STATISTICS IN TRANSITION new series, September 2023
Vol. 24, No. 4, pp. 169–178, https://doi.org/10.59170/stattrans-2023-058
Received – 15.09.2022; accepted – 23.03.2023

An empirical study of hierarchical Bayes small area estimators using different priors for model variances

Yong You[1]

Abstract

In this paper, we study hierarchical Bayes (HB) estimators based on different priors for small area estimation. In particular, we use inverse gamma and flat priors for variance components in the HB small area models of You and Chapman (2006) and You (2021). We evaluate the HB estimators through a simulation study and real data analysis. Our results indicate that using the inverse gamma prior for the variance components in the HB models can be very effective.

Key words: CPO, flat prior, inverse gamma prior, relative error, variance component.

1. Introduction

Small area estimation is very popular and important in survey data analysis due to growing demand for reliable small area estimates. Model-based estimates have been widely used to provide reliable indirect estimates. Various area level models have been proposed in the literature to improve direct survey estimates, see Rao and Molina (2015). In this paper, we use the well-known Fay-Herriot model (Fay and Herriot, 1979) as a basic model and present the Fay-Herriot model in hierarchical Bayes (HB) framework of You and Chapman (2006) and You (2016, 2021). The Fay-Herriot model has two components, namely a sampling model for the direct survey estimates and a linking model for small area parameters of interest. The sampling model assumes that a direct estimator y_i is design unbiased for a small area parameter θ_i such that

$$y_i = \theta_i + e_i,\ i = 1,\ldots,m, \tag{1}$$

where e_i is the sampling error and m is the number of small areas. It is customary to assume that e_i's are independently distributed normal random variables with mean $E(e_i|\theta_i) = 0$ and variance $Var(e_i|\theta_i) = \sigma_i^2$. The linking model assumes that the small

[1] ESMD, Statistics Canada, Ottawa, K1A 0T6, Canada. E-mail: yongyou@statcan.gc.ca.
ORCID: https://orcid.org/0009-0000-8030-1484.

area parameter θ_i is related to area level auxiliary variables $x_i = (x_{i1}, \dots, x_{ip})'$ through a linear regression model

$$\theta_i = x_i{}'\beta + v_i,\ i = 1,\dots,m, \tag{2}$$

where $\beta = (\beta_1, \dots, \beta_p)'$ is a $p \times 1$ vector of regression coefficients and v_i's are random effects assumed to be independent and normally distributed with $E(v_i) = 0$ and $Var(v_i) = \sigma_v^2$. The model variance σ_v^2 is unknown and needs to be estimated. Combining models (1) and (2) leads to a linear mixed area level model given as

$$y_i = x_i{}'\beta + v_i + e_i,\ i = 1,\dots,m. \tag{3}$$

Model (3) involves both design-based random errors e_i and model-based random effects v_i. For the Fay-Herriot model, the sampling variance σ_i^2 is assumed to be known in model (3). This is a very strong assumption. Generally smoothed estimators of the sampling variances are used in the Fay-Herriot model and then treated as known. Alternatively, the sampling variance σ_i^2 can be modelled together with the small area parameter θ_i. Let s_i^2 denote a direct estimator for σ_i^2. We consider a commonly used model for s_i^2 as $d_i s_i^2 \sim \sigma_i^2 \chi_{d_i}^2$, where $d_i = n_i - 1$ and n_i is the sample size for the i-th area. We combine the sampling variance model $d_i s_i^2 \sim \sigma_i^2 \chi_{d_i}^2$ with the small area model (3) to construct an integrated model in the HB framework. The integrated model borrows strength for small area estimates and sampling variance estimates simultaneously. This integrated HB modelling approach has been widely used in practice, for example, see You and Chapman (2006), Dass, Maiti, Ren and Sinha (2012), Sugasawa, Tamae and Kubokawa (2017), Ghosh, Myung and Moura (2018), Hidiroglou, Beaumont and Yung (2019) and You (2008, 2021).

In Section 2, we present two HB small area models and consider two priors for variance components, namely inverse gamma (IG) prior and flat prior. In Section 3, we conduct a simulation study to evaluate the impact of priors on small area estimation. In Section 4, we apply the models to a real data application. And in Section 5, we offer some concluding remarks.

2. Hierarchical Bayes small area models

In this section, we present two HB models with sampling variance modelling. The first model is considered in You and Chapman (2006), in which an inverse gamma model is used for the sampling variance σ_i^2 with known vague values. The second model is considered in You (2016, 2021), where a log-linear random error model is used for σ_i^2.

HB Model 1: You-Chapman Model (You and Chapman, 2006), denoted as YCM:

- $y_i | \theta_i, \sigma_i^2 \sim ind\ N(\theta_i, \sigma_i^2)$, $i = 1,\dots,m$;
- $d_i s_i^2 | \sigma_i^2 \sim ind\ \sigma_i^2 \chi_{d_i}^2$, $d_i = n_i - 1$, $i = 1,\dots,m$;

- $\theta_i|\beta,\sigma_v^2 \sim ind\ N(x_i{}'\beta,\sigma_v^2)$, $i = 1,\ldots,m$;
- $\sigma_i^2 \sim IG(a_i,b_i)$, where $a_i = 0.0001$, $b_i = 0.0001$, $i = 1,\ldots,m$;
- priors for unknown parameters:$\pi(\beta) \propto 1, \pi(\sigma_v^2) \sim IG(a_v,b_v)$, where a_v, b_v are chosen to be very small constants (0.0001) to reflect vague knowledge on σ_v^2.

The full conditional distributions for the Gibbs sampling procedure under YCM can be found in You and Chapman (2006).

HB Model 2: You (2016, 2021) log-linear model on sampling variances, denoted as YLLM:

- $y_i|\theta_i,\sigma_i^2 \sim ind\ N(\theta_i,\sigma_i^2)$, $i = 1,\ldots,m$;
- $d_i s_i^2|\sigma_i^2 \sim ind\ \sigma_i^2\chi_{d_i}^2,\ d_i = n_i - 1$, $i = 1,\ldots,m$;
- $\theta_i|\beta,\sigma_v^2 \sim indN(x_i{}'\beta,\sigma_v^2)$, $i = 1,\ldots,m$;
- $log(\sigma_i^2) \sim N(\delta_1 + \delta_2\ log(n_i),\tau^2)$, $i = 1,\ldots,m$;
- priors for unknown parameters:$\pi(\beta) \propto 1$, $\pi(\delta_1,\delta_2) \propto 1$, $\pi(\sigma_v^2) \sim IG(a_v,b_v)$, $\pi(\tau^2) \sim IG(a_\tau,b_\tau)$, where $a_v,\ b_v,\ a_\tau,\ b_\tau$ are chosen to be very small constants (say, 0.0001).

The full conditional distributions for the Gibbs sampling procedure under YLLM are given in the Appendix.

For both YCM and YLLM, we use IG priors with very small constant parameters for the variance components σ_v^2 and τ^2. Ghosh, Myung and Moura (2018) used an IG prior with some fixed values for the model variance σ_v^2. IG prior is a proper prior and conditionally conjugate for the variance components. IG prior is widely used in Bayesian literature (e.g. Gelman, Carlin, Stern and Rubin, 2004) and Bayesian software packages (e.g. WinBUGS, Lunn, Thomas, Best and Spiegelhalter, 2000). Alternatively, flat priors $\pi(\sigma_v^2) \propto 1$ and $\pi(\tau^2) \propto 1$ can be used for the model variances σ_v^2 and τ^2 in the YCM and YLLM models. Flat prior is used as a non-informative prior in the literature (e.g. Gelman, 2006). You (2021) compared the models of YCM and YLLM with the model of Sugasawa, Tamae and Kubokawa (2017) using flat priors on the variance components. In this paper we use YCM and YLLM as two studying models and compare the HB estimators using IG and flat priors through simulation study and real data analysis.

3. Simulation study

In this section, we estimate model variance σ_v^2 and small area means through a simulation study. We generate $\theta_i = \boldsymbol{x}'_i\beta + v_i = \beta_0 + x_i\beta_1 + v_i$ with $\beta_0 = 3.5$ and $\beta_1 = 1.5$ fixed through the simulation. The single covariate x_i is generated from an exponential distribution with mean equal to 1, and then fixed for the simulation study. Random effect v_i is generated from $v_i \sim N(0,\sigma_v^2)$. Following Lahiri and Rao (1995) and

Rivest and Vandal (2002), we let the number of small areas m = 30. These 30 areas are divided into five groups with different sampling variances. The true sampling variance is set at σ_i^2 = 1, 0.75, 0.5, 0.25, and 0.1 for each grouped areas, with the corresponding sample size n_i= 4, 6, 8, 10 and 12. That is, for areas from 1 to 6, σ_i^2 = 1 (i = 1,…,6) and the corresponding sample size n_i= 4 for each area (i = 1,…,6). For areas 7 to 12, σ_i^2 = 0.75 (i = 7,…,12) and the corresponding sample size n_i= 6 for each area (i = 7,…,12). And so on for other areas. We consider three choices of the model variance: the true σ_v^2 is set to be 1, 0.5 and 0.1, respectively. The direct sampling variance estimate is generated as $s_i^2 = (d_i)^{-1}\sigma_i^2\chi_{d_i}^2$, where $d_i = n_i - 1$ (e.g. Ghosh, Myung and Moura, 2018; You, 2021). For each case, we perform 5000 simulation runs. For each run, the Gibbs sampling procedure consists of 1000 burn-in period and 5000 more iterations for each simulation run.

We first compare the estimates of the model variance σ_v^2 based on YCM and YLLM using IG and flat priors on σ_v^2. Table 1 presents the estimates of σ_v^2 when the true σ_v^2 is 1, 0.5 and 0.1. It is clear from Table 1 that both YCM and YLLM lead to almost unbiased estimates of the model variance under IG prior. However, when flat prior is used, both YCM and YLLM lead to over-estimation of the model variance. The over estimation is substantially large when the true model variance is small. For example, if the true σ_v^2 is 1, under flat prior, both YCM and YLLM lead to about 22% over-estimation; if the true σ_v^2 is 0.1, the over-estimation could be more than 100%. The result in Table 1 indicates that IG prior performs much better than the flat prior for the model variance estimation.

Table 1: Estimates of model variance under YCM and YLLM using IG and flat priors

True σ_v^2	YCM		YLLM	
	IG prior	Flat prior	IG prior	Flat prior
1	1.025	1.217	1.027	1.228
0.5	0.518	0.683	0.512	0.672
0.1	0.119	0.251	0.096	0.225

To compare the small area HB estimators, we consider the average absolute relative bias (ARB) for the HB estimator $\hat{\theta}_i$ of the simulated small area mean θ_i as $\overline{ARB} = (\sum_{i=1}^m ARB_i)/m$, where

$$ARB_i = \left|\frac{1}{R}\sum_{r=1}^R \frac{(\hat{\theta}_i^{(r)} - \theta_i^{(r)})}{\theta_i^{(r)}}\right|,$$

and $\hat{\theta}_i^{(r)}$ is the HB estimate and $\theta_i^{(r)}$ is the true mean based on the r-th simulated sample, R = 5000. The estimated average coefficient of variation (ACV) is computed as $\overline{ACV} = (\sum_{i=1}^m CV_i)/m$, where

$$CV_i = \frac{1}{R}\sum_{r=1}^R CV_i^{(r)} \text{and } CV_i^{(r)} = \frac{\sqrt{var(\hat{\theta}_i^{(r)})}}{\hat{\theta}_i^{(r)}}$$

where $var(\hat{\theta}_i^{(r)})$ is estimated posterior variance of the HB estimator $\hat{\theta}_i^{(r)}$. We also compare the average simulation relative root MSE (RRMSE), and the RRMSE is computed as $\overline{RRMSE} = (\sum_{i=1}^{m} RRMSE_i)/m$, where

$$RRMSE_i = \frac{1}{R}\sum_{r=1}^{R} RRMSE_i^{(r)}, \text{ and } RRMSE_i^{(r)} = \frac{\sqrt{(\hat{\theta}_i^{(r)} - \theta_i^{(r)})^2}}{\theta_i^{(r)}}.$$

Table 2 presents the comparison results of ARB, ACV and RRMSE under models YCM and YLLM using IG and flat priors. The HB estimator $\hat{\theta}_i$ should be unbiased for the small area parameter θ_i following the conditional posterior distribution of θ_i given in the Appendix. When the true σ_v^2= 1, the average ARB is around 1.7% to 1.8% for both models YCM and YLLM, and the average ARB becomes much smaller when the true σ_v^2= 0.1. The results of ARB also indicate that the posterior HB estimators are unbiased for the small area parameter θ_i under both the IG and flat priors. However, both YCM and YLLM have smaller average CVs and RRMSE using IG prior, and particularly, using the flat prior leads to much larger average CVs for both YCM and YLLM. For example, when the true σ_v^2= 0.1, the average CV using flat prior is 8.15% under YCM and 7.88% under YLLM, the average CV using IG prior is 5.87% under YCM and 5.56% under YLLM. The results in Table 2 indicate that both IG and flat priors lead to similar performance of the HB estimator. However, using IG prior in both YCM and YLLM leads to smaller CV and RRMSE for the HB estimator. The results in Table 2 also demonstrate that YLLM performs slightly better than YCM in terms of CV and RRMSE. This simulation result is consistent with the results shown in You (2021).

Table 2: Comparison of average ARB%, average CV (ACV%) and RRMSE%

Specification	$\hat{\theta}_i^{HB}$	YCM		YLLM	
		IG prior	Flat prior	IG prior	Flat prior
σ_v^2=1	ARB	1.83	1.78	1.73	1.77
	ACV	12.31	12.69	12.07	12.48
	RRMSE	10.49	10.46	10.24	10.25
σ_v^2=0.5	ARB	1.15	1.14	1.14	1.16
	ACV	9.99	10.93	9.87	10.71
	RRMSE	8.76	8.87	8.62	8.78
σ_v^2=0.1	ARB	0.22	0.35	0.28	0.31
	ACV	5.87	8.15	5.56	7.88
	RRMSE	5.68	5.72	5.45	5.56

4. Data analysis

In this section, we compare YCM and YLLM using IG and flat priors through a real data application. Following Hidiroglou, Beaumont and Yung (2019) and You (2021), we apply both the YCM and YLLM to a Canadian Labour Force Survey (LFS) data and compare the HB estimates of unemployment rates with the census estimates. We apply both the YCM and YLLM to the May 2016 unemployment rate estimates for the Census Metropolitan Areas (CMAs) and Census Agglomerations (CAs), and then we compare the HB estimates and the direct estimates with the census estimates. For both the YCM and YLLM, the local area employment insurance monthly beneficiary rate is used as an auxiliary variable in the model, same as in Hidiroglou, Beaumont and Yung (2019) and You (2021). We compute the absolute relative error (ARE) of the direct and HB estimates with respect to the census estimates for each CMA/CA as follows:

$$\text{ARE}_i = \left|\frac{\theta_i^{Census} - \theta_i^{Est}}{\theta_i^{Census}}\right|,$$

where θ_i^{Est} is the direct or HB estimate and θ_i^{Census} is the corresponding census value of the LFS unemployment rate. Then we take the average of AREs over CMA/CAs. For CV, we compute the average CVs of the direct and model-based estimates. We prefer a model with smaller ARE and smaller CV. We first apply both models YCM and YLLM to all the 117 CMA/CAs with sample size ≥ 2, and then apply to 92 CMA/CAs with sample size ≥ 5, and 79 CMA/CAs with sample size ≥ 7, respectively. Table 3 presents the average ARE and the corresponding average CV (in brackets) for the YCM and YLLM using IG and flat priors.

Table 3: Comparison of average ARE and average CV (in parenthesis)

CMA/CAs	Direct LFS	YCM IG prior	YCM Flat prior	YLLM IG prior	YLLM Flat prior
Average over 117 CMA/CAs (sample size ≥ 2)	0.263 (0.329)	0.149 (0.127)	0.148 (0.136)	0.135 (0.116)	0.135 (0.123)
Average over 92 CMA/CAs (sample size ≥ 5)	0.216 (0.262)	0.132 (0.115)	0.132 (0.121)	0.126 (0.112)	0.125 (0.117)
Average over 79 CMA/CAs (sample size ≥ 7)	0.181 (0.232)	0.123 (0.112)	0.122 (0.115)	0.119 (0.109)	0.118 (0.114)

It is clear from Table 3 that both the YCM and YLLM improve the direct LFS estimates substantially by reducing the average ARE and CV, and YLLM performs better than YCM. For both the YCM and YLLM, using IG and flat priors leads to about the same ARE. However, using IG prior in the models can lead to smaller CV as shown in Table 3. For example, for YCM, the average CV over 117 CMA/CAs is 0.127 under IG prior and 0.136 under flat prior. For YLLM, the average CV is 0.116 under IG prior and 0.123 under flat prior. Thus, in our application, for the point estimation, there is no difference using IG or flat prior. However, using IG prior in both the YCM and YLLM can lead to smaller CV. This result is consistent with the simulation result reported in Table 2.

Now we present a Bayesian model comparison using conditional predictive ordinate (CPO) for both the YCM and YLLM with IG and flat priors. CPOs are the observed likelihoods based on the cross-validation predictive density $f(y_i|y_{obs(i)})$. We compute the CPO value $\text{CPO}_i = f(y_{i,obs}|y_{obs(i)})$ for each observed data point $y_{i,obs}$, and larger CPO_i indicates a better model fit. For model choice, we can compute the CPO ratio of model A against model B. If this ratio is greater than 1, then $y_{i,obs}$ supports model A. We compute the CPO ratios for YCM IG/Flat and YLLM IG/Flat, and count the number of the CPO ratios that are larger than 1. We can also plot the CPO values or summarize the CPO values by taking the average of the estimated CPOs. For more detail on applications of CPO, see for example, Gilks, Richardson and Spiegelhalter (1996), You and Rao (2000), Molina, Nandram and Rao (2014) and You (2021). Table 4 presents the average CPO values and # of CPO ratios larger than 1 over the 117, 92 and 79 CMA/CAs for the YCM and YLLM with IG vs flat priors.

Table 4: Summary of the average CPO values and # of CPO ratios larger than 1

CMA/CAs	YCM			YLLM		
	IG prior	Flat prior	# of CPO ratio >1	IG prior	Flat prior	# of CPO ratio >1
117	0.1228	0.1222	76	0.1253	0.1242	73
92	0.1412	0.1392	61	0.1419	0.1398	59
79	0.1516	0.1491	50	0.1526	0.1517	52

It is clear from Table 4 that for both the YCM and YLLM models, IG prior has larger CPO values than flat prior, and more than half of the observations support the model with IG prior. For example, over 117 CMA/CAs, for YCM, the average CPO under IG prior is 0.1228, and 0.1222 under flat prior, and 76 observations support YCM with IG prior. For YLLM, the average CPO is 0.1253 under IG prior and 0.1242 under flat prior, and 73 observations support YLLM with IG prior. We also note that the YLLM model is better than the YCM with larger CPO values for both IG and flat priors.

5. Concluding remarks

In this paper, we have studied the performance of HB small area models using IG and flat priors on variance components through a simulation study and real data analysis. Our results indicate that both the YCM (You and Chapman 2006) and YLLM (You, 2021) models using IG and flat priors perform very well. However, using IG prior in both the YCM and YLLM leads to slightly better results (smaller CV) and better model fit. Our simulation study and real data analysis demonstrate that proper IG prior should be used in the HB small area models for variance components. Flat prior for the model variance should be avoided as using the flat prior has no advantage over the IG prior with respect to the final HB estimates. For future work, informative priors such as IG prior with parameter values based on previous survey data could also be used in the model to improve the HB small area estimators. It is also interesting to compare the HB estimators using informative priors.

Acknowledgments

I would like to thank the Editor and two referees for their careful reading of the manuscript and suggestions to improve the paper.

References

Dass, S. C., Maiti,T., Ren, H. and Sinha, S., (2012). Confidence interval estimation of small area parameters shrinking both means and variances. *Survey Methodology*, 38, pp. 173–187.

Fay, R. E., 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, pp. 269–277.

Gelman, A., (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1, pp. 515–533.

Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B., (2004). *Bayesian Data Analysis*. 2nd Edition, Chapman & Hall/CRC.

Ghosh, M., Myung, J. and Moura, F. A. S., (2018). Robust Bayesian small area estimation. *Survey Methodology*, 44, pp. 101–115.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J., (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC.

Hidiroglou, M. A., Beaumont, J.-F. and Yung, W., (2019). Development of a small area estimation system at Statistics Canada. *Survey Methodology*, 45, pp. 101–126.

Lahiri, P., Rao, J. N. K., (1995). Robust estimation of mean squared error of small area estimators. *Journal of the American Statistical Association*, 82, pp. 758–766.

Lunn, D. J., Thomas, A., Best, N. and Spiegelhalter, D. J., (2000). WinBUGS – A Bayesian modeling framework: concepts, structure and extensibility. *Statistics and Computing*, 10, pp. 325–337.

Molina, I, Nandram, B. and Rao, J. N. K., (2014). Small area estimation of general parameters with application to poverty indicators: A hierarchical Bayes approach, *Annals of Applied Statistics*, 8, pp. 852–885.

Rao, J. N. K., Molina, I., (2015). *Small Area Estimation*, 2nd Edition. John Wiley & Sons, New York.

Rivest, L. P., Vandal, N. (2002). Mean squared error estimation for small areas when the small area variances are estimated. *Proceedings of the International Conference on Recent Advances in Survey Sampling*, July 10–13, 2002, Ottawa, Canada.

Sugasawa, S., Tamae, H. and Kubokawa, T., (2017). Bayesian estimators for small area models shrinking both means and variances. *Scandinavian Journal of Statistics*, 44, pp. 150–167.

You, Y., (2008). An integrated modeling approach to unemployment rate estimation for sub-provincial areas of Canada. *Survey Methodology*, 34, pp. 19–27.

You, Y., (2016). *Hierarchical Bayes sampling variance modeling for small area estimation based on area level models with applications*. Methodology branch working paper, ICCSMD-2016-03-E, Statistics Canada, Ottawa, Canada.

You, Y., (2021). Small area estimation using Fay-Herriot model with sampling variance smoothing and modeling. *Survey Methodology*, 47, pp. 361–370.

You, Y., Chapman, B., (2006). Small area estimation using area level models and estimated sampling variances. *Survey Methodology*, 32, pp. 97–103.

You, Y., Rao, J. N. K., (2000). Hierarchical Bayes estimation of small area means using multi-level models. *Survey Methodology*, 26, pp. 173–181.

Appendix

Full conditional distributions and sampling procedure for the YLLM model:

- $[\theta_i|y,\beta,\sigma_i^2,\sigma_v^2] \sim N\left(\gamma_i y_i + (1-\gamma_i)x_i{}'\beta, \gamma_i\sigma_i^2\right)$, where $\gamma_i = \frac{\sigma_v^2}{\sigma_v^2+\sigma_i^2}$, $i = 1,\ldots,m$;
- $[\beta|y,\theta,\sigma_i^2,\sigma_v^2] \sim N_p((\sum_{i=1}^m x_i\, x_i{}')^{-1}(\sum_{i=1}^m x_i\theta_i), \sigma_v^2(\sum_{i=1}^m x_i x_i{}')^{-1})$;
- $[\sigma_v^2|y,\theta,\beta,\sigma_i^2] \sim I\,G\left(a_v + \frac{m}{2}, b_v + \frac{1}{2}\sum_{i=1}^m(\theta_i - x_i{}'\beta)^2\right)$;
- $\left[\sigma_i^2\middle|y,\theta,\beta,\sigma_v^2,\delta,\tau^2\right] \propto f(\sigma_i^2)\cdot h(\sigma_i^2)$, where $f(\sigma_i^2)$ and $h(\sigma_i^2)$ are $f(\sigma_i^2) \sim I\,G\left(\frac{d_i+1}{2}, \frac{(y_i-\theta_i)^2+d_i s_i^2}{2}\right)$, and $h(\sigma_i^2) = exp\left(-\frac{(log(\sigma_i^2)-z'_i\delta)^2}{2\tau^2}\right)$;
- $[\delta|y,\theta,\beta,\sigma_i^2,\sigma_v^2,\tau^2] \sim N_2\left((\sum_{i=1}^m z_i z'_i)^{-1}(\sum_{i=1}^m z_i\, log(\sigma_i^2)), \tau^2(\sum_{i=1}^m z_i z'_i)^{-1}\right)$;
- $[\tau^2|y_i,\theta\ \ ,\beta,\sigma_i^2,\sigma_v^2,\delta] \sim I\,G\left(a_\tau + \frac{m}{2}, b_\tau + \frac{1}{2}\sum_{i=1}^m(log(\sigma_i^2) - z'_i\delta)^2\right)$.

We use Metropolis-Hastings rejection step to update σ_i^2:

(1) Draw ${\sigma_i^2}^*$ from $IG\left(\frac{d_i+1}{2}, \frac{(y_i-\theta_i)^2+d_i s_i^2}{2}\right)$;

(2) Compute the acceptance probability $\alpha\left({\sigma_i^2}^*, {\sigma_i^2}^{(k)}\right) = min\left\{h({\sigma_i^2}^*)/h({\sigma_i^2}^{(k)}),1\right\}$;

(3) Generate u from Uniform(0,1), if $u < \alpha\left({\sigma_i^2}^*, {\sigma_i^2}^{(k)}\right)$, the candidate ${\sigma_i^2}^*$ is accepted, ${\sigma_i^2}^{(k+1)} = {\sigma_i^2}^*$; otherwise ${\sigma_i^2}^*$ is rejected, and set ${\sigma_i^2}^{(k+1)} = {\sigma_i^2}^{(k)}$.