathbf{B}_r \mathbf{A}_r^t, \end{aligned}$$

where $\mathbf{H}_r = \text{diag}[\hat{p}_{ij}(1 - \hat{p}_{ij})]_{j \in U_i \setminus s_i}$ for $i = 1, \dots, d$, $\mathbf{V}_s = \sigma_u^2 \mathbf{Z}_s \mathbf{Z}_s^t + \mathbf{B}_s^{-1}$, $\mathbf{h}_k^t = (h_{k1}, \dots, h_{kd}) = \mathbf{a}_{rk}^t \mathbf{Z}_r \mathbf{T}$, $\mathbf{a}_{rk}^t$ denotes the k -th row of $\mathbf{A}_r \mathbf{H}_r$, and $\mathbf{h}_k^{t'} = (\partial h_{k1} / \partial \sigma_u^2, \dots, \partial h_{kd} / \partial \sigma_u^2)$. Observe that the approximation in (6) depends on how well the linear mixed model in the linearized response variable approximates the original generalized linear mixed model, and how good is the first-order Taylor approximation. An estimator for $MSE(\hat{\mathbf{y}}^{bench})$ can then be obtained by $mse(\hat{\mathbf{y}}^{bench}) = \hat{\mathbf{G}}_1(\hat{\sigma}_u^2) + \hat{\mathbf{G}}_2(\hat{\sigma}_u^2) + \hat{\mathbf{G}}_3(\hat{\sigma}_u^2) + \hat{\mathbf{G}}_4(\hat{\sigma}_u^2)$, where hats denote substitution with estimated values from the output of the algorithm. For more details in the un-benchmarked case see Saei and Chambers (2003) and González-Manteiga *et al.* (2007).

4 Simulation study and some concluding remarks

We conduct a simulation study (i) to investigate the effect of benchmarking on the performance of the estimators, (ii) to compare the performance of estimators based on linear mixed models and that of those based on logistic mixed models to estimate small area counts and (iii) to compare the performance of estimators based on alternative algorithms that estimate logistic mixed models. To these ends we start

from real data coming from the Italian Multipurpose household survey. The sample from the Region Umbria has been used as our finite population made of $N = 4,879$ units divided into $d = 51$ small area given by municipalities. To simulate a complex sampling design, units are randomly assigned to four strata of dimension 500, 800, 1,500 and 2,079, respectively. The response variable is simulated and constructed using age, sex and stratum membership as covariates in the following linear predictor

$$\eta_{ij} = 0.4\text{str}_{1ij} + 0.1\text{str}_{2ij} - 1.7\text{str}_{3ij} - 2\text{str}_{4ij} - 0.16\text{age}_{ij} + 0.9\text{sex}_{ij} + u_i,$$

with $d = 51$ random effects u_i drawn from a zero mean normal variable with variance 3. This model allows to build an informative stratification. A bernoulli experiment is then drawn once with probabilities $p_{ij} = \exp(\eta_{ij}) / (1 + \exp(\eta_{ij}))$ to obtain binary response variable y_{ij} . An adjustment is then made to have at least 3 values equal to 1 for each small area. This allows to use overall measures of performance based on relative bias (see later in this section) to ease the comparison of the different estimators. The overall mean of the response variable is 7.8%.

Five hundreds stratified random samples of dimension $n = 500$ have been selected using a disproportionate allocation. In particular, we select 112 elements from the first stratum, 120 from the second, 112 from the third and 156 from the fourth. This gets sampling rates within strata of 22.4%, 15.0%, 7.5% and 7.5%, respectively. The benchmark is the Horvitz-Thompson estimator of the total.

For each sample several small area estimators are computed. All use models based on the same two covariates – age and sex – and a normal random area effect. Firstly, a set of three estimators based on a unit level linear mixed model are considered: the classical Battese *et al.* (1988) estimator – BHF – that is not benchmarked, then its ratio adjustment as in equation (1) – BHFr – and, finally, the Pfeffermann and Barnard (1991) estimator – PB – as in equation (2). Secondly, a set of seven estimators based on a logistic mixed model is considered. It is well known that, differently from the normal case, maximum likelihood estimates of the parameters and of the variance components in a logistic-normal mixed model is hindered by the presence of a d dimensional integral, so that direct calculation is intractable and well known computational issues arise. Therefore, more attention should be paid when computing these estimators by, possibly, comparing results coming from different softwares that use different algorithms. Consequently, we consider the estimator RPQL obtained using the function `glmPQL` of R, and its ratio adjustment RPQLr, the estimator RML obtained using the function `glmML` of R, and its ratio adjustment RMLr. Finally, three estimators are considered that are obtained using the Saei and Chambers (2003) algorithm: the first one is the un-benchmarked one – SC –, then its ratio adjustment – SCr – and, finally, the proposed self-benchmarked estimator – SCsb.

Let $\hat{Y}_i^r$ be the value of a small area estimator of the total at replicate r , for $r = 1, \dots, 500$; then the following evaluation criteria have been computed and reported in Table 1:

- % Relative Bias for small area i : $RB_i = \frac{1}{R} \left[\sum_{r=1}^R \frac{\hat{Y}_i^r - Y_i}{Y_i} \right] 100$;
- Average Absolute RB: $AARB = \frac{1}{d} \sum_{i=1}^d |RB_i|$;
- Maximum Absolute RB: $MARB = \max_i |RB_i|$;
- % Relative Root MSE for small area i : $RRMSE_i = \sqrt{\frac{1}{R} \left[\sum_{r=1}^R \left(\frac{\hat{Y}_i^r - Y_i}{Y_i} \right)^2 \right]} 100$;
- Average RRMSE: $ARRMSE = \frac{1}{d} \sum_{i=1}^d RRMSE_i$;
- Maximum RRMSE: $MRRMSE = \max_i RRMSE_i$.

Table 1 Simulation results: Average and Maximum Absolute % Relative Bias, Average and Maximum % Relative Root Mean Squared Error for all estimators.

	BHF	BHFr	PB	RPQL	RPQLr	RML	RMLr	SC	SCr	SCsb
AARB	47.8	40.8	39.3	56.3	48.4	51.4	39.0	49.5	42.5	39.5
MARB	221.5	185.8	178.3	291.3	246.1	232.8	173.9	230.9	192.9	168.3
ARRMSE	71.5	63.3	65.7	80.9	70.6	80.8	64.2	69.9	61.7	63.1
MRRMSE	243.3	207.9	205.4	327.4	281.4	266.1	208.3	255.8	218.3	199.6

First we can note that estimators with the benchmarking property all have a smaller bias and MSE than their corresponding un-benchmarked ones. This can be explained by the fact that benchmarking model based estimates to the unbiased Horvitz-Thompson estimator helps in decreasing the bias of the estimates. Secondly, we compare the performance of estimators based on linear mixed models and that of those based on a logistic mixed model. In fact, if in classical statistics using normal models for a binary variable instead of the appropriate logistic ones is usually deprecated, for small area statistics there is not a clear-cut evidence of the superiority in terms of performance of the latter over the former. In fact, in this case the performance is similar, both in terms of bias and error. However, note that normal based models produce negative estimates for the areas with a low population count. This has happened 60 times for BHF and BHFr and 268 times for PB in the 500×51 estimates. Thirdly, when comparing estimators based on a logistic mixed model computed using different algorithms, we can note that the Saei and Chambers (2003) algorithm has a good performance with respect to the two others, both in terms of bias and error. Finally, it can be noted that the self-benchmarked proposed estimator shows a very good performance in terms of overall bias, and it also decreases significantly the bias for the most biased area – it has the smallest MARB value. This is true also for the MSE.

Decrease of bias shown by the benchmarked estimators can be likely explained by the fact that the sampling design is, in this case, informative. It will be interesting,

therefore, to evaluate the effect of weighting the likelihood equations provided in Section 3 to account for the sampling design. In addition, we envision the study of the effect of benchmarking to GREG-type estimators based on auxiliary information that may or may not coincide with that used in the small area models. Also, it will be of interest to compare the performance of the different estimators for different values of the population proportion of the attribute and of the variance of the area effects. Finally, evaluation of the analytic MSE estimator is also on the agenda.

Acknowledgements Work supported by the project PRIN 2007 *Efficient use of auxiliary information at the design and at the estimation stage of complex surveys: methodological aspects and applications for producing official statistics*.

References

1. Battese, G.E., Harter, R.M., Fuller, W.A., An error-components model for prediction of county crop areas using survey and satellite data. *J Am Stat Ass.* **83**, 28–36 (1988).
2. McCulloch, C.E., Searle, S.R., *Generalized, Linear, and Mixed Model*, Wiley, New York (2001)
3. González-Manteiga, W., Lombardía, M.J., Molina, I., Morales, D., Santamaría, L., Estimation of the mean squared error of predictors of small area linear parameters under a logistic mixed model. *Comput Stat Data An.* **51**, 2720–2733 (2007)
4. Pfeffermann, D., Barnard, C.H., Some New Estimators for Small-Area Means With Application to the Assessment of Farmland Values, *J Bus Econ Stat.* **90**, 73–84 (1991)
5. Prasad, N.G.N., Rao, J.N.K., The estimation of the mean squared error of small area estimators. *J Am Stat Ass.* **85**, 163–171 (1990)
6. Rao, J.N.K., *Small area estimation*. Wiley, New Jersey (2003)
7. Saei, A., Chambers, R., *Small Area Estimation Under Linear and Generalized Linear Model With Time and Area Effects*. Working Paper M03/15, Southampton Statistical Sciences Research Institute, University of Southampton (2003).
8. Wang, J., Fuller, W.A., Qu, Y., Small area estimation under a restriction. *Surv Methodol.* **34**, 29–36 (2008)
9. You, Y., Rao, J.N.K., A pseudo-Empirical Best Linear Unbiased Prediction Approach to Small Area Estimation Using Survey Weights. *Can J Stat.* **30**, 431–439 (2002).
10. You, Y., Rao, J.N.K., Dick, P., Benchmarking hierarchical bayes small area estimators in the Canadian census undercoverage estimation. *Statistics in Transition*, **6**, 631–640 (2004)