Accurate Confidence Interval Estimation of Small Area Parameters under the Fay-Herriot Model

L. Diao* T. Maiti* J.D. Opsomer †

March 5, 2008

Abstract

Small area estimation has long been a popular and important research topic due to its growing demand in public and private sectors. We consider here the basic area level model, popularly known as Fay-Herriot model. While much of current research is predominantly focused on finding second order correct mean squared prediction errors and their estimators, we concentrate on developing confidence intervals for the small area parameters that are second order correct. The findings are illustrated with a simulation study.

1 Introduction

The purpose of this note is to develop confidence intervals for small area means when only area level summary statistics are available. Small-area estimation is important in survey applications, particularly in those fields of official statistics where legislative mandates require socioeconomic estimates within jurisdictions narrower than can accurately be described by direct estimates from national surveys. Prediction based on the **?**) model (FH) is one of the most popular techniques in small area

*Department of Statistics, Iowa State University, Ames, IA 50011.
†Department of Statistics, Colorado State University, Fort Collins, CO 80523-1877.

estimation. Model estimates and predictions are simple, well studied and easy to implement via standard software like SAS. Another important advantage of the FH model is that it only requires summary data, not element-level data that might be unavailable to the analyst because of confidentiality concerns.

The Fay-Herriot model has two stochastic variables, one for the sampling error and the other one representing small area specific random effects. These two random variables and a fixed effect linear regression are additively related to the design based estimators of small area means. The best linear unbiased predictor (BLUP) is commonly used as the small area estimator, and construction of the BLUP does not require any distributional assumption. However, the BLUP is a function of model parameters, and thus their estimated values are plugged-in before the BLUP can be used. The resulting plug-in estimators are known as empirical or estimated best linear unbiased predictor (EBLUP).

There are several methods of model parameter estimation available and the inference is sensitive to the method of estimation. ?) found that the parameter estimation method proposed in ?) performed best in terms of mean squared prediction error for a wide range of parametric model specifications. However, they did not study the construction of accurate confidence intervals for the small area parameters. In this article, we concentrate on finding accurate confidence intervals for small area parameters using the point estimators and their mean squared prediction error (MSPE) estimators under the Fay-Herriot method. While MSPE is a useful measure for the uncertainty of the predictions, confidence intervals (CI) are often much more informative and useful from the practitioners' perspective. However, the literature on CI for small area estimation is limited.

Naive confidence intervals for the small area parameters are easy to construct, but their coverage accuracy is questionable. Recently, ?) and ?) proposed methods based on parametric bootstrap. Although the ?) results are applicable in a general setting and second order correct, they are based on double bootstrap calibrated sample and are not centered around the EBLUP, which is often a desirable property in practice. ?) proposed a method based on a single bootstrap and normality assumptions, and the resulting confidence intervals are centered around the EBLUP and $O(d^5 n^{-3/2})$ order correct, where d is the number of model parameters and n is the total sample size. Their bootstrap based intervals can be calibrated one or more times, resulting in a coverage accuracy of $O(d^5 n^{-5/2})$ or higher if needed. The computational issue of bootstrap methods has been discussed by ?), ?), ?) and others. In addition to the computational difficulties especially in the case of calibration, these methods have other issues, for example, choosing between equal tail or shortest interval quantile points, etc. Thus, a simple closed-form plug-in formula with equivalent theoretically accurate confidence interval and which could be implemented with the existing software without any further programming would be very useful, at least for standard models.

Another drawback of the bootstrap inference approach is that it is not clear how to construct the confidence intervals for a difference of two small area parameters or a linear combination of small area parameters using the above resampling methods. To this end, ?) constructed empirical Bayes (EB) confidence intervals with coverage accuracy $O(n^{-3/2})$ for the special case of equal sampling variances across small areas, a situation which is not common in practical applications. In this article, we develop closed form confidence intervals for small area parameters without the equal sampling error variance assumption, and coverage probability is correct to $O(n^{-2})$. We follow the previous authors in assuming normality of the random components, but will investigate the robustness of the method in simulations. We extend our results to finding accurate confidence intervals for the difference of two small area parameters. This can be easily further generalized to a linear combination of small area parameters. We note that, since the small area estimators are not independent, this extension from a single prediction confidence interval to one for differences (or more generally, linear combinations) of predictions is not immediate.

The remainder of the article is organized as follows. We review the FH model and the Fay-Herriot method of model parameter estimation in Section 2. The main results on constructing improved confidence intervals are given in Section 3, where a simulation study is also discussed. Section 4 describes the extension to the interval for the difference between two small area estimators.

2 The FH Model and Small Area Estimation

Let $y_1, y_2, \ldots, y_n$ be observations (the direct survey estimates) for the n small areas, and $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be fixed predictors. Then the FH model is defined as

$$y_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + u_i + \varepsilon_i, i = 1, \ldots, n, \tag{1}$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ is a $p \times 1$ vector of regression coefficients. The area specific random effects u_i are assumed to be independent and identically distributed (iid) with $\mathrm{E}(u_i) = 0$ and $\mathrm{Var}(u_i) = \sigma_u^2 (\geq 0)$. The sampling errors ϵ_i are also independently distributed with mean zero and variance D_i, and are independent of the u_i.

Small area estimates (SAE) for the FH model are statistics designed to estimate the parameters

$$\theta_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + u_i, i = 1, \ldots, n. \tag{2}$$

In a typical survey application, the values y_i are direct survey estimators of the target small-area parameters θ_i in the sampled areas but may be unacceptably variable because of small sample size in some or all the small areas. The D_i represent the sampling variance of the y_i and are required to be known from an outside source, for instance the statistical agency responsible for collecting the survey data. If the remaining parameters σ_u^2 and $\boldsymbol{\beta}$ were also known, the SAE would be the BLUP given by

$$\tilde{\theta}_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + \gamma_i (y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}), i = 1, \cdots, n. \tag{3}$$

where $\gamma_i = \frac{\sigma_u^2}{\sigma_u^2 + D_i}$. Since $\sigma_u^2, \boldsymbol{\beta}$ are unknown, the BLUP are not usable until we estimate these model parameters.

A number of estimation procedures exist for the model parameters of the FH model. For example, the method of moments used by **?**), maximum-likelihood (ML) and restricted maximum likelihood (REML) used by **?**), and the Fay-Herriot method used by **?**) and then studied by **?**). We focus on the Fay-Herriot method in this paper. The estimator $\widehat{\sigma}_u^2$ is obtained by solving

$$\frac{1}{n-p}\sum_{i=1}^{n}\frac{(y_i-\boldsymbol{x}_i^T\tilde{\boldsymbol{\beta}})^2}{\sigma_u^2+D_i}-1=0 \tag{4}$$

where $\tilde{\boldsymbol{\beta}}=(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\sum_{i=1}^n \frac{\boldsymbol{x}_i y_i}{\sigma_u^2+D_i}$. The estimator $\widehat{\boldsymbol{\beta}}$ is the same as $\tilde{\boldsymbol{\beta}}$ with $\widehat{\sigma}_u^2$ replacing σ_u^2. Then the SAE considered in this paper are defined as

$$\widehat{\theta}_i=\boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}+\frac{\widehat{\sigma}_u^2}{\widehat{\sigma}_u^2+D_i}(y_i-\boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}),\ i=1,\cdots,n.$$

The mean square prediction error (MSPE) for small area estimation is defined as

$$\mathrm{MSPE}(\widehat{\theta}_i) \quad = \quad \mathrm{E}(\widehat{\theta}_i-\theta_i)^2$$

and is the most commonly used measure of the uncertainty of the SAE $\widehat{\theta}_i$. The estimation of the MSPE is an integral part of small area estimation research. We refer to **?**) for an extensive overview and list of references. MSPE estimation depends on the method of model parameter estimation, but also on the assumptions made about the distributions of the random model components. The results for the Fay-Herriot method were shown by **?**) under the assumption of normal errors and small area effects. They obtained

$$\begin{aligned}\mathrm{MSPE}(\widehat{\theta}_i) \quad &= \quad \frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}+\frac{D_i^2}{(\sigma_u^2+D_i)^2}\boldsymbol{x}_i^T(\sum_{i=1}^{n}\frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i\\ &\quad +\frac{2nD_i^2}{(\sigma_u^2+D_i)^3}(\sum_{i=1}^{n}\frac{1}{\sigma_u^2+D_i})^{-2}+O(n^{-2})\end{aligned}$$

for the MSPE, and proposed

$$
\begin{aligned}
mspe(\widehat{\theta}_i) &= \frac{D_i\widehat{\sigma}_u^2}{\widehat{\sigma}_u^2+D_i} + (\frac{D_i}{\widehat{\sigma}_u^2+D_i})^2\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\widehat{\sigma}_u^2+D_i})^{-1}\boldsymbol{x}_i \\
&\quad + \frac{4nD_i^2}{(\widehat{\sigma}_u^2+D_i)^3}(\sum_{i=1}^n \frac{1}{\widehat{\sigma}_u^2+D_i})^{-2} + 2(\frac{D_i}{\widehat{\sigma}_u^2+D_i})^2(\sum_{i=1}^n \frac{1}{\widehat{\sigma}_u^2+D_i})^{-1} \\
&\quad - 2n(\frac{D_i}{\widehat{\sigma}_u^2+D_i})^2(\sum_{i=1}^n \frac{1}{\widehat{\sigma}_u^2+D_i})^{-3}\sum_{i=1}^n \frac{1}{(\widehat{\sigma}_u^2+D_i)^2}.
\end{aligned}
$$

as an estimator for MSPE($\widehat{\theta}_i$).

As mentioned in the introduction, while results on MSPE approximation and estimation has been widely developed for a range of modeling scenarios, the construction of accurate confidence intervals is still relatively uncommon. ?) and ?) proposed bootstrap approximation intervals which can be callibrated one or more times with the coverage accuracy of $O(d^5n^{-5/2})$ or higher if needed. But there is still no traditional closed form interval estimates with the coverage probability corrected to $O(n^{-2})$. The closest work is done by ?), but they only considered the case with equal D_i (sampling error variance) and when the two stochastic variables are both normally distributed (i.e. the sampling error and the small area specific random effect). They considered the unconditional and conditional coverage probabilities which are both corrected to achieve $O(n^{-3/2})$ accuracy. In the next section, we provide closed form confidence intervals for the small area parameters with unequal sampling variances and with coverage probabilities that are correct up to $O(n^{-2})$.

3 Confidence interval with corrected coverage probability

The traditional closed form confidence interval is of the form EBLUP $\pm z_{\alpha/2}\sqrt{mspe}$, with $z_{\alpha/2}$ denoting the upper $100(1-\alpha/2)\%$ percentile of the standard normal distribution. It is well known that this interval has coverage error $O(n^{-1})$. Thus, it is not appropriate for small to moderate sample size small area estimation problems. ?) improved the accuracy of the confidence interval to $O(n^{-3/2})$ under the special case of equal sampling variances and normality of the errors. We have expanded their correction to $O(n^{-2})$ under normality of random component distributions and unequal sampling variances. The second order correction is the generally accepted standard for the precision of small area estimation.

Let $\phi()$ and $\Phi()$ denote the pdf and df, respectively, of the $N(0,1)$ distribution. In order to derive our results, we make the following standard assumptions:

A1 *The matrix* $(\frac{1}{n}\boldsymbol{X}^T\boldsymbol{X})^{-1}$ *is* $O(1)$ *element-wise, and*

$$\sup_i \tilde{h}_{ii} = \boldsymbol{x}_i^T(\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i^T)^{-1}\boldsymbol{x}_i = O(n^{-1}).$$

A2 *The quantities* $\boldsymbol{x}_i, \sigma_u^2, D_i$ *are bounded, and* $D_i > \gamma_1 \geq 0$, $\sigma_u^2 > \gamma_2 \geq 0$.

A3 *The small area random effects* u_i *are independent and identically normally distributed with mean 0 and variance* σ_u^2. *The errors* e_i *are independent and normally distributed with mean 0 and variance* D_i, *and are independent of the* u_i.

Theorem 3.1 *Under Assumption A??, A?? and A??, for any real* t,

$$P\left[\theta_i \in \widehat{\theta}_i \pm t \times \sqrt{mspe(\widehat{\theta}_i)}\right] = 2\Phi(t) - 1 - \frac{t\phi(t)}{2}\frac{(t^2+1)nD_i^2}{\sigma_u^4(\sigma_u^2+D_i)^2}(\sum_{i=1}^{n}\frac{1}{\sigma_u^2+D_i})^{-2} + O(n^{-2}).$$

Corollary 3.1 *Setting* $t_\alpha = z_{\alpha/2}\left[(z_{\alpha/2}^2+1)\frac{nD_i^2}{4\sigma_u^4(\sigma_u^2+D_i)^2}(\sum_{i=1}^{n}\frac{1}{\sigma_u^2+D_i})^{-2}+1\right]$, *a corrected* $(1-\alpha)$ *confidence interval for* θ_i *defined as*

$$P\left[\theta_i \in \widehat{\theta}_i \pm t_\alpha \times \sqrt{mspe(\widehat{\theta}_i)}\right]$$

achieves a confidence level equal to $1-\alpha+O(n^{-2})$. *When* σ_u^2 *is unknown and estimated by the* $\widehat{\sigma}_u^2$ *which solves (??),* $\widehat{t}_\alpha = z_{\alpha/2}\left[(z_{\alpha/2}^2+1)\frac{nD_i^2}{4\widehat{\sigma}_u^4(\widehat{\sigma}_u^2+D_i)^2}(\sum_{i=1}^{n}\frac{1}{\widehat{\sigma}_u^2+D_i})^{-2}+1\right]$ *is used instead of* t_α *and the confidence level still equals* $1-\alpha+O(n^{-2})$.

Corollary **??** follows from Theorem **??** by Taylor expansion.

Proof of Theorem **??**. The proof will follow the approach of proof used in **?**). Note first that

$$\mathrm{P}\left[\frac{\theta_i-\widehat{\theta}_i}{\sqrt{mspe(\widehat{\theta}_i)}}\le t\right] \;=\; \mathrm{P}\left[\frac{\theta_i-\tilde{\theta}_i}{\sqrt{\mathrm{E}(\theta_i-\tilde{\theta}_i)^2}}\le\frac{\widehat{\theta}_i-\tilde{\theta}_i+t\times\sqrt{mspe(\widehat{\theta}_i)}}{\sqrt{\mathrm{E}(\theta_i-\tilde{\theta}_i)^2}}\right],$$

and since u_i and e_i are independent normally distributed with variances σ_u^2 and D_i, respectively,

$$\begin{aligned}
\mathrm{P}\left[\frac{\theta_i-\widehat{\theta}_i}{\sqrt{mspe(\widehat{\theta}_i)}}\le t\right] &= \mathrm{P}\left[\frac{\frac{D_iu_i}{\sigma_u^2+D_i}-\frac{\sigma_u^2e_i}{\sigma_u^2+D_i}}{\sqrt{\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}}}\le\frac{\widehat{\theta}_i-\tilde{\theta}_i+t\times\sqrt{mspe(\widehat{\theta}_i)}}{\sqrt{\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}}}\right]\\
&= \mathrm{E}\left(\Phi(\frac{\widehat{\theta}_i-\tilde{\theta}_i+t\times\sqrt{mspe(\widehat{\theta}_i)}}{\sqrt{\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}}})\right)\\
&= \mathrm{E}\left(\Phi(t+\frac{t\times(\sqrt{mspe(\widehat{\theta}_i)}-\sqrt{\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}})+\widehat{\theta}_i-\tilde{\theta}_i}{\sqrt{\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}}})\right)\\
&= \mathrm{E}\left(\Phi(t+Z(t))\right),
\end{aligned}$$

where $Z(t)=\frac{t\times(\sqrt{mspe(\widehat{\theta}_i)}-\sqrt{\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}})+\widehat{\theta}_i-\tilde{\theta}_i}{\sqrt{\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}}}$. With a Taylor expansion, we get

$$\Phi(t+Z(t))-\Phi(t)=Z(t)\phi(t)+\frac{1}{2}Z^2(t)\phi'(t)+\frac{1}{6}\phi''(t)Z^3(t)+\frac{1}{24}\phi'''(t^*)Z^4(t) \quad (5)$$

where t^* is between t and $t+Z(t)$. We compute the expectation for the right side term by term.

For the first term in (**??**),

$$\begin{aligned}
\mathrm{E}Z(t) &= t\frac{1}{2}\left(\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}\right)^{-1}\mathrm{E}\left(mspe(\widehat{\theta}_i)-\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}\right)\\
&\quad -t\frac{1}{8}\left(\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}\right)^{-2}\mathrm{E}\left(mspe(\widehat{\theta}_i)-\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}\right)^2\\
&\quad +t\frac{3}{16}\left(\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}\right)^{-3}\mathrm{E}\left(mspe(\widehat{\theta}_i)-\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}\right)^3\\
&\quad +t\frac{3}{16}\times\frac{5}{48}\left(\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}\right)^{-1/2}\mathrm{E}\int_{\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}}^{mspe(\widehat{\theta}_i)}x^{-7/2}\left(mspe(\widehat{\theta}_i)-\frac{\sigma_u^2D_i}{\sigma_u^2+D_i}\right)^3,
\end{aligned}$$

where we have used the fact that $\mathrm{E}\theta_i - \widehat{\theta}_i = 0$ by the properties mentioned in **?**) that the FH variance estimator is even and translation invariant, and hence that $\mathrm{E}(\theta_i - \widehat{\theta}_i - \theta_i + \tilde{\theta}_i) = 0$, which means $\mathrm{E}(\widehat{\theta}_i - \tilde{\theta}_i) = 0$. The 3rd and 4th terms in $\mathrm{E}Z(t)$ are shown to be negligible (see Appendix), and by Taylor expansion we obtain

$$\begin{aligned}
\mathrm{E}Z(t) &= \frac{1}{2}t\left(\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^{-1}\left(\frac{D_i^2}{(\sigma_u^2+D_i)^2}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i\right. \\
&\left. + \frac{2nD_i^2}{(\sigma_u^2+D_i)^3}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2}\right) \\
&- \frac{1}{8}t\left(\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^{-2}\frac{2nD_i^4}{(\sigma_u^2+D_i)^4}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2} + O(n^{-2}).
\end{aligned}$$

For the second term in (**??**),

$$\begin{aligned}
Z(t)^2 &= \frac{(t(\sqrt{mspe(\widehat{\theta}_i)} - \sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}))^2 + (\widehat{\theta}_i - \tilde{\theta}_i)^2}{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}} \\
&+ \frac{2(t(\sqrt{mspe(\widehat{\theta}_i)} - \sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}))(\widehat{\theta}_i - \tilde{\theta}_i)}{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}
\end{aligned}$$

by simple expansions, we get

$$\begin{aligned}
\mathrm{E}Z(t)^2 &= \frac{D_i}{\sigma_u^2(\sigma_u^2+D_i)}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i + \frac{2nD_i}{\sigma_u^2(\sigma_u^2+D_i)^2}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2} \\
&+ \frac{t^2 nD_i^2}{2\sigma_u^4(\sigma_u^2+D_i)^2}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2} + O(n^{-2}).
\end{aligned}$$

Combining this with the fact that $\phi'(t) = -t\phi(t)$, we get

$$\begin{aligned}
\mathrm{E}Z(t)^2\phi'(t)/2 &= -t\phi(t)\left\{\frac{D_i}{2\sigma_u^2(\sigma_u^2+D_i)}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i\right. \\
&+ \frac{nD_i}{\sigma_u^2(\sigma_u^2+D_i)^2}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2} \\
&\left. + \frac{t^2 nD_i^2}{4\sigma_u^4(\sigma_u^2+D_i)^2}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2}\right\} + O(n^{-2}).
\end{aligned}$$

For the third term in (**??**),

$$
\begin{aligned}
Z(t)^3 &= \frac{(t(\sqrt{mspe(\widehat{\theta}_i)}-\sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}))^3+(\widehat{\theta}_i-\tilde{\theta}_i)^3}{(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i})^{3/2}} \\
&+\frac{3(t(\sqrt{mspe(\widehat{\theta}_i)}-\sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}))^2(\widehat{\theta}_i-\tilde{\theta}_i)}{(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i})^{3/2}} \\
&+\frac{3(t(\sqrt{mspe(\widehat{\theta}_i)}-\sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}))(\widehat{\theta}_i-\tilde{\theta}_i)^2}{(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i})^{3/2}}.
\end{aligned}
$$

After simple Taylor expansion and by Assumption A**??**, A**??**, A**??** and Lemma **??**, we can get $\mathrm{E}Z^3(t)=O(n^{-2})$.

Finally, for the remainder in (**??**), $\phi'''(t^*)=3t^*\phi(t^*)-(t^*)^3\phi(t^*)$ and $|3t^*\phi(t^*)-(t^*)^3\phi(t^*)|<6\phi(\sqrt{2})$, and by Assumption A**??**, A**??** A**??**, Taylor expansion and Lemma **??** we get

$$
\begin{aligned}
\mathrm{E}\frac{1}{24}\phi'''(t^*)Z^4(t) &\leq \mathrm{E}|\frac{1}{24}\phi'''(t^*)Z^4(t)| \\
&\leq \mathrm{E}|\frac{1}{24}\phi'''(t^*)|Z^4(t) \\
&\leq \mathrm{E}\frac{\phi(\sqrt{2})}{4}Z^4(t)=O(n^{-2}).
\end{aligned}
$$

Then observe that

$$
\begin{aligned}
\mathrm{P}\left[\frac{\theta_i-\widehat{\theta}_i}{\sqrt{mspe(\widehat{\theta}_i)}}\leq t\right] &= \mathrm{E}\left(\Phi(t+Z(t))\right) \\
\mathrm{P}\left[\frac{\theta_i-\widehat{\theta}_i}{\sqrt{mspe(\widehat{\theta}_i)}}\leq -t\right] &= \mathrm{E}\left(\Phi(-t+Z(-t))\right).
\end{aligned}
$$

Combining all above, we conclude

$$\begin{aligned}
\mathrm{P}\left[\theta_i \in \widehat{\theta}_i \pm t \times \sqrt{mspe(\widehat{\theta}_i)}\right] &= \mathrm{E}\left(\Phi(t+Z(t))\right) - \mathrm{E}\left(\Phi(-t+Z(-t))\right) \\
&= 2\Phi(t) - 1 - \frac{t\phi(t)}{2}\frac{nD_i^2}{\sigma_u^4(\sigma_u^2+D_i)^2}(\sum_{i=1}^{n}\frac{1}{\sigma_u^2+D_i})^{-2} \\
&\quad +O(n^{-2}).
\end{aligned}$$

■

3.1 Finite Sample Performance

Now we describe a simulation study investigating the finite sample performance of the proposed confidence intervals. We adopt the simulation setup of Datta *et al.* (2005). We consider the FH model (1) under two distributional scenarios: (i) both u and ϵ are normally distributed and (ii) both u and ϵ are distributed as chi-square variates. The Chi-square distribution is considered to assess the robustness of the procedure to misspecification of the random component distributions. The sample sizes considered are 15 and 60. The covariate x_i is generated from a normal distribution with mean 0 and variance 1, and the FH model mean is taken to be $\boldsymbol{x_i\beta} = (1, x_i)(0, 0)^T = 0$ for each small area. This means that we generate data using $y_i = u_i + \varepsilon_i$, but we will continue to estimate the full FH model that includes a linear mean component.

The small areas are divided into 5 equal sized groups and the D_i remain the same within each group. The sampling errors ε_i are generated from $N(0, D_i)$ for case (i), and from a centered Chi-square distribution with mean 0 and variance D_i for case (ii), where the (possibly fractional) degrees of freedom are set to achieve the desired variances. We consider three different pattern for the D_i, namely (a) (0.7, 0.6, 0.5, 0.4, 0.3) (b) (2.0, 0.6, 0.5, 0.4, 0.2) (c) (4.0, 0.6, 0.5, 0.4, 0.1). The u_i are generated from $N(0, \sigma_u^2)$ for case (i) and from a centered Chi-square distribution with variance $\sigma_u^2 = 1$ for case (ii). Note that for pattern (a) the sampling variances are all smaller than the model variance and evenly distributed; for patterns (b) and

(c) the sampling variances are not evenly distributed over the areas and some are larger than the model variance. All reported results are based on 10,000 replicates.

[Table 1, 2 and 3 will be here]

Tables **??**, **??** and **??** report coverage probabilities and average interval length of confidence intervals for θ_i, for different patterns of D_i. In each table, the first line corresponds to our proposed method, and the second line (within parentheses) corresponds to the "naive" FH method where the CI is $\widehat{\theta}_i \pm z_{\alpha/2} \times \sqrt{mspe(\widehat{\theta}_i)}$. For all three patterns, our method results in coverage probabilities that are higher than those for the naive method. When n equals 15, our method achieves results that are closer to the nominal level of 0.95, especially for patterns (b) and (c). As n increases to 60, the coverage probabilities are close to 0.95 for both the naive and the proposed method, although the proposed method continues to be closer and hence shows its consistency. It appears that for pattern (a) and under normality of the random components, there is not much need of the improved method. However, the proposed method improves the coverage probabilities when the random components are not normal. For patterns (b) and (c) and particularly for large sampling variances, the improved method clearly shows its utility for both normal and non-normal situations.

[Table 4 will be here]

In order to compare our method with existing bootstrap-based corrections to confidence intervals, we adopt the simulation setup of **?**). We consider the FH model with errors distributed normally. The sample size is taken as 15, and $\boldsymbol{x}_i\boldsymbol{\beta} = 0$ as before. Two different patterns of D_i's are considered, namely (a) (4.0, 0.6, 0.5, 0.4, 0.2) with $\sigma_u^2 = 1$ (b) (8.0, 1.2, 1.0, 0.8, 0.4) with $\sigma_u^2 = 2$. Table **??** shows the coverage probabilities and average interval length of confidence intervals for θ_i. For both patterns, the coverage probabilities under the proposed method are comparable to the results reported in **?**). However, our proposed method is computationally much simpler, because it does not require resampling.

4 Confidence Interval for the Difference of Two Small Areas

We extend the results of the previous section to build a confidence interval for $\theta_i - \theta_j$ where $i \neq j$, to allow comparison between two different small areas. While the interval for an individual small area can be obtained using either our method or one of the resampling-based approaches, this is not the case here, as finding this type of interval using existing resampling methods is not obvious because of the correlation between the estimators.

The following theorem provides the asymptotic confidence intervals. The proof is given in the Appendix.

Theorem 4.1 *Under Assumption A??, A?? and A?? ,*

$$\begin{aligned}
& P\left[\theta_i - \theta_j \in \widehat{\theta}_i - \widehat{\theta}_j \pm t \times \sqrt{mspe(\widehat{\theta}_i - \widehat{\theta}_j)}\right] \\
= \; & 2\Phi(t) - 1 - \frac{t\phi(t)n}{2}(t^2+1)(\frac{D_i\sigma_u^2}{\sigma_u^2+D_i} + \frac{D_j\sigma_u^2}{\sigma_u^2+D_j})^{-2} \\
& \times(\frac{D_i^2}{(\sigma_u^2+D_i)^2} + \frac{D_j^2}{(\sigma_u^2+D_j)^2})^2(\sum_{i=1}^{n}\frac{1}{\sigma_u^2+D_i})^{-2} + O(n^{-2})
\end{aligned}$$

where

$$mspe(\widehat{\theta}_i - \widehat{\theta}_j) \;=\; mspe(\widehat{\theta}_i) + mspe(\widehat{\theta}_j) - \frac{2D_iD_j}{(\widehat{\sigma}_u^2+D_i)(\widehat{\sigma}_u^2+D_j)}\boldsymbol{x}_i^T(\sum_{i=1}^{n}\frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\widehat{\sigma}_u^2+D_i})^{-1}\boldsymbol{x}_j.$$

Corollary 4.1 *Setting*

$$\begin{aligned}
t_\alpha \;=\; & z_{\alpha/2}\left[(z_{\alpha/2}^2+1)\frac{n}{4}(\frac{D_i\sigma_u^2}{\sigma_u^2+D_i} + \frac{D_j\sigma_u^2}{\sigma_u^2+D_j})^{-2}(\frac{D_i^2}{(\sigma_u^2+D_i)^2} + \frac{D_j^2}{(\sigma_u^2+D_j)^2})^2\right. \\
& \left.(\sum_{i=1}^{n}\frac{1}{\sigma_u^2+D_i})^{-2} + 1\right],
\end{aligned}$$

a corrected $(1-\alpha)$ *level confidence interval for* $\theta_i - \theta_j$ *defined as*

$$P\left[\theta_i - \theta_j \in \widehat{\theta}_i - \widehat{\theta}_j \pm t_\alpha \times \sqrt{mspe(\widehat{\theta}_i - \widehat{\theta}_j)}\right]$$

achieves a confidence level equal to $1-\alpha+O(n^{-2})$. *If* σ_u^2 *is unknown and estimated by the* $\widehat{\sigma}_u^2$ *which solves (**??**),*

$$\widehat{t}_\alpha = z_{\alpha/2}\left[(z_{\alpha/2}^2+1)\frac{n}{4}(\frac{D_i\widehat{\sigma}_u^2}{\widehat{\sigma}_u^2+D_i}+\frac{D_j\widehat{\sigma}_u^2}{\widehat{\sigma}_u^2+D_j})^{-2}(\frac{D_i^2}{(\widehat{\sigma}_u^2+D_i)^2}+\frac{D_j^2}{(\widehat{\sigma}_u^2+D_j)^2})^2 (\sum_{i=1}^{n}\frac{1}{\widehat{\sigma}_u^2+D_i})^{-2}+1\right]$$

is used instead of t_α *and the confidence level still equals* $1-\alpha+O(n^{-2})$.

Proof of Corollary **??** can be easily proven by Taylor expansion.

4.1 *Finite Sample Performance*

We adopt the simulation setup of **?**) with $\boldsymbol{x_i\beta}=0$ as in the previous section, and we examine the difference for $\theta_1-\theta_2$ for each pattern of D_i. In each area, there is $n/5$ elements in each pattern, and we calculated the difference between the first two 2 elements in each. Table **??**, **??**, **??** show the coverage probabilities and confidence interval lengths for $\theta_1-\theta_2$ for the three different patterns of D_i.

[Table 5, 6 and 7 will go here]

Like in the previous section, the naive confidence intervals are reasonable when the sampling variances are really small compared to the model variance. However, when the sampling variance is high or the sampling errors are not normally distributed, the naive method underestimates the coverage probability severely. The proposed method works well under normality and improves upon the naive method significantly under non-normality.

5 Conclusion

In this article, we developed accurate confidence intervals for small area parameters and provided a closed form expression that can be readily computed. The intervals remain centered around the EBLUP. Substantial improvements in coverage relative to naive intervals which use unadjusted critical values are illustrated in the simulation study. A major advantage of the proposed method is computational. Similarly to the MSPE estimation using Taylor expansion, the model parameters need to be estimated only once using standard software, and can then be used in the construction of plug-in type confidence intervals. The procedure is implemented in R and is available upon request.

A Proofs

Lemma A.1 *Under Assumption A??, A?? and A?? ,*

$$\begin{aligned}
E\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right) &= \frac{D_i^2}{(\sigma_u^2 + D_i)^2}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2 + D_i})^{-1}\boldsymbol{x}_i \\
&\quad + \frac{2nD_i^2}{(\sigma_u^2 + D_i)^3}(\sum_{i=1}^n \frac{1}{\sigma_u^2 + D_i})^{-2} + O(n^{-2}) \\
E\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^2 &= \frac{2nD_i^4}{(\sigma_u^2 + D_i)^4}(\sum_{i=1}^n \frac{1}{\sigma_u^2 + D_i})^{-2} + O(n^{-2}) \\
E\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^3 &= O(n^{-2}) \\
E\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^4 &= O(n^{-2}).
\end{aligned}$$

Proof of Lemma **??**.

The leading term of $\widehat{\sigma}_u^2 - \sigma_u^2$ is $(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-1}\left\{\sum_{i=1}^n \frac{(u_i+\varepsilon_i)^2}{\sigma_u^2+D_i} - n\right\} + \frac{p}{n}(\frac{1}{n}\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-1}$, and $\mathrm{E}(\widehat{\sigma}_u^2 - \sigma_u^2)^{2s} = O(n^{-s})$ for all $s = 1, 2, \cdots$, by Assumption A??, A??, A??, Taylor expansion and similar steps as in **?**), where p is the rank for the $\boldsymbol{X} = (\boldsymbol{x}_1^T, \ldots, \boldsymbol{x}_n^T)^T$. Then the proof follows easily by Taylor expansion.

■

Additional steps for the proof of Theorem 3.1

We also need to compute

$$\begin{aligned}
&\mathrm{E}\int_{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}^{mspe(\widehat{\theta}_i)} x^{-7/2}\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^3 \\
&\quad = \mathrm{E}\int_{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}^{mspe(\widehat{\theta}_i)} x^{-7/2}\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^3 I_{\{mspe(\widehat{\theta}_i)\geq \frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\}} \\
&\qquad +\mathrm{E}\int_{mspe(\widehat{\theta}_i)}^{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}} x^{-7/2}\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^3 I_{\{mspe(\widehat{\theta}_i)\leq \frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\}} \\
&\quad = \mathrm{E}I_1 + \mathrm{E}I_2.
\end{aligned}$$

We compute $\mathrm{E}I_1 = O(n^{-2})$ and $\mathrm{E}I_2 = O(n^{-2})$ separately. For $\mathrm{E}I_1$, we get

$$\begin{aligned}
\mathrm{E}I_1 &\leq \mathrm{E}\int_{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}^{mspe(\widehat{\theta}_i)} (\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i})^{-7/2}\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^3 dx I_{\{mspe(\widehat{\theta}_i)\geq \frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\}} \\
&= \frac{1}{4}(\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i})^{-7/2}\mathrm{E}\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^4 I_{\{mspe(\widehat{\theta}_i)\geq \frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\}} \\
&\leq \frac{1}{4}(\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i})^{-7/2}\mathrm{E}\left(mspe(\widehat{\theta}_i) - \frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i}\right)^4 \\
&= O(n^{-2})
\end{aligned}$$

by assumption A**??**, A**??**, A**??** and Taylor expansion.

For the second part $\mathrm{E}I_2$, choose a sequence of constants $\eta = n^{-1/8}$ and write

$$I_{\{mspe(\widehat{\theta}_i)\leq \frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\}} = I_{\{mspe(\widehat{\theta}_i)\leq \frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta\}} + I_{\{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta\leq mspe(\widehat{\theta}_i)\leq \frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\}}.$$

Then we get $I_2 = I_{21} + I_{22}$ respectively. For $\mathrm{E}I_{21}$

$$\begin{aligned}
|\mathrm{E}I_{21}| &= \mathrm{E}\int_0^{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-mspe(\widehat{\theta}_i)} x^3\left(mspe(\widehat{\theta}_i)+x\right)^{-7/2} dx I_{\{mspe(\widehat{\theta}_i)\le\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta\}} \\
&\le \mathrm{E}\int_0^{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-mspe(\widehat{\theta}_i)} x^{-1/2} dx I_{\{mspe(\widehat{\theta}_i)\le\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta\}} \\
&= 2\mathrm{E}\left(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-mspe(\widehat{\theta}_i)\right)^{1/2} I_{\{mspe(\widehat{\theta}_i)\le\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta\}} \\
&\le 2\sqrt{\mathrm{E}\left|\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-mspe(\widehat{\theta}_i)\right|}\sqrt{\mathrm{P}(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-mspe(\widehat{\theta}_i)\ge\eta)} \\
&\le O(n^{-1/2})\sqrt{\frac{\mathrm{E}(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-mspe(\widehat{\theta}_i))^4}{\eta^4}} \\
&= O(n^{-2})
\end{aligned}$$

by assumption A**??**, A**??**, A**??** and Taylor expansion. For $\mathrm{E}I_{22}$, we compute

$$\begin{aligned}
|\mathrm{E}I_{22}| &= \mathrm{E}\int_{mspe(\widehat{\theta}_i)}^{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}} x^{-7/2}\left(x-mspe(\widehat{\theta}_i)\right)^3 dx I_{\{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta\le mspe(\widehat{\theta}_i)\le\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\}} \\
&\le (\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta)^{-7/2}\mathrm{E}\int_{mspe(\widehat{\theta}_i)}^{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}\left(x-mspe(\widehat{\theta}_i)\right)^3 dx I_{\{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta\le mspe(\widehat{\theta}_i)\le\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\}} \\
&\le (\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta)^{-7/2}\frac{1}{4}\mathrm{E}\left(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-mspe(\widehat{\theta}_i)\right)^4 I_{\{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta\le mspe(\widehat{\theta}_i)\le\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\}} \\
&\le (\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\eta)^{-7/2}\frac{1}{4}\mathrm{E}\left(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-mspe(\widehat{\theta}_i)\right)^4 \\
&= O(n^{-2})
\end{aligned}$$

by assumption A**??**, A**??**, A**??** and Taylor expansion. Combining above, we get $\mathrm{E}I_1 + \mathrm{E}I_2 = O(n^{-2})$.

Next, we need the following quantities for $\mathrm{E}Z(t)^2$.

$$\mathrm{E}(\widehat{\theta}_i - \tilde{\theta}_i)^2 = \frac{D_i^2}{(\sigma_u^2+D_i)^2}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i + \frac{2nD_i^2}{(\sigma_u^2+D_i)^3}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2} + O(n^{-2})$$

by assumption A??, A??, A?? and Taylor expansion. Finally, we obtain the approximations

$$\mathrm{E}\frac{(t(\sqrt{mspe(\widehat{\theta}_i)} - \sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}))^2}{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}} = \frac{t^2 n D_i^2}{2\sigma_u^4(\sigma_u^2+D_i)^2}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2} + O(n^{-2})$$

$$\mathrm{E}\frac{2(t(\sqrt{mspe(\widehat{\theta}_i)} - \sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}}))(\widehat{\theta}_i - \tilde{\theta}_i)}{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}} = O(n^{-2})$$

by assumption A??, A??, A?? and Taylor expansion.

■

Lemma A.2 *Under Assumption A??, A?? and A?? ,*

$$E\left(mspe(\widehat{\theta}_i - \widehat{\theta}_j) - (\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i} + \frac{\sigma_u^2 D_j}{\sigma_u^2+D_j})\right)$$
$$= \frac{D_i^2}{(\sigma_u^2+D_i)^2}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i + \frac{2nD_i^2}{(\sigma_u^2+D_i)^3}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2}$$
$$+\frac{D_j^2}{(\sigma_u^2+D_j)^2}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i + \frac{2nD_j^2}{(\sigma_u^2+D_j)^3}(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2}$$
$$-\frac{2D_iD_j}{(\sigma_u^2+D_i)(\sigma_u^2+D_j)}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_j + O(n^{-2})$$

$$E\left(mspe(\widehat{\theta}_i - \widehat{\theta}_j) - (\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i} + \frac{\sigma_u^2 D_j}{\sigma_u^2+D_j})\right)^2$$
$$= 2n\left(\frac{D_i^2}{(\sigma_u^2+D_i)^2} + \frac{D_j^2}{(\sigma_u^2+D_j)^2}\right)^2(\sum_{i=1}^n \frac{1}{\sigma_u^2+D_i})^{-2} + O(n^{-2})$$

$$E\left(mspe(\widehat{\theta}_i - \widehat{\theta}_j) - (\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i} + \frac{\sigma_u^2 D_j}{\sigma_u^2+D_j})\right)^3 = O(n^{-2})$$

$$E\left(mspe(\widehat{\theta}_i - \widehat{\theta}_j) - (\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i} + \frac{\sigma_u^2 D_j}{\sigma_u^2 + D_j})\right)^4 = O(n^{-2}).$$

The proof of Lemma **??** can be derived from A**??**, A**??**, A**??** and Taylor expansion and is similar to the proof of Lemma **??**.

Proof of Theorem **??**. As in the proof of Theorem **??**,

$$\begin{aligned}
&\mathrm{P}\left[\frac{\theta_i - \theta_j - (\widehat{\theta}_i - \widehat{\theta}_i)}{\sqrt{mspe(\widehat{\theta}_i - \widehat{\theta}_j)}} \leq t\right] \\
&= \mathrm{E}\left(\Phi(t + \frac{t \times (\sqrt{mspe(\widehat{\theta}_i - \widehat{\theta}_j)} - \sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i} + \frac{\sigma_u^2 D_j}{\sigma_u^2 + D_j}}) + \widehat{\theta}_i - \tilde{\theta}_i - (\widehat{\theta}_j - \tilde{\theta}_j)}{\sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i} + \frac{\sigma_u^2 D_j}{\sigma_u^2 + D_j}}})\right) \\
&= \mathrm{E}\left(\Phi(t + Z(t))\right)
\end{aligned}$$

where $Z(t) = \frac{t \times (\sqrt{mspe(\widehat{\theta}_i - \widehat{\theta}_j)} - \sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i} + \frac{\sigma_u^2 D_j}{\sigma_u^2 + D_j}}) + \widehat{\theta}_i - \tilde{\theta}_i - (\widehat{\theta}_j - \tilde{\theta}_j)}{\sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2 + D_i} + \frac{\sigma_u^2 D_j}{\sigma_u^2 + D_j}}}$. As before, we again write

$$\Phi(t + Z(t)) = \Phi(t) + Z(t)\phi(t) + \frac{1}{2}Z^2(t)\phi'(t) + \frac{1}{6}\phi''(t)Z^3(t) + \frac{1}{24}\phi'''(t^*)Z^4(t)$$

where t^* is between t and $t+Z(t)$. Similarly to the proof of Theorem **??**, we compute

$$\begin{aligned}
\mathrm{E}Z(t) &= t\frac{1}{2}\left(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}+\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}\right)^{-1}\mathrm{E}\left(mspe(\widehat{\theta}_i-\widehat{\theta}_j)-\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}\right)\\
&\quad -t\frac{1}{8}\left(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}+\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}\right)^{-2}\mathrm{E}\left(mspe(\widehat{\theta}_i-\widehat{\theta}_j)-\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}-\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}\right)^2\\
&\quad +O(n^{-2})\\
&= \frac{1}{2}t\left(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}+\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}\right)^{-1}\left(\frac{D_i^2}{(\sigma_u^2+D_i)^2}\boldsymbol{x}_i^T(\sum_{i=1}^n\frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i\right.\\
&\quad +\frac{2nD_i^2}{(\sigma_u^2+D_i)^3}(\sum_{i=1}^n\frac{1}{\sigma_u^2+D_i})^{-2}+\frac{D_j^2}{(\sigma_u^2+D_j)^2}\boldsymbol{x}_i^T(\sum_{i=1}^n\frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i\\
&\quad \left.+\frac{2nD_j^2}{(\sigma_u^2+D_j)^3}(\sum_{i=1}^n\frac{1}{\sigma_u^2+D_i})^{-2}-\frac{2D_iD_j}{(\sigma_u^2+D_i)(\sigma_u^2+D_j)}\boldsymbol{x}_i^T(\sum_{i=1}^n\frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_j\right)\\
&\quad -\frac{1}{8}t\left(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}+\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}\right)^{-2}2n\left(\frac{D_i^2}{(\sigma_u^2+D_i)^2}+\frac{D_j^2}{(\sigma_u^2+D_j)^2}\right)^2\\
&\quad \times(\sum_{i=1}^n\frac{1}{\sigma_u^2+D_i})^{-2}\\
&\quad +O(n^{-2}).
\end{aligned}$$

Next we consider $\mathrm{E}Z(t)^2$. Write

$$\begin{aligned}
Z(t)^2 &= \frac{(t(\sqrt{mspe(\widehat{\theta}_i-\widehat{\theta}_j)}-\sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}}))^2+(\widehat{\theta}_i-\tilde{\theta}_i-(\widehat{\theta}_j-\tilde{\theta}_j))^2}{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}+\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}}\\
&\quad +\frac{2(t(\sqrt{mspe(\widehat{\theta}_i-\widehat{\theta}_j)}-\sqrt{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}}))(\widehat{\theta}_i-\tilde{\theta}_i-(\widehat{\theta}_j-\tilde{\theta}_j))}{\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}+\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}}.
\end{aligned}$$

and hence

$$
\begin{aligned}
\mathrm{E}Z(t)^2 \quad = \quad & \left(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}+\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}\right)^{-1}\left(\frac{D_i^2}{(\sigma_u^2+D_i)^2}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i\right.\\
& +\frac{2nD_i^2}{(\sigma_u^2+D_i)^3}(\sum_{i=1}^n\frac{1}{\sigma_u^2+D_i})^{-2}+\frac{D_j^2}{(\sigma_u^2+D_j)^2}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_i\\
& \left.+\frac{2nD_j^2}{(\sigma_u^2+D_j)^3}(\sum_{i=1}^n\frac{1}{\sigma_u^2+D_i})^{-2}-\frac{2D_iD_j}{(\sigma_u^2+D_i)(\sigma_u^2+D_j)}\boldsymbol{x}_i^T(\sum_{i=1}^n \frac{\boldsymbol{x}_i\boldsymbol{x}_i^T}{\sigma_u^2+D_i})^{-1}\boldsymbol{x}_j\right)\\
& +\frac{1}{4}t^2\left(\frac{\sigma_u^2 D_i}{\sigma_u^2+D_i}+\frac{\sigma_u^2 D_j}{\sigma_u^2+D_j}\right)^{-2}2n\left(\frac{D_i^2}{(\sigma_u^2+D_i)^2}+\frac{D_j^2}{(\sigma_u^2+D_j)^2}\right)^2\\
& \times(\sum_{i=1}^n\frac{1}{\sigma_u^2+D_i})^{-2}\\
& +O(n^{-2}).
\end{aligned}
$$

Combining the above results and the fact the $\phi'(t)=-t\phi(t)$, the theorem is proven.

■

Table 1: Coverage probabilities (CP) and coverage length (CL) for nominal 95% confidence intervals for pattern (a). The first number in each cell is the result for our proposed method and the second one (in parentheses) is the naive FH method.

		Both Normal Distribution					Both Centered Chi-squared Distribution				
		0.7	0.6	0.5	0.4	0.3	0.7	0.6	0.5	0.4	0.3
n=15	CP	0.967 (0.941)	0.964 (0.941)	0.965 (0.946)	0.964 (0.949)	0.963 (0.952)	0.957 (0.912)	0.959 (0.918)	0.963 (0.930)	0.967 (0.940)	0.971 (0.952)
	CL	3.016 (2.675)	2.833 (2.526)	2.764 (2.450)	2.503 (2.232)	2.246 (2.002)	4.218 (2.314)	4.039 (2.200)	4.448 (2.248)	4.182 (2.083)	4.195 (1.971)
n=60	CP	0.951 (0.948)	0.950 (0.948)	0.950 (0.949)	0.951 (0.950)	0.951 (0.950)	0.936 (0.922)	0.937 (0.924)	0.941 (0.929)	0.943 (0.933)	0.950 (0.942)
	CL	2.582 (2.559)	2.464 (2.446)	2.314 (2.299)	2.134 (2.125)	1.915 (1.909)	2.843 (2.356)	2.781 (2.263)	2.659 (2.137)	2.526 (1.989)	2.399 (1.809)

Table 2: Coverage probabilities (CP) and coverage length (CL) for nominal 95% confidence intervals for pattern (b). The first number in each cell is the result for our proposed method and the second one (in parentheses) is the naive FH method.

		Both Normal Distribution					Both Centered Chi-squared Distribution				
		2.0	0.6	0.5	0.4	0.2	2.0	0.6	0.5	0.4	0.2
n=15	CP	0.969	0.966	0.966	0.965	0.961	0.949	0.960	0.968	0.966	0.977
		(0.917)	(0.937)	(0.944)	(0.945)	(0.955)	(0.874)	(0.911)	(0.932)	(0.932)	(0.966)
	CL	3.989	2.855	2.813	2.525	1.953	4.221	3.662	4.143	3.761	3.878
		(3.402)	(2.523)	(2.462)	(2.231)	(1.708)	(2.745)	(2.148)	(2.227)	(2.027)	(1.801)
n=60	CP	0.951	0.950	0.950	0.951	0.951	0.937	0.938	0.941	0.944	0.959
		(0.945)	(0.948)	(0.949)	(0.950)	(0.951)	(0.909)	(0.922)	(0.927)	(0.933)	(0.952)
	CL	3.346	2.470	2.318	2.138	1.622	3.355	2.716	2.598	2.464	2.159
		(3.267)	(2.449)	(2.302)	(2.127)	(1.619)	(2.953)	(2.251)	(2.126)	(1.979)	(1.554)

Table 3: Coverage probabilities (CP) and coverage length (CL) for nominal 95% confidence intervals for pattern (c). The first number in each cell is the result for our proposed method and the second one (in parentheses) is the naive FH method.

		Both Normal Distribution					Both Centered Chi-squared Distribution				
		4.0	0.6	0.5	0.4	0.1	4.0	0.6	0.5	0.4	0.1
n=15	CP	0.934	0.949	0.960	0.957	0.958	0.768	0.930	0.959	0.955	0.984
		(0.878)	(0.910)	(0.926)	(0.921)	(0.951)	(0.723)	(0.848)	(0.903)	(0.894)	(0.978)
	CL	4.275	2.689	2.688	2.378	1.454	3.580	2.519	2.955	2.496	2.471
		(3.639)	(2.453)	(2.416)	(2.171)	(1.268)	(2.831)	(1.928)	(2.064)	(1.808)	(1.446)
n=60	CP	0.950	0.950	0.950	0.951	0.951	0.911	0.936	0.938	0.943	0.974
		(0.940)	(0.946)	(0.947)	(0.948)	(0.949)	(0.893)	(0.903)	(0.908)	(0.915)	(0.966)
	CL	3.695	2.472	2.321	2.141	1.203	3.410	2.419	2.299	2.153	1.498
		(3.570)	(2.446)	(2.300)	(2.124)	(1.190)	(3.204)	(2.218)	(2.096)	(1.949)	(1.163)

Table 4: Coverage probabilities (CP) and coverage length (CL) for nominal 95% confidence intervals for the proposed method for the scenarios considered in ?).

		Pattern a					Pattern b				
		4.0	0.6	0.5	0.4	0.2	8.0	1.2	1.0	0.8	0.4
n=15	CP	0.965	0.968	0.968	0.967	0.964	0.974	0.967	0.967	0.966	0.963
	CL	4.522	2.977	3.008	2.699	2.222	6.911	4.870	5.109	4.596	4.105
n=60	CP	0.951	0.950	0.951	0.951	0.951	0.951	0.950	0.951	0.951	0.951
	CL	3.715	2.492	2.341	2.162	1.655	5.261	3.531	3.317	3.063	2.345

Table 5: Coverage probabilities (CP) and coverage length (CL) for nominal 95% confidence intervals the difference between two small area means for pattern (a). The first number in each cell is the result for our proposed method and the second one (in parentheses) is the naive FH method.

		Both Normal Distribution					Both Centered Chi-squared Distribution				
		0.7	0.6	0.5	0.4	0.3	0.7	0.6	0.5	0.4	0.3
n=15	CP	0.964 (0.934)	0.965 (0.941)	0.963 (0.941)	0.963 (0.947)	0.964 (0.953)	0.953 (0.874)	0.956 (0.892)	0.955 (0.895)	0.960 (0.927)	0.967 (0.949)
	CL	4.062 (3.639)	3.936 (3.522)	3.706 (3.323)	3.497 (3.126)	3.136 (2.803)	5.154 (3.032)	5.424 (3.027)	5.418 (2.924)	5.749 (2.895)	5.751 (2.734)
n=60	CP	0.952 (0.950)	0.948 (0.945)	0.951 (0.949)	0.953 (0.951)	0.952 (0.951)	0.924 (0.900)	0.927 (0.907)	0.933 (0.915)	0.935 (0.919)	0.943 (0.929)
	CL	3.618 (3.586)	3.463 (3.437)	3.254 (3.233)	3.030 (3.016)	2.699 (2.690)	3.843 (3.284)	3.817 (3.168)	3.640 (2.992)	3.645 (2.831)	3.327 (2.542)

Table 6: Coverage probabilities (CP) and coverage length (CL) for nominal 95% confidence intervals the difference between two small area means for pattern (b). The first number in each cell is the result for our proposed method and the second one (in parentheses) is the naive FH method.

		Both Normal Distribution					Both Centered Chi-squared Distribution				
		2.0	0.6	0.5	0.4	0.2	2.0	0.6	0.5	0.4	0.2
n=15	CP	0.956	0.965	0.964	0.963	0.964	0.867	0.957	0.960	0.960	0.969
		(0.903)	(0.935)	(0.936)	(0.944)	(0.957)	(0.752)	(0.879)	(0.8890)	(0.921)	(0.963)
	CL	5.192	3.959	3.729	3.532	2.734	4.372	4.893	4.879	5.208	5.376
		(4.514)	(3.516)	(3.318)	(3.128)	(2.397)	(3.388)	(2.948)	(2.846)	(2.828)	(2.512)
n=60	CP	0.951	0.947	0.950	0.953	0.953	0.922	0.927	0.935	0.939	0.951
		(0.945)	(0.945)	(0.949)	(0.952)	(0.952)	(0.887)	(0.903)	(0.913)	(0.921)	(0.940)
	CL	4.661	3.470	3.258	3.035	2.288	4.391	3.719	3.541	3.555	3.001
		(4.551)	(3.440)	(3.235)	(3.019)	(2.284)	(4.086)	(3.149)	(2.974)	(2.817)	(2.186)

Table 7: Coverage probabilities (CP) and coverage length (CL) for nominal 95% confidence intervals the difference between two small area means for pattern (c). The first number in each cell is the result for our proposed method and the second one (in parentheses) is the naive FH method.

		Both Normal Distribution					Both Centered Chi-squared Distribution				
		4.0	0.6	0.5	0.4	0.1	4.0	0.6	0.5	0.4	0.1
n=15	CP	0.916	0.941	0.945	0.956	0.960	0.662	0.896	0.918	0.951	0.973
		(0.865)	(0.904)	(0.909)	(0.920)	(0.954)	(0.629)	(0.788)	(0.821)	(0.869)	(0.968)
	CL	5.517	3.710	3.496	3.336	2.045	4.212	3.284	3.237	3.496	3.462
		(4.758)	(3.408)	(3.219)	(3.048)	(1.748)	(3.573)	(2.607)	(2.515)	(2.537)	(2.030)
n=60	CP	0.948	0.947	0.950	0.953	0.953	0.881	0.925	0.922	0.940	0.967
		(0.941)	(0.943)	(0.947)	(0.950)	(0.952)	(0.870)	(0.894)	(0.891)	(0.910)	(0.954)
	CL	5.130	3.470	3.258	3.041	1.698	4.666	3.312	3.126	3.098	2.095
		(4.958)	(3.435)	(3.231)	(3.016)	(1.680)	(4.435)	(3.100)	(2.925)	(2.774)	(1.638)