Some Contributions to Small Area Estimation

by

Mahmoud Torabi, B.Sc., M.Sc.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

School of Mathematics and Statistics
Ottawa-Carleton Institute for Mathematics and Statistics

Carleton University
Ottawa, Ontario, Canada

December, 2006

© Copyright
2006, Mahmoud Torabi

Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-23302-3

Our file Notre référence

ISBN: 978-0-494-23302-3

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

This dissertation consists of nine chapters. Chapter 1 is an introduction to small area estimation. It presents a review of small area problem and related estimation approaches including empirical best linear unbiased prediction (EBLUP) approach and Bayes approach. EBLUP approach includes EBLUP estimators and mean squared error (MSE) approximation. Bayes approach includes empirical Bayes (EB) analysis and MSE approximation by using the jackknife method. Chapters 2 to 8 contain several developments in the topic of small area estimation. Chapter 9 gives some topics for future research.

In Chapter 2, a new generalized regression (GREG) estimator of a small area mean under a two-level model is studied. We derive the MSE of the new GREG estimator under the model and compare it to the MSE of the BLUP estimator. We also provide empirical results on the relative efficiency of the estimators when the model parameters are estimated.

In Chapter 3, we derive a second order approximation to the MSE of the pseudo-EBLUP estimator, introduced by You and Rao (2002), of a small area mean. Using this approximation, an estimator of MSE that is approximately unbiased is derived. Empirical results on the performance of the proposed approximation are also presented.

In Chapter 4, a sub-area level small area model is introduced. Not only may this model be used to estimate small area means by borrowing strength from related area levels, but also it can borrow strength from sub-area levels to obtain more efficient sub-area estimators. We obtain second order approximations to the MSE

of the model-based small area estimators and then derive MSE estimators unbiased to second order. Results of a simulation study on the accuracy of the second order approximation to the MSE and the bias of the estimators of the MSE are also reported.

In Chapter 5, a conditional MSE estimation for the nested error linear regression model is studied. We derive a second order approximation to the conditional MSE given the area-specific response data. We then derive a nearly unbiased estimator of the conditional MSE. A data analysis on the performance of our approach is also provided.

In Chapter 6, we derive second order approximations of the MSE and the conditional MSE given $y_i = (y_{i1}, \dots, y_{in_i})'$ of $\hat{\mu}_{iw}$ under the nested error linear regression model. Moreover, we derive the estimators of MSE and conditional MSE of $\hat{\mu}_{iw}$ that are nearly unbiased.

In Chapter 7, a structural measurement error model with unit-specific covariates x_{ij} in the nested error regression model is studied. We study the asymptotic behavior of Bayes risk of EB estimators, along the lines of Ghosh and Sinha (2006), and establish asymptotic optimality of EB estimators. Moreover, we find a second order approximation to the MSE of EB estimators and obtain a MSE estimator which is nearly unbiased. A simulation study on the performance of the proposed method is also provided.

In Chapter 8, we study the functional measurement error nested error regression model with unit level covariates x_{ij} . We study the asymptotic behavior of Bayes risk of pseudo-EB (PEB) estimators and establish the asymptotic optimality of the pseudo-EB estimators. Moreover, we find a second order approximation to the MSE

of PEB estimators and obtain a MSE estimator which is nearly unbiased. Empirical results on the performance of the proposed pseudo-EB estimators are also presented.

Finally, some suggestions for future research are presented in Chapter 9.

Acknowledgements

First of all, I would like to express my sincere gratitude to Professor J.N.K. Rao, my thesis supervisor, for his constant and patient guidance and especially for the invaluable research experience I learned in the research of this thesis, as well as for his kindness and support during these years.

I would like to thank Professors G.S. Datta, M.Alvo, M. Mojirsheibani and D.R. Thomas for serving as members of my thesis examining committee. I also thank Professor M. Mojirsheibani who conducted the two comprehensive exams for my Ph.D. program and taught me some statistics courses.

I am grateful to the School of Mathematics and Statistics, Carleton University, for considering me as the most outstanding and deserving graduate (Ph.D.) student and awarding me D. Chand and R.D. Marwah award. I am also grateful to my thesis examining committee to consider my research as extraordinary and outstanding, and to recommend me to the university for the Senate medal for outstanding academic achievement.

I would also like to thank the School of Mathematics and Statistics, Carleton University, for providing the necessary research facilities for completing this thesis. My thanks also go to the Faculty members of the Ottawa-Carleton institute, who provided the excellent academic environment with many interesting courses and seminars. I also thank my friend, Qunshu Ren, for his friendship and for very helpful discussions that we had during the Ph.D. program.

I wish to thank my parents for their love and concerns. I wish to give my heartfelt thanks to my wife, Elahe. I am forever grateful for her continuous support, encour-

agement and understanding and without whom this thesis may never have seen the light.

Contents

Title Page	i
Acceptance Sheet	ii
Abstract	iii
Acknowledgements	vi
Contents	viii
List of Tables	xvii
1 An Overview of Small Area Estimation	1
1.1 Introduction	1
1.2 General Linear Mixed Model with a Block Diagonal Covariance Structure	4
1.3 Small Area Models	5

1.3.1	Area Level Model	5
1.3.2	Unit Level Model	7
1.4	Model-based Estimation	8
1.4.1	EB Approach	8
1.4.2	EBLUP Approach	9
1.4.3	MSE Estimation of EB Estimator	10
1.4.4	MSE Estimation of EBLUP	11
1.5	Some Background	15
1.6	Thesis Outline	16
2	Small Area Estimation Under a Two-level Model	19
2.1	Introduction	19
2.2	BLUP and GREG Estimators	20
2.3	Mean Squared Error	22
2.4	Empirical Results	27
2.4.1	Empirical Comparison of MSE Values	27
2.4.2	Simulation Study	29
2.5	Conclusion	32
3	Pseudo EBLUP and Survey Weights	36

3.1	Introduction	37
3.2	Pseudo EBLUP	37
3.3	MSE Estimation of Pseudo EBLUP	40
3.3.1	You and Rao Method	40
3.3.2	The Proposed Method	42
3.4	Empirical Results	46
3.5	Summary	48
3.6	Proofs	49
3.6.1	Proof of Theorem 3.1	49
3.6.2	Proof of Theorem 3.2	52
3.6.3	Proof of Theorem 3.3	53
4	On Small Area Estimation Under a Sub-area Level Model	58
4.1	Introduction	59
4.2	EBLUP Estimation	61
4.2.1	Estimation of Variance Components	62
4.3	MSE Approximation	64
4.3.1	MSE Approximation of $\hat{\mu}_{ij}$	64
4.3.2	MSE Approximation of $\hat{\mu}_i$	66

4.4	Variances and Covariance of $\tilde{\sigma}_\nu^2$ and $\tilde{\sigma}_u^2$	67
4.5	MSE Estimation	74
4.6	Bias of $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_u^2$	76
4.7	Simulation Study	79
4.8	Summary	83
4.9	Proofs	83
4.9.1	Proof of (4.12)	83
4.9.2	Proof that the reminder term in $MSE(\hat{\mu}_{ij})$ is $o(m^{-1})$	85
4.9.3	Proof of $(n - p)^{-1}y'A_1^1y = O_p(1)$	88
4.9.4	Proof of $(n - p)^{-1}y'A_1^2y = O_p(1)$	89
4.9.5	Proof of $(m - p)^{-1}\bar{y}'A_2^1\bar{y} = O_p(1)$	89
4.9.6	Proof of $(m - p)^{-1}\bar{y}'A_2^2\bar{y} = O_p(1)$	89
4.9.7	Proof that $E[mse(\hat{\mu}_{ij})] - MSE(\hat{\mu}_{ij}) = o(m^{-1})$	90
4.9.8	Some Lemmas	92
5	Estimation of Conditional MSE in Small Area Estimation	109
5.1	Introduction	109
5.2	Second Order Approximation to Conditional MSE	111
5.3	Estimation of Conditional MSE	117

5.4	Conditional Bias of $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$	118
5.5	Data Analysis	120
5.6	Summary	122
5.7	Proofs	122
5.7.1	Proof of $\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} = O_p(m^{-1/2})$	122
5.7.2	Proof of $\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = O_p(m^{-1/2})$	123
5.7.3	Proof of $E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 y_i] = E(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 + o_p(m^{-1})$	124
5.7.4	Proof of $E[g_{2i}(\hat{\delta}) y_i] = g_{2i}(\delta) + o_p(m^{-1})$	125
5.7.5	Proof of $E[g_{3i}^c(\hat{\theta}) y_i] = g_{3i}^c(\theta) + o_p(m^{-1})$	126
6	Estimation of MSE of Weighted Predictors	133
6.1	Introduction	134
6.2	Second Order Approximation of the MSE of Weighted Predictors . .	135
6.3	MSE Estimation for Weighted Predictors	139
6.4	Conditional MSE Approximation of Weighted Predictors	140
6.5	Conditional MSE Estimation of Weighted Predictors	146
6.6	Summary	152
7	Structural Measurement Error Models Under Nested Error Linear	

Regression Model (EB Approach) 153

7.1	Introduction	154
7.2	Empirical Bayes	155
7.3	MSE Approximation	157
7.4	Variances and Covariances of Model Parameter Estimators	160
7.5	MSE Estimation	166
7.6	Asymptotic Result on $m^{-1} \sum_{i=1}^m E(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B)^2$	168
7.7	Simulation Study	170
7.8	Summary	174
7.9	Appendix	175

8 Functional Measurement Error Models Under Nested Error Linear

Regression Model (EB Approach) 178

8.1	Introduction	179
8.2	Empirical Bayes	180
8.3	MSE Approximation	182
8.4	Variances and Covariances of Model Parameter Estimators	185
8.5	MSE Estimation	192
8.6	Asymptotic Result on $m^{-1} \sum_{i=1}^m E(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})^2$	194
8.7	Simulation Study	194

8.8 Summary	200
9 Future Research	204
Bibliography	207

List of Tables

2.1	Parameter estimates based on Moura and Holt's data set (1999), where σ_0^2 , σ_1^2 and σ_2^2 are the diagonal elements and σ_{01} , σ_{02} and σ_{12} are the off-diagonal elements of the covariance matrix Σ_ν	33
2.2	Comparison of small area estimators: relative efficiency (EFF) and average MSE ($\overline{MSE}$) for the case of diagonal covariance matrix based on Moura and Holt's data set (1999)	34
2.3	Comparison of small area estimators: relative efficiency (EFF) and average MSE ($\overline{MSE}$) for the case of a general covariance matrix based on Moura and Holt's data set (1999)	34
2.4	Comparison of small area estimators: average MSE ($\overline{MSE}_1$) and average absolute relative error ($\overline{ARE}_1$) under a model-based framework	34
2.5	Comparison of small area estimators: average MSE ($\overline{MSE}_2$) and average absolute relative error ($\overline{ARE}_2$) under a finite population framework ($n_i = 5$)	35

2.6	Comparison of small area estimators: average MSE ($\overline{MSE}_2$) and average absolute relative error ($\overline{ARE}_2$) under a finite population framework ($n_i=10$)	35
3.1	MSE estimators and cross-product term $2(C_{1iw} + \tilde{C}_{2iw}) = 2C_{iw}$ for corn when weights are based on $x_{1ij} + x_{2ij}$	55
3.2	The cross-product terms and the estimators of MSE for corn where weights are based on x_{1ij} only	56
3.3	The cross-product terms and the estimators of MSE for corn based on adhoc weights	57
4.1	Average mean squared error of sub-area EBLUP	104
4.2	Average mean squared error of area EBLUP	104
4.3	Percent relative bias of variance component estimators ($\hat{\sigma}_v^2, \hat{\sigma}_u^2$)	105
4.4	Percent average relative bias of sub-area naive MSE estimator	105
4.5	Percent average relative bias of area naive MSE estimator	106
4.6	Percent average relative bias of sub-area MSE estimator	106
4.7	Percent average relative bias of area MSE estimator	106
4.8	Monte carlo probabilities ($pr(\tilde{\sigma}_v^2 \leq 0), pr(\tilde{\sigma}_u^2 \leq 0)$)	107
4.9	Percent average relative efficiency of sub-area EBLUP over direct estimator	107
4.10	Percent average relative efficiency of area EBLUP over direct estimator	107

4.11	Percent average relative efficiency of sub-area EBLUP over the Fay-Herriot estimator	108
5.1	The components of conditional MSE estimation of predicted hectares of corn by Battese et al. data set (1988) for our approach and Booth and Hobert approach (1998)	132
7.1	Percent average relative bias and percent absolute relative error of model parameters based on the proposed, naive and true methods for $m = 768$	177
7.2	Standard error of EB estimators for the proposed, naive and true methods for different number of areas	177
8.1	Percent average relative bias and percent absolute relative error of model parameters based on the proposed, naive and true methods for $m = 768$	201
8.2	Empirical MSE of PEB estimators for the proposed method and similar estimators for the naive and true methods	201
8.3	Leading terms of EMSE of PEB, naive and true methods	202
8.4	Empirical values of $EMSE_{2p}$, $EMSE_{2n}$ and $EMSE_{2t}$	202
8.5	Empirical values of the cross-product terms $EMSE_{3p}$, $EMSE_{3n}$ and $EMSE_{3t}$	203

Chapter 1

An Overview of Small Area Estimation

1.1 Introduction

Sample surveys have been widely used to obtain information on wide-ranging topics of interest at frequent intervals over time. Sample surveys are generally designed to provide reliable direct estimates for the total population of interest as well as a variety of subpopulations (domains). A direct domain estimate is based on the domain specific sample data. Increasingly, such surveys are also being used to provide estimates for small areas, which are needed particularly for policy making to allocate government funds in regional programs. In this context, a domain (area) is regarded as large if the domain-specific sample is large enough to provide direct estimates of adequate precision. On the other hand, a domain is regarded as small if the domain-

specific sample is not large enough to support direct estimates of adequate precision. Consequently, the term "small area" is commonly used to denote any domain which cannot produce direct estimates of adequate precision. This makes it necessary to use "indirect" estimators that "borrow strength" from related areas to find more accurate estimates for a given area or, simultaneously, for several areas. Such estimates are based on either implicit or explicit models that provide a link to related small areas through supplementary data such as administrative records, auxiliary information and recent census counts. Mainly, these small areas estimators may be classified into two types : (1) Traditional indirect estimators based on implicit linking models, and (2) Model-based indirect estimators based on explicit linking models.

Traditional indirect estimators proposed in the literature include synthetic and composite estimators. These estimators are generally design-based and their design variances are usually small relative to the design variances of direct estimators. However, the design bias of such indirect estimators will not decrease as the overall sample size increases. If the implicit linking model is close to the true model, then the design bias will be small, leading to significantly smaller design mean squared error (MSE) compared to the MSE of a direct estimator. In fact, having smaller MSE is the main reason for using indirect estimators.

Model-based indirect estimators define the way that the related areas are incorporated in the estimation process. These estimators include empirical best linear unbiased prediction (EBLUP), parametric empirical Bayes (EB) and parametric hierarchical Bayes (HB). The model-based approach to small area estimation has some advantages such as, optimal estimators can be derived under the assumed model. Moreover, area specific measures of variability can be associated with each estimator

unlike global measures often used with traditional indirect estimators. In addition, the models can be validated from the sample data. Also, different models can be studied depending on the nature of the response variables and the complexity of data structures. These are main reasons that in this research we focus on model-based indirect estimators.

In recent years, model-based estimators have widely been used in small area estimation due to the development of the EBLUP and EB theories. The major progress for the EBLUP method is the work on the mean squared error (MSE) approximation of EBLUP estimators by Kackar and Harville (1984) and followed by Prasad and Rao (1990). After the important work of Prasad and Rao (1990), several small area models associated with EBLUP estimators and their MSE estimators have been proposed. These models include correlated sampling errors (Datta et al., 1992), time series models (Rao and Yu, 1992, 1994), spatial models (Cressie, 1991), random error variance linear mixed model (Kleffe and Rao, 1992), two-fold nested error regression model (Stukel and Rao, 1999), and two-level model (Moura and Holt, 1999).

In the rest of this Chapter, we review some basic small area models. We classify small area models into two broad types: (i) Area level model that relates small area direct estimators to area-specific covariates. (ii) Unit level model that relates the unit values of a study variable to unit-specific covariates. We then study these two models with their properties. Before that, we need to introduce the more general model where two models mentioned above are special cases of this model.

1.2 General Linear Mixed Model with a Block Diagonal Covariance Structure

Suppose that the sample data obey the general linear mixed model

$$y = X\beta + Z\nu + e, \quad (1.1)$$

where y is the $n \times 1$ vector of sample observations, X and Z are known $n \times p$ and $n \times h$ matrices of full rank, and ν and e are independent random vectors with means 0 and covariance matrices G and R . The covariance matrices G and R may depend on some variance parameters $\delta = (\delta_1, \delta_2, \dots, \delta_q)'$. We further assume that δ belongs to a specified subset of Euclidean q -space such that $Var(y) = V = V(\delta) = R + ZGZ'$ is nonsingular for all δ belonging to the subset, where $Var(y)$ denotes the variance-covariance matrix of y . We are interested in estimating a linear combination of fixed effects β and realized value of ν , $\mu = l'\beta + h'\nu$, for specified vectors, l and h , of constants.

A special case of the general linear mixed model (1.1) with a block diagonal covariance structure covers many small area models considered in the literature including area level model and unit level model based on independent sampling across areas. For this model

$$y = col_{1 \leq i \leq m}(y_i), \quad X = col_{1 \leq i \leq m}(X_i), \quad Z = diag_{1 \leq i \leq m}(Z_i), \quad \nu = col_{1 \leq i \leq m}(\nu_i),$$

$$e = col_{1 \leq i \leq m}(e_i),$$

where m is the number of small areas, X_i is $n_i \times p$, Z_i is $n_i \times h_i$ and y_i is an $n_i \times 1$ vector such that $\sum_{i=1}^m n_i = n$ and $\sum_{i=1}^m h_i = h$. Moreover, $Var(y)$ has a block diagonal covariance structure: $V = Var(y) = diag_{1 \leq i \leq m}(V_i)$ with $V_i = R_i + Z_i G_i Z_i'$ where $R = diag_{1 \leq i \leq m}(R_i)$, $G = diag_{1 \leq i \leq m}(G_i)$.

Hence, the model (1.1) may be decomposed into m submodels

$$y_i = X_i\beta + Z_iv_i + e_i, i = 1, \dots, m. \quad (1.2)$$

Under (1.2), we are interested in estimating linear combinations of β and v_i :

$$\mu_i = l_i'\beta + h_i'v_i, i = 1, \dots, m. \quad (1.3)$$

The two-level model is a special case of (1.2) where $X_i = X_iZ_i$ and $Z_i = X_i$. In this thesis, we focus on the model (1.2) with the block diagonal covariance structure.

1.3 Small Area Models

1.3.1 Area Level Model

For the area level model, only area-specific auxiliary data $x_i = (x_{i1}, \dots, x_{ip})'$ are available for the areas $i = 1, \dots, m$. Using this auxiliary data, we are interested in estimating, for example, the small area mean μ_i which is assumed to be related to x_i through a linear model with random area effects v_i :

$$\mu_i = x_i'\beta + v_i, i = 1, \dots, m \quad (1.4)$$

where $\beta = (\beta_1, \dots, \beta_p)'$ is the $p \times 1$ vector of regression coefficients. Further, the v_i 's are assumed to be independent and identically distributed (iid) with mean zero and variance σ_v^2 . Normality of the random effects v_i is also often assumed. The parameter σ_v^2 is a measure of homogeneity of the areas after accounting for the covariates x_i . Moreover, we assume a model of the form (1.4) for the population, that is, $\mu_i =$

$x_i'\beta + v_i, i = 1, \dots, M$, where we have M areas in the population and only m areas are selected in the sample. We further assume that the sample areas obey the population model, that is, the bias in the sample selection of areas is absent so that (1.4) holds for the sampled areas.

For making inference about the μ_i 's, we assume that direct survey estimators, $\hat{\mu}_i$, are available and it is customary to assume that

$$\hat{\mu}_i = \mu_i + e_i, i = 1, \dots, m \quad (1.5)$$

where the sampling errors e_i are independent with mean zero and known variance ψ_i . If the sampling variances ψ_i are unknown, we may estimate ψ_i from unit-level sample data and then smooth the estimated variances $\hat{\psi}_i$ to get more stable estimates of ψ_i . Smoothed estimates are then taken as ψ_i .

Combining the sampling model (1.5) with the linking model (1.4), we get the small area level model as

$$\hat{\mu}_i = x_i'\beta + v_i + e_i, i = 1, \dots, m \quad (1.6)$$

Note that model (1.6) involves design-based errors e_i as well as model-based errors v_i . We assume that v_i and e_i are independent. It is easily seen that models (1.4) and (1.6) are special cases of (1.3) and (1.2), respectively. The area level model (1.6) is also known as the Fay-Herriot model (Fay and Herriot, 1979).

1.3.2 Unit Level Model

For the unit level model, unit-specific auxiliary data $x_{ij} = (x_{ij1}, \dots, x_{ijp})'$ are available for the population units. It is often sufficient to assume that only the population means $\bar{X}_i$ are known. Further, unit values y_{ij} are assumed to be related to x_{ij} through a nested error regression model

$$y_{ij} = x_{ij}'\beta + v_i + e_{ij}; j = 1, \dots, N_i; i = 1, \dots, m \quad (1.7)$$

where N_i is the number of population units in the i -th area and the area-specific effects v_i are assumed to be iid random variables with mean zero and variance σ_v^2 . In addition, the sampling errors e_{ij} are assumed to be iid random variables with mean zero and variance σ_e^2 . We assume that v_i and e_{ij} are independent. Also, normality of the v_i 's and e_{ij} 's is often assumed. The mean for the i -th area may be written as

$$\mu_i = \bar{X}_i'\beta + v_i, i = 1, \dots, m \quad (1.8)$$

It is seen that models (1.7) and (1.8) are special cases of (1.2) and (1.3), respectively. We further assume that a sample of size n_i is taken from the N_i units in the i -th area ($i=1, \dots, m$) and that the sample values also obey the assumed model (1.7). The latter assumption is satisfied under simple random sampling (SRS) from each area. We also assume that the population size N_i is large and $\bar{X}_i$ is the known mean of the x_{ij} for the i -th area. Noting that $\mu_i = \bar{X}_i'\beta + v_i + \bar{E}_i$ and $\bar{E}_i \approx 0$, where $\bar{E}_i$ is the mean of the N_i errors e_{ij} . It should be mentioned that the estimation of μ_i in model (1.8) is equivalent to the estimation of a linear combination of β and the realization of the random variable v_i .

1.4 Model-based Estimation

We now present the empirical best (or Bayes), EB, and the empirical best linear unbiased prediction, EBLUP, approaches to small area estimation. Then, we consider the corresponding measures of variability of the estimators for the two basic models, area level and unit level models. Rao (2003) gives a detailed treatment of model-based small area estimation.

1.4.1 EB Approach

In the EB approach, the fundamental tool is the Bayes theorem, but no prior distribution on model parameters is assumed. EB approach is applicable more generally than the EBLUP approach in the sense of handling models for binary and count data as well as normal linear mixed models. In the latter case, EB and EBLUP estimators are identical.

In the EB approach, the posterior density, $f(\mu|y, \lambda)$, of the parameters of interest, μ , given the data and model parameters λ is first obtained using the conditional density, $f(y|\mu, \lambda_1)$, of y given μ and the density $f(\mu|\lambda_2)$ of μ , assuming that the model parameters $\lambda' = (\lambda'_1, \lambda'_2)$ are known. The model parameters are estimated from the marginal density, $f(y|\lambda)$, of y , and inferences are then based on the estimated posterior density. The density of μ is often interpreted as prior density on μ , but it is actually a part of the assumed model on (y, μ) and it can be validated from the data. In this sense the EB approach is essentially frequentist, and EB inferences refer to averaging over the joint distribution of y and μ .

In the parametric empirical Bayes (PEB) approach, a parametric form is assumed for the distribution of μ . On the other hand, nonparametric empirical Bayes (NPEB) methods do not specify the form of the distribution of μ . We refer to Martiz and Lwin (1989) and Carlin and Louis (2000) for excellent accounts of the EB methodology.

1.4.2 EBLUP Approach

For the general linear mixed model (1.2) with a block diagonal covariance structure, the best linear unbiased prediction (BLUP) estimator of $\mu_i = l_i'\beta + h_i'\nu_i$ is given by

$$\tilde{\mu}_i(\delta, y_i) = l_i'\tilde{\beta} + h_i'G_iZ_i'V_i^{-1}(y_i - X_i\tilde{\beta}), \quad (1.9)$$

where

$$\tilde{\beta} = \left(\sum_{i=1}^m X_i'V_i^{-1}X_i \right)^{-1} \sum_{i=1}^m X_i'V_i^{-1}y_i,$$

(Henderson, 1975). The EBLUP estimator of μ_i is then given by $\hat{\mu}_i = \tilde{\mu}_i(\hat{\delta}, y_i)$.

Using (1.9), the BLUP estimator of $\mu_i = x_i'\beta + v_i$ under the area level model (1.6) is given by

$$\tilde{\mu}_i(\delta, y_i) = x_i'\tilde{\beta}_1 + \gamma_{i1}(y_i - x_i'\tilde{\beta}_1),$$

where $\gamma_{i1} = \sigma_v^2/(\sigma_v^2 + \psi_i)$, y_i is area-specific response data and $\tilde{\beta}_1 = [\sum_{i=1}^m x_i x_i' / (\psi_i + \sigma_v^2)]^{-1} [\sum_{i=1}^m x_i y_i / (\psi_i + \sigma_v^2)]$. The EBLUP estimator of μ_i is given by $\hat{\mu}_i = \tilde{\mu}_i(\hat{\delta}, y_i)$.

Similarly, the BLUP estimator of $\mu_i = \bar{X}_i'\beta + v_i$ under the unit level model (1.7) is given by

$$\tilde{\mu}_i(\delta, \bar{y}_i) = \bar{X}_i'\tilde{\beta}_2 + \gamma_{i2}(\bar{y}_i - \bar{x}_i'\tilde{\beta}_2),$$

where $\gamma_{i2} = \sigma_v^2/(\sigma_v^2 + \sigma_e^2/n_i)$, $\tilde{\beta}_2 = (\sum_{i=1}^m X_i' V_i^{-1} X_i)^{-1} \sum_{i=1}^m X_i' V_i^{-1} y_i$, and $\bar{X}_i$ and $\bar{x}_i$ are the population mean and the sample mean for the i -th area, respectively. The EBLUP estimator of μ_i is given by $\hat{\mu}_i = \tilde{\mu}_i(\hat{\delta}, \bar{y}_i)$. Note that $EBLUP = EB$ under normality.

1.4.3 MSE Estimation of EB Estimator

A measure of variability associated with the EB estimator $\hat{\mu}_i^{EB}$ of μ_i is given by the mean squared error, $MSE(\hat{\mu}_i) = E(\hat{\mu}_i^{EB} - \mu_i)^2$. We may write the MSE of EB estimator of i -th area mean, $\hat{\mu}_i^{EB}$, as follows:

$$MSE(\hat{\mu}_i^{EB}) = E(\hat{\mu}_i^{EB} - \mu_i)^2 = E(\hat{\mu}_i^{EB} - \tilde{\mu}_i^B)^2 + E(\tilde{\mu}_i^B - \mu_i)^2, \quad (1.10)$$

where $\tilde{\mu}_i^B = E(\mu_i | y_i, \beta, \delta)$ and noting that $E(\hat{\mu}_i^{EB} - \tilde{\mu}_i^B)(\tilde{\mu}_i^B - \mu_i) = E[(\hat{\mu}_i^{EB} - \tilde{\mu}_i^B)E(\tilde{\mu}_i^B - \mu_i | y_i)] = 0$. In fact, we have an exact expression for the last term of the right hand side of (1.10) as $E(\tilde{\mu}_i^B - \mu_i)^2 \equiv g_{1i}(\delta)$. In addition, we can find a second order approximation to $E(\hat{\mu}_i^{EB} - \tilde{\mu}_i^B)^2 \equiv g_i(\delta)$ based on a Taylor series expansion.

Jiang, Lahiri, and Wan (2002) proposed a jackknife method of estimating the MSE of EB estimators. Let $\hat{\mu}_i^{EB} = k_i(y_i, \hat{\beta}, \hat{\delta})$ be the EB estimator of μ_i expressed as a function of y_i and the parameter estimators $\hat{\beta}$ and $\hat{\delta}$. We calculate the delete-area estimators $\hat{\beta}_{-l}$ and $\hat{\delta}_{-l}$ by deleting the l -th area data set (y_l, x_l) from the full data set $\{(y_l, x_l); l = 1, \dots, m\}$. This calculation is done for each l in turn to get m estimators of β and δ : $\{(\hat{\beta}_{-l}, \hat{\delta}_{-l}); l = 1, \dots, m\}$ which, in turn, provide m estimators of μ_i : $\{\hat{\mu}_{i,-l}^{EB}; l = 1, \dots, m\}$, where $\hat{\mu}_{i,-l}^{EB} = k_i(y_i, \hat{\beta}_{-l}, \hat{\delta}_{-l})$.

We now define the Jiang et al. (2002) jackknife estimator of $MSE(\hat{\mu}_i^{EB})$. A

jackknife estimator of $g_i(\theta)$ is given by

$$\widehat{g_i(\delta)} = \frac{m-1}{m} \sum_{l=1}^m (\hat{\mu}_{i,-l}^{EB} - \hat{\mu}_i^{EB})^2.$$

Moreover, a jackknife estimator of $g_{1i}(\delta)$ is given by

$$\widehat{g_{1i}(\delta)} = g_{1i}(\hat{\delta}) - \frac{m-1}{m} \sum_{l=1}^m [g_{1i}(\hat{\delta}_{-l}) - g_{1i}(\hat{\delta})].$$

Hence, the jackknife estimator of $MSE(\hat{\mu}_i^{EB})$ is

$$mse_J(\hat{\mu}_i^{EB}) = \widehat{g_i(\delta)} + \widehat{g_{1i}(\delta)}.$$

Note that $\widehat{g_{1i}(\delta)}$ tracks the MSE when the model parameters are known, while $\widehat{g_i(\delta)}$ estimates the additional variability due to estimating the model parameters.

The jackknife estimator is applicable to ML, REML, or moment estimators of model parameters β and δ . Jiang et al. (2002) proved that it is nearly unbiased. Also, we should point out that the jackknife method is applicable to the EB estimator of any function $\psi_i = h(\mu_i)$. The jackknife approach is not studied in this dissertation to find MSE estimators.

1.4.4 MSE Estimation of EBLUP

We may write the MSE of the EBLUP estimator $\hat{\mu}_i$, in general, as

$$MSE(\hat{\mu}_i) = E(\hat{\mu}_i - \mu_i)^2 = MSE(\tilde{\mu}_i) + E(\hat{\mu}_i - \tilde{\mu}_i)^2 + 2E(\hat{\mu}_i - \tilde{\mu}_i)(\tilde{\mu}_i - \mu_i), \quad (1.11)$$

where $\tilde{\mu}_i = \tilde{\mu}_i(\delta, y_i)$ and $MSE(\tilde{\mu}_i) = E(\tilde{\mu}_i - \mu_i)^2$. Under normality of the random effects ν_i and e_i defined in the model (1.2), the cross-product term in (1.11) is zero

provided $\hat{\delta}$ is translation invariant, i.e., $\hat{\delta}$ remains unchanged when y is changed to $-y$ or to $y-Xa$ for all y and a . Then $MSE(\hat{\mu}_i)$ may be written as

$$MSE(\hat{\mu}_i) = MSE(\tilde{\mu}_i) + E(\hat{\mu}_i - \tilde{\mu}_i)^2.$$

For the first term, $MSE(\tilde{\mu}_i)$, an exact expression is given by

$$MSE(\tilde{\mu}_i) = g_{1i}(\delta) + g_{2i}(\delta), \quad (1.12)$$

where

$$g_{1i}(\delta) = h_i'(G_i - G_i Z_i' V_i^{-1} Z_i G_i) h_i, \quad (1.13)$$

and

$$g_{2i}(\delta) = (l_i - X_i' V_i^{-1} Z_i G_i' h_i)' \left(\sum_{i=1}^m X_i' V_i^{-1} X_i \right)^{-1} (l_i - X_i' V_i^{-1} Z_i G_i' h_i). \quad (1.14)$$

The term $g_{2i}(\delta)$ is due to the estimation of β .

For the last term $E(\hat{\mu}_i - \tilde{\mu}_i)^2$ in (1.11), following Kackar and Harville's Taylor approximation (1984), Prasad and Rao (1990) and Datta and Lahiri (2000) proposed a further approximation under the model (1.2) with a block diagonal covariance structure. They assumed the following regularity conditions:

- (C1). The elements of X_i and Z_i are uniformly bounded such that $\sum_{i=1}^m X_i' V_i^{-1} X_i = O(m)$ for large m ,
- (C2). $l_i - X_i' V_i^{-1} Z_i G_i' h_i = [O(1)]_{p \times 1}$,
- (C3). $\sup_i n_i < \infty$ and $\sup_i b_i < \infty$,
- (C4). $\frac{\partial}{\partial \delta_j} [X_i' V_i^{-1} Z_i G_i' h_i] = [O(1)]_{p \times 1}$ for $j = 1, \dots, q$,
- (C5). $R_i(\delta) = \sum_{j=0}^q \delta_j D_{ij} D_{ij}'$ and $G_i(\delta) = \sum_{j=0}^q \delta_j F_{ij} F_{ij}'$, where $\delta_0 = 1$, D_{ij} and F_{ij} ($i = 1, \dots, m; j = 0, \dots, q$) are known matrices of order $n_i \times b_i$ and $b_i \times b_i$ respectively, and the elements are uniformly bounded known constants such that $R_i(\delta)$ and

$G_i(\delta)(i = 1, \dots, m)$ are all positive definite matrices. In special cases, some of D_{ij} and F_{ij} may be null matrices.

(C6). $\hat{\delta}$ is an estimator of δ which satisfies (i) $\hat{\delta} - \delta = O_p(m^{-1/2})$, (ii) $\hat{\delta} - \tilde{\delta} = O_p(m^{-1})$, (iii) $\hat{\delta}(-y) = \hat{\delta}(y)$, and (iv) $\hat{\delta}(y + Xl) = \hat{\delta}(y)$ for any $l \in R^p$ and for all y , where $\tilde{\delta}$ is the ML estimator of δ .

Condition (C5) reduces (1.2) to a generalization of the usual variance components model (Rao, 2001). Under (C5), we can write (1.2) as $y_i = X_i\beta + \sum_{j=0}^q E_{ij}\xi_{ij}$, where $E_{ij} = [Z_i F_{ij}, D_{ij}]$, $\xi_{ij} \stackrel{ind.}{\sim} N(0, \delta I_{2b_i})$, ($i = 1, \dots, m; j = 0, \dots, q$). Unlike the usual variance components model, we do not restrict matrices E_{ij} to 0 and 1 and entries for $E'_{ij}E_{ij}$ to be diagonal for all $i = 1, \dots, m; j = 0, \dots, q$. Condition (C6) is generally satisfied by the ML and the REML estimators of variance components.

The approximation to $E(\hat{\mu}_i - \tilde{\mu}_i)^2$ is given by $E(\hat{\mu}_i - \tilde{\mu}_i)^2 = g_{3i}(\delta) + o(m^{-1})$ under regularity conditions (C1) – (C6), where

$$g_{3i}(\delta) = tr[(\frac{db'_i}{d\delta})V_i(\frac{db'_i}{d\delta})'V_a(\hat{\delta})], \quad (1.15)$$

with $b'_i = h'_i G_i Z'_i V_i^{-1}$ and $V_a(\hat{\delta})$ is the asymptotic variance of $\hat{\delta}$. It may be noted that $g_{2i}(\delta)$ is $O(m^{-1})$ and $g_{1i}(\delta)$ is $O(1)$ for large m . Now using (1.13), (1.14) and (1.15), we get

$$MSE(\hat{\mu}_i) = g_{1i}(\delta) + g_{2i}(\delta) + g_{3i}(\delta) + o(m^{-1}). \quad (1.16)$$

A second order (or nearly) unbiased MSE estimator, $mse(\hat{\mu}_i)$, is given by

$$mse(\hat{\mu}_i) = g_{1i}(\hat{\delta}) - (\partial g_{1i} / \partial \delta') b_{\hat{\delta}}(\delta) |_{\delta=\hat{\delta}} + g_{2i}(\hat{\delta}) + 2g_{3i}(\hat{\delta}),$$

in the sense that $E[mse(\hat{\mu}_i)] = MSE(\hat{\mu}_i) + o(m^{-1})$, where $b_{\hat{\delta}}(\delta) = E(\hat{\delta} - \delta)$ is zero in the case of the REML estimator and some moment estimators.

Using (1.13)- (1.15), the MSE of EBLUP of $\hat{\mu}_i$ under the area level model may be written as

$$MSE(\hat{\mu}_i) = g_{1i}(\sigma_\nu^2) + g_{2i}(\sigma_\nu^2) + g_{3i}(\sigma_\nu^2) + o(m^{-1}),$$

where

$$\begin{aligned} g_{1i}(\sigma_\nu^2) &= \gamma_{i1}\psi_i, \\ g_{2i}(\sigma_\nu^2) &= (1 - \gamma_{i1})^2 x_i' \left[\sum_{i=1}^m x_i x_i' / (\psi_i + \sigma_\nu^2) \right]^{-1} x_i, \\ g_{3i}(\sigma_\nu^2) &= \psi_i^2 (\psi_i + \sigma_\nu^2)^{-3} V_a(\hat{\sigma}_\nu^2). \end{aligned}$$

The MSE estimator of $\hat{\mu}_i$ may be written as

$$mse(\hat{\mu}_i) = g_{1i}(\hat{\sigma}_\nu^2) - (\partial g_{1i} / \partial \sigma_\nu^2) b_{\hat{\sigma}_\nu^2}(\sigma_\nu^2) |_{\sigma_\nu^2 = \hat{\sigma}_\nu^2} + g_{2i}(\hat{\sigma}_\nu^2) + 2g_{3i}(\hat{\sigma}_\nu^2).$$

Similarly, the MSE of EBLUP estimator for $\hat{\mu}_i$ in unit level model is given by

$$MSE(\hat{\mu}_i) = g_{1i}(\delta) + g_{2i}(\delta) + g_{3i}(\delta) + o(m^{-1}),$$

where

$$\begin{aligned} g_{1i}(\delta) &= \gamma_{i2}\sigma_e^2/n_i, \\ g_{2i}(\delta) &= (\bar{X}_i - \gamma_{i2}\bar{x}_i)' \left(\sum_{i=1}^m X_i' V_i^{-1} X_i \right)^{-1} (\bar{X}_i - \gamma_{i2}\bar{x}_i), \\ g_{3i}(\delta) &= n_i^{-2} (\sigma_v^2 + \sigma_e^2/n_i)^{-3} [\sigma_e^4 V_a(\hat{\sigma}_v^2) + \sigma_v^4 V_a(\hat{\sigma}_e^2) - 2\sigma_v^2 \sigma_e^2 Cov_a(\hat{\sigma}_v^2, \hat{\sigma}_e^2)], \end{aligned}$$

where $Cov_a(\hat{\sigma}_v^2, \hat{\sigma}_e^2)$ is the asymptotic covariance of $\hat{\sigma}_v^2$ and $\hat{\sigma}_e^2$. Also, second order unbiased MSE estimator, $mse(\hat{\mu}_i)$, is given by

$$mse(\hat{\mu}_i) = g_{1i}(\hat{\delta}) - (\partial g_{1i} / \partial \delta') b_{\hat{\delta}}(\delta) |_{\delta = \hat{\delta}} + g_{2i}(\hat{\delta}) + 2g_{3i}(\hat{\delta}).$$

1.5 Some Background

In Chapter 2, we study a new generalized regression (GREG) estimator of a small area mean proposed by Lehtonen and Veijanen (1999), under a two-level model (Moura and Holt, 1999). We derive the mean squared error (MSE) of the new GREG estimator under the model and compare it to the MSE of the BLUP estimator.

In Chapter 3, we derive a second order approximation to the mean squared error (MSE) of the survey weighted pseudo-EBLUP estimators introduced by You and Rao (2002). Using this approximation, we obtain estimators of MSE that are nearly unbiased. Our approximation accounts for the cross-product terms in the MSE, which are ignored in You and Rao (2002).

In Chapter 4, we introduce a sub-area level small area model. This model is an extension of the Fay and Herriot (1979) area level model to sub-area level. We extend the work of Datta et al. (2005) from area level model to sub-area level model.

Most of the research on small area estimation has focused on unconditional MSE estimation under an assumed small area model. Datta, Kubokawa, and Rao (2004) studied conditional MSE for the Fay-Herriot area level model. In Chapter 5, we follow their approach to study conditional MSE estimation for the nested error linear regression model. Datta, Kubokawa, and Rao (2004) also studied MSE and conditional MSE of fixed weighted predictors for the Fay-Herriot area level model. In Chapter 6, we use this idea to study the MSE and the conditional MSE of predictors with fixed weights for the nested error regression model.

Ghosh, Sinha, and Kim (2006) studied the nested error linear regression model

when the covariates are subject to structural measurement error. Ghosh, Sinha, and Kim (2006) used area level covariates x_i . Ghosh, Sinha, and Kim (2006) only studied the asymptotic behavior of Bayes risk as measure of variability of EB estimators. In Chapter 7, we study structural measurement error nested error regression model with unit level covariates x_{ij} . We study the asymptotic behavior of Bayes risk of EB estimators. Moreover, we obtain a second order approximation to the MSE of EB estimators and obtain a MSE estimator to EB estimators which is nearly unbiased.

Ghosh and Sinha (2006) also studied the nested error linear regression model when the covariates are subject to functional measurement error, using area level covariates x_i . Ghosh and Sinha (2006) only studied the asymptotic behavior of Bayes risk as measure of variability of pseudo-EB (PEB) estimators. In Chapter 8, we study the functional measurement error nested error regression model with unit level covariates x_{ij} . We study the asymptotic behavior of Bayes risk of PEB estimators. Moreover, we obtain a second order approximation to the MSE of PEB estimators and obtain a MSE estimator which is nearly unbiased.

1.6 Thesis Outline

In Chapter 2, we study the estimation of small area means using a two-level model-assisted GREG estimator. We further derive the mean squared error (MSE) of the new GREG estimator. A simulation study shows that the new GREG performs better than the traditional GREG in the sense of MSE and average absolute relative error. But it is significantly less efficient than the BLUP estimator. In Chapter 3, we consider a pseudo empirical best linear unbiased prediction (pseudo-EBLUP)

estimator proposed by You and Rao (2002) to estimate small area means using survey weights under a nested error linear regression model. We derive a second order approximation to the mean squared error (MSE) of the pseudo-EBLUP estimator of a small area mean. An estimator of MSE that is approximately unbiased is also derived. Empirical results on the performance of the proposed MSE estimator are also presented. In Chapter 4, we introduce a sub-area level small area model. We consider area and sub-area means, and obtain the corresponding EBLUPs, using an iterative method of estimating the model variances based on weighted residual sum of squares. We also obtain a second order approximation to the MSE of the EBLUPs and derive MSE estimators unbiased to second order. Accuracy of the approximation to the MSE is justified, by showing that the neglected terms are of order $o(m^{-1})$, for large m , where m is the number of small areas. Results of a simulation study on the accuracy of the second order approximation to the MSE and the bias of the estimators of the MSE are also reported. In Chapter 5, we study the conditional MSE given the area-specific response data under a nested error linear regression model. We derive a second order approximation to the conditional MSE of the EBLUP estimator of a small area mean. We then derive an estimator of the conditional MSE that is conditionally unbiased to second order. We justify the accuracy of approximation to the conditional MSE by showing that the neglected terms are of order $o(m^{-1})$, for large m . Empirical results on the performance of our approach are also provided.

In Chapter 6, we study the unconditional and conditional MSE of small area predictors with fixed weights under a nested error linear regression model. We derive a second order approximation to the unconditional and conditional MSE of the fixed weighted predictor and then obtain estimators of MSE that are nearly unbiased. In

Chapter 7, we study the nested error regression model, when the covariates in the model are random and subject to measurement errors (i.e, structural measurement errors). We develop empirical Bayes (EB) estimators of area means and derive a second order approximation to the mean squared error of EB estimators. Estimators of MSE that are approximately unbiased are then derived. In addition, we show that the EB estimators are asymptotically optimal in the sense of Bayes risk. Results of a simulation study on the performance of the proposed method are also provided. In Chapter 8, we study the nested error regression model when the covariates in the model are fixed but subject to measurement errors (i.e, functional measurement errors). In particular, pseudo-empirical Bayes (PEB) estimators of area means are developed. We derive a second order approximation to the mean squared error of the PEB estimators, and then obtain an estimator of MSE that is nearly unbiased. In addition, we show that the PEB estimators are asymptotically optimal in the sense of Bayes risk. Empirical results on the performance of the proposed method are also presented. Finally, in Chapter 9, we present some possible future research directions and application areas.

Chapter 2

Small Area Estimation Under a Two-level Model

In this Chapter, we study a new generalized regression (GREG) estimator of a small area mean under a two-level model. We derive the mean squared error (MSE) of the new GREG estimator under the model and compare it to the MSE of the BLUP estimator. We also provide empirical results on the relative efficiency of the estimators when the model parameters are estimated.

2.1 Introduction

In the context of linear mixed models, a two-level model is given by

$$y_{ij} = x'_{ij}\beta_i + e_{ij}; \quad \beta_i = Z_i\beta + \nu_i; j = 1, \dots, N_i; i = 1, \dots, m \quad (2.1)$$

where N_i is the number of units in the i -th area, y_{ij} is the response and x_{ij} is a $p \times 1$ vector of unit level covariates attached to the j -th unit in the i -th area, Z_i is a $p \times q$ matrix of area level covariates, β is a $q \times 1$ vector of regression parameters,

ν_i 's are independent random vectors with mean zero and covariance Σ_ν and e_{ij} 's are independent random variables with mean zero and variance σ_e^2 . We can express the mean $\bar{y}_i$ of i -th area as

$$\bar{y}_i \approx \mu_i = \bar{X}_i'(Z_i\beta + \nu_i),$$

assuming N_i is large. The sample values $\{(y_{ij}, x_{ij}); j = 1, \dots, n_i; i = 1, \dots, m\}$ are assumed to obey the model (2.1), that is, there is no sample selection bias. The model for the sample is then given by

$$y_{ij} = x'_{ij}(Z_i\beta + \nu_i) + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m. \quad (2.2)$$

In matrix notation, (2.2) may be written as

$$y_i = X_i(Z_i\beta + \nu_i) + e_i, i = 1, \dots, m$$

with variance $Var(y_i) = V_i = X_i\Sigma_\nu X_i' + \sigma_e^2 I_{n_i}$, where y_i is a $n_i \times 1$ vector and X_i is a $n_i \times p$ matrix. The two-level model (2.2) was first introduced by Moura and Holt (1999) in the context of small area estimation. This model effectively integrates the use of unit level and area level covariates into a single model, by modeling the random slopes β_i in terms of area level covariates Z_i . In Section 2.2, we introduce the two-level model assisted GREG estimator of μ_i and derive the mean squared error of the new GREG estimator under the model.

2.2 BLUP and GREG Estimators

The two-level model (2.2) is a special case of a general linear mixed model with block diagonal covariance structure introduced in Chapter 1. Therefore, we may calculate the BLUP estimator of μ_i as

$$\tilde{\mu}_i^B = \bar{X}_i'(Z_i\beta + \tilde{\nu}_i), \quad (2.3)$$

where

$$\tilde{\nu}_i = \Sigma_\nu X_i' V_i^{-1} (y_i - X_i Z_i \beta),$$

and the superscript B on $\tilde{\mu}_i$ stands for BLUP estimator. Similarly, the BLUP estimator of a non-sampled observation j in i -th area can be written as

$$\tilde{\mu}_{ij}^B = x_{ij}' (Z_i \beta + \tilde{\nu}_i). \quad (2.4)$$

On the other hand, in the context of survey sampling, a model-assisted GREG estimator of the mean $\bar{y}_i$ is given by

$$\tilde{\mu}_i^G = \frac{1}{N_i} \left[\sum_{j=1}^{N_i} \tilde{y}_{ij} + \sum_{j=1}^{n_i} w_{ij} (y_{ij} - \tilde{y}_{ij}) \right]; i = 1, \dots, m \quad (2.5)$$

where $\tilde{y}_{ij} = x_{ij}' Z_i \beta$ is the predictor of y_{ij} under the fixed effects model $y_{ij} = x_{ij}' Z_i \beta + e_{ij}$ and w_{ij} is the survey weight which equals $\frac{N_i}{n_i}$ in the case of simple random sample (SRS) within areas. We focus on SRS within areas in this Chapter. Using (2.5), we can write the GREG estimator of the mean $\bar{y}_i$ as

$$\tilde{\mu}_i^G = \bar{y}_i + (\bar{X}_i - \bar{x}_i)' Z_i \beta, \quad (2.6)$$

and superscript G on $\tilde{\mu}_i$ denotes GREG (Särndal et al., 1992), where $\bar{y}_i$ is sample mean of y_{ij} in the i -th area, $\bar{X}_i$ and $\bar{x}_i$ are population and sample mean of x_{ij} in i -th area, respectively.

Using the predictor (2.4) based on the two-level model (2.2) in (2.5), we get a new GREG estimator as

$$\tilde{\mu}_i^{NG} = [\bar{X}_i' (Z_i \beta + \tilde{\nu}_i) + (\bar{y}_i - \bar{x}_i' (Z_i \beta + \tilde{\nu}_i))] = \bar{y}_i + (\bar{X}_i - \bar{x}_i)' (Z_i \beta + \tilde{\nu}_i), \quad (2.7)$$

where the superscript NG on $\tilde{\mu}_i$ denotes that it is a new GREG (Lehtonen and Veijanen, 1999). In practice, we replace the parameters β, Σ_ν and σ_e^2 in (2.3), (2.6)

and (2.7) by suitable estimators. For simplicity, we assume that the parameters are known and derive the mean squared error of GREG (2.6) and the new GREG (2.7) in Section 2.3 under the two-level model (2.2).

2.3 Mean Squared Error

The mean squared error (MSE) of the GREG estimator $\tilde{\mu}_i^G$ under the two-level model can be written as

$$\begin{aligned} MSE(\tilde{\mu}_i^G) &= E(\tilde{\mu}_i^G - \mu_i)^2 = E[\bar{y}_i + (\bar{X}_i - \bar{x}_i)'Z_i\beta - \bar{X}_i'(Z_i\beta + \nu_i)]^2 \\ &= E[(\bar{x}_i - \bar{X}_i)'\nu_i + \bar{e}_i]^2 \\ &= (\bar{X}_i - \bar{x}_i)'\Sigma_\nu(\bar{X}_i - \bar{x}_i) + \frac{\sigma_e^2}{n_i}. \end{aligned}$$

Theorem 2.1. *Assume the model parameters are known. Then the MSE of the GREG estimator (2.6) is given by*

$$MSE(\tilde{\mu}_i^G) = (\bar{X}_i - \bar{x}_i)'\Sigma_\nu(\bar{X}_i - \bar{x}_i) + \frac{\sigma_e^2}{n_i}. \quad (2.8)$$

In addition, we may write the MSE of the BLUP estimator $\tilde{\mu}_i^B$ as follows :

$$MSE(\tilde{\mu}_i^B) = E(\tilde{\mu}_i^B - \mu_i)^2 = E[\bar{X}_i'(\tilde{\nu}_i - \nu_i)]^2 = \bar{X}_i'(\Sigma_\nu - \Sigma_\nu X_i' V_i^{-1} X_i \Sigma_\nu) \bar{X}_i.$$

Theorem 2.2. *Assume the model parameters are known. Then the MSE of the BLUP estimator (2.3) is given by*

$$MSE(\tilde{\mu}_i^B) = \bar{X}_i'(\Sigma_\nu - \Sigma_\nu X_i' V_i^{-1} X_i \Sigma_\nu) \bar{X}_i. \quad (2.9)$$

We now derive the MSE of the new GREG estimator. We have

$$\begin{aligned}
MSE(\tilde{\mu}_i^{NG}) &= E(\tilde{\mu}_i^{NG} - \mu_i)^2 \\
&= E[\bar{y}_i + (\bar{X}_i - \bar{x}_i)'(Z_i\beta + \tilde{\nu}_i) - \bar{X}_i'(Z_i\beta + \nu_i)]^2 = E[\bar{y}_i + \bar{X}_i'(\tilde{\nu}_i - \nu_i) - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i)]^2 \\
&= E[\bar{X}_i'(\tilde{\nu}_i - \nu_i)]^2 + E[\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i)]^2 + 2E[(\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i))\bar{X}_i'(\tilde{\nu}_i - \nu_i)], \quad (2.10)
\end{aligned}$$

where the first term on the right hand side of (2.10) is the MSE of the BLUP estimator under the two-level model, given by (2.9). Moreover, we may write

$$E[\bar{y}_i - \bar{x}_i'(Z_i\beta + \tilde{\nu}_i)]^2 = E[\bar{y}_i - \bar{x}_i'Z_i\beta]^2 + E(\bar{x}_i'\tilde{\nu}_i)^2 - 2E[(\bar{y}_i - \bar{x}_i'Z_i\beta)(\bar{x}_i'\tilde{\nu}_i)], \quad (2.11)$$

where

$$E[\bar{y}_i - \bar{x}_i'Z_i\beta]^2 = Var(\bar{y}_i) = \bar{x}_i'\Sigma_\nu\bar{x}_i + \frac{\sigma_e^2}{n_i},$$

and

$$\begin{aligned}
E(\bar{x}_i'\tilde{\nu}_i)^2 &= Var(\bar{x}_i'\tilde{\nu}_i) + [E(\bar{x}_i'\tilde{\nu}_i)]^2 \\
&= Var[\bar{x}_i'\Sigma_\nu X_i'V_i^{-1}(y_i - X_iZ_i\beta)] + [E(\bar{x}_i'\Sigma_\nu X_i'V_i^{-1}(y_i - X_iZ_i\beta))]^2 \\
&= \bar{x}_i'\Sigma_\nu X_i'V_i^{-1}X_i\Sigma_\nu\bar{x}_i.
\end{aligned}$$

In addition,

$$\begin{aligned}
E[(\bar{y}_i - \bar{x}_i'Z_i\beta)(\bar{x}_i'\tilde{\nu}_i)] &= E[\bar{y}_i\bar{x}_i'\Sigma_\nu X_i'V_i^{-1}(y_i - X_iZ_i\beta)] \\
&\quad - E[\bar{x}_i'Z_i\beta\bar{x}_i'\Sigma_\nu X_i'V_i^{-1}(y_i - X_iZ_i\beta)],
\end{aligned}$$

where the second term is zero. Therefore, we may write

$$E[(\bar{y}_i - \bar{x}_i'Z_i\beta)(\bar{x}_i'\tilde{\nu}_i)] = E[\bar{y}_i\bar{x}_i'\Sigma_\nu X_i'V_i^{-1}y_i] - \bar{x}_i'Z_i\beta\bar{x}_i'\Sigma_\nu X_i'V_i^{-1}X_iZ_i\beta,$$

where the first term can be written as

$$E[\frac{1'}{n_i}y_i\bar{x}_i'\Sigma_\nu X_i'V_i^{-1}y_i] = \frac{1'}{n_i}(X_i\Sigma_\nu\bar{x}_i + X_iZ_i\beta\bar{x}_i'\Sigma_\nu X_i'V_i^{-1}X_iZ_i\beta),$$

using the following Lemma:

Lemma 2.1. *If y is a $n \times 1$ vector with mean μ and variance-covariance matrix Σ and b is a $n \times 1$ vector, then $E(yb'y) = \Sigma b + \mu b'\mu$.*

Hence,

$$\begin{aligned} E[(\bar{y}_i - \bar{x}_i' Z_i \beta)(\bar{x}_i' \tilde{\nu}_i)] &= \bar{x}_i' \Sigma_\nu \bar{x}_i + \bar{x}_i' Z_i \beta \bar{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta \\ &\quad - \bar{x}_i' Z_i \beta \bar{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta \\ &= \bar{x}_i' \Sigma_\nu \bar{x}_i. \end{aligned}$$

Then we may write (2.11) as follows

$$\begin{aligned} E[\bar{y}_i - \bar{x}_i'(Z_i \beta + \tilde{\nu}_i)]^2 &= \bar{x}_i' \Sigma_\nu \bar{x}_i + \frac{\sigma_e^2}{n_i} + \bar{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i \Sigma_\nu \bar{x}_i - 2\bar{x}_i' \Sigma_\nu \bar{x}_i \\ &= \bar{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i \Sigma_\nu \bar{x}_i - \bar{x}_i' \Sigma_\nu \bar{x}_i + \frac{\sigma_e^2}{n_i}. \end{aligned} \quad (2.12)$$

Finally, we need to find the cross-product term of (2.10). We have

$$E[(\bar{y}_i - \bar{x}_i'(Z_i \beta + \tilde{\nu}_i))\bar{X}_i'(\tilde{\nu}_i - \nu_i)] = E[\bar{y}_i \bar{X}_i'(\tilde{\nu}_i - \nu_i)] - E[\bar{x}_i'(Z_i \beta + \tilde{\nu}_i)\bar{X}_i'(\tilde{\nu}_i - \nu_i)], \quad (2.13)$$

where the first term on the right side of (2.13) may be written as

$$\begin{aligned} E[\bar{y}_i \bar{X}_i'(\tilde{\nu}_i - \nu_i)] &= E[\bar{y}_i \bar{X}_i'(\Sigma_\nu X_i' V_i^{-1}(y_i - X_i Z_i \beta) - \nu_i)] \\ &= E[\bar{y}_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} y_i] - E[\bar{y}_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta] \\ &\quad - E[\bar{y}_i \bar{X}_i' \nu_i]. \end{aligned} \quad (2.14)$$

The first term of (2.14) is

$$\begin{aligned} E[\bar{y}_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} y_i] &= E[\frac{1'}{n_i} y_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} y_i] \\ &= \frac{1'}{n_i} X_i (\Sigma_\nu \bar{X}_i + Z_i \beta \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta), \end{aligned} \quad (2.15)$$

the second term of (2.14) can be written as

$$E[\bar{y}_i \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta] = \bar{x}_i' Z_i \beta \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta, \quad (2.16)$$

and the third term can be expressed as

$$E[\bar{y}_i \bar{X}_i' \nu_i] = E[(\bar{x}_i' Z_i \beta + \bar{x}_i' \nu_i + e_i) \bar{X}_i' \nu_i] = E[\bar{x}_i' \nu_i \bar{X}_i' \nu_i] = \bar{x}_i' \Sigma_\nu \bar{X}_i. \quad (2.17)$$

Therefore, substituting (2.15), (2.16) and (2.17) in (2.14), we have

$$\begin{aligned} E[\bar{y}_i \bar{X}_i' (\tilde{\nu}_i - \nu_i)] &= \bar{x}_i' \Sigma_\nu \bar{X}_i + \bar{x}_i' Z_i \beta \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta - \bar{x}_i' Z_i \beta \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} X_i Z_i \beta - \bar{x}_i' \Sigma_\nu \bar{X}_i \\ &= 0. \end{aligned}$$

We now turn to the second term of (2.13). We have

$$\begin{aligned} E[\bar{x}_i' (Z_i \beta + \tilde{\nu}_i) \bar{X}_i' (\tilde{\nu}_i - \nu_i)] &= E[\bar{x}_i' (Z_i \beta + \tilde{\nu}_i) \bar{X}_i' \tilde{\nu}_i] - E[\bar{x}_i' (Z_i \beta + \tilde{\nu}_i) \bar{X}_i' \nu_i] \\ &= E[\bar{x}_i' Z_i \beta \bar{X}_i' \tilde{\nu}_i] + E[\bar{x}_i' \tilde{\nu}_i \bar{X}_i' \tilde{\nu}_i] - E[\bar{x}_i' Z_i \beta \bar{X}_i' \nu_i] - E[\bar{x}_i' \tilde{\nu}_i \bar{X}_i' \nu_i]. \end{aligned} \quad (2.18)$$

Then, we obtain the following expression for the four terms on the right side of (2.18):

$$E[\bar{x}_i' Z_i \beta \bar{X}_i' \tilde{\nu}_i] = E[\bar{x}_i' Z_i \beta \bar{X}_i' \Sigma_\nu X_i' V_i^{-1} (y_i - X_i Z_i \beta)] = 0, \quad (2.19)$$

$$\begin{aligned} E[\bar{x}_i' \tilde{\nu}_i \bar{X}_i' \tilde{\nu}_i] &= \bar{x}_i' \Sigma_\nu \bar{X}_i + \bar{x}_i' E(\tilde{\nu}_i) \bar{X}_i' E(\tilde{\nu}_i) \\ &= \bar{x}_i' \text{Var}[\Sigma_\nu X_i' V_i^{-1} (y_i - X_i Z_i \beta)] \bar{X}_i \\ &\quad + \bar{x}_i' E(\Sigma_\nu X_i' V_i^{-1} (y_i - X_i Z_i \beta)) \bar{X}_i' E(\Sigma_\nu X_i' V_i^{-1} (y_i - X_i Z_i \beta)) \\ &= \bar{x}_i' \Sigma_\nu X_i' V_i^{-1} X_i \Sigma_\nu \bar{X}_i, \end{aligned} \quad (2.20)$$

$$E[\bar{x}_i' Z_i \beta \bar{X}_i' \nu_i] = 0, \quad (2.21)$$

and

$$\begin{aligned}
E[\bar{x}'_i \tilde{\nu}_i \bar{X}'_i \nu_i] &= E[\bar{x}'_i \Sigma_\nu X'_i V_i^{-1} (y_i - X_i Z_i \beta) \bar{X}'_i \nu_i] \\
&= E[\bar{x}'_i \Sigma_\nu X'_i V_i^{-1} y_i \bar{X}'_i \nu_i] - E[\bar{x}'_i \Sigma_\nu X'_i V_i^{-1} X_i Z_i \beta \bar{X}'_i \nu_i] \\
&= E[\bar{x}'_i \Sigma_\nu X'_i V_i^{-1} (X_i Z_i \beta + X_i \nu_i + e_i) \bar{X}'_i \nu_i] \\
&= E[\bar{x}'_i \Sigma_\nu X'_i V_i^{-1} X_i \nu_i \bar{X}'_i \nu_i] \\
&= \bar{x}'_i \Sigma_\nu X'_i V_i^{-1} X_i \Sigma_\nu \bar{X}_i.
\end{aligned} \tag{2.22}$$

Therefore, substituting (2.19)- (2.22) in (2.18), we get

$$E[\bar{x}'_i (Z_i \beta + \tilde{\nu}_i) \bar{X}'_i (\tilde{\nu}_i - \nu_i)] = \bar{x}'_i \Sigma_\nu X'_i V_i^{-1} X_i \Sigma_\nu \bar{X}_i - \bar{x}'_i \Sigma_\nu X'_i V_i^{-1} X_i \Sigma_\nu \bar{X}_i = 0.$$

Hence,

$$E[(\bar{y}_i - \bar{x}'_i (Z_i \beta + \tilde{\nu}_i)) \bar{X}'_i (\tilde{\nu}_i - \nu_i)] = 0. \tag{2.23}$$

It now follows from (2.10), (2.12) and (2.23) that

$$\begin{aligned}
MSE(\tilde{\mu}_i^{NG}) &= MSE(\tilde{\mu}_i^B) + E[\bar{y}_i - \bar{x}'_i (Z_i \beta + \tilde{\nu}_i)]^2 \\
&= MSE(\tilde{\mu}_i^B) + \bar{x}'_i \Sigma_\nu X'_i V_i^{-1} X_i \Sigma_\nu \bar{x}_i - \bar{x}'_i \Sigma_\nu \bar{x}_i + \frac{\sigma_e^2}{n_i}.
\end{aligned}$$

Theorem 2.3. *Assume the model parameters are known. Then the MSE of the new GREG estimator (2.7) is given by*

$$MSE(\tilde{\mu}_i^{NG}) = MSE(\tilde{\mu}_i^B) + \bar{x}'_i \Sigma_\nu X'_i V_i^{-1} X_i \Sigma_\nu \bar{x}_i - \bar{x}'_i \Sigma_\nu \bar{x}_i + \frac{\sigma_e^2}{n_i}. \tag{2.24}$$

By definition, we have $MSE(\tilde{\mu}_i^B) \leq MSE(\tilde{\mu}_i^{NG})$ and (2.24) gives an explicit expression for the increase in MSE of $\tilde{\mu}_i^{NG}$ relative to the MSE of the BLUP estimator $\tilde{\mu}_i^B$.

2.4 Empirical Results

2.4.1 Empirical Comparison of MSE Values

In order to study the efficiency of the new GREG estimator, we used the data from Moura and Holt (1999) based on 38,740 households in the enumeration districts (small areas) in one county in Brazil. The income of household's head was treated as the response variable y . Two unit level independent variables were identified as the educational attainment of household's head (ordinal scale of 0-5) and the number of rooms in the household (1-11+). The following two-level model was assumed for this data:

$$y_{ij} = \beta_{i0} + \beta_{i1}x_{1ij} + \beta_{i2}x_{2ij} + \varepsilon_{ij}; j = 1, \dots, n_i; i = 1, \dots, m \quad (2.25)$$

with

$$\beta_{i0} = \beta_0 + \nu_{i0}; \beta_{i1} = \beta_1 + \nu_{i1}; \beta_{i2} = \beta_2 + \nu_{i2}, \quad (2.26)$$

where $\nu = (\nu_{i0}, \nu_{i1}, \nu_{i2}) \stackrel{i.i.d.}{\sim} N_3(0, \Sigma_\nu)$ and x_1 and x_2 respectively represent the number of rooms and the educational attainment of household's head (centered about their respective population means). Note that the model (2.26) for the random β_i -coefficients does not contain area level covariates Z .

Moura and Holt (1999) also studied another model with an area level covariate Z for modeling β_i 's in (2.25). For this data, average number of cars per household in each area was used as a covariate z for modeling the random coefficients β_{i1} and β_{i2} corresponding to the variables x_1 and x_2 , but not for the random intercept term, β_{i0} . The model (2.25) with the area level covariate z is given by

$$y_{ij} = \beta_{i0} + \beta_{i1}x_{1ij} + \beta_{i2}x_{2ij} + \varepsilon_{ij}; j = 1, \dots, n_i; i = 1, \dots, m \quad (2.27)$$

with

$$\beta_{i0} = \beta_0 + \nu_{i0}; \beta_{i1} = \beta_1 + \alpha_1 z_i + \nu_{i1}; \beta_{i2} = \beta_2 + \alpha_2 z_i + \nu_{i2}. \quad (2.28)$$

Moura and Holt (1999) fitted models (2.25) and (2.27) to the full data set mentioned above. We summarized their results in Table 2.1.

The area means of x_1 and x_2 were calculated from the whole data and treated as the known population means $\bar{X}_{1i}$ and $\bar{X}_{2i}$. A random subsample of 10% of the records was selected from each small area. The overall sample size $n = 3,876$ and the number of small areas $m = 140$. Using the sample values on x_1, x_2 and z and the population means $\bar{X}_{1i}$ and $\bar{X}_{2i}$, we computed $MSE(\tilde{\mu}_i^G)$, $MSE(\tilde{\mu}_i^B)$ and $MSE(\tilde{\mu}_i^{NG})$ using (2.8), (2.9) and (2.24), respectively and treating the estimates of regression parameters, Σ_ν and σ_e^2 in Table 2.1 for the full data as true values. We then calculated the average MSE values over the areas: $\overline{MSE}_G = \frac{1}{m} \sum_{i=1}^m MSE(\tilde{\mu}_i^G)$, $\overline{MSE}_B = \frac{1}{m} \sum_{i=1}^m MSE(\tilde{\mu}_i^B)$ and $\overline{MSE}_{NG} = \frac{1}{m} \sum_{i=1}^m MSE(\tilde{\mu}_i^{NG})$. We define the relative efficiency of $\tilde{\mu}^B$ over $\tilde{\mu}^G$ as EFF_B and the relative efficiency of $\tilde{\mu}^{NG}$ over $\tilde{\mu}^G$ as EFF_{NG} , where

$$EFF_B = \frac{\overline{MSE}_G}{\overline{MSE}_B}; \quad EFF_{NG} = \frac{\overline{MSE}_G}{\overline{MSE}_{NG}}.$$

We summarize the results in Tables 2.2 and 2.3.

Tables 2.2 and 2.3 reveal that the new GREG estimator is slightly more efficient than the usual GREG estimator in terms of average MSE: $EFF_{NG} \leq 112\%$. However, the new GREG estimator is substantially less efficient than the BLUP estimator, under the assumed small area model. For example, for the model with z and diagonal covariance matrix (Table 2.2), $EFF_B = 292\%$ compared to $EFF_{NG} = 106\%$, and $\overline{MSE}_B = 0.62$ compared to $\overline{MSE}_{NG} = 1.72$.

2.4.2 Simulation Study

Simulation Study Under Model-based Framework

In order to investigate the efficiency of the new GREG estimator with estimated model parameters, a small simulation study based on the two-level model (2.25) was undertaken. We only considered a diagonal covariance structure Σ_ν with diagonal elements σ_0^2, σ_1^2 and σ_2^2 . We again used the data from Moura and Holt (1999). The parameter estimates in $\beta_0, \beta_1, \beta_2, \sigma_0^2, \sigma_1^2, \sigma_2^2$ and σ_e^2 reported in Table 2.1 are treated as true values.

In our simulation study, we took (x_{1ij}, x_{2ij}, z_i) from Moura and Holt (1999) and then generated y_{ij} based on the model given by (2.25). By using the generated samples $(y_{ij}^{(b)}, x_{1ij}, x_{2ij}, z_i)$, $b=1, \dots, B=1000$, we calculated $\hat{\beta}^{(b)}$ by generalized least squared estimation for the new GREG method as well as for the BLUP method. In the GREG method we used ordinary least squared estimation to estimate β as $\hat{\beta}_{ols}^{(b)}$. In addition, $\hat{\Sigma}_\nu^{(b)}$ and $\hat{\sigma}_e^{2(b)}$ were computed based on the restricted maximum likelihood (REML) method. For each generated sample, we calculated

$$\mu_i^{(b)} = \bar{X}_i'(Z_i\beta + \nu_i^{(b)}); i = 1, \dots, m; b = 1, \dots, B.$$

We computed the new GREG estimator of $\mu_i^{(b)}$ as $\hat{\mu}_i^{NG(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i)'(Z_i\hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})$, the GREG estimator of $\mu_i^{(b)}$ as $\hat{\mu}_i^{G(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i)'Z_i\hat{\beta}_{ols}^{(b)}$ and the empirical BLUP (EBLUP) estimator of $\mu_i^{(b)}$ as $\hat{\mu}_i^{B(b)} = \bar{X}_i'(Z_i\hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})$, where $\hat{\nu}_i^{(b)} = \hat{\Sigma}_\nu^{(b)} X_i' \hat{V}_i^{-1(b)} (y_i^{(b)} - X_i Z_i \hat{\beta}^{(b)})$.

We then computed the average mean squared error ($\overline{MSE}_1$) and average absolute

relative error ($\overline{ARE}_1$)

$$\overline{MSE}_1 = \frac{1}{m} \sum_i^m MSE_{1i} \quad \text{where} \quad MSE_{1i} = B^{-1} \sum_{b=1}^B (\hat{\mu}_i^{(b)} - \mu_i^{(b)})^2,$$

$$\overline{ARE}_1 = \frac{1}{m} \sum_i^m ARE_{1i} \quad \text{where} \quad ARE_{1i} = B^{-1} \sum_{b=1}^B |\hat{\mu}_i^{(b)} - \mu_i^{(b)}| / \mu_i^{(b)},$$

where $\hat{\mu}_i^{(b)}$ denotes $\hat{\mu}_i^{NG(b)}$, $\hat{\mu}_i^{G(b)}$ or $\hat{\mu}_i^{B(b)}$. We report the results in Table 2.4. Both models with area level covariate z and without z have slightly smaller values of $\overline{MSE}_1$ and $\overline{ARE}_1$ for the new GREG estimator relative to the GREG estimator. However, $\overline{MSE}_1$ and $\overline{ARE}_1$ are significantly smaller for the BLUP estimator due to borrowing strength from related areas. Moreover, comparing Tables 2.2 and 2.4, we can see that the values of $\overline{MSE}_1$ in Table 2.4 are slightly larger than corresponding values in Table 2.2 due to estimating model parameters.

Simulation Study Under Finite Population Framework

To study the performance of estimators under a finite population framework, we created a synthetic finite population. By duplicating the sample of (x_{1ij}, x_{2ij}, z_i) -values used by Moura and Holt (1999) five times, we treated the new (x_1, x_2, z) -data of size 19,380 as our real population. We then generated the y_{ij} based on model (2.25) by using the 19,380 values (x_{1ij}, x_{2ij}, z_i) and parameter values reported in Table 2.1 with diagonal covariance structure Σ_ν . This results in a synthetic finite population $(y_{ij}, x_{1ij}, x_{2ij}, z_i)$ of size 19,380.

We generated 500 independent samples ($B=500$), each of size $n=700$ and $n=1400$, by taking simple random samples of size $n_i = 5$ and $n_i = 10$ in each area $i = 1, \dots, 140$. As before for each generated sample, we calculated $\hat{\beta}^{(b)}$ for the new GREG and the

BLUP methods and $\hat{\beta}_{ols}^{(b)}$ for the GREG method. In addition, $\hat{\Sigma}_\nu^{(b)}$ and $\hat{\sigma}_e^{2(b)}$ were calculated based on the REML method. We also computed the population mean of y_{ij} for each area i as

$$\bar{Y}_i = \sum_{j=1}^{N_i} y_{ij} / N_i; i = 1, \dots, 140,$$

where N_i is the population size in i -th area. Further, for each sample $b = 1, \dots, B$, we calculated the new GREG estimate of i -th area mean as $\hat{\mu}_i^{NG(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i^{(b)})'(Z_i \hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})$, the GREG estimate as $\hat{\mu}_i^{G(b)} = \bar{y}_i^{(b)} + (\bar{X}_i - \bar{x}_i^{(b)})' Z_i \hat{\beta}_{ols}^{(b)}$ and the EBLUP estimate as $\hat{\mu}_i^{B(b)} = \bar{X}_i' (Z_i \hat{\beta}^{(b)} + \hat{\nu}_i^{(b)})$, where $\hat{\nu}_i^{(b)} = \hat{\Sigma}_\nu^{(b)} X_i' \hat{V}_i^{-1(b)} (y_i^{(b)} - X_i Z_i \hat{\beta}^{(b)})$.

We then computed the average mean squared error ($\overline{MSE}_2$) and average absolute relative error ($\overline{ARE}_2$)

$$\begin{aligned} \overline{MSE}_2 &= \frac{1}{m} \sum_i^m MSE_{2i} \quad \text{where} \quad MSE_{2i} = B^{-1} \sum_{b=1}^B (\hat{\mu}_i^{(b)} - \bar{Y}_i)^2, \\ \overline{ARE}_2 &= \frac{1}{m} \sum_i^m ARE_{2i} \quad \text{where} \quad ARE_{2i} = B^{-1} \sum_{b=1}^B |\hat{\mu}_i^{(b)} - \bar{Y}_i| / \bar{Y}_i, \end{aligned}$$

where $\hat{\mu}_i^{(b)}$ denotes $\hat{\mu}_i^{NG(b)}$, $\hat{\mu}_i^{G(b)}$ or $\hat{\mu}_i^{B(b)}$. We report the results in Tables 2.5 and 2.6 for $n_i = 5$ and $n_i = 10$, respectively. Both models with area level covariate z and without z are considered with diagonal covariance structure Σ_ν .

Tables 2.5 and 2.6 show that the new GREG estimator exhibits better performance relative to the GREG estimator in the sense of smaller $\overline{MSE}_2$ and $\overline{ARE}_2$. However, the EBLUP estimator gives significantly smaller $\overline{MSE}_2$ and $\overline{ARE}_2$ than the GREG and the new GREG due to borrowing strength from related small areas. For example, for the model without z and $n_i = 10$, $\overline{MSE}_2 = 5.19, 5.51$ and 2.72 for the new GREG, the GREG and the EBLUP, respectively.

2.5 Conclusion

In this Chapter, we studied the estimation of small area means using a two-level model assisted GREG estimator, and derived the model mean squared error (MSE) of the new GREG estimator. In addition, we used a data set of Moura and Holt (1999) to demonstrate empirically that the BLUP estimator is substantially more efficient than the new GREG in terms of model MSE, while the new GREG is only slightly more efficient than the standard GREG based on the regression model $y_i = X_i Z_i \beta + e_i; i = 1, \dots, m$. Moreover, using a simulation study under a model-based framework, we have shown that the new GREG estimator has consistently better performance relative to the GREG estimator in terms of average MSE, $\overline{MSE}$ and average relative error $\overline{ARE}$. However, due to borrowing strength from related small areas, the EBLUP estimator exhibited significantly better performance relative to the GREG and the new GREG estimators. In addition, we considered a simulation study under a synthetic finite population framework and have shown that the EBLUP estimator outperforms the new GREG and the GREG estimators in terms of $\overline{MSE}$ and $\overline{ARE}$.

Parameter	Diagonal Covariance: model (2.28) with z	Diagonal Covariance: model (2.26) without z	General Covariance: model (2.26) without z
β_0	8.442	8.688	8.456
β_1	0.451	1.321	1.223
β_2	0.744	2.636	2.596
α_1	3.779	-	-
α_2	1.659	-	-
σ_0^2	0.745	0.637	1.385
σ_1^2	0.237	0.471	0.234
σ_2^2	0.700	1.472	0.926
σ_{01}	-	-	0.354
σ_{02}	-	-	0.492
σ_{12}	-	-	0.333
σ_e^2	44.00	44.01	47.74

Table 2.1: Parameter estimates based on Moura and Holt's data set (1999), where σ_0^2 , σ_1^2 and σ_2^2 are the diagonal elements and σ_{01} , σ_{02} and σ_{12} are the off-diagonal elements of the covariance matrix Σ_ν

Quality Measure	Model without z			Model with z		
	GREG	New GREG	BLUP	GREG	New GREG	BLUP
EFF	100%	112%	306%	100%	106%	292%
$\overline{MSE}$	1.92	1.71	0.62	1.83	1.72	0.62

Table 2.2: Comparison of small area estimators: relative efficiency (EFF) and average MSE ($\overline{MSE}$) for the case of diagonal covariance matrix based on Moura and Holt's data set (1999)

Quality Measure	Model without z		
	GREG	New GREG	BLUP
EFF	100%	108%	253%
$\overline{MSE}$	2.02	1.87	0.80

Table 2.3: Comparison of small area estimators: relative efficiency (EFF) and average MSE ($\overline{MSE}$) for the case of a general covariance matrix based on Moura and Holt's data set (1999)

Quality Measure	Model without z			Model with z		
	New GREG	GREG	EBLUP	New GREG	GREG	EBLUP
$\overline{MSE}_1$	1.73	1.93	0.67	1.73	1.84	0.73
$\overline{ARE}_1$	0.13	0.14	0.08	0.12	0.13	0.08

Table 2.4: Comparison of small area estimators: average MSE ($\overline{MSE}_1$) and average absolute relative error ($\overline{ARE}_1$) under a model-based framework

Quality	Model without z			Model with z		
Measure	New GREG	GREG	EBLUP	New GREG	GREG	EBLUP
$\overline{MSE}_2$	9.35	10.32	3.22	10.73	11.16	4.85
$\overline{ARE}_2$	0.33	0.34	0.20	0.30	0.31	0.18

Table 2.5: Comparison of small area estimators: average MSE ($\overline{MSE}_2$) and average absolute relative error ($\overline{ARE}_2$) under a finite population framework ($n_i = 5$)

Quality	Model without z			Model with z		
Measure	New GREG	GREG	EBLUP	New GREG	GREG	EBLUP
$\overline{MSE}_2$	5.19	5.51	2.72	6.47	6.57	4.17
$\overline{ARE}_2$	0.24	0.25	0.18	0.22	0.23	0.16

Table 2.6: Comparison of small area estimators: average MSE ($\overline{MSE}_2$) and average absolute relative error ($\overline{ARE}_2$) under a finite population framework ($n_i = 10$)

Chapter 3

Pseudo EBLUP and Survey Weights

Using survey weights, You and Rao (2002) proposed a pseudo empirical best linear unbiased prediction (pseudo-EBLUP) estimator to estimate small area means under a nested error linear regression model. This estimator borrows strength across areas through the linking model, and makes use of survey weights to preserve benchmarking property in the sense that the estimators add up to a reliable direct estimator of a large area covering the small areas. In this Chapter, we derive a second order approximation to the mean squared error (MSE) of the pseudo-EBLUP estimator of a small area mean. Using this approximation, an estimator of MSE that is approximately unbiased is derived. Empirical results on the performance of the proposed approximation are also presented.

3.1 Introduction

Unit level random effects models, particularly nested error linear regression models, are often used in small area estimation to obtain efficient model-based estimators of small area means. A nested error linear regression model was used by Battese, Harter, and Fuller (1988) and Prasad and Rao (1990) for small area estimation, but they did not make use of the unit level survey weights. As a result, the estimators are not design consistent as the area sample size becomes large, unless the survey weights are equal within areas. On the other hand, direct estimators are design consistent, but fail to borrow strength from related small areas. You and Rao (2002) proposed a pseudo-EBLUP estimator for a small area mean that makes uses of survey weights. In addition, the method of You and Rao has a nice benchmarking property in the sense that the estimators automatically add up to a reliable direct estimator of a large area covering the small areas. However, they ignored the cross-product terms in the MSE, in deriving an estimator of MSE. As a result, this MSE estimator is not nearly unbiased. In this Chapter, we take account of this cross-product terms and obtain an estimator of MSE of the pseudo-EBLUP that is nearly unbiased. We also present some empirical results.

3.2 Pseudo EBLUP

The nested error linear regression model of Battese et al. (1988) is given by

$$y_{ij} = x'_{ij}\beta + \nu_i + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m \quad (3.1)$$

where y_{ij} and x_{ij} are the values of the response variable y and covariates x attached to the j -th sampled unit in the i -th area, β is a $p \times 1$ vector, $\nu_i \stackrel{i.i.d.}{\sim} N(0, \sigma_\nu^2)$, $e_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma_e^2)$, n_i is the sample size within i -th area and m is the number of areas. The i -th area mean μ_i is given by $\mu_i = \bar{X}_i' \beta + \nu_i$, assuming that the number of population units, N_i , in each area i is large, where $\bar{X}_i$ is the known population mean of x_{ij} 's.

We have the BLUP of i -th area mean μ_i as

$$\tilde{\mu}_i = \bar{X}_i' \tilde{\beta} + \gamma_i (\bar{y}_i - \bar{x}_i' \tilde{\beta}),$$

where $\tilde{\beta} = (\sum_{i=1}^m X_i' V_i^{-1} X_i)^{-1} (\sum_{i=1}^m X_i' V_i^{-1} y_i)$ with $V_i^{-1} = \sigma_e^{-2} [I_{n_i} - \frac{\gamma_i}{n_i} J_{n_i}]$ and $\gamma_i = \sigma_\nu^2 / (\sigma_\nu^2 + \sigma_e^2 / n_i)$. The EBLUP of i -th area mean μ_i is given by

$$\hat{\mu}_i = \bar{X}_i' \hat{\beta} + \hat{\gamma}_i (\bar{y}_i - \bar{x}_i' \hat{\beta}),$$

where $\hat{\beta}$ and $\hat{\gamma}_i$ are obtained by substituting estimators $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$ in $\tilde{\beta}$ and γ_i .

You and Rao (2002) proposed a pseudo-EBLUP estimator of μ_i under model (3.1) that depends on the survey weights $\tilde{w}_{ij}$ and satisfies the design consistency property. First assuming the parameters β, σ_ν^2 and σ_e^2 are known, a pseudo-Bayes estimator of μ_i is obtained as

$$\tilde{\mu}_{iw}^B = \bar{X}_i' \beta + \gamma_{iw} (\bar{y}_{iw} - \bar{x}_{iw}' \beta), \quad (3.2)$$

where $\bar{y}_{iw} = \sum_{j=1}^{n_i} w_{ij} y_{ij}$, $\bar{x}_{iw} = \sum_{j=1}^{n_i} w_{ij} x_{ij}$, $\gamma_{iw} = \sigma_\nu^2 / (\sigma_\nu^2 + \sigma_e^2 \alpha_i^2)$ with $\alpha_i^2 = \sum_{j=1}^{n_i} w_{ij}^2$, $w_{ij} = \tilde{w}_{ij} / \sum_j \tilde{w}_{ij}$ and $\sum_j w_{ij} = 1$. The estimator (3.2) is obtained from the aggregate data model $\bar{y}_{iw} = \bar{x}_{iw}' \beta + \nu_i + \bar{e}_{iw}$ with $\bar{e}_{iw} \stackrel{ind.}{\sim} N(0, \sigma_e^2 \alpha_i^2)$ as $\tilde{\mu}_{iw}^B = E(\mu_i | \bar{y}_{iw}, \beta, \sigma_\nu^2, \sigma_e^2)$.

The regression parameter β is then estimated for fixed $\delta = (\sigma_\nu^2, \sigma_e^2)'$ by solving the following survey-weighted estimating equation for β :

$$\sum_{i=1}^m \sum_{j=1}^{n_i} \tilde{w}_{ij} x_{ij} (y_{ij} - x'_{ij} \beta - \tilde{\nu}_{iw}(\beta, \delta)) = 0, \quad (3.3)$$

where $\tilde{\nu}_{iw}(\beta, \delta) = \gamma_{iw}(\bar{y}_{iw} - \bar{x}'_{iw} \beta)$. The solution to (3.3) is given by

$$\tilde{\beta}_w(\delta) = \left\{ \sum_{i=1}^m \sum_{j=1}^{n_i} \tilde{w}_{ij} x_{ij} (x_{ij} - \gamma_{iw} \bar{x}_{iw})' \right\}^{-1} \left\{ \sum_{i=1}^m \sum_{j=1}^{n_i} \tilde{w}_{ij} (x_{ij} - \gamma_{iw} \bar{x}_{iw}) y_{ij} \right\}.$$

Given δ , the estimator $\tilde{\beta}_w(\delta)$ is model-unbiased for β .

To estimate σ_ν^2 and σ_e^2 , You and Rao (2002) used the well known method of fitting-of-constants (Henderson, 1953) that can be simplified through a transformation method of Battese et al. (1988). In this method, we first calculate the residual sum of squares SSE_1 with ν_1 degrees of freedom by regression ordinary least squares (OLS) through the origin the y-deviations $(y_{ij} - \bar{y}_i)$ on the nonzero x -deviations $(x_{ij} - \bar{x}_i)$ for those areas with $n_i > 1$. This leads to an unbiased estimator of σ_e^2 as

$$\hat{\sigma}_e^2 = \frac{SSE_1}{\nu_1},$$

where $\nu_1 = n - m - p_1$ and p_1 is the number of nonzero x -deviations. We then calculate the residual sum of squares SSE_2 by regressing ordinary least squares y_{ij} on x_{ij} . An unbiased estimator of σ_ν^2 is given by

$$\tilde{\sigma}_\nu^2 = [SSE_2 - (n - p)\hat{\sigma}_e^2]/\eta_1,$$

where p is the dimension of β and $\eta_1 = \sum_{i=1}^m n_i [1 - n_i \bar{x}'_i (\sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} x'_{ij})^{-1} \bar{x}_i]$. Note that $E[SSE_2] = \eta_1 \sigma_\nu^2 + (n - p) \sigma_e^2$. Since $\tilde{\sigma}_\nu^2$ can take negative values, we truncate estimator $\tilde{\sigma}_\nu^2$ to zero whenever it is negative. The truncated estimated $\hat{\sigma}_\nu^2 = \max(\tilde{\sigma}_\nu^2, 0)$ is no longer unbiased but it is consistent as m increases.

In this Chapter, we estimate δ by the restricted maximum likelihood (REML) method, instead of the fitting-of-constants method, to get the REML estimators $\hat{\delta} = (\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)'$. The resulting estimator of β is given by $\hat{\beta}_w = \tilde{\beta}(\hat{\delta})$. Substituting $\hat{\beta}_w, \hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$ for β, σ_ν^2 and σ_e^2 in (3.2), a pseudo-EBLUP estimator of μ_i is given by

$$\hat{\mu}_{iw} = \bar{X}_i' \hat{\beta}_w + \hat{\gamma}_{iw}(\bar{y}_{iw} - \bar{x}_{iw}' \hat{\beta}_w). \quad (3.4)$$

The pseudo-EBLUP estimator (3.4) is design consistent as n_i becomes large. Also, the pseudo-EBLUP estimators automatically satisfy the benchmarking property when the estimators $\hat{\mu}_{iw}$ are aggregated over i , assuming that the weights $\tilde{w}_{ij}$ are calibrated to agree with the known population total N_i , i.e., $\sum_j \tilde{w}_{ij} = \tilde{w}_i = N_i$. We have $\sum_i N_i \hat{\mu}_{iw} = \hat{Y}_w + (X - \hat{X}_w)' \hat{\beta}_w$, where $\hat{Y}_w = \sum_i \sum_j \tilde{w}_{ij} y_{ij}$ and $\hat{X}_w = \sum_i \sum_j \tilde{w}_{ij} x_{ij}$ are the direct estimators of the large area totals Y and X respectively, and $\hat{Y}_w + (X - \hat{X}_w)' \hat{\beta}_w$ is the direct survey regression estimator of Y , assuming X is known. Hence, the pseudo-EBLUP estimators $\hat{\mu}_{iw}$ satisfy the benchmarking property without any adjustment.

3.3 MSE Estimation of Pseudo EBLUP

3.3.1 You and Rao Method

Ignoring the cross-product terms, $C_{1iw}(\delta) = E(\tilde{\mu}_{iw}^B - \mu_i)(\tilde{\mu}_{iw} - \tilde{\mu}_{iw}^B)$ and $C_{2iw}(\delta) = E(\hat{\mu}_{iw} - \tilde{\mu}_{iw})(\tilde{\mu}_{iw} - \mu_i)$, in $MSE(\hat{\mu}_{iw})$ where $\tilde{\mu}_{iw} = \bar{X}_i' \tilde{\beta}_w + \gamma_{iw}(\bar{y}_{iw} - \bar{x}_{iw}' \tilde{\beta}_w)$, You and Rao (2002) obtained the following MSE approximation, along the lines of Prasad and Rao (1990):

$$MSE(\hat{\mu}_{iw}) = g_{1iw}(\delta) + g_{2iw}(\delta) + g_{3iw}(\delta), \quad (3.5)$$

where

$$g_{1iw}(\delta) = (1 - \gamma_{iw})\sigma_\nu^2, \quad (3.6)$$

$$g_{2iw}(\delta) = (\bar{X}_i - \gamma_{iw}\bar{x}_{iw})'\Phi_w(\bar{X}_i - \gamma_{iw}\bar{x}_{iw}), \quad (3.7)$$

$$g_{3iw}(\delta) = \gamma_{iw}(1 - \gamma_{iw})^2\sigma_e^{-4}\sigma_\nu^{-2}h(\delta),$$

$$\begin{aligned} \Phi_w = & \sigma_e^2 \left(\sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} t'_{ij} \right)^{-1} \left(\sum_{i=1}^m \sum_{j=1}^{n_i} t_{ij} t'_{ij} \right) \left(\sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} t'_{ij} \right)^{-T} \\ & + \sigma_\nu^2 \left(\sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} t'_{ij} \right)^{-1} \left[\sum_{i=1}^m \left(\sum_{j=1}^{n_i} t_{ij} \right) \left(\sum_{j=1}^{n_i} t_{ij} \right)' \right] \left(\sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} t'_{ij} \right)^{-T}, \end{aligned}$$

with $t_{ij} = \tilde{w}_{ij}(x_{ij} - \gamma_{iw}\bar{x}_{iw})$, $(\sum_i \sum_j x_{ij} t'_{ij})^{-T} = \left((\sum_i \sum_j x_{ij} t'_{ij})^{-1} \right)'$ and $h(\delta)$ is given by $h(\delta) = \sigma_e^4 Var(\hat{\sigma}_\nu^2) + \sigma_\nu^4 Var(\hat{\sigma}_e^2) - 2\sigma_\nu^2 \sigma_e^2 Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)$.

In terms of REML estimators $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$, the estimator of MSE, based on the MSE approximation (3.5), can be defined as

$$mse_{YR1}(\hat{\mu}_{iw}) = g_{1iw}(\hat{\delta}) + g_{2iw}(\hat{\delta}) + 2g_{3iw}(\hat{\delta}). \quad (3.8)$$

If $C_{1iw}(\delta) = C_{2iw}(\delta) = 0$, the MSE estimator (3.8) is approximately unbiased in the sense that its bias is of order $o(m^{-1})$ for large m . This is true in the self-weighting case: $n_i \tilde{w}_{ij} = \tilde{w}_i$ or $w_{ij} = 1/n_i$, where $w_{ij} = \tilde{w}_{ij} / \sum_j \tilde{w}_{ij}$ are the normalized weights.

An area-specific version of MSE estimator, based on the MSE approximation (3.5), is given by

$$mse_{YR2}(\hat{\mu}_{iw}) = g_{1iw}(\hat{\delta}) + g_{2iw}(\hat{\delta}) + 2g_{3iw}^*(\hat{\delta}, \bar{y}_{iw}), \quad (3.9)$$

where $g_{3iw}^*(\delta, \bar{y}_{iw}) = \gamma_{iw}^2(1 - \gamma_{iw})^2\sigma_e^{-4}\sigma_\nu^{-4}h(\delta)(\bar{y}_{iw} - \bar{x}'_{iw}\tilde{\beta}_w)^2$.

3.3.2 The Proposed Method

We now derive an approximation to $MSE(\hat{\mu}_{iw})$ taking account of the cross-product terms $C_{1iw}(\delta)$ and $C_{2iw}(\delta)$. We have

$$MSE(\hat{\mu}_{iw}) = E(\hat{\mu}_{iw} - \mu_i)^2 = E(\hat{\mu}_{iw} - \tilde{\mu}_{iw})^2 + E(\tilde{\mu}_{iw} - \mu_i)^2 + 2E(\hat{\mu}_{iw} - \tilde{\mu}_{iw})(\tilde{\mu}_{iw} - \mu_i), \quad (3.10)$$

where

$$\begin{aligned} E(\tilde{\mu}_{iw} - \mu_i)^2 &= E(\tilde{\mu}_{iw}^B - \mu_i)^2 + E(\tilde{\mu}_{iw} - \tilde{\mu}_{iw}^B)^2 + 2E(\tilde{\mu}_{iw} - \tilde{\mu}_{iw}^B)(\tilde{\mu}_{iw}^B - \mu_i) \\ &= g_{1iw}(\delta) + g_{2iw}(\delta) + 2C_{1iw}(\delta), \end{aligned}$$

with

$$C_{1iw}(\delta) = \sigma_e^2 \tilde{w}_i \gamma_{iw} (\bar{X}_i - \gamma_{iw} \bar{x}_{iw})' \left\{ \sum_i \sum_j \tilde{w}_{ij} x_{ij} (x_{ij} - \gamma_{iw} \bar{x}_{iw})' \right\}^{-1} \cdot \left[\sum_j w_{ij}^2 x_{ij} - \bar{x}_{iw} \alpha_i^2 \right]. \quad (3.11)$$

In addition, $E(\hat{\mu}_{iw} - \tilde{\mu}_{iw})^2 = g_{3iw}(\delta) + o(m^{-1})$. To obtain an approximation for the cross-product term in (3.10), $C_{2iw}(\delta)$, we make a further assumption on variance component parameters δ . We assume that the estimator $\hat{\delta}$ of δ is a solution to the estimating equation of the form

$$M(\delta) = m^{-1} \sum_{i=1}^m a_i(y_{i1}, \dots, y_{in_i}, \delta) = 0, \quad (3.12)$$

where $a_i(\cdot)$ is a 2-vector $m^{-1} \sum_i E[a_i(y_{i1}, \dots, y_{in_i}, \delta)] = 0$ provided that δ is the true vector of parameters. This assumption is satisfied by the REML method.

Following Jiang and Lahiri (2006), we approximate the cross-product term

$C_{2iw}(\delta)$, based on a Taylor expansion, as (see Section 3.6.1):

$$\begin{aligned}
C_{2iw}(\delta) &\approx E \left\{ \left[-2m^{-1} \left(\frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} \right) A^{-1} a_i(\cdot) + \left(\frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} \right) A^{-1} \left(\frac{\partial M}{\partial \delta'} \right) A^{-1} M(\delta) \right. \right. \\
&\quad \left. - \frac{1}{2} \left(\frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} \right) A^{-1} \text{col}_{1 \leq j \leq 2} \{ [M'(\delta) A^{-T} E \left(\frac{\partial^2 M_j}{\partial \delta \partial \delta'} \right)]' \} A^{-1} M(\delta) \right. \\
&\quad \left. + \frac{1}{2} M'(\delta) A^{-T} \left(\frac{\partial^2 \tilde{\mu}_{iw}}{\partial \delta \partial \delta'} \right) A^{-1} M(\delta) \right] \left((\bar{X}_i - \gamma_{iw} \bar{x}_{iw})' (\tilde{\beta}_w - \beta) + \tilde{\nu}_{iw} - \nu_i \right) \right\} \\
&\equiv \tilde{C}_{2iw}(\delta), \tag{3.13}
\end{aligned}$$

where $M(\hat{\delta}) = \frac{1}{m} \sum_{i=1}^m a_i(y_{i1}, \dots, y_{in_i}, \hat{\delta}) = 0$, $A = E(\partial M(\delta)/\partial \delta')$, $\text{col}_{1 \leq j \leq 2} \{c_j\}$ represents a 2×2 matrix with j -th column as c_j , M_j denotes the j -th component of M . We should point out that we calculated the above expectation and it is very complicated, so we omit it here. However, we used this expression (exact expectation) for empirical results at the end of this Chapter.

Note that equations (3.12) and (3.13) are different from Jiang and Lahiri (2006) in the sense that our equations involve only the variance components σ_v^2 and σ_e^2 such that the matrices involved are at most with 2-dimension. On the other hand, Jiang and Lahiri (2006) have a similar equation with matrices involving dimension $p + 2$ involving the regression parameters β as well as variance components.

The restricted maximum likelihood estimator of δ satisfies the estimating equations (3.12) based on the REML score functions. Score functions are given by

$$\begin{aligned}
S_1(\delta) &= -\frac{1}{2} \text{tr}(P J_n) + \frac{1}{2} y' P J_n P y, \\
S_2(\delta) &= -\frac{1}{2} \text{tr}(P) + \frac{1}{2} y' P P y,
\end{aligned}$$

where $P = V^{-1} - V^{-1} X (X' V^{-1} X)^{-1} X' V^{-1}$, $V^{-1} = \text{diag}(V_1^{-1}, \dots, V_m^{-1})$, J_n is a $n \times n$ matrix of one's and y is a $n \times 1$ vector with elements $y_{ij} (j = 1, \dots, n_i; i = 1, \dots, m)$.

Therefore, we may express the components a_i corresponding to σ_ν^2 and σ_e^2 respectively as

$$a_{1i} = y_i' P_i J_{n_i} P_i y_i - \text{tr}(P_i J_{n_i}), \quad (3.14)$$

$$a_{2i} = y_i' P_i P_i y_i - \text{tr}(P_i), \quad (3.15)$$

where J_{n_i} is a $n_i \times n_i$ matrix of one's, P_i and y_i are i -th elements of P and y respectively. Hence, (3.13) is well-defined using (3.14) and (3.15) for a_{1i} and a_{2i} .

Theorem 3.1. *A second order approximation of the mean squared error (MSE) of the pseudo-EBLUP estimator $\hat{\mu}_{iw}$ is given by*

$$MSE(\hat{\mu}_{iw}) \approx g_{1iw}(\delta) + g_{2iw}(\delta) + g_{3iw}(\delta) + 2C_{1iw}(\delta) + 2\tilde{C}_{2iw}(\delta), \quad (3.16)$$

where $C_{1iw}(\delta)$ and $\tilde{C}_{2iw}(\delta)$ are given by (3.11) and (3.13) respectively.

Proof : See Section 3.6.1.

In addition, we can show that the cross-product terms (3.11) and (3.13) are zero for the EBLUP estimator $\hat{\mu}_i$. In the case of the EBLUP estimator $\hat{\mu}_i$, we may write (3.11) as

$$C_{1i}(\delta) = \sigma_e^2 \gamma_i N_i (\bar{X}_i - \gamma_i \bar{x}_i)' \left\{ \sum_i \sum_j \frac{N_i}{n_i} x_{ij} (x_{ij} - \gamma_i \bar{x}_i)' \right\}^{-1} \left\{ n_i^{-2} \sum_j x_{ij} - n_i^{-1} \bar{x}_i \right\} = 0,$$

since $w_{ij} = 1/n_i$ and $n_i^{-2} \sum_j x_{ij} = n_i^{-1} \bar{x}_i$. We can also express $C_{2i}(\delta) = E(\hat{\mu}_i - \tilde{\mu}_i)(\tilde{\mu}_i - \mu_i)$ as $E \left\{ f_i(y, \delta) \left((\bar{X}_i - \gamma_i \bar{x}_i)' (\tilde{\beta} - \beta) + \tilde{\nu}_i - \nu_i \right) \right\}$, where $E(\mu_i | \bar{y}_i) = \bar{X}_i' \beta + \tilde{\nu}_i$.

Hence, $C_{2i}(\delta) = E \left\{ f_i(y, \delta) \left((\bar{X}_i - \gamma_i \bar{x}_i)' (\tilde{\beta} - \beta) + \tilde{\nu}_i - \nu_i \right) \right\} = E \left\{ f_i(y, \delta) \left[(\bar{X}_i - \gamma_i \bar{x}_i)' E(\tilde{\beta} - \beta | A'y) + E(\nu_i | \bar{y}_i) - E(\nu_i | y) \right] \right\} = 0$, noting that $E(\nu_i | \bar{y}_i) = E(\nu_i | y_i) =$

$E(\nu_i|y) = \tilde{\nu}_i$ and $E(\tilde{\beta} - \beta|A'y) = 0$, since $\tilde{\beta} - \beta$ is independent of $A'y$ provided that $A'X = 0$. On the other hand, $C_{1iw}(\delta) = C_{2iw}(\delta) = 0$ for the pseudo-EBLUP only in the self-weighting case, $n_i \tilde{w}_{ij} = \tilde{w}_i$.

Using the MSE approximation (3.16), we propose two MSE estimators for the pseudo-EBLUP. The first MSE estimator is given by

$$mse_{TR1}(\hat{\mu}_{iw}) = g_{1iw}(\hat{\delta}) + g_{2iw}(\hat{\delta}) + 2g_{3iw}(\hat{\delta}) + 2C_{1iw}(\hat{\delta}) + 2\tilde{C}_{2iw}(\hat{\delta}). \quad (3.17)$$

The estimator (3.17) is nearly unbiased in the sense of Theorem 3.2.

Theorem 3.2.

$$E[mse_{TR1}(\hat{\mu}_{iw})] = MSE(\hat{\mu}_{iw}) + o(m^{-1}).$$

Proof : See Section 3.6.2.

Following Jiang and Lahiri (2006), we propose a second MSE estimator, an area-specific MSE estimator, as

$$mse_{TR2}(\hat{\mu}_{iw}) = g_{1iw}(\hat{\delta}) + g_{2iw}(\hat{\delta}) + 2\widehat{g_{3iw}(\delta)} + 2C_{1iw}(\hat{\delta}) + 2\tilde{C}_{2iw}(\hat{\delta}), \quad (3.18)$$

where $\widehat{g_{3iw}(\delta)} = tr\{\widehat{V_{\hat{\delta}}(\delta)}E[(\frac{\partial \hat{\mu}_{iw}^B}{\partial \delta})(\frac{\partial \hat{\mu}_{iw}^B}{\partial \delta})']|_{\delta=\hat{\delta}}\}$ with $\widehat{V_{\hat{\delta}}(\delta)} = m^{-2}\hat{A}^{-1}(\hat{\delta})\{\sum_{i=1}^m \hat{a}_i(.) \cdot \hat{a}_i'(.)\}(\hat{A}^{-1}(\hat{\delta}))'$, $\hat{A}(\hat{\delta}) = m^{-1}\sum_{i=1}^m [\partial a_i(.) / \partial \delta']|_{\delta=\hat{\delta}}$ and $\widehat{V_{\hat{\delta}}(\delta)}$ is an estimator of $V_{\hat{\delta}}(\delta) = E(\hat{\delta} - \delta)(\hat{\delta} - \delta)'$. The estimator $mse_{TR2}(\hat{\mu}_{iw})$ is nearly unbiased in the sense of Theorem 3.3.

Theorem 3.3.

$$E[mse_{TR2}(\hat{\mu}_{iw})] = MSE(\hat{\mu}_{iw}) + o(m^{-1}).$$

Proof : See Section 3.6.3.

3.4 Empirical Results

In this Section, we study the performance of the You-Rao MSE estimators mse_{YR1} and mse_{YR2} and the proposed MSE estimators mse_{TR1} and mse_{TR2} for the pseudo-EBLUP, by applying them to a real data set given by Battese, Harter, and Fuller (1988). We consider the estimation of mean hectares of corn per segment for 12 counties (area) in north-central Iowa. The total sample size, $n = \sum_i n_i$, for the 12 counties is 36, and the sample size n_i within each county ranged from 1 to 5. The population size N_i within each county ranged from 402 to 965. The nested error model used here is given by

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \nu_i + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m$$

where y_{ij} is the number of hectares of corn in the j -th segment of the i -th county, x_{1ij} and x_{2ij} are the number of pixels classified as corn and soybeans respectively, in the j -th segment of the i -th county. The small area mean $\mu_i = \beta_0 + \beta_1 \bar{X}_{1i} + \beta_2 \bar{X}_{2i} + \nu_i$, where $\bar{X}_{1i}$ and $\bar{X}_{2i}$ are the known population mean number of pixels classified as corn and soybeans per segment respectively in the i -th county, by using LANDSAT satellite readings. We considered two scenarios. In each scenario, we constructed a sampling design to compare the performance of the proposed MSE estimators.

For the first scenario, the weights w_{ij} were calculated based on $x_{1ij} + x_{2ij}$, the sum of number of pixels classified as corn x_{1ij} and soybeans x_{2ij} :

$$w_{ij} = \frac{x_{1ij} + x_{2ij}}{\sum_{j=1}^{n_i} (x_{1ij} + x_{2ij})}; j = 1, \dots, n_i; i = 1, \dots, m.$$

Table 3.1 reports the weights w_{ij} for different areas. The variability of the above weights for each i is fairly small, because the area segments within a county do not vary significantly with respect to their sizes.

In the second scenario, we calculated weights w_{ij} based on x_{1ij} , the number of pixels classified as corn, as

$$w_{ij} = \frac{x_{1ij}}{\sum_{j=1}^{n_i} x_{1ij}}; j = 1, \dots, n_i; i = 1, \dots, m.$$

The weights w_{ij} for different counties (areas) are shown in Table 3.2. Variability of the weights for each i is now significant, unlike in the case of scenario 1.

The variance components estimators $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$ are 140.02 and 147.27 respectively, by using the restricted maximum likelihood method. We compute the estimators of MSE , using mse_{YR1} , mse_{YR2} , mse_{TR1} and mse_{TR2} given by (3.8), (3.9), (3.17) and (3.18).

Tables 3.1 and 3.2 suggest that the effect of estimating the cross-product terms, $C_{1iw} + \tilde{C}_{2iw}$, on MSE estimator is fairly small when comparing mse_{YR1} to mse_{TR1} . However, the area-specific version mse_{TR2} seems to be significantly smaller than the corresponding mse_{YR2} for some counties; for example, for the county Worth in Table 3.1, $mse_{YR2} = 109.50$ compared to $mse_{TR2} = 78.66$. In addition, in Tables 3.1 and 3.2, mse_{TR2} is consistently smaller than mse_{TR1} across areas. Moreover, since the weights using corn only (Table 3.2) are more variable than using both corn and soybeans (Table 3.1), the absolute values of the cross-products using corn only are larger than using both corn and soybeans for most counties.

On the other hand, if we choose adhoc weights as in Table 3.3, it can lead to non-ignorable cross-product terms. In fact, ignoring the cross-product terms in this case leads to biased MSE estimator. For example, for the county Kossuth in Table 3.3, $mse_{YR1} = 42.52$ compared to $mse_{TR1} = 45.09$, where $2(C_{1iw} + \tilde{C}_{2iw}) = 2.57$. Table 3.3 strengthens the case for estimating the covariance terms in calculating the

MSE estimator.

3.5 Summary

In this Chapter, we derived a second order approximation to the mean squared error (MSE) of the survey weighted pseudo-EBLUP estimators introduced by You and Rao (2002). Using this approximation, we obtained estimators of MSE that are nearly unbiased. Our approximation accounts for the cross-product terms in the MSE, which are ignored in You and Rao (2002). Empirical results are also provided on the performance of the proposed approximation.

3.6 Proofs

3.6.1 Proof of Theorem 3.1

We can write $MSE(\hat{\mu}_{iw})$ as

$$MSE(\hat{\mu}_{iw}) = E(\tilde{\mu}_{iw} - \mu_i)^2 + E(\hat{\mu}_{iw} - \tilde{\mu}_{iw})^2 + 2E(\tilde{\mu}_{iw} - \mu_i)(\hat{\mu}_{iw} - \tilde{\mu}_{iw}). \quad (3.19)$$

Further, we have shown that $E(\tilde{\mu}_{iw} - \mu_i)^2 = g_{1iw}(\delta) + g_{2iw}(\delta) + 2C_{1iw}(\delta)$, where $g_{1iw}(\delta)$, $g_{2iw}(\delta)$ and $C_{1iw}(\delta)$ are given by (3.6), (3.7) and (3.11) respectively. We first evaluate the cross-product term of (3.19).

Under regularity conditions (Jiang and Lahiri, 2006), we have

$M(\delta) = [O_p(m^{-\frac{1}{2}})]_{2 \times 1}$, $\frac{\partial M}{\partial \delta'} = [O_p(1)]_{2 \times 2}$, $\frac{\partial^2 M_j}{\partial \delta \partial \delta'} = [O_p(1)]_{2 \times 2}$, ($j = 1, 2$), $\frac{\partial M}{\partial \delta'} - A = [O_p(m^{-\frac{1}{2}})]_{2 \times 2}$, $\hat{\delta} - \delta = [O_p(m^{-\frac{1}{2}})]_{2 \times 1}$, where δ is the true parameter. We now have a Taylor expansion of $M(\hat{\delta})$ as follows:

$$0 = M(\hat{\delta}) = M(\delta) + \frac{\partial M}{\partial \delta'}(\hat{\delta} - \delta) + \frac{1}{2} \text{col}_{1 \leq j \leq 2} \left\{ \frac{\partial^2 M_j}{\partial \delta \partial \delta'}(\hat{\delta} - \delta) \right\} (\hat{\delta} - \delta) + [o_p(m^{-1})]_{2 \times 1}.$$

Then, we may write

$$\hat{\delta} - \delta = - \left\{ \frac{\partial M}{\partial \delta'} + \frac{1}{2} \text{col}_{1 \leq j \leq 2} \left\{ \frac{\partial^2 M_j}{\partial \delta \partial \delta'}(\hat{\delta} - \delta) \right\} \right\}^{-1} M(\delta) + [o_p(m^{-1})]_{2 \times 1}. \quad (3.20)$$

We let $\frac{\partial M}{\partial \delta'} + \frac{1}{2} \text{col}_{1 \leq j \leq 2} \left\{ \frac{\partial^2 M_j}{\partial \delta \partial \delta'}(\hat{\delta} - \delta) \right\} \equiv B$. Therefore, $B = A + [O_p(m^{-1/2})]_{2 \times 2}$, noting that $\partial M / \partial \delta' = A + [O_p(m^{-1/2})]_{2 \times 2}$, $\hat{\delta} - \delta = [O_p(m^{-1/2})]_{2 \times 1}$ and $\partial^2 M_j / \partial \delta \partial \delta'$ is $[O_p(1)]_{2 \times 2}$.

On the other hand, we have $B^{-1} = A^{-1} - A^{-1}(B - A)B^{-1} = A^{-1} + [O_p(m^{-1/2})]_{2 \times 2}$.

Then we may express (3.20) as

$$\hat{\delta} - \delta = -A^{-1}M(\delta) + [O_p(m^{-1})]_{2 \times 1}, \quad (3.21)$$

Again, we may write B^{-1} as follows

$$\begin{aligned} B^{-1} &= A^{-1} - A^{-1}(B - A)B^{-1} \\ &= A^{-1} - A^{-1}(B - A)[A^{-1} - A^{-1}(B - A)B^{-1}] \\ &= A^{-1} - A^{-1}(B - A)A^{-1} + \{A^{-1}(B - A)\}^2 B^{-1} \\ &= A^{-1} - A^{-1} \left\{ \frac{\partial M}{\partial \delta'} + \frac{1}{2} \text{col}_{1 \leq j \leq 2} \left\{ \frac{\partial^2 M_j}{\partial \delta \partial \delta'} (\hat{\delta} - \delta) \right\} - A \right\} A^{-1} + [O_p(m^{-1})]_{2 \times 2} \\ &= A^{-1} - A^{-1} \left(\frac{\partial M}{\partial \delta} - A \right) A^{-1} - \frac{1}{2} A^{-1} \text{col}_{1 \leq j \leq 2} \left\{ \frac{\partial^2 M_j}{\partial \delta \partial \delta'} (\hat{\delta} - \delta) \right\} A^{-1} + [O_p(m^{-1})]_{2 \times 2} \\ &= A^{-1} - A^{-1} \left(\frac{\partial M}{\partial \delta} - A \right) A^{-1} + \frac{1}{2} A^{-1} \text{col}_{1 \leq j \leq 2} \left\{ [M' A^{-T} \frac{\partial^2 M_j}{\partial \delta \partial \delta'}]' \right\} A^{-1} + [O_p(m^{-1})]_{2 \times 2}, \end{aligned}$$

by using (3.21). Then, we may write (3.20) as

$$\begin{aligned} \hat{\delta} - \delta &= -A^{-1}M + A^{-1} \left(\frac{\partial M}{\partial \delta'} - A \right) A^{-1}M - \frac{1}{2} A^{-1} \text{col}_{1 \leq j \leq 2} \left\{ [M' A^{-T} E(\frac{\partial^2 M_j}{\partial \delta \partial \delta'})]' \right\} A^{-1}M \\ &\quad + [o_p(m^{-1})]_{2 \times 1}. \end{aligned} \quad (3.22)$$

Using Taylor expansion of $\hat{\mu}_{iw}(\hat{\delta})$, we obtain

$$\hat{\mu}_{iw}(\hat{\delta}) = \tilde{\mu}_{iw}(\delta) + \left(\frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} \right) (\hat{\delta} - \delta) + \frac{1}{2} (\hat{\delta} - \delta)' \frac{\partial^2 \tilde{\mu}_{iw}}{\partial \delta \partial \delta'} (\hat{\delta} - \delta) + o_p(m^{-1}).$$

Hence, we may write

$$\begin{aligned} E(\hat{\mu}_{iw} - \tilde{\mu}_{iw})(\tilde{\mu}_{iw} - \mu_i) &= E \left\{ \left[\left(\frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} \right) (\hat{\delta} - \delta) + \frac{1}{2} (\hat{\delta} - \delta)' \frac{\partial^2 \tilde{\mu}_{iw}}{\partial \delta \partial \delta'} (\hat{\delta} - \delta) \right] \right. \\ &\quad \left. \cdot \left((\bar{X}_i - \gamma_{iw} \bar{x}_{iw})' (\tilde{\beta}_w - \beta) + \tilde{\nu}_{iw} - \nu_i \right) \right\} + o(m^{-1}). \end{aligned}$$

We know that $E[a_i(\tilde{\mu}_{jw} - \mu_j)] = 0, \quad \forall i \neq j.$

Therefore, by using (3.22), we may obtain

$$\begin{aligned}
E(\hat{\mu}_{iw} - \tilde{\mu}_{iw})(\tilde{\mu}_{iw} - \mu_i) &= E \left\{ \left[-\frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} A^{-1} M + \frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} A^{-1} \left(\frac{\partial M}{\partial \delta'} - A \right) A^{-1} M - \frac{1}{2} \frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} \right. \right. \\
&\quad \cdot A^{-1} \text{col}_{1 \leq j \leq 2} \left\{ [M' A^{-T} E(\frac{\partial^2 M_j}{\partial \delta \partial \delta'})]' \right\} A^{-1} M + \frac{1}{2} M' A^{-T} \\
&\quad \cdot \frac{\partial^2 \tilde{\mu}_{iw}}{\partial \delta \partial \delta'} A^{-1} M \left. \right] \left((\bar{X}_i - \gamma_{iw} \bar{x}_{iw})' (\tilde{\beta}_w - \beta) + \tilde{\nu}_{iw} - \nu_i \right) \right\} + o(m^{-1}) \\
&= m^{-1} E \left\{ \left[-\frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} A^{-1} a_i + \frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} A^{-1} \left(\frac{\partial M}{\partial \delta'} - A \right) A^{-1} a_i \right. \right. \\
&\quad - \frac{m}{2} \frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} A^{-1} \cdot \text{col}_{1 \leq j \leq 2} \left\{ [M' A^{-T} E(\frac{\partial^2 M_j}{\partial \delta \partial \delta'})]' \right\} A^{-1} M \\
&\quad + \frac{m}{2} M' A^{-T} \frac{\partial^2 \tilde{\mu}_{iw}}{\partial \delta \partial \delta'} A^{-1} M \left. \right] \left((\bar{X}_i - \gamma_{iw} \bar{x}_{iw})' (\tilde{\beta}_w - \beta) + \tilde{\nu}_{iw} - \nu_i \right) \right\} \\
&\quad + o(m^{-1}) \\
&\equiv \tilde{C}_{2iw}(\delta) + o(m^{-1}).
\end{aligned}$$

Next, we can write the second term of right side of (3.19) as

$$\begin{aligned}
E(\hat{\mu}_{iw} - \tilde{\mu}_{iw})^2 &= E \left\{ \left[\left(\frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} \right) (\hat{\delta} - \delta) + \frac{1}{2} (\hat{\delta} - \delta)' \frac{\partial^2 \tilde{\mu}_{iw}}{\partial \delta \partial \delta'} (\hat{\delta} - \delta) \right]^2 \right\} + O(m^{-2}) \\
&= E \left[\left(\frac{\partial \tilde{\mu}_{iw}}{\partial \delta'} \right) V_{\hat{\delta}}(\delta) \left(\frac{\partial \tilde{\mu}_{iw}}{\partial \delta} \right) \right] + o(m^{-1}) \\
&= g_{3iw}(\delta) + o(m^{-1}), \tag{3.23}
\end{aligned}$$

where

$$V_{\hat{\delta}}(\delta) = E(\hat{\delta} - \delta)(\hat{\delta} - \delta)',$$

$$g_{3iw}(\delta) = \gamma_{iw}(1 - \gamma_{iw})^2 \sigma_e^{-4} \sigma_\nu^{-2} h(\delta),$$

where $h(\delta) = \sigma_e^4 \text{Var}(\hat{\sigma}_\nu^2) + \sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2) - 2\sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)$, with $\text{Var}(\hat{\sigma}_\nu^2) = (I_{\nu\nu} I_{ee} - I_{\nu e}^2)^{-1} I_{ee}$, $\text{Var}(\hat{\sigma}_e^2) = (I_{\nu\nu} I_{ee} - I_{\nu e}^2)^{-1} I_{\nu\nu}$ and $\text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) = -(I_{\nu\nu} I_{ee} - I_{\nu e}^2)^{-1} I_{\nu e}$ where

$I_{\nu\nu} = \frac{1}{2} \sum_i n_i^2 (\sigma_e^2 + n_i \sigma_\nu^2)^{-2}$, $I_{\nu e} = \frac{1}{2} \sum_i n_i (\sigma_e^2 + n_i \sigma_\nu^2)^{-2}$ and $I_{ee} = \frac{1}{2} \sum_i \left[(n_i - 1) \sigma_e^{-4} + (\sigma_e^2 + n_i \sigma_\nu^2)^{-2} \right]$. Note that the all neglected terms in (3.23) are of lower order $o(m^{-1})$ for large m . Proof of Theorem 3.1 is complete.

3.6.2 Proof of Theorem 3.2

We must show that $mse_{TR1}(\hat{\mu}_{iw})$ is correct to $O_p(m^{-1})$. Under regularity conditions and using Taylor expansion of $g_{1iw}(\hat{\delta})$, we obtain

$$g_{1iw}(\hat{\delta}) = g_{1iw}(\delta) + \left(\frac{\partial g_{1iw}}{\partial \delta'} \right) (\hat{\delta} - \delta) + \frac{1}{2} (\hat{\delta} - \delta)' \frac{\partial^2 g_{1iw}}{\partial \delta \partial \delta'} (\hat{\delta} - \delta) + o_p(m^{-1}),$$

and then take its expectation to get

$$E[g_{1iw}(\hat{\delta})] = g_{1iw}(\delta) + \frac{1}{2} tr \left\{ \frac{\partial^2 g_{1iw}}{\partial \delta \partial \delta'} V_{\hat{\delta}}(\delta) \right\} + o(m^{-1}).$$

It is easy to show that $\frac{1}{2} tr \left\{ \frac{\partial^2 g_{1iw}}{\partial \delta \partial \delta'} V_{\hat{\delta}}(\delta) \right\} = -g_{3iw}(\delta)$. Hence,

$$E[g_{1iw}(\hat{\delta})] = g_{1iw}(\delta) - g_{3iw}(\delta) + o(m^{-1}).$$

Furthermore, since $g_{2iw}(\delta)$ and $g_{3iw}(\delta)$ are of order $O(m^{-1})$, it follows that

$$E[g_{2iw}(\hat{\delta})] = g_{2iw}(\delta) + o(m^{-1}),$$

$$E[g_{3iw}(\hat{\delta})] = g_{3iw}(\delta) + o(m^{-1}).$$

Therefore, $g_{2iw}(\hat{\delta})$ and $g_{3iw}(\hat{\delta})$ are the correct estimators of $g_{2iw}(\delta)$ and $g_{3iw}(\delta)$ respectively, each with bias $o(m^{-1})$.

As a final step, we need to show that $C_{1iw}(\hat{\delta})$ and $C_{2iw}(\hat{\delta})$ are the correct estimators of $C_{1iw}(\delta)$ and $C_{2iw}(\delta)$ respectively, with bias $o(m^{-1})$. We may write

$E[C_{1iw}(\hat{\delta})] - C_{1iw}(\delta) = o(m^{-1})$, since $C_{1iw}(\delta) = O(m^{-1})$. Similarly, we may write $E[\tilde{C}_{2iw}(\hat{\delta})] - \tilde{C}_{2iw}(\delta) = o(m^{-1})$, since we assume $\tilde{C}_{2iw}(\delta) = O(m^{-1})$. Therefore, $C_{1iw}(\hat{\delta})$ and $\tilde{C}_{2iw}(\hat{\delta})$ are the correct estimators of $C_{1iw}(\delta)$ and $\tilde{C}_{2iw}(\delta)$ with bias $o(m^{-1})$. Combining above results, we have $E[mse_{TRL}(\hat{\mu}_{iw})] = MSE(\hat{\mu}_{iw}) + o(m^{-1})$. The proof is complete.

3.6.3 Proof of Theorem 3.3

We have shown in Theorem 3.2 that $E[g_{1iw}(\hat{\delta})] = g_{1iw}(\delta) - g_{3iw}(\delta) + o(m^{-1})$. Also, we have $E[g_{2iw}(\hat{\delta})] = g_{2iw}(\delta) + o(m^{-1})$.

On the other hand, we have

$$g_{3iw}(\delta) = tr \left\{ V_{\hat{\delta}}(\delta) E \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right) \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right)' \right\}.$$

To have an-area specific estimator of $g_{3iw}(\delta)$, we may use (3.22) to write

$$\begin{aligned} V_{\hat{\delta}}(\delta) &= E(\hat{\delta} - \delta)(\hat{\delta} - \delta)' = E \left[A^{-1} M M' A^{-T} \right] + o(m^{-1}) \\ &= m^{-2} E \left[A^{-1} \left[\sum_{i=1}^m a_i(.) a_i'(.) \right] A^{-T} \right] + o(m^{-1}), \end{aligned} \quad (3.24)$$

then the area-specific estimator of $V_{\hat{\delta}}(\delta)$ is as follows

$$\widehat{V_{\hat{\delta}}(\delta)} = m^{-2} \hat{A}^{-1} \left[\sum_{i=1}^m \hat{a}_i(.) \hat{a}_i'(.) \right] \hat{A}^{-T},$$

where $\hat{A} = m^{-1} \sum_{i=1}^m \left[\partial a_i(.) / \partial \delta' \right]_{\delta=\hat{\delta}}$. We also use plug-in estimator for $E(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta})$ as $E(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta})(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta})'|_{\delta=\hat{\delta}}$. That means, we first take expectation $E(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta})(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta})'$, then we substitute δ with $\hat{\delta}$ to get $E(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta})(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta})'|_{\delta=\hat{\delta}}$. Therefore we have an area-specific estimator of $g_{3iw}(\delta)$ as follows

$$\widehat{g_{3iw}(\delta)} = tr \left\{ \widehat{V_{\hat{\delta}}(\delta)} E \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right) \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right)' \Big|_{\delta=\hat{\delta}} \right\}.$$

In next step, we show that $\widehat{g_{3iw}(\delta)}$ is estimator of order $O_p(m^{-1})$. Under regularity conditions, we may write

$$\begin{aligned}\widehat{g_{3iw}(\delta)} - g_{3iw}(\delta) &= tr \left\{ \widehat{V_{\hat{\delta}}(\delta)} \left[E \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right) \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right)' \Big|_{\delta=\hat{\delta}} - E \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right) \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right)' \right] \right\} \\ &\quad + tr \left\{ (\widehat{V_{\hat{\delta}}(\delta)} - V_{\hat{\delta}}(\delta)) E \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right) \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right)' \right\},\end{aligned}$$

and then take its expectation to get $E[\widehat{g_{3iw}(\delta)}] - g_{3iw}(\delta) = o(m^{-1})$, where $E \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right) \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right)' \Big|_{\delta=\hat{\delta}} - E \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right) \left(\frac{\partial \tilde{\mu}_{iw}^B}{\partial \delta} \right)' = [O_p(m^{-1/2})]_{2 \times 2}$ and $E[\widehat{V_{\hat{\delta}}(\delta)}] - V_{\hat{\delta}}(\delta) = o(m^{-1})$ since $V_{\hat{\delta}}(\delta)$ is of order $O(m^{-1})$ based on (3.24). Therefore, $\widehat{g_{3iw}(\delta)}$ is the correct estimator of $g_{3iw}(\delta)$ with bias $o(m^{-1})$.

In addition, we have shown in Theorem 3.2 that $E[C_{1iw}(\hat{\delta})] - C_{1iw}(\delta) = o(m^{-1})$ and $E[\tilde{C}_{2iw}(\hat{\delta})] - \tilde{C}_{2iw}(\delta) = o(m^{-1})$. Hence, combining above results, the proof is complete, namely, $E[mse_{TR2}(\hat{\mu}_{iw})] = MSE(\hat{\mu}_{iw}) + o(m^{-1})$.

County	Sample Size	Weight	mse_{YR1}	mse_{TR1}	mse_{YR2}	mse_{TR2}	$2C_{iw}$
Cerro	1	1	101.79	101.79	84.46	84.73	0
Hami.	1	1	98.06	98.06	83.09	80.96	0
Worth	1	1	95.96	95.96	109.50	78.66	0
Humb.	2	(.49,.51)	70.99	71.14	68.03	60.74	0.15
Fran.	3	(.33,.32,.35)	44.98	44.94	52.00	38.16	-0.04
Poca.	3	(.29,.37,.34)	45.42	45.66	47.79	38.85	0.14
Winn.	3	(.36,.33,.31)	45.41	45.42	44.13	38.66	0.01
Wrig.	3	(.33,.32,.35)	47.20	47.16	40.89	40.42	-0.04
Webs.	4	(.27,.25,.20,.28)	35.76	35.60	38.60	30.83	-0.16
Hanc.	5	(.20,.20,.20,.19,.21)	30.05	30.03	26.70	26.65	-0.02
Koss.	5	(.15,.22,.23,.21,.19)	29.64	30.22	32.20	26.73	0.58
Hard.	5	(.19,.20,.19,.21,.21)	34.04	34.02	33.88	30.69	-0.02

Table 3.1: MSE estimators and cross-product term $2(C_{1iw} + \tilde{C}_{2iw}) = 2C_{iw}$ for corn when weights are based on $x_{1ij} + x_{2ij}$

County	Sample Size	Weight	mse_{YR1}	mse_{TR1}	mse_{YR2}	mse_{TR2}	$2C_{iw}$
Cerro	1	1	102.49	102.49	85.53	85.47	0
Hami.	1	1	97.93	97.93	83.32	80.83	0
Worth	1	1	95.83	95.83	108.92	78.53	0
Humb.	2	(.54,.46)	72.57	71.92	68.87	61.50	-0.65
Fran.	3	(.36,.28,.36)	45.57	45.34	50.40	38.48	-0.23
Poca.	3	(.31,.47,.22)	48.67	49.19	52.71	41.67	0.52
Winn.	3	(.41,.34,.25)	46.89	46.79	45.86	39.74	-0.10
Wrig.	3	(.44,.27,.29)	51.53	52.31	45.14	45.21	0.78
Webs.	4	(.24,.28,.19,.29)	36.12	35.92	39.55	31.09	-0.20
Hanc.	5	(.22,.17,.25,.19,.17)	29.99	30.28	26.70	26.78	0.29
Koss.	5	(.14,.23,.22,.22,.19)	29.95	30.51	32.97	26.98	0.56
Hard.	5	(.16,.26,.19,.14,.25)	33.21	33.63	32.86	29.99	0.42

Table 3.2: The cross-product terms and the estimators of MSE for corn where weights are based on x_{1ij} only

County	Sample Size	Weight	mse_{YR1}	mse_{TR1}	mse_{YR2}	mse_{TR2}	$2C_{iw}$
Cerro	1	1	102.48	102.48	85.16	85.26	0
Hami.	1	1	99.55	99.55	82.33	82.26	0
Worth	1	1	95.75	95.75	113.75	78.37	0
Humb.	2	(.10,.90)	89.18	90.65	108.71	75.08	1.47
Fran.	3	(.20,.75,.05)	70.77	70.26	104.78	57.70	-0.51
Poca.	3	(.30,.15,.55)	55.83	54.69	50.93	45.98	-1.14
Winn.	3	(.40,.30,.30)	46.17	46.17	47.16	39.22	0
Wrig.	3	(.80,.10,.10)	86.77	86.34	85.89	73.11	-0.43
Webs.	4	(.40,.10,.45,.05)	49.73	49.66	46.12	41.85	-0.07
Hanc.	5	(.30,.20,.40,.03,.07)	40.45	40.68	36.58	34.83	0.23
Koss.	5	(.40,.25,.30,.03,.02)	42.52	45.09	44.67	38.78	2.57
Hard.	5	(.03,.30,.04,.30,.33)	44.47	44.30	49.63	38.61	-0.17

Table 3.3: The cross-product terms and the estimators of MSE for corn based on adhoc weights

Chapter 4

On Small Area Estimation Under a Sub-area Level Model

In this Chapter, we introduce a sub-area level small area model. A detailed description of the model is given in Section 4.1. This model is an extension of the Fay and Herriot (1979) area level model to sub-area level model. This model not only may be used to estimate small area means by borrowing strength from related areas, but also it can borrow strength from sub-areas to obtain more efficient sub-area estimators. Model-based estimators are obtained from the best linear unbiased prediction (BLUP) estimators by replacing the unknown model parameters by suitable estimators, using an iterative method of estimating the model parameters based on weighted residual sum of squares. We obtain second order approximations to the mean squared error (MSE) of the model-based small area estimators and then derive MSE estimators unbiased to second order. Results of a small simulation study on the performance of the proposed estimators are also provided.

4.1 Introduction

In the context of linear mixed models, we propose the following sub-area level linking model for the sub-area means μ_{ij} :

$$\mu_{ij} = x'_{ij}\beta + \nu_i + u_{ij}; j = 1, \dots, N_i; i = 1, \dots, m, \quad (4.1)$$

where x_{ij} is a $p \times 1$ vector of sub-area level auxiliary variables, β is a $p \times 1$ vector of regression parameters, $\nu_i \stackrel{i.i.d.}{\sim} N(0, \sigma_\nu^2)$ and $u_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma_u^2)$. We assume that n_i sub-areas are independently sampled from the N_i sub-areas in the i -th area. Let y_{ij} be the direct estimator of the sub-area mean μ_{ij} . For example, j may denote a county (sub-area) j within a state (area) i .

On the other hand, the sampling model is given by

$$y_{ij} = \mu_{ij} + e_{ij},$$

where the sampling errors e_{ij} are independent $N(0, \sigma_{eij}^2)$ with known sampling variances σ_{eij}^2 . Assuming no sample selection bias, the combined model for the sample is given by

$$y_{ij} = x'_{ij}\beta + \nu_i + u_{ij} + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m. \quad (4.2)$$

Fuller and Goyeneche (1998) proposed the sub-area model (4.2) in the context of Small Area Income and Poverty Estimation (SAIPE) project in the United States.

In matrix notation, the combined model (4.2) can be written as

$$y_i = X_i\beta + \nu_i 1_{n_i} + u_i + e_i, i = 1, \dots, m,$$

where $y_i = (y_{i1}, \dots, y_{in_i})'$ is a $n_i \times 1$ vector, X_i is a $n_i \times p$ matrix, $u_i = (u_{i1}, \dots, u_{in_i})'$ and $e_i = (e_{i1}, \dots, e_{in_i})'$.

Equivalently, we have

$$y_i = X_i\beta + Z_iv_i + e_i, i = 1, \dots, m, \quad (4.3)$$

where Z_i is a $n_i \times (n_i + 1)$ matrix given by

$$Z_i = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & & \\ 1 & 0 & 0 & \dots & 1 \end{pmatrix},$$

and $v_i = (\nu_i, u_{i1}, \dots, u_{in_i})'$ is a $(n_i + 1) \times 1$ vector.

Model (4.2) accounts for the sub-area level effect u_{ij} as well as the area level effect ν_i . It helps us to estimate both small area means and sub-area means by borrowing strength from related areas as well as sub-areas. We are interested in estimating the sub-area means μ_{ij} and the area means μ_i , where μ_{ij} is given by (4.1) and

$$\mu_i = \bar{X}_i'\beta + \nu_i,$$

where $\bar{X}_i = \sum_{j=1}^{N_i} N_{ij}x_{ij} / \sum_j N_{ij}$ is the known population mean of x for the i -th area and N_{ij} the number of units in the j -th sub-area of the i -th area. The mean μ_i is based on the assumption of large N_i so that $\bar{U}_i = \sum_j N_{ij}u_{ij} / \sum_j N_{ij} \approx 0$ and $\mu_i \approx \sum_{j=1}^{N_i} N_{ij}\mu_{ij} / \sum_j N_{ij}$. In Section 4.2, we derive the EBLUP estimators of μ_{ij} and μ_i .

4.2 EBLUP Estimation

We first derive the BLUP estimators of the two means μ_{ij} and μ_i . We may write μ_{ij} as $\mu_{ij} = x'_{ij}\beta + m'_{ij}v_i$, where $m_{ij} = (1, 0, \dots, 0, 1, 0, \dots, 0)' = (1, \text{col}_{1 \leq l \leq n_i}(\delta_{il})')'$ with

$$\delta_{il} = \begin{cases} 1 & l = j \\ 0 & \text{o.w.} \end{cases}$$

It follows from Henderson (1975) that for known variance components $\delta = (\sigma_\nu^2, \sigma_u^2)'$, the BLUP estimator of μ_{ij} can be written as $\tilde{\mu}_{ij}(\delta) = x'_{ij}\tilde{\beta}(\delta) + m'_{ij}\tilde{v}_i = x'_{ij}\tilde{\beta}(\delta) + m'_{ij}G_i Z'_i V_i^{-1}(y_i - X_i \tilde{\beta}(\delta))$, where $\tilde{\beta}(\delta) = (\sum_{i=1}^m X'_i V_i^{-1} X_i)^{-1} (\sum_{i=1}^m X'_i V_i^{-1} y_i)$ and $V_i = R_i + Z_i G_i Z'_i$ with $R_i = \text{diag}(\sigma_{ei1}^2, \dots, \sigma_{ein_i}^2)$ and

$$G_i = \begin{pmatrix} \sigma_\nu^2 & 0 \\ 0 & \sigma_u^2 I_{n_i} \end{pmatrix}.$$

For simplicity of notation, we let $\tilde{\beta}(\delta) = \tilde{\beta}$ and $\tilde{\mu}_{ij}(\delta) = \tilde{\mu}_{ij}$.

We can write V_i as

$$V_i = \sigma_\nu^2 J_{n_i} + \text{diag}(\sigma_u^2 + \sigma_{ei1}^2, \dots, \sigma_u^2 + \sigma_{ein_i}^2). \quad (4.4)$$

Using (4.4), we employ the following Lemma to obtain V_i^{-1} .

Lemma 4.1. *Let A be a $k \times k$ nonsingular matrix and u and v be two $k \times 1$ vectors such that $A + uv'$ is nonsingular. Then we have*

$$(A + uv')^{-1} = A^{-1} - A^{-1}uv'A^{-1}(1 + v'A^{-1}u)^{-1}.$$

We have

$$V_i^{-1} = \text{diag}_j((\sigma_u^2 + \sigma_{eij}^2)^{-1}) - \frac{\sigma_u^2 \gamma_i}{\gamma_i} w_i w'_i, \quad (4.5)$$

where $\gamma_{ij} = \sigma_u^2 / (\sigma_u^2 + \sigma_{eij}^2)$; $\gamma_{i.} = \sum_{j=1}^{n_i} \gamma_{ij}$; $\gamma_i = \sigma_v^2 / (\sigma_v^2 + \sigma_u^2 / \gamma_{i.})$; $w_{ij} = (\sigma_u^2 + \sigma_{eij}^2)^{-1}$; $w_i = (w_{i1}, \dots, w_{in_i})'$.

Using (4.5), the BLUP of μ_{ij} , after simplification, is given by

$$\begin{aligned}\tilde{\mu}_{ij} &= x'_{ij}\tilde{\beta} + \frac{\gamma_{i.}}{\sigma_u^2} \sigma_v^2 (\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) + \gamma_{ij}(y_{ij} - x'_{ij}\tilde{\beta}) - \frac{\gamma_i \sigma_v^2 (\sigma_v^2 \gamma_{i.} + \sigma_u^2 \gamma_{ij})}{\sigma_u^2 (\sigma_u^2 + \sigma_v^2 \gamma_{i.})} (\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) \\ &= x'_{ij}\tilde{\beta} + \gamma_i \sigma_v^2 \frac{1 - \gamma_{ij}}{\sigma_u^2 + \sigma_v^2 \gamma_{i.}} (\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) + \gamma_{ij}(y_{ij} - x'_{ij}\tilde{\beta}) \\ &= x'_{ij}\tilde{\beta} + \gamma_i (1 - \gamma_{ij}) (\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) + \gamma_{ij}(y_{ij} - x'_{ij}\tilde{\beta}),\end{aligned}$$

where $\bar{y}_{i\gamma} = \frac{1}{\gamma_{i.}} \sum_{j=1}^{n_i} \gamma_{ij} y_{ij}$; $\bar{x}_{i\gamma} = \frac{1}{\gamma_{i.}} \sum_{j=1}^{n_i} \gamma_{ij} x_{ij}$.

We now derive the BLUP estimator of $\mu_i = \bar{X}'_i \beta + m'_i v_i$, where $m_i = (1, 0, 0, \dots, 0)'$ $= (1, 0')'$. The BLUP of μ_i may be written as $\tilde{\mu}_i(\delta) = \bar{X}'_i \tilde{\beta} + m'_i G_i Z'_i V_i^{-1} (y_i - X_i \tilde{\beta})$.

We write $\tilde{\mu}_i(\delta) = \tilde{\mu}_i$ for simplicity, and $\tilde{\mu}_i$ simplifies to

$$\tilde{\mu}_i = \bar{X}'_i \tilde{\beta} + \gamma_i (\bar{y}_{i\gamma} - \bar{x}'_{i\gamma} \tilde{\beta}). \quad (4.6)$$

4.2.1 Estimation of Variance Components

At this stage, we need to estimate the variance components σ_v^2 and σ_u^2 to obtain the EBLUP estimators of μ_{ij} and μ_i . We write the model (4.3) in matrix form as

$$y = X\beta + \varepsilon,$$

where y is a $n \times 1$ vector, X is a $n \times p$ matrix, ε is a $n \times 1$ vector and $n = \sum_{i=1}^m n_i$.

We then use generalized least squares to estimate β as

$$\tilde{\beta}(\delta) = (X'V^{-1}X)^{-1}(X'V^{-1}y) = \left(\sum_{i=1}^m X'_i V_i^{-1} X_i\right)^{-1} \left(\sum_{i=1}^m X'_i V_i^{-1} y_i\right),$$

where $V = \text{diag}(V_1, \dots, V_m)$.

We now extend the Fay-Herriot (1979) iterative method of estimating the variance component in the area level model to the sub-area level model (4.2). Based on weighted residual sum of squares, we have

$$(\mathbf{y} - X\tilde{\beta})'V^{-1}(\mathbf{y} - X\tilde{\beta}) \sim \chi^2(n - p).$$

The method of moments is then used to form an estimating equation as follows :

$$(\mathbf{y} - X\tilde{\beta})'V^{-1}(\mathbf{y} - X\tilde{\beta}) = (n - p). \quad (4.7)$$

On the other hand, we can take the sample mean of the model (4.2) as

$$\bar{y}_i = \bar{x}_i'\beta + \nu_i + \bar{u}_i + \bar{e}_i, \quad (4.8)$$

where $\bar{y}_i = n_i^{-1} \sum_j y_{ij}$ and so on. In matrix notation, we can write (4.8) as $\bar{\mathbf{y}} = \bar{\mathbf{x}}'\beta + \varepsilon_1$ with $\bar{\mathbf{y}} = (\bar{y}_1, \dots, \bar{y}_m)'$ and ε_1 has 0 mean and covariance matrix $V^* = \text{diag}(V_1^*, \dots, V_m^*)$ with $V_i^* = \sigma_\nu^2 + \sigma_u^2/n_i + n_i^{-2} \sum_{j=1}^{n_i} \sigma_{eij}^2 = n_i^{-1}(n_i\sigma_\nu^2 + \sigma_u^2 + \sigma_{ei}^2)$ where $\sigma_{ei}^2 = \sum_{j=1}^{n_i} \sigma_{eij}^2/n_i$. We then estimate β as $\tilde{\beta} = (\bar{\mathbf{x}}'V^{*-1}\bar{\mathbf{x}})^{-1}(\bar{\mathbf{x}}'V^{*-1}\bar{\mathbf{y}})$ where $\tilde{\beta} = \tilde{\beta}(\delta)$. Now, based on weighted residual sum of squares, we have

$$\sum_{i=1}^m (\bar{y}_i - \bar{x}_i'\tilde{\beta})^2/V_i^* \sim \chi^2(m - p).$$

We then use the method of moments to form the estimating equation

$$\sum_{i=1}^m (\bar{y}_i - \bar{x}_i'\tilde{\beta})^2/V_i^* = m - p. \quad (4.9)$$

(4.7) and (4.9) give two equations in the two parameters σ_ν^2 and σ_u^2 . The estimators $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_u^2$ are obtained by solving the two equations (4.7) and (4.9) iteratively. Since $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_u^2$ can take negative values, we define $\hat{\sigma}_\nu^2 = \max(0, \tilde{\sigma}_\nu^2)$ and $\hat{\sigma}_u^2 = \max(0, \tilde{\sigma}_u^2)$.

Using $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_u^2$, the EBLUP estimators of μ_{ij} and μ_i can be expressed as

$$\hat{\mu}_{ij} = x'_{ij}\hat{\beta} + \hat{\gamma}_i(1 - \hat{\gamma}_{ij})(\hat{y}_{i\gamma} - \hat{x}'_{i\gamma}\hat{\beta}) + \hat{\gamma}_{ij}(y_{ij} - x'_{ij}\hat{\beta}),$$

and

$$\hat{\mu}_i = \bar{X}'_i\hat{\beta} + \hat{\gamma}_i(\hat{y}_{i\gamma} - \hat{x}'_{i\gamma}\hat{\beta}),$$

where $\hat{\mu}_{ij} = \tilde{\mu}_{ij}(\hat{\delta})$; $\hat{\mu}_i = \tilde{\mu}_i(\hat{\delta})$ and $\hat{\beta} = \tilde{\beta}(\hat{\delta})$.

In Section 4.3, we obtain a second order approximation to the mean squared error (MSE) of $\hat{\mu}_{ij}$ and $\hat{\mu}_i$.

4.3 MSE Approximation

In this Section, we obtain a second order approximation to the MSE of the EBLUP estimators $\hat{\mu}_{ij}$ and $\hat{\mu}_i$, in the sense that the neglected terms are of order $o(m^{-1})$ for large m . We assume normality of ν_i, u_{ij} and e_{ij} .

4.3.1 MSE Approximation of $\hat{\mu}_{ij}$

Under normality of the random effects ν_i, u_{ij} and e_{ij} , we can write $MSE(\hat{\mu}_{ij})$ as

$$MSE(\hat{\mu}_{ij}) = MSE(\tilde{\mu}_{ij}) + E(\hat{\mu}_{ij} - \tilde{\mu}_{ij})^2, \quad (4.10)$$

where $MSE(\tilde{\mu}_{ij}) = E(\tilde{\mu}_{ij} - \mu_{ij})^2$. Further, an exact expression for $MSE(\tilde{\mu}_{ij})$ is given by

$$MSE(\tilde{\mu}_{ij}) = g_{1ij}(\delta) + g_{2ij}(\delta), \quad (4.11)$$

where

$$g_{1ij}(\delta) = (1 - \gamma_{ij})^2(\sigma_u^2 + (1 - \gamma_i)\sigma_\nu^2) + \gamma_{ij}^2\sigma_{eij}^2, \quad (4.12)$$

$$g_{2ij}(\delta) = (1 - \gamma_{ij})^2(x_{ij} - \gamma_i\bar{x}_{i\gamma})'\Phi(x_{ij} - \gamma_i\bar{x}_{i\gamma}), \quad (4.13)$$

where $\Phi = Var(\tilde{\beta}) = (\sum_{i=1}^m X_i'V_i^{-1}X_i)^{-1}$, see Section 4.9.1.

The second term $g_{2ij}(\delta)$ of (4.11), due to estimating β , is of order $O(m^{-1})$, while the first term $g_{1ij}(\delta)$, given by (4.12), is $O(1)$. We can therefore interpret $g_{1ij}(\delta)$ as the *MSE* when all parameters are known, and $g_{2ij}(\delta)$ as the inflation of MSE due to estimating β . It can be shown that the leading term $g_{1ij}(\delta) \leq \sigma_{eij}^2$, the variance of the direct estimator y_{ij} .

It remains to evaluate the second term $E(\hat{\mu}_{ij} - \tilde{\mu}_{ij})^2$ in (4.10). Following Das, Jiang, and Rao (2004), we propose the following approximation, based on a Taylor series expansion:

$$\begin{aligned} E(\hat{\mu}_{ij} - \tilde{\mu}_{ij})^2 &\approx E\left[\frac{d\tilde{\mu}_{ij}^B}{d\sigma_\nu^2}\right]^2 Var(\tilde{\sigma}_\nu^2) + E\left[\frac{d\tilde{\mu}_{ij}^B}{d\sigma_u^2}\right]^2 Var(\tilde{\sigma}_u^2) \\ &\quad + 2E\left\{\left(\frac{d\tilde{\mu}_{ij}^B}{d\sigma_\nu^2}\right)\left(\frac{d\tilde{\mu}_{ij}^B}{d\sigma_u^2}\right)\right\} Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \equiv g_{3ij}(\delta), \end{aligned}$$

where $\tilde{\mu}_{ij}^B$ is the best (Bayes) estimator of μ_{ij} defined as

$$\tilde{\mu}_{ij}^B = x'_{ij}\beta + \gamma_i(1 - \gamma_{ij})(\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\beta) + \gamma_{ij}(y_{ij} - x'_{ij}\beta).$$

After considerable simplification, $g_{3ij}(\delta)$ is obtained as

$$\begin{aligned} g_{3ij}(\delta) &= (1 - \gamma_{ij})^2 \left\{ \left[\frac{\sigma_u^2}{\gamma_i} \right]^2 \left(\frac{\gamma_i}{\sigma_\nu^2} \right)^3 Var(\tilde{\sigma}_\nu^2) + \left[\frac{\sigma_\nu^2}{\gamma_i} \left[\frac{\gamma_i}{\sigma_u^2 \gamma_i} \sum_j \gamma_{ij}^2 \left(1 - \frac{\gamma_i \sigma_u^2}{\gamma_i \sigma_\nu^2} \right) - \frac{\gamma_i \gamma_{ij}}{\sigma_u^2} \right] \left[-\frac{\gamma_i}{\sigma_u^2 \gamma_i} \right. \right. \right. \\ &\quad \left. \left. \left. \sum_j \gamma_{ij}^2 \left(1 + \frac{\gamma_i \sigma_u^2}{\gamma_i \sigma_\nu^2} \right) + \frac{\gamma_{ij}}{\sigma_u^2} (2 - \gamma_i) \right] + \sigma_u^{-4} \left[\gamma_{ij}^2 \sigma_\nu^2 + \sigma_u^2 \gamma_{ij} + \frac{\gamma_i^2}{\gamma_i^2} (\sigma_u^2 \sum_j \gamma_{ij}^3 + \sigma_\nu^2 (\sum_j \gamma_{ij}^2)^2) \right] \right\} \end{aligned}$$

$$\left. -2\frac{\gamma_i\gamma_{ij}}{\gamma_{i.}}(\sigma_\nu^2 \sum_j \gamma_{ij}^2 + \sigma_u^2 \gamma_{ij}) \right] \left. Var(\tilde{\sigma}_u^2) + 2\frac{\gamma_i}{\gamma_{i.}\sigma_\nu^2} \left[\gamma_{ij}(1-\gamma_i) - \frac{\gamma_i^2\sigma_u^2}{\gamma_{i.}\sigma_\nu^2} \sum_j \gamma_{ij}^2 \right] Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \right\}. \quad (4.14)$$

The term $g_{3ij}(\delta)$ is of the same order as $g_{2ij}(\delta)$. Combining (4.11) and (4.14), we get a second order approximation to $MSE(\hat{\mu}_{ij})$ as

$$MSE(\hat{\mu}_{ij}) \approx g_{1ij}(\delta) + g_{2ij}(\delta) + g_{3ij}(\delta). \quad (4.15)$$

In Section 4.9.2, we show that the neglected terms in the approximation (4.15) are of lower order, $o(m^{-1})$, for large m . Furthermore, $Var(\tilde{\sigma}_\nu^2)$, $Var(\tilde{\sigma}_u^2)$ and $Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2)$ are derived in Section 4.4.

4.3.2 MSE Approximation of $\hat{\mu}_i$

Similar to Section 4.3.1, we follow Das, Jiang, and Rao (2004) to find $MSE(\hat{\mu}_i)$. We have

$$MSE(\hat{\mu}_i) = MSE(\tilde{\mu}_i) + E(\hat{\mu}_i - \tilde{\mu}_i)^2,$$

where

$$MSE(\tilde{\mu}_i) = g_{1i}(\delta) + g_{2i}(\delta), \quad (4.16)$$

with

$$g_{1i}(\delta) = \frac{\gamma_i}{\gamma_{i.}} \sigma_u^2,$$

$$g_{2i}(\delta) = (\bar{X}_i - \gamma_i \bar{x}_{i\gamma})' \Phi (\bar{X}_i - \gamma_i \bar{x}_{i\gamma}).$$

The terms $g_{1i}(\delta)$ and $g_{2i}(\delta)$ are of order $O(1)$ and $O(m^{-1})$ respectively.

Further, we derive $E(\hat{\mu}_i - \tilde{\mu}_i)^2$ as follows

$$E(\hat{\mu}_i - \tilde{\mu}_i)^2 \approx E\left[\frac{d\tilde{\mu}_i^B}{d\sigma_\nu^2}\right]^2 Var(\tilde{\sigma}_\nu^2) + E\left[\frac{d\tilde{\mu}_i^B}{d\sigma_u^2}\right]^2 Var(\tilde{\sigma}_u^2)$$

$$+2E\left\{\left(\frac{d\tilde{\mu}_i^B}{d\sigma_\nu^2}\right)\left(\frac{d\tilde{\mu}_i^B}{d\sigma_u^2}\right)\right\}Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \equiv g_{3i}(\delta),$$

where $\tilde{\mu}_i^B$ is the best (Bayes) estimator of μ_i :

$$\tilde{\mu}_i^B = \bar{X}_i' \beta + \gamma_i(\bar{y}_{i\gamma} - \bar{x}_{i\gamma}' \beta).$$

Then, after some calculation we have that

$$\begin{aligned} g_{3i}(\delta) = & \left(\frac{\gamma_i}{\gamma_{i.}}\right)^2 \left\{ \frac{\sigma_u^4 \gamma_i}{\sigma_\nu^6} Var(\tilde{\sigma}_\nu^2) + \sigma_u^{-4} \sum_j \gamma_{ij}^4 \left[(1 - \gamma_i) \sigma_\nu^2 + \sigma_u^2 \left(1 - 2 \frac{\gamma_i}{\gamma_{i.}}\right) + \sigma_{eij}^2 \right] Var(\tilde{\sigma}_u^2) \right. \\ & \left. + 2\sigma_\nu^{-4} \gamma_i \sum_j \gamma_{ij}^2 \left[(\gamma_i - 1) \sigma_\nu^2 + \frac{\sigma_u^2}{\gamma_{i.}} (\gamma_i - \gamma_{ij}) - \frac{\gamma_{ij}}{\gamma_{i.}} \sigma_{eij}^2 \right] Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \right\}, \end{aligned} \quad (4.17)$$

where $g_{3i}(\delta)$ is of order $O(m^{-1})$. Combining (4.16) and (4.17), we may write $MSE(\hat{\mu}_i)$ as follows

$$MSE(\hat{\mu}_i) \approx g_{1i}(\delta) + g_{2i}(\delta) + g_{3i}(\delta). \quad (4.18)$$

The neglected terms in the approximation (4.18) are $o(m^{-1})$ for large m , similar to Section 4.3.1.

4.4 Variances and Covariance of $\tilde{\sigma}_\nu^2$ and $\tilde{\sigma}_u^2$

We are interested in finding expressions for the variances of $\tilde{\sigma}_\nu^2$ and $\tilde{\sigma}_u^2$ as well as the covariance between these estimators, under model (4.2). This enables the evaluation of the terms $g_{3ij}(\delta)$ and $g_{3i}(\delta)$ given by expressions (4.14) and (4.17) respectively. Recall that from (4.7) and (4.9), we might write

$$(\mathbf{y} - X\tilde{\beta})'V^{-1}(\mathbf{y} - X\tilde{\beta}) = \mathbf{y}'A_1(\delta)\mathbf{y},$$

and

$$\sum_{i=1}^m (\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 / V_i^* = \bar{y}' A_2(\delta) \bar{y},$$

with

$$A_1(\delta) = V^{-1}(\delta) - V^{-1}(\delta) X (X' V^{-1}(\delta) X)^{-1} X' V^{-1}(\delta),$$

$$A_2(\delta) = V^{*-1}(\delta) - V^{*-1}(\delta) \bar{x} (\bar{x}' V^{*-1}(\delta) \bar{x})^{-1} \bar{x}' V^{*-1}(\delta),$$

$$\text{and } \bar{y} = (\bar{y}_1, \dots, \bar{y}_m)'; \quad \bar{x}' = (\bar{x}_1, \dots, \bar{x}_m).$$

Therefore, we may write

$$y' A_1(\tilde{\delta}) y - (n - p) = 0,$$

$$\bar{y}' A_2(\tilde{\delta}) \bar{y} - (m - p) = 0.$$

Then by a Taylor series expansion, we have that

$$\begin{aligned} & \begin{pmatrix} y' A_1(\tilde{\delta}) y / (n - p) \\ \bar{y}' A_2(\tilde{\delta}) \bar{y} / (m - p) \end{pmatrix} = \begin{pmatrix} y' A_1(\delta) y / (n - p) \\ \bar{y}' A_2(\delta) \bar{y} / (m - p) \end{pmatrix} \\ & + \begin{pmatrix} y' A_1^1(\delta) y / (n - p) & y' A_1^2(\delta) y / (n - p) \\ \bar{y}' A_2^1(\delta) \bar{y} / (m - p) & \bar{y}' A_2^2(\delta) \bar{y} / (m - p) \end{pmatrix} \begin{pmatrix} \tilde{\sigma}_\nu^2 - \sigma_\nu^2 \\ \tilde{\sigma}_u^2 - \sigma_u^2 \end{pmatrix} + [O(\|\tilde{\delta} - \delta\|^2)]_{2 \times 1}. \end{aligned}$$

Hence,

$$\begin{aligned} & \begin{pmatrix} \frac{y' A_1(\delta) y}{(n-p)} - 1 \\ \frac{\bar{y}' A_2(\delta) \bar{y}}{(m-p)} - 1 \end{pmatrix} + \begin{pmatrix} y' A_1^1(\delta) y / (n - p) & y' A_1^2(\delta) y / (n - p) \\ \bar{y}' A_2^1(\delta) \bar{y} / (m - p) & \bar{y}' A_2^2(\delta) \bar{y} / (m - p) \end{pmatrix} \begin{pmatrix} \tilde{\sigma}_\nu^2 - \sigma_\nu^2 \\ \tilde{\sigma}_u^2 - \sigma_u^2 \end{pmatrix} \\ & + [O(\|\tilde{\delta} - \delta\|^2)]_{2 \times 1} = 0, \end{aligned} \tag{4.19}$$

where $A_i^j(\delta) = \frac{dA_i(\delta)}{d\delta_j}$, with $\delta_1 = \sigma_\nu^2$ and $\delta_2 = \sigma_u^2$; $i = 1, 2$.

Therefore, we may write that

$$\begin{pmatrix} \tilde{\sigma}_v^2 - \sigma_v^2 \\ \tilde{\sigma}_u^2 - \sigma_u^2 \end{pmatrix} = \begin{pmatrix} y'A_1^1(\delta)y/(n-p) & y'A_1^2(\delta)y/(n-p) \\ \bar{y}'A_2^1(\delta)\bar{y}/(m-p) & \bar{y}'A_2^2(\delta)\bar{y}/(m-p) \end{pmatrix}^{-1} \begin{pmatrix} 1 - \frac{y'A_1(\delta)y}{(n-p)} \\ 1 - \frac{\bar{y}'A_2(\delta)\bar{y}}{(m-p)} \end{pmatrix} + [O(\|\tilde{\delta} - \delta\|^2)]_{2 \times 1}. \quad (4.20)$$

On the other hand $(n-p)^{-1}y'A_1(\delta)y = (n-p)^{-1}E(y'A_1(\delta)y) + O_p(m^{-1/2}) = 1 + O_p(m^{-1/2})$, since $Var[(n-p)^{-1}y'A_1y] = O(m^{-1})$, (Bishop et al., 1975). Similarly $(m-p)^{-1}\bar{y}'A_2(\delta)\bar{y} = (m-p)^{-1}E(\bar{y}'A_2(\delta)\bar{y}) + O_p(m^{-1/2}) = 1 + O_p(m^{-1/2})$, since $Var[(m-p)^{-1}\bar{y}'A_2\bar{y}] = O(m^{-1})$. That means, $(n-p)^{-1}y'A_1(\delta)y - 1$ and $(m-p)^{-1}\bar{y}'A_2(\delta)\bar{y} - 1$ are order $O_p(m^{-1/2})$.

Moreover, we know that $V_i = Var(y_i) = \sigma_v^2 J_{n_i} + \sigma_u^2 I_{n_i} + R_i$, where $V = diag(V_1, V_2, \dots, V_m)$ such that $\frac{dV}{d\sigma_v^2} = diag(J_{n_1}, \dots, J_{n_m}) = \tilde{J}_n$, say, and J_{n_i} is a matrix of ones with dimension $n_i \times n_i$ and $\frac{dV}{d\sigma_u^2} = diag(I_{n_1}, \dots, I_{n_m}) = I_n$. Hence, we can calculate $A_1^1(\delta)$ as

$$\begin{aligned} A_1^1(\delta) &= -V^{-1}\tilde{J}_nV^{-1} + V^{-1}\tilde{J}_nV^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} \\ &\quad \cdot \tilde{J}_nV^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} + V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}\tilde{J}_nV^{-1} \\ &= -V^{-1}\tilde{J}_nV^{-1} + V^{-1}\tilde{J}_n(V^{-1} - A_1) - (V^{-1} - A_1)\tilde{J}_n(V^{-1} - A_1) + (V^{-1} - A_1)\tilde{J}_nV^{-1} \\ &= V^{-1}\tilde{J}_nV^{-1} - V^{-1}\tilde{J}_nA_1 - V^{-1}\tilde{J}_nV^{-1} + V^{-1}\tilde{J}_nA_1 + A_1\tilde{J}_nV^{-1} - A_1\tilde{J}_nA_1 - A_1\tilde{J}_nV^{-1} \\ &= -A_1\tilde{J}_nA_1. \end{aligned}$$

Then, $y'A_1^1(\delta)y = -y'A_1(\delta)\tilde{J}_nA_1(\delta)y$.

In addition,

$$\begin{aligned} E(y'A_1^1(\delta)y) &= -E(y'A_1(\delta)\tilde{J}_nA_1(\delta)y) \\ &= -E\{(A_1y)'\tilde{J}_n(A_1y)\} \\ &= -tr\{\tilde{J}_nE[(A_1y)(A_1y)']\}, \end{aligned}$$

where $E[(A_1Y)(A_1Y)'] = Var(A_1Y) + E(A_1Y)E(A_1Y)'$ with $Var(A_1Y) = A_1VA_1 = A_1$ and $E(A_1Y) = E(V^{-1}y - V^{-1}X\tilde{\beta}) = V^{-1}X\beta - V^{-1}X\beta = 0$, so $E[(A_1Y)(A_1Y)'] = A_1$.

We may further write $E(y'A_1^1(\delta)y) = -tr(\tilde{J}_nA_1) \approx -tr(\tilde{J}_nV^{-1}) = -\sum_{i=1}^m tr(\tilde{J}_{n_i}V_i^{-1}) = O(m)$, then

$$(n-p)^{-1}y'A_1^1(\delta)y = -(n-p)^{-1}tr(\tilde{J}_nV^{-1}) + O_p(m^{-1/2}), \quad (4.21)$$

i.e., $(n-p)^{-1}y'A_1^1(\delta)y = O_p(1)$, (see Section 4.9.3).

Similarly, $E(y'A_1^2(\delta)y) \approx -tr(V^{-1}) = -\sum_{i=1}^m tr(V_i^{-1}) = O(m)$, then

$$(n-p)^{-1}y'A_1^2(\delta)y = -(n-p)^{-1}tr(V^{-1}) + O_p(m^{-1/2}), \quad (4.22)$$

i.e., $(n-p)^{-1}y'A_1^2(\delta)y = O_p(1)$, (see Section 4.9.4).

Moreover, we know that $V_i^* = \sigma_\nu^2 + \frac{\sigma_u^2}{n_i} + \frac{\sigma_{\epsilon_i}^2}{n_i}$, where $V^* = diag(V_1^*, \dots, V_m^*)$ such that $\frac{dV^*}{d\sigma_u^2} = I_m$ and $\frac{dV^*}{d\sigma_{\epsilon_i}^2} = diag(n_1^{-1}, \dots, n_m^{-1}) = \tilde{N}^{-1}$, say. As before, we may write $E(\bar{y}'A_2^1(\delta)\bar{y}) \approx -tr(V^{*-1}) = -\sum_{i=1}^m tr(V_i^{*-1}) = O(m)$, then

$$(m-p)^{-1}\bar{y}'A_2^1(\delta)\bar{y} = -(m-p)^{-1}tr(V^{*-1}) + O_p(m^{-1/2}), \quad (4.23)$$

i.e., $(m-p)^{-1}\bar{y}'A_2^1(\delta)\bar{y} = O_p(1)$, (see Section 4.9.5).

Similarly, $E(\bar{y}'A_2^2(\delta)\bar{y}) \approx -tr(\tilde{N}^{-1}V^{*-1}) = -\sum_{i=1}^m tr(\tilde{N}_i^{-1}V_i^{*-1}) = O(m)$, then

$$(m-p)^{-1}\bar{y}'A_2^2(\delta)\bar{y} = -(m-p)^{-1}tr(\tilde{N}^{-1}V^{*-1}) + O_p(m^{-1/2}), \quad (4.24)$$

i.e., $(m-p)^{-1}\bar{y}'A_2^2(\delta)\bar{y} = O_p(1)$, (see Section 4.9.6).

Hence, it follows from (4.20) that $\tilde{\sigma}_\nu^2 - \sigma_\nu^2 = O_p(m^{-1/2})$ and $\tilde{\sigma}_u^2 - \sigma_u^2 = O_p(m^{-1/2})$.

Therefore, by (4.21)-(4.24), we may write (4.20) as

$$\begin{pmatrix} \tilde{\sigma}_\nu^2 - \sigma_\nu^2 \\ \tilde{\sigma}_u^2 - \sigma_u^2 \end{pmatrix} = - \begin{pmatrix} \text{tr}(\tilde{J}_n V^{-1})/(n-p) & \text{tr}(V^{-1})/(n-p) \\ \text{tr}(V^{*-1})/(m-p) & \text{tr}(\tilde{N}^{-1} V^{*-1})/(m-p) \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 - \frac{\mathbf{y}' A_1(\delta) \mathbf{y}}{(n-p)} \\ 1 - \frac{\bar{\mathbf{y}}' A_2(\delta) \bar{\mathbf{y}}}{(m-p)} \end{pmatrix} + [O_p(m^{-1})]_{2 \times 1}. \quad (4.25)$$

Then, the asymptotic variance-covariance of $\tilde{\delta}$ is given by

$$\begin{pmatrix} \text{Var}(\tilde{\sigma}_\nu^2) & \text{Cov}(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \\ \text{Cov}(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) & \text{Var}(\tilde{\sigma}_u^2) \end{pmatrix} = \begin{pmatrix} \frac{\text{tr}(\tilde{J}_n V^{-1})}{n-p} & \frac{\text{tr}(V^{-1})}{n-p} \\ \frac{\text{tr}(V^{*-1})}{m-p} & \frac{\text{tr}(\tilde{N}^{-1} V^{*-1})}{m-p} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \text{Var}(\frac{\mathbf{y}' A_1(\delta) \mathbf{y}}{n-p}) & \text{Cov}(\frac{\mathbf{y}' A_1(\delta) \mathbf{y}}{n-p}, \frac{\bar{\mathbf{y}}' A_2(\delta) \bar{\mathbf{y}}}{m-p}) \\ \text{Cov}(\frac{\mathbf{y}' A_1(\delta) \mathbf{y}}{n-p}, \frac{\bar{\mathbf{y}}' A_2(\delta) \bar{\mathbf{y}}}{m-p}) & \text{Var}(\frac{\bar{\mathbf{y}}' A_2(\delta) \bar{\mathbf{y}}}{m-p}) \end{pmatrix} \begin{pmatrix} \frac{\text{tr}(\tilde{J}_n V^{-1})}{n-p} & \frac{\text{tr}(V^{-1})}{n-p} \\ \frac{\text{tr}(V^{*-1})}{m-p} & \frac{\text{tr}(\tilde{N}^{-1} V^{*-1})}{m-p} \end{pmatrix}^{-T},$$

where

$$\text{Var}(\mathbf{y}' A_1(\delta) \mathbf{y} / (n-p)) = 2(n-p)^{-1}, \quad (4.26)$$

$$\text{Var}(\bar{\mathbf{y}}' A_2(\delta) \bar{\mathbf{y}} / (m-p)) = 2(m-p)^{-1}, \quad (4.27)$$

$$\begin{aligned} \text{Cov}(\mathbf{y}' A_1(\delta) \mathbf{y}, \bar{\mathbf{y}}' A_2(\delta) \bar{\mathbf{y}}) &= \text{Cov}(\mathbf{y}' A_1(\delta) \mathbf{y}, \mathbf{y}' \tilde{N}_1^{-T} A_2(\delta) \tilde{N}_1^{-1} \mathbf{y}) \\ &= 2\text{tr}(A_1(\delta) V \tilde{N}_1^{-T} A_2(\delta) \tilde{N}_1^{-1} V), \end{aligned} \quad (4.28)$$

and $\tilde{N}_1^{-1} \mathbf{y} = \bar{\mathbf{y}}$ with

$$\tilde{N}_1^{-1} = \begin{pmatrix} n_1^{-1} \mathbf{1}_{n_1} & \mathbf{0}_{1n_1} & \dots & \mathbf{0}_{1n_1} \\ \mathbf{0}_{1n_2} & n_2^{-1} \mathbf{1}_{n_2} & \dots & \mathbf{0}_{1n_2} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}_{1n_m} & \mathbf{0}_{1n_m} & \dots & n_m^{-1} \mathbf{1}_{n_m} \end{pmatrix}'.$$

In fact, to get (4.26)- (4.28) we used the following Lemma.

Lemma 4.2. (Searle, 1971): If $x \sim N(\mu, V)$ where x is a vector, $x'Px$ and $x'Qx$ are two quadratic forms such that P and Q are square matrices, then we have that

$$\text{Cov}(x'Px, x'Qx) = 2\text{tr}(PVQV) + 4\mu'PVQ\mu.$$

Hence, combining (4.26)-(4.28), the asymptotic variance-covariance of $\tilde{\delta}$ is well-defined. To further simplify, we first note that

$$V_i^{-1} = \text{diag}_j((\sigma_u^2 + \sigma_{ej}^2)^{-1}) - \frac{\sigma_\nu^2 \sigma_u^2}{\sigma_u^2 + \sigma_\nu^2 \gamma_i} w_i w_i'.$$

Hence,

$$\begin{aligned} \text{tr}(V^{-1}) &= \sum_{i=1}^m \sum_{j=1}^{n_i} (\sigma_u^2 + \sigma_{ej}^2)^{-1} - \sum_{i=1}^m \left[\frac{\sigma_\nu^2}{1 + \sigma_\nu^2 \sum_{j=1}^{n_i} (\sigma_u^2 + \sigma_{ej}^2)^{-1}} \sum_{j=1}^{n_i} (\sigma_u^2 + \sigma_{ej}^2)^{-2} \right] \\ &= \sigma_u^{-2} \sum_{i=1}^m \gamma_i - \sum_{i=1}^m \left[\frac{\sigma_\nu^2 \sigma_u^{-2}}{\sigma_u^2 + \sigma_\nu^2 \gamma_i} \sum_{j=1}^{n_i} \gamma_{ij}^2 \right] = \frac{1}{\sigma_u^2} \sum_{i=1}^m [\gamma_i - \frac{\gamma_i}{\gamma_i} \sum_{j=1}^{n_i} \gamma_{ij}^2]. \end{aligned} \quad (4.29)$$

In addition, we may write that

$$\begin{aligned} \text{tr}(\tilde{J}_n V^{-1}) &= \sum_{i=1}^m \sum_{j=1}^{n_i} (\sigma_u^2 + \sigma_{ej}^2)^{-1} - \sum_{i=1}^m \left[\frac{\sigma_u^2 \sigma_\nu^2}{\sigma_u^2 + \sigma_\nu^2 \gamma_i} \left(\sum_{j=1}^{n_i} (\sigma_u^2 + \sigma_{ej}^2)^{-1} \right)^2 \right] \\ &= \sum_{i=1}^m \sum_{j=1}^{n_i} (\sigma_u^2 + \sigma_{ej}^2)^{-1} - \sum_{i=1}^m \left(\frac{\sigma_u^2 \sigma_\nu^2}{\sigma_u^2 + \sigma_\nu^2 \gamma_i} \sigma_u^{-4} \gamma_i^2 \right). \end{aligned}$$

Hence,

$$\text{tr}(\tilde{J}_n V^{-1}) = \sigma_u^{-2} \sum_{i=1}^m \gamma_i - \frac{\sigma_\nu^2}{\sigma_u^2} \sum_{i=1}^m \left(\frac{\gamma_i^2}{\sigma_u^2 + \sigma_\nu^2 \gamma_i} \right) = \frac{1}{\sigma_u^2} \sum_{i=1}^m \gamma_i (1 - \gamma_i). \quad (4.30)$$

Furthermore, we have that

$$\text{tr}(V^{*-1}) = \sum_{i=1}^m \left(\sigma_\nu^2 + \frac{\sigma_u^2}{n_i} + \frac{\sigma_{ei}^2}{n_i} \right)^{-1} = \sum_{i=1}^m n_i (n_i \sigma_\nu^2 + \sigma_u^2 + \sigma_{ei}^2)^{-1}. \quad (4.31)$$

Also,

$$\begin{aligned} \text{tr}(\tilde{N}^{-1}V^{*-1}) &= \text{tr}(\text{diag}(n_1^{-1}V_1^{*-1}, \dots, n_m^{-1}V_m^{*-1})) = \sum_{i=1}^m n_i^{-1}(\sigma_\nu^2 + \frac{\sigma_u^2}{n_i} + \frac{\sigma_{ei}^2}{n_i})^{-1} \\ &= \sum_{i=1}^m (n_i\sigma_\nu^2 + \sigma_u^2 + \sigma_{ei}^2)^{-1}. \end{aligned} \quad (4.32)$$

Therefore, we may write the variances and the covariance of $\tilde{\sigma}_\nu^2$ and $\tilde{\sigma}_u^2$ as

$$\begin{aligned} \text{Var}(\tilde{\sigma}_\nu^2) &= 2 \left[\text{tr}(\tilde{J}_n V^{-1}) \text{tr}(\tilde{N}^{-1}V^{*-1}) - \text{tr}(V^{*-1}) \text{tr}(V^{-1}) \right]^{-2} \left[n(\text{tr}(\tilde{N}^{-1}V^{*-1}))^2 \right. \\ &\quad \left. + m(\text{tr}(V^{-1}))^2 - 2\text{tr}(V^{-1}) \text{tr}(\tilde{N}^{-1}V^{*-1}) \text{tr}(A_1 V \tilde{N}_1^{-T} A_2 \tilde{N}_1^{-1} V) \right] + o(m^{-1}), \end{aligned}$$

$$\begin{aligned} \text{Var}(\tilde{\sigma}_u^2) &= 2 \left[\text{tr}(\tilde{J}_n V^{-1}) \text{tr}(\tilde{N}^{-1}V^{*-1}) - \text{tr}(V^{*-1}) \text{tr}(V^{-1}) \right]^{-2} \left[n(\text{tr}(V^{*-1}))^2 \right. \\ &\quad \left. + m(\text{tr}(\tilde{J}_n V^{-1}))^2 - 2\text{tr}(V^{*-1}) \text{tr}(\tilde{J}_n V^{-1}) \text{tr}(A_1 V \tilde{N}_1^{-T} A_2 \tilde{N}_1^{-1} V) \right] + o(m^{-1}), \end{aligned}$$

and

$$\begin{aligned} \text{Cov}(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) &= 2 \left[\text{tr}(\tilde{J}_n V^{-1}) \text{tr}(\tilde{N}^{-1}V^{*-1}) - \text{tr}(V^{*-1}) \text{tr}(V^{-1}) \right]^{-2} \left[-n \text{tr}(V^{*-1}) \right. \\ &\quad \left. \cdot \text{tr}(\tilde{N}^{-1}V^{*-1}) - m \text{tr}(V^{-1}) \text{tr}(\tilde{J}_n V^{-1}) + \left(\text{tr}(V^{-1}) \text{tr}(V^{*-1}) + \text{tr}(\tilde{N}^{-1}V^{*-1}) \text{tr}(\tilde{J}_n V^{-1}) \right) \right. \\ &\quad \left. \cdot \text{tr}(A_1 V \tilde{N}_1^{-T} A_2 \tilde{N}_1^{-1} V) \right] + o(m^{-1}), \end{aligned}$$

where $\text{tr}(V^{-1})$, $\text{tr}(\tilde{J}_n V^{-1})$, $\text{tr}(V^{*-1})$ and $\text{tr}(\tilde{N}^{-1}V^{*-1})$ are given by (4.29)-(4.32) respectively.

4.5 MSE Estimation

Using the second order MSE approximations (4.15) and (4.18), we now obtain MSE estimators that are approximately unbiased in the sense that

$$E[mse(\hat{\mu}_{ij})] - MSE(\hat{\mu}_{ij}) = o(m^{-1}),$$

and

$$E[mse(\hat{\mu}_i)] - MSE(\hat{\mu}_i) = o(m^{-1}).$$

Since $g_{2ij}(\delta)$ and $g_{3ij}(\delta)$ are of order $O(m^{-1})$, it follows that

$$E[g_{2ij}(\hat{\delta})] = g_{2ij}(\delta) + o(m^{-1}),$$

$$E[g_{3ij}(\hat{\delta})] = g_{3ij}(\delta) + o(m^{-1}).$$

Therefore $g_{2ij}(\hat{\delta})$ and $g_{3ij}(\hat{\delta})$ are the correct estimators of $g_{2ij}(\delta)$ and $g_{3ij}(\delta)$ respectively, each with bias $o(m^{-1})$. However, $g_{1ij}(\delta)$ is of order $O(1)$, so $g_{1ij}(\hat{\delta})$ is not the correct estimator of $g_{1ij}(\delta)$, since the bias is of order $O(m^{-1})$. Thus, the bias in $g_{1ij}(\delta)$ must be estimated to the correct order. Following Prasad and Rao (1990) and Datta et al. (2005), we can write

$$E[g_{1ij}(\hat{\delta}) - (\hat{\delta} - \delta)' \nabla g_{1ij}(\delta)|_{\delta=\hat{\delta}} + g_{3ij}(\hat{\delta})] = g_{1ij}(\delta) + o(m^{-1}).$$

Therefore, combining above results, a second order unbiased estimator of $MSE(\hat{\mu}_{ij}(\hat{\delta}))$ is given by

$$mse(\hat{\mu}_{ij}(\hat{\delta})) = g_{1ij}(\hat{\delta}) + g_{2ij}(\hat{\delta}) + 2g_{3ij}(\hat{\delta}) - \frac{dg_{1ij}(\delta)}{d\sigma_\nu^2} \Big|_{\delta=\hat{\delta}} \hat{b}(\hat{\sigma}_\nu^2) - \frac{dg_{1ij}(\delta)}{d\sigma_u^2} \Big|_{\delta=\hat{\delta}} \hat{b}(\hat{\sigma}_u^2), \quad (4.33)$$

where $\hat{b}(\hat{\sigma}_\nu^2)$ and $\hat{b}(\hat{\sigma}_u^2)$ are estimators of biases $b(\hat{\sigma}_\nu^2) = E(\hat{\sigma}_\nu^2 - \sigma_\nu^2)$ and $b(\hat{\sigma}_u^2) = E(\hat{\sigma}_u^2 - \sigma_u^2)$. Section 4.6 gives expressions for $b(\hat{\sigma}_\nu^2)$ and $b(\hat{\sigma}_u^2)$.

Similarly,

$$mse(\hat{\mu}_i(\hat{\delta})) = g_{1i}(\hat{\delta}) + g_{2i}(\hat{\delta}) + 2g_{3i}(\hat{\delta}) - \frac{dg_{1i}(\delta)}{d\sigma_\nu^2}|_{\delta=\hat{\delta}}\hat{b}(\hat{\sigma}_\nu^2) - \frac{dg_{1i}(\delta)}{d\sigma_u^2}|_{\delta=\hat{\delta}}\hat{b}(\hat{\sigma}_u^2). \quad (4.34)$$

We now find $\frac{dg_{1ij}(\delta)}{d\sigma_\nu^2}$ and $\frac{dg_{1ij}(\delta)}{d\sigma_u^2}$ in (4.33). Recall that $g_{1ij}(\delta) = (1 - \gamma_{ij})^2(\sigma_u^2 + \sigma_\nu^2(1 - \gamma_i)) + \gamma_{ij}^2\sigma_{eij}^2$, so that

$$\frac{dg_{1ij}(\delta)}{d\sigma_\nu^2} = (1 - \gamma_{ij})^2(1 - \gamma_i - \frac{\gamma_i^2\sigma_u^2}{\gamma_i\sigma_\nu^2}),$$

and that

$$\begin{aligned} \frac{dg_{1ij}(\delta)}{d\sigma_u^2} &= -2(1 - \gamma_{ij})^2\frac{\gamma_{ij}}{\sigma_u^2}[\sigma_u^2 + (1 - \gamma_i)\sigma_\nu^2] + (1 - \gamma_{ij})^2[1 + \frac{\gamma_i^2}{\gamma_i^2} \sum_{j=1}^{n_i} \gamma_{ij}^2] + \frac{2\gamma_{ij}^2(1 - \gamma_{ij})\sigma_{eij}^2}{\sigma_u^2} \\ &= (1 - \gamma_{ij}) \left[\frac{-2\gamma_{ij}(1 - \gamma_{ij})}{\sigma_u^2}(\sigma_u^2 + (1 - \gamma_i)\sigma_\nu^2) + (1 - \gamma_{ij})(1 + \frac{\gamma_i^2 \sum_{j=1}^{n_i} \gamma_{ij}^2}{\gamma_i^2}) + \frac{2\gamma_{ij}^2\sigma_{eij}^2}{\sigma_u^2} \right]. \end{aligned}$$

Combining above results, we obtain $mse(\hat{\mu}_{ij}(\hat{\delta}))$ from (4.33).

Theorem 4.1. *The second order unbiased estimator of the $MSE(\hat{\mu}_{ij}(\hat{\delta}))$ is given by $mse(\hat{\mu}_{ij}(\hat{\delta}))$ with the property $E[mse(\hat{\mu}_{ij}(\hat{\delta}))] = MSE(\hat{\mu}_{ij}(\hat{\delta})) + o(m^{-1})$.*

We now find $\frac{dg_{1i}(\delta)}{d\sigma_\nu^2}$ and $\frac{dg_{1i}(\delta)}{d\sigma_u^2}$ in (4.34). Recall that $g_{1i}(\delta) = \gamma_i \frac{\sigma_u^2}{\gamma_i}$, so that

$$\frac{dg_{1i}(\delta)}{d\sigma_\nu^2} = \frac{\gamma_i^2\sigma_u^4}{\gamma_i^2\sigma_\nu^4}, \quad (4.35)$$

and that

$$\frac{dg_{1i}(\delta)}{d\sigma_u^2} = \frac{\gamma_i \sum_{j=1}^{n_i} \gamma_{ij}^2}{\gamma_i^3\sigma_\nu^2}(\gamma_i\sigma_\nu^2 - \gamma_i\sigma_u^2) = (\frac{\gamma_i}{\gamma_i})^2 \sum_{j=1}^{n_i} \gamma_{ij}^2. \quad (4.36)$$

Using (4.35) and (4.36) and the two bias expressions $b(\hat{\sigma}_\nu^2)$ and $b(\hat{\sigma}_u^2)$ in (4.34), we obtain $mse(\hat{\mu}_i(\hat{\delta}))$.

Theorem 4.2. *The second order unbiased estimator of the MSE of $\hat{\mu}_i(\hat{\delta})$ is given by $mse(\hat{\mu}_i(\hat{\delta}))$ with the property $E[mse(\hat{\mu}_i(\hat{\delta}))] = MSE(\hat{\mu}_i(\hat{\delta})) + o(m^{-1})$.*

4.6 Bias of $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_u^2$

We now find expressions for bias terms $b(\hat{\sigma}_\nu^2)$ and $b(\hat{\sigma}_u^2)$. We have $b(\hat{\delta}_i) = E(\hat{\delta}_i - \delta_i) = E(\tilde{\delta}_i - \delta_i) + E(\hat{\delta}_i - \tilde{\delta}_i)$ for $i = 1, 2$, where $E(\hat{\delta}_i - \tilde{\delta}_i) = -E(\tilde{\delta}_i I(\tilde{\delta}_i \leq 0)) \leq (E(\tilde{\delta}_i^2))^{1/2} (Pr(\tilde{\delta}_i \leq 0))^{1/2} = O(m^{-2})$ (see the body of proof of Lemma 4.3). To get $E(\tilde{\delta}_i - \delta_i) = b(\tilde{\delta}_i)$, $E(\tilde{\sigma}_\nu^2 - \sigma_\nu^2)$ and $E(\tilde{\sigma}_u^2 - \sigma_u^2)$, we follow Datta et al. (2005) to derive the bias of $\tilde{\sigma}_\nu^2$ and $\tilde{\sigma}_u^2$. By (4.19), letting $A_k(\delta) = A_k$, $A_k^1(\delta) = A_k^1$ and $A_k^2(\delta) = A_k^2$ for $k=1,2$, we have

$$\begin{aligned} & \frac{1}{n-p} \left[y' A_1 y - (n-p) + y' A_1^1 y (\tilde{\sigma}_\nu^2 - \sigma_\nu^2) + y' A_1^2 y (\tilde{\sigma}_u^2 - \sigma_u^2) + \frac{1}{2} y' A_1^{11} y (\tilde{\sigma}_\nu^2 - \sigma_\nu^2)^2 \right. \\ & \quad \left. + \frac{1}{2} y' A_1^{22} y (\tilde{\sigma}_u^2 - \sigma_u^2)^2 + y' A_1^{12} y (\tilde{\sigma}_\nu^2 - \sigma_\nu^2)(\tilde{\sigma}_u^2 - \sigma_u^2) \right] + o_p(m^{-1}) = 0, \end{aligned}$$

and

$$\begin{aligned} & \frac{1}{m-p} \left[\bar{y}' A_2 \bar{y} - (m-p) + \bar{y}' A_2^1 \bar{y} (\tilde{\sigma}_\nu^2 - \sigma_\nu^2) + \bar{y}' A_2^2 \bar{y} (\tilde{\sigma}_u^2 - \sigma_u^2) + \frac{1}{2} \bar{y}' A_2^{11} \bar{y} (\tilde{\sigma}_\nu^2 - \sigma_\nu^2)^2 \right. \\ & \quad \left. + \frac{1}{2} \bar{y}' A_2^{22} \bar{y} (\tilde{\sigma}_u^2 - \sigma_u^2)^2 + \bar{y}' A_2^{12} \bar{y} (\tilde{\sigma}_\nu^2 - \sigma_\nu^2)(\tilde{\sigma}_u^2 - \sigma_u^2) \right] + o_p(m^{-1}) = 0, \end{aligned}$$

where $A_k^{ij}(\delta) = \frac{d^2 A_k(\delta)}{d\delta_i d\delta_j}$, $k = 1, 2$. We may rewrite the above two equations as

$$\begin{aligned} & \frac{1}{n-p} \left[y' A_1 y - (n-p) + (y' A_1^1 y + tr(\tilde{J}_n A_1))(\tilde{\sigma}_\nu^2 - \sigma_\nu^2) + (y' A_1^2 y + tr(A_1))(\tilde{\sigma}_u^2 - \sigma_u^2) \right. \\ & \quad - (\tilde{\sigma}_\nu^2 - \sigma_\nu^2) tr(\tilde{J}_n A_1) - (\tilde{\sigma}_u^2 - \sigma_u^2) tr(A_1) + \frac{1}{2} tr(V A_1^{11}) (\tilde{\sigma}_\nu^2 - \sigma_\nu^2)^2 + \frac{1}{2} tr(V A_1^{22}) (\tilde{\sigma}_u^2 - \sigma_u^2)^2 \\ & \quad \left. + tr(V A_1^{12}) (\tilde{\sigma}_\nu^2 - \sigma_\nu^2)(\tilde{\sigma}_u^2 - \sigma_u^2) \right] + o_p(m^{-1}) = 0 \end{aligned} \quad (4.37)$$

and

$$\begin{aligned} & \frac{1}{m-p} \left[\bar{y}' A_2 \bar{y} - (m-p) + (\bar{y}' A_2^1 \bar{y} + tr(A_2))(\tilde{\sigma}_\nu^2 - \sigma_\nu^2) + (\bar{y}' A_2^2 \bar{y} + tr(\tilde{N}^{-1} A_2))(\tilde{\sigma}_u^2 - \sigma_u^2) \right. \\ & \quad - (\tilde{\sigma}_\nu^2 - \sigma_\nu^2) tr(A_2) - (\tilde{\sigma}_u^2 - \sigma_u^2) tr(\tilde{N}^{-1} A_2) + \frac{1}{2} tr(V^* A_2^{11}) (\tilde{\sigma}_\nu^2 - \sigma_\nu^2)^2 + \frac{1}{2} tr(V^* A_2^{22}) (\tilde{\sigma}_u^2 - \sigma_u^2)^2 \\ & \quad \left. + tr(V^* A_2^{12}) (\tilde{\sigma}_\nu^2 - \sigma_\nu^2)(\tilde{\sigma}_u^2 - \sigma_u^2) \right] + o_p(m^{-1}) = 0 \end{aligned}$$

$$+tr(V^*A_2^{12})(\tilde{\sigma}_\nu^2 - \sigma_\nu^2)(\tilde{\sigma}_u^2 - \sigma_u^2)\Big] + o_p(m^{-1}) = 0. \quad (4.38)$$

We then take expectations of (4.37) and (4.38) to get

$$\begin{aligned} & \frac{1}{n-p} \left[Cov(y'A_1^1y, \tilde{\sigma}_\nu^2) + Cov(y'A_1^2y, \tilde{\sigma}_u^2) - tr(\tilde{J}_n A_1)b(\tilde{\sigma}_\nu^2) - tr(A_1)b(\tilde{\sigma}_u^2) \right. \\ & \left. + \frac{1}{2}tr(VA_1^{11})Var(\tilde{\sigma}_\nu^2) + \frac{1}{2}tr(VA_1^{22})Var(\tilde{\sigma}_u^2) + tr(VA_1^{12})Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \right] + o(m^{-1}) = 0, \end{aligned} \quad (4.39)$$

and

$$\begin{aligned} & \frac{1}{m-p} \left[Cov(\bar{y}'A_2^1\bar{y}, \tilde{\sigma}_\nu^2) + Cov(\bar{y}'A_2^2\bar{y}, \tilde{\sigma}_u^2) - tr(A_2)b(\tilde{\sigma}_\nu^2) - tr(\tilde{N}^{-1}A_2)b(\tilde{\sigma}_u^2) \right. \\ & \left. + \frac{1}{2}tr(V^*A_2^{11})Var(\tilde{\sigma}_\nu^2) + \frac{1}{2}tr(V^*A_2^{22})Var(\tilde{\sigma}_u^2) + tr(V^*A_2^{12})Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \right] + o(m^{-1}) = 0. \end{aligned} \quad (4.40)$$

Therefore, we can find approximations to $b(\tilde{\sigma}_\nu^2)$ and $b(\tilde{\sigma}_u^2)$ from (4.39) and (4.40) as

$$\begin{aligned} b(\tilde{\sigma}_\nu^2) = & \left[\frac{tr(A_1)tr(A_2)}{tr(\tilde{N}^{-1}A_2)} - tr(\tilde{J}_n A_1) \right]^{-1} \left\{ -Cov(y'A_1^1y, \tilde{\sigma}_\nu^2) - Cov(y'A_1^2y, \tilde{\sigma}_u^2) \right. \\ & - tr(A_1\tilde{J}_n A_1\tilde{J}_n)Var(\tilde{\sigma}_\nu^2) - tr(A_1A_1)Var(\tilde{\sigma}_u^2) - 2tr(A_1A_1\tilde{J}_n)Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \\ & + \frac{tr(A_1)}{tr(\tilde{N}^{-1}A_2)} \left[Cov(\bar{y}'A_2^1\bar{y}, \tilde{\sigma}_\nu^2) + tr(A_2A_2)Var(\tilde{\sigma}_\nu^2) + Cov(\bar{y}'A_2^2\bar{y}, \tilde{\sigma}_u^2) \right. \\ & \left. \left. + tr(A_2\tilde{N}^{-1}A_2\tilde{N}^{-1})Var(\tilde{\sigma}_u^2) + 2tr(A_2A_2\tilde{N}^{-1})Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \right] \right\}, \end{aligned} \quad (4.41)$$

and

$$\begin{aligned} b(\tilde{\sigma}_u^2) = & \left[\frac{tr(A_1)tr(A_2)}{tr(\tilde{J}_n A_1)} - tr(\tilde{N}^{-1}A_2) \right]^{-1} \left\{ -Cov(\bar{y}'A_2^1\bar{y}, \tilde{\sigma}_\nu^2) - Cov(\bar{y}'A_2^2\bar{y}, \tilde{\sigma}_u^2) \right. \\ & - tr(A_2A_2)Var(\tilde{\sigma}_\nu^2) - tr(A_2\tilde{N}^{-1}A_2\tilde{N}^{-1})Var(\tilde{\sigma}_u^2) - 2tr(A_2A_2\tilde{N}^{-1})Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \\ & + \frac{tr(A_2)}{tr(\tilde{J}_n A_1)} \left[Cov(y'A_1^1y, \tilde{\sigma}_\nu^2) + tr(A_1\tilde{J}_n A_1\tilde{J}_n)Var(\tilde{\sigma}_\nu^2) + Cov(y'A_1^2y, \tilde{\sigma}_u^2) \right. \end{aligned}$$

$$+tr(A_1 A_1)Var(\tilde{\sigma}_u^2) + 2tr(A_1 A_1 \tilde{J}_n)Cov(\tilde{\sigma}_\nu^2, \tilde{\sigma}_u^2) \Big] \Big\}, \quad (4.42)$$

where $Cov(y' A_1^1 y, \tilde{\sigma}_\nu^2)$, $Cov(y' A_1^2 y, \tilde{\sigma}_u^2)$, $Cov(\bar{y}' A_2^1 \bar{y}, \tilde{\sigma}_\nu^2)$ and $Cov(\bar{y}' A_2^2 \bar{y}, \tilde{\sigma}_u^2)$ are calculated below.

By (4.25), we may write

$$\tilde{\sigma}_\nu^2 - \sigma_\nu^2 = \frac{tr(\tilde{N}^{-1} V^{*-1})(y' A_1 y - (n - p)) - tr(V^{-1})(\bar{y}' A_2 \bar{y} - (m - p))}{tr(\tilde{J}_n V^{-1})tr(\tilde{N}^{-1} V^{*-1}) - tr(V^{-1})tr(V^{*-1})} + O_p(m^{-1}).$$

Now using Lemma 4.2, we get

$$Cov(y' A_1^1 y, \tilde{\sigma}_\nu^2) = 2 \left[tr(\tilde{J}_n V^{-1})tr(\tilde{N}^{-1} V^{*-1}) - tr(V^{-1})tr(V^{*-1}) \right]^{-1} \\ \cdot \left[-tr(\tilde{N}^{-1} V^{*-1})tr(A_1 \tilde{J}_n) - tr(V^{-1})tr(A_1^1 V \tilde{N}_1^{-T} A_2 \tilde{N}_1^{-1} V) \right] + o(m^{-1}),$$

and

$$Cov(\bar{y}' A_2^1 \bar{y}, \tilde{\sigma}_\nu^2) = 2 \left[tr(\tilde{J}_n V^{-1})tr(\tilde{N}^{-1} V^{*-1}) - tr(V^{-1})tr(V^{*-1}) \right]^{-1} \left[tr(\tilde{N}^{-1} V^{*-1}) \right. \\ \left. \cdot tr(\tilde{N}_1^{-T} A_2^1 \tilde{N}_1^{-1} V A_1 V) - tr(V^{-1})tr(\tilde{N}_1^{-T} A_2^1 \tilde{N}_1^{-1} V \tilde{N}_1^{-T} A_2 \tilde{N}_1^{-1} V) \right] + o(m^{-1}).$$

Similarly, from (4.25), we have

$$\tilde{\sigma}_u^2 - \sigma_u^2 = \frac{-tr(V^{*-1})(y' A_1 y - (n - p)) + tr(\tilde{J}_n V^{-1})(\bar{y}' A_2 \bar{y} - (m - p))}{tr(\tilde{J}_n V^{-1})tr(\tilde{N}^{-1} V^{*-1}) - tr(V^{-1})tr(V^{*-1})} + O_p(m^{-1}).$$

Again using Lemma 4.2, we have

$$Cov(y' A_1^2 y, \tilde{\sigma}_u^2) = 2 \left[tr(\tilde{J}_n V^{-1})tr(\tilde{N}^{-1} V^{*-1}) - tr(V^{-1})tr(V^{*-1}) \right]^{-1} \\ \cdot \left[tr(V^{*-1})tr(A_1) + tr(\tilde{J}_n V^{-1})tr(A_1^2 V \tilde{N}_1^{-T} A_2 \tilde{N}_1^{-1} V) \right] + o(m^{-1}),$$

and

$$Cov(\bar{y}' A_2^2 \bar{y}, \tilde{\sigma}_u^2) = 2 \left[tr(\tilde{J}_n V^{-1})tr(\tilde{N}^{-1} V^{*-1}) - tr(V^{-1})tr(V^{*-1}) \right]^{-1} \left[-tr(V^{*-1}) \right. \\ \left. \cdot tr(\tilde{N}_1^{-T} A_2^2 \tilde{N}_1^{-1} V A_1 V) + tr(\tilde{J}_n V^{-1})tr(\tilde{N}_1^{-T} A_2^2 \tilde{N}_1^{-1} V \tilde{N}_1^{-T} A_2 \tilde{N}_1^{-1} V) \right] + o(m^{-1}).$$

Finally, combining above results, we obtain bias approximations $b(\hat{\sigma}_\nu^2)$ and $b(\hat{\sigma}_u^2)$ from (4.41) and (4.42) respectively.

4.7 Simulation Study

A small simulation study was undertaken in order to investigate the performance of the proposed estimators under the sub-area model. For simplicity, we considered the simple balanced, sub-area model with a common mean and equal error variances $\sigma_{eij}^2 = \sigma_e^2$. It is given by

$$y_{ij} = \mu + \nu_i + u_{ij} + e_{ij}; j = 1, \dots, \bar{n}; i = 1, \dots, m. \quad (4.43)$$

Model (4.43) is a special case of the sub-area level small area model (4.2) with $x_{ij} = 1, \beta = \mu, p = 1, n_i = \bar{n}$.

In our simulation study, we used $\bar{n} = 3$ sample sub-areas in each area, $m = 60$ small areas and consequently $n = 180$ in each sample. Without loss of generality, we set $\mu = 0$. We fixed $\sigma_{eij}^2 = \sigma_e^2 = 200$ as variance of sampling error, while σ_ν^2 and σ_u^2 took values 40, 100, 140, 200 and 400, leading to 25 combinations of $(\sigma_\nu^2, \sigma_u^2)$.

We generated $B = 500$ independent sets of normally distributed random variables $\{\nu_i; i = 1, \dots, 60\}, \{u_{ij}; j = 1, 2, 3; i = 1, \dots, 60\}, \{e_{ij}; j = 1, 2, 3; i = 1, \dots, 60\}$ having mean zero and specified variances σ_ν^2, σ_u^2 and σ_e^2 . From these generated data sets the observations $\{y_{ij}; j = 1, 2, 3; i = 1, \dots, 60\}$ were obtained using the model $y_{ij} = \nu_i + u_{ij} + e_{ij}$. By using the generated samples $y_{ij}^{(b)}, b=1, \dots, B=500$, we calculated $\hat{\sigma}_\nu^{2(b)}$ and $\hat{\sigma}_u^{2(b)}$ by weighted residual sum of squares iteratively as described in Section 4.2. For each generated sample, we calculated

$$\mu_{ij}^{(b)} = \nu_i^{(b)} + u_{ij}^{(b)}; j = 1, 2, 3; i = 1, \dots, 60; b = 1, \dots, B,$$

$$\mu_i^{(b)} = \nu_i^{(b)}; i = 1, \dots, 60; b = 1, \dots, B.$$

We computed the EBLUP estimates $\hat{\mu}_{ij}^{(b)}$ and $\hat{\mu}_i^{(b)}$ for each generated sample b . We then calculated the average mean squared error ($\overline{MSE}$) for sub-area and area EBLUP as

$$\overline{MSE}_1 = \frac{1}{n} \sum_i \sum_j MSE_{ij} \quad \text{where} \quad MSE_{ij} = B^{-1} \sum_{b=1}^B (\hat{\mu}_{ij}^{(b)} - \mu_{ij}^{(b)})^2$$

and

$$\overline{MSE}_2 = \frac{1}{m} \sum_i MSE_i \quad \text{where} \quad MSE_i = B^{-1} \sum_{b=1}^B (\hat{\mu}_i^{(b)} - \mu_i^{(b)})^2$$

respectively. We report the results in Tables 4.1 and 4.2. In both Tables, the value of average mean squared error is increased with increasing σ_ν^2 or σ_u^2 . In addition, $\overline{MSE}_2$ is significantly smaller than $\overline{MSE}_1$, as expected.

In addition, we computed the percent relative bias of variance component estimators $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_u^2$ as $RB_\nu = 100 \left[B^{-1} \sum_{b=1}^B \hat{\sigma}_\nu^{2(b)} / \sigma_\nu^2 - 1 \right]$ and $RB_u = 100 \left[B^{-1} \sum_{b=1}^B \hat{\sigma}_u^{2(b)} / \sigma_u^2 - 1 \right]$. The results are reported in Table 4.3. As shown in Table 4.3, with increasing σ_u^2 , the relative bias (RB) of $\hat{\sigma}_u^2$ is decreased. However, when σ_ν^2 is increased, the RB of $\hat{\sigma}_\nu^2$ is increased for $\sigma_\nu^2 \geq \sigma_u^2$ while the RB of $\hat{\sigma}_u^2$ remains constant. Although the RB of $\hat{\sigma}_\nu^2$ is increased for $\sigma_\nu^2 \geq \sigma_u^2$, but it is still negligible.

We also calculated the percent average relative bias of the naive MSE estimator of sub-area mean and naive MSE estimator of area mean as

$$\overline{AB}_{1.N} = 100 \left[\frac{1}{n} \sum_i \sum_j RB_{1ij.N} \right] \quad \text{where} \quad RB_{1ij.N} = B^{-1} \sum_{b=1}^B mse_{ij.N}^{(b)} / MSE_{ij} - 1$$

and

$$\overline{AB}_{2.N} = 100 \left[\frac{1}{m} \sum_i RB_{1i.N} \right] \quad \text{where} \quad RB_{1i.N} = B^{-1} \sum_{b=1}^B mse_{i.N}^{(b)} / MSE_i - 1,$$

where the naive MSE estimator of sub-area mean is $mse_{ij.N}^{(b)} = g_{1ij}^{(b)}(\hat{\delta}) + g_{2ij}^{(b)}(\hat{\delta})$ and the naive MSE estimator of area mean is $mse_{i.N}^{(b)} = g_{1i}^{(b)}(\hat{\delta}) + g_{2i}^{(b)}(\hat{\delta})$. We report the results

in Tables 4.4 and 4.5. As expected, the naive estimators, $mse_{ij.N}^{(b)}$ and $mse_{i.N}^{(b)}$, lead to underestimation. For sub-area level as shown in Table 4.4, the underestimation is not significant ($< 5\%$) when $\sigma_u^2 \geq 100$ and $\sigma_v^2 \geq 100$. However, for area level as shown in Table 4.5, there is significant underestimation in some regions. For example, $RB = -19.0\%$ if $\sigma_v^2 = 40$ and $\sigma_u^2 = 140$.

Similarly, we computed the percent average relative bias of the second-order correct MSE estimator of sub-area mean and the second-order correct MSE estimator of area mean as

$$\overline{AB}_1 = 100 \left[\frac{1}{n} \sum_i \sum_j RB_{2ij} \right] \quad \text{where} \quad RB_{2ij} = B^{-1} \sum_{b=1}^B mse_{ij}^{(b)} / MSE_{ij} - 1$$

and

$$\overline{AB}_2 = 100 \left[\frac{1}{m} \sum_i RB_{2i} \right] \quad \text{where} \quad RB_{2i} = B^{-1} \sum_{b=1}^B mse_i^{(b)} / MSE_i - 1$$

where $mse_{ij}^{(b)}$ and $mse_i^{(b)}$ are given by (4.33) and (4.34) for each generated sample. We report the results in Tables 4.6 and 4.7. Tables 4.6 and 4.7 show that $mse(\hat{\mu}_{ij})$ and $mse(\hat{\mu}_i)$ generally overestimate the true $MSE(\hat{\mu}_{ij})$ and $MSE(\hat{\mu}_i)$ respectively. For sub-area level as shown in Table 4.6, for a given value of σ_v^2 , RB decreases as σ_u^2 increases, while RB generally increases with $\sigma_v^2 \geq 140$ for a given σ_u^2 . The overestimation is quite small ($\leq 3\%$) for all combinations (σ_v^2, σ_u^2) .

On the other hand, for area level as shown in Table 4.7, for a given value of $\sigma_v^2 \geq 100$, RB increases as $\sigma_v^2 \geq \sigma_u^2$. The overestimation is small ($\leq 4\%$) in a wide region of practical interest, where $\sigma_v^2 \geq 100$. It may be noted that the Monte Carlo probability of getting a negative $\tilde{\sigma}_v^2$ is very small ($\leq 4\%$) in this region as shown in Table 4.8. As we can see from Tables 4.6 and 4.7, the behavior of relative bias under sub-area and area models depend on both σ_v^2 and σ_u^2 for a given σ_e^2 .

Furthermore, to evaluate the efficiency of sub-area and area EBLUP relative to direct estimators, we computed the percent average relative efficiency (EFF) of sub-area EBLUP over sub-area direct estimator and area EBLUP over area direct estimator as

$$EFF_1 = 100(\overline{MSE}_{1.dir}/\overline{MSE}_1)^{1/2} \quad \text{and} \quad EFF_2 = 100(\overline{MSE}_{2.dir}/\overline{MSE}_2)^{1/2}$$

respectively, where $\overline{MSE}_{1.dir} = \frac{1}{n} \sum_i \sum_j MSE_{ij.dir}$, $MSE_{ij.dir} = B^{-1} \sum_{b=1}^B (y_{ij}^{(b)} - \mu_{ij}^{(b)})^2$, $\overline{MSE}_{2.dir} = \frac{1}{m} \sum_i MSE_{i.dir}$, $MSE_{i.dir} = B^{-1} \sum_{b=1}^B (\bar{y}_i^{(b)} - \mu_i^{(b)})^2$ with $\bar{y}_i^{(b)} = n_i^{-1} \sum_j y_{ij}^{(b)}$. The results are reported in Tables 4.9 and 4.10. As expected, in Table 4.9 with increasing σ_ν^2 or σ_u^2 , the relative efficiency of sub-area EBLUP is decreased. The rate of decrease is less with increasing σ_ν^2 rather than with increasing σ_u^2 . However, as shown in Table 4.10, for a given value of σ_ν^2 , EFF_2 increases as σ_u^2 increases while EFF_2 decreases with σ_ν^2 for a given σ_u^2 .

Furthermore, if we take $\sigma_\nu^2 = 0$, our model (4.2) reduces to the Fay-Herriot area-level model. To investigate the lose of efficiency by using the Fay-Herriot model wrongly, we computed the percent average relative efficiency of sub-area EBLUP over Fay-Herriot estimator as $EFF_3 = 100(\overline{MSE}_{3.F-H}/\overline{MSE}_1)^{1/2}$ where $\overline{MSE}_{3.F-H} = \frac{1}{n} \sum_i \sum_j MSE_{ij.F-H}$, $MSE_{ij.F-H} = B^{-1} \sum_{b=1}^B (\hat{\mu}_{ij.F-H}^{(b)} - \mu_{ij}^{(b)})^2$ with $\hat{\mu}_{ij.F-H}^{(b)}$ denoting the Fay-Herriot estimator for the b -th simulated sample. As expected from Table 4.11, for a given σ_u^2 , EFF_3 increases as σ_ν^2 increases while EFF_3 decreases with σ_u^2 for a given σ_ν^2 . In fact, for a large σ_ν^2 , using the Fay-Herriot estimator leads to significant loss in efficiency.

4.8 Summary

We have studied the estimation of sub-area and area means under the sub-area level model which is the extension of the Fay-Herriot model from area level model to sub-area level. We have obtained an empirical best linear unbiased prediction (EBLUP) estimators of sub-area level and area level means by replacing the unknown variance components in the BLUP by suitable estimators, using an iterative method of estimating the model variances based on weighted residual sum of squares. We have also derived second order approximations to the mean squared error (MSE) of the model-based small area estimators and then derived MSE estimators unbiased to second order. Results of a limited simulation study based on a simple, balanced sub-area level model are also reported. As shown in simulation study, the relative bias of the MSE estimators, $mse(\hat{\mu}_{ij})$ and $mse(\hat{\mu}_i)$, depend on σ_ν^2 and σ_u^2 as well as σ_e^2 .

4.9 Proofs

4.9.1 Proof of (4.12)

We may write $g_{1ij}(\delta)$ as follows

$$\begin{aligned} g_{1ij}(\delta) &= E \left[\gamma_i(1 - \gamma_{ij})(\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\beta) + \gamma_{ij}(y_{ij} - x'_{ij}\beta) - \nu_i - u_{ij} \right]^2 \\ &= \gamma_i^2(1 - \gamma_{ij})^2[\sigma_\nu^2 + \frac{\sigma_u^2 \sum_{j=1}^{n_i} \gamma_{ij}^2}{\gamma_i^2} + \frac{\sum_{j=1}^{n_i} \gamma_{ij}^2 \sigma_{eij}^2}{\gamma_i^2}] + \gamma_{ij}^2(\sigma_\nu^2 + \sigma_u^2 + \sigma_{eij}^2) \\ &\quad + (\sigma_\nu^2 + \sigma_u^2) - 2\gamma_{ij}(\sigma_\nu^2 + \sigma_u^2) - 2\gamma_i(1 - \gamma_{ij})(\sigma_\nu^2 + \frac{\sigma_u^2 \gamma_{ij}}{\gamma_i}) \end{aligned}$$

$$\begin{aligned}
& +2\gamma_i\gamma_{ij}(1-\gamma_{ij})(\sigma_\nu^2 + \frac{\sigma_u^2\gamma_{ij}}{\gamma_i} + \frac{\gamma_{ij}\sigma_{ej}^2}{\gamma_i}) \\
& = \gamma_i^2(1-\gamma_{ij})^2(\sigma_\nu^2 + \frac{\sigma_u^2}{\gamma_i}) + \gamma_{ij}^2(\sigma_\nu^2 + \sigma_u^2 + \sigma_{ej}^2) + (\sigma_\nu^2 + \sigma_u^2)(1-2\gamma_{ij}) \\
& \quad -2\gamma_i(1-\gamma_{ij})^2(\sigma_\nu^2 + \frac{\sigma_u^2\gamma_{ij}}{\gamma_i}) + 2\frac{\gamma_i}{\gamma_i}(1-\gamma_{ij})\gamma_{ij}^2\sigma_{ej}^2 \\
& = \sigma_\nu^2\gamma_i(1-\gamma_{ij})^2 + \gamma_{ij}^2(\sigma_\nu^2 + \sigma_u^2 + \sigma_{ej}^2) + (\sigma_\nu^2 + \sigma_u^2)(1-2\gamma_{ij}) \\
& \quad -2\gamma_i(1-\gamma_{ij})^2(\sigma_\nu^2 + \frac{\sigma_u^2\gamma_{ij}}{\gamma_i}) + 2\frac{\gamma_i}{\gamma_i}(1-\gamma_{ij})\gamma_{ij}^2\sigma_{ej}^2 \\
& = \sigma_\nu^2\gamma_i(1-\gamma_{ij})^2 + (1-\gamma_{ij})^2(\sigma_\nu^2 + \sigma_u^2) + \gamma_{ij}^2\sigma_{ej}^2 \\
& \quad -2\gamma_i(1-\gamma_{ij})^2(\sigma_\nu^2 + \frac{\sigma_u^2\gamma_{ij}}{\gamma_i}) + 2\frac{\gamma_i}{\gamma_i}(1-\gamma_{ij})\gamma_{ij}^2\sigma_{ej}^2 \\
& = (1-\gamma_{ij})\left[\sigma_\nu^2\gamma_i(1-\gamma_{ij}) + (1-\gamma_{ij})(\sigma_\nu^2 + \sigma_u^2) - 2\gamma_i(1-\gamma_{ij})(\sigma_\nu^2 + \frac{\sigma_u^2\gamma_{ij}}{\gamma_i}) + 2\frac{\gamma_i\gamma_{ij}^2\sigma_{ej}^2}{\gamma_i}\right] + \gamma_{ij}^2\sigma_{ej}^2 \\
& = (1-\gamma_{ij})\left[(1-\gamma_{ij})((1+\gamma_i)\sigma_\nu^2 + \sigma_u^2 - 2\gamma_i(\sigma_\nu^2 + \frac{\sigma_u^2\gamma_{ij}}{\gamma_i})) + 2\frac{\gamma_i\gamma_{ij}^2\sigma_{ej}^2}{\gamma_i}\right] + \gamma_{ij}^2\sigma_{ej}^2 \\
& = (1-\gamma_{ij})\left[(1+\gamma_i)\sigma_\nu^2 + \sigma_u^2 - 2\gamma_i(\sigma_\nu^2 + \frac{\sigma_u^2\gamma_{ij}}{\gamma_i}) - \sigma_\nu^2\gamma_i\gamma_{ij} - \sigma_\nu^2\gamma_{ij} - \sigma_u^2\gamma_{ij} \right. \\
& \quad \left. + 2\gamma_i\gamma_{ij}(\sigma_\nu^2 + \frac{\sigma_u^2\gamma_{ij}}{\gamma_i}) + 2\frac{\gamma_i\gamma_{ij}^2\sigma_{ej}^2}{\gamma_i}\right] + \gamma_{ij}^2\sigma_{ej}^2 \\
& = (1-\gamma_{ij})\left[\gamma_{ij}^2(2\frac{\gamma_i}{\gamma_i}(\sigma_u^2 + \sigma_{ej}^2)) + \gamma_{ij}(-2\frac{\sigma_u^2\gamma_i}{\gamma_i} - \sigma_u^2 - (1-\gamma_i)\sigma_\nu^2) + \sigma_u^2 + (1-\gamma_i)\sigma_\nu^2\right] + \gamma_{ij}^2\sigma_{ej}^2.
\end{aligned}$$

The last expression satisfies to

$$g_{1ij}(\delta) = (1-\gamma_{ij})^2(\sigma_u^2 + (1-\gamma_i)\sigma_\nu^2) + \gamma_{ij}^2\sigma_{ej}^2.$$

4.9.2 Proof that the reminder term in $MSE(\hat{\mu}_{ij})$ is $o(m^{-1})$

In this Section, we intend to verify that $MSE(\hat{\mu}_{ij})$ is correct to terms of order $O(m^{-1})$. Very similar approach can be used for $MSE(\hat{\mu}_i)$. We may write

$$MSE(\hat{\mu}_{ij}) = E(\hat{\mu}_{ij} - \mu_{ij})^2 = E(\tilde{\mu}_{ij} - \mu_{ij})^2 + E(\hat{\mu}_{ij} - \tilde{\mu}_{ij})^2 = MSE(\tilde{\mu}_{ij}) + E(\hat{\mu}_{ij} - \tilde{\mu}_{ij})^2,$$

where $MSE(\tilde{\mu}_{ij}) = g_{1ij}(\delta) + g_{2ij}(\delta)$ with $\delta = (\sigma_v^2, \sigma_u^2)'$ and $g_{1ij}(\delta)$ and $g_{2ij}(\delta)$ are given by (4.12) and (4.13) respectively. It remains to show that $E(\hat{\mu}_{ij} - \tilde{\mu}_{ij})^2 = g_{3ij}(\delta) + o(m^{-1})$, where $g_{3ij}(\delta)$ is given by (4.14).

We make a Taylor series expansion of $\hat{\mu}_{ij}$ about δ as follows :

$$\hat{\mu}_{ij} - \tilde{\mu}_{ij} = (\hat{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta} + \frac{1}{2}(\hat{\delta} - \delta)' \frac{d^2\tilde{\mu}_{ij}}{d\delta^* d\delta^{*'}} (\hat{\delta} - \delta),$$

where $|\delta^* - \delta| \leq |\hat{\delta} - \delta|$ and $\frac{d^2\tilde{\mu}_{ij}}{d\delta^* d\delta^{*'}} = \frac{d^2\tilde{\mu}_{ij}}{d\delta d\delta'}|_{\delta=\delta^*}$. Squaring and applying the expectation operator to the expansion yields

$$\begin{aligned} E(\hat{\mu}_{ij} - \tilde{\mu}_{ij})^2 &= E((\hat{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta})^2 + \frac{1}{4}E[(\hat{\delta} - \delta)' \frac{d^2\tilde{\mu}_{ij}}{d\delta^* d\delta^{*'}} (\hat{\delta} - \delta)]^2 \\ &+ E\left\{[(\hat{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta}][(\hat{\delta} - \delta)' \frac{d^2\tilde{\mu}_{ij}}{d\delta^* d\delta^{*'}} (\hat{\delta} - \delta)]\right\} \equiv E((\hat{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta})^2 + R_1 + R_2. \end{aligned} \quad (4.44)$$

Now in Lemma 4.5, Section 4.9.8, we will show that $E[(\hat{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta}]^2 = E[(\tilde{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta}]^2 + o(m^{-1})$. In addition, $E[(\tilde{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta}]^2 = g_{3ij}(\delta) + o(m^{-1})$ upon direct application of Theorem A.1 of Prasad and Rao (1990).

It remains to show that both R_1 and R_2 are $o(m^{-1})$. We first show that R_2 is $o(m^{-1})$.

The reminder term R_2 can be written as

$$\begin{aligned} R_2 &= E\left\{[(\hat{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta}][(\hat{\delta} - \delta)' \frac{d^2\tilde{\mu}_{ij}}{d\delta^* d\delta^{*'}} (\hat{\delta} - \delta)]\right\} \\ &= E\left\{[(\hat{\delta}_1 - \delta_1) \frac{d\tilde{\mu}_{ij}}{d\delta_1} + (\hat{\delta}_2 - \delta_2) \frac{d\tilde{\mu}_{ij}}{d\delta_2}][\sum_{k=1}^2 \sum_{l=1}^2 (\hat{\delta}_k - \delta_k)(\hat{\delta}_l - \delta_l) \frac{d^2\tilde{\mu}_{ij}}{d\delta_k^* d\delta_l^{*'}}]\right\} \equiv E(T_2). \end{aligned}$$

We require the following Lemma, Serfling (2001).

Lemma 4.3. *If $\{X_n\}$ is a sequence of random variables such that $X_n = o_p(1)$ and $\sup_{n \geq 1} E|X_n|^{1+\alpha} < \infty$ for some $0 < \alpha < 1$ then $E(X_n) = o(1)$.*

We first show that $m^{1+\alpha} E\{|\hat{\delta}_j - \delta_j|^{1+\alpha} |\frac{d\tilde{\mu}_{ij}}{d\delta_j}|^{1+\alpha} |\hat{\delta}_k - \delta_k|^{1+\alpha} |\hat{\delta}_l - \delta_l|^{1+\alpha} |\frac{d^2\tilde{\mu}_{ij}}{d\delta_k^* d\delta_l^*}|^{1+\alpha}\} \leq c$ for some $c < \infty$ and $j, k, l = 1, 2$. We shall deal with the case $j = k = l$, but the other cases follow similarly. Thus, we need only show that $m^{1+\alpha} E\{|\hat{\delta}_j - \delta_j|^{3(1+\alpha)} |\frac{d\tilde{\mu}_{ij}}{d\delta_j}|^{1+\alpha} |\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}}|^{1+\alpha}\} \leq c$. Now appealing to Holder's inequality we have that

$$\begin{aligned} & m^{1+\alpha} E\{|\hat{\delta}_j - \delta_j|^{3(1+\alpha)} |\frac{d\tilde{\mu}_{ij}}{d\delta_j}|^{1+\alpha} |\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}}|^{1+\alpha}\} \\ & \leq m^{1+\alpha} (E|\hat{\delta}_j - \delta_j|^{6(1+\alpha)})^{1/2} \cdot [E|\frac{d\tilde{\mu}_{ij}}{d\delta_j}|^{2(1+\alpha)} |\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}}|^{2(1+\alpha)}]^{1/2} \\ & \leq m^{1+\alpha} (E|\hat{\delta}_j - \delta_j|^{6(1+\alpha)})^{1/2} \cdot (E|\frac{d\tilde{\mu}_{ij}}{d\delta_j}|^{3(1+\alpha)})^{1/3} \cdot (E|\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}}|^{6(1+\alpha)})^{1/6}. \end{aligned}$$

Choosing $\alpha = 1/3$, we need only show that

$$m^{4/3} (E|\hat{\delta}_j - \delta_j|^8)^{1/2} \cdot (E|\frac{d\tilde{\mu}_{ij}}{d\delta_j}|^4)^{1/3} \cdot (E|\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}}|^8)^{1/6} = O(1). \quad (4.45)$$

We first show that $E|\hat{\delta}_j - \delta_j|^8 = O(m^{-4})$. We have

$$\begin{aligned} E|\hat{\delta}_j - \delta_j|^8 &= E[|\tilde{\delta}_j - \delta_j|^8 I(\tilde{\delta}_j > 0)] + E[|\delta_j|^8 I(\tilde{\delta}_j \leq 0)] \\ &\leq (\delta_j)^8 Pr(\tilde{\delta}_j \leq 0) + E|\tilde{\delta}_j - \delta_j|^8. \end{aligned} \quad (4.46)$$

Further,

$$\begin{aligned} Pr(\tilde{\delta}_j \leq 0) &= Pr(\tilde{\delta}_j - \delta_j \leq -\delta_j) \\ &\leq Pr(|\tilde{\delta}_j - \delta_j| \geq \delta_j) \\ &\leq \frac{E|\tilde{\delta}_j - \delta_j|^8}{(\delta_j)^8} \quad \text{by Markov's inequality} \\ &= O(m^{-4}), \end{aligned}$$

by Lemma 4.8, Section 4.9.8. By using Lemma 4.8 again, $E|\tilde{\delta}_j - \delta_j|^8 = O(m^{-4})$, so we may write that $E|\hat{\delta}_j - \delta_j|^8 = O(m^{-4})$ by (4.46).

In fact, the Holder's and Markov's inequalities used in above results are defined in the following Proposition.

Proposition 4.1. Holder's inequality: *Let x and y be random variables. Then $E|xy| \leq (E|x^r|)^{1/r} (E|y^s|)^{1/s}$ for $r^{-1} + s^{-1} = 1$ and $r, s > 1$.*

Markov's inequality: *Let x be a random variable. Then $Pr(|x| > a) \leq \frac{E|x|}{a}$ for $a > 0$.*

Returning to (4.45), we need to show that $E|\frac{d\tilde{\mu}_{ij}}{d\delta_j}|^4$ and $E|\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}}|^8$ are bounded. We show these facts by Lemma 4.6 and Lemma 4.7 respectively. Therefore, by (4.45), since $m^{4/3}(E|\hat{\delta}_j - \delta_j|^8)^{1/2} = O(m^{-2/3})$, $E|\frac{\tilde{\mu}_{ij}}{d\delta_j}|^4 < \infty$ and $E|\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}}|^8 < \infty$, we get $E|mT_2|^{1+\alpha} < \infty$.

Next, we need to show that T_2 is $o_p(m^{-1})$. That means, we must show that $(\hat{\delta}_j - \delta_j)(\hat{\delta}_k - \delta_k)(\hat{\delta}_l - \delta_l)\frac{d\tilde{\mu}_{ij}}{d\delta_j}\frac{d^2\tilde{\mu}_{ij}}{d\delta_k^*d\delta_l^*} = o_p(m^{-1}); j, k, l = 1, 2$. As before, we shall deal with the case $j = k = l$, and the other cases are similar to this. Then we must show that

$$(\hat{\delta}_j - \delta_j)^3 \frac{d\tilde{\mu}_{ij}}{d\delta_j} \frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}} = o_p(m^{-1}) \quad \text{for } j = 1, 2. \quad (4.47)$$

We can show that $\hat{\delta}_j - \delta_j = o_p(m^{-\epsilon})$ for $0 < \epsilon < 1/2$ and $j = 1, 2$ (Stukel, 1991, Lemma A.2(C)). Then $(\hat{\delta}_j - \delta_j)^3 = o_p(m^{-1})$ by choosing $\epsilon = 1/3$. Furthermore, we have shown in Lemma 4.6 and Lemma 4.7 that $E|\frac{d\tilde{\mu}_{ij}}{d\delta_j}|^4 = O(1)$ and $E|\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}}|^4 = O(1)$ for $s = 2$. Therefore $E|\frac{d\tilde{\mu}_{ij}}{d\delta_j}| = O(1)$ and $E|\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}}| = O(1)$, since any lower order moment is also bounded. Appealing to the Lemma 4.4 below, we have that $\frac{d\tilde{\mu}_{ij}}{d\delta_j} = O_p(1)$ and $\frac{d^2\tilde{\mu}_{ij}}{d\delta_j^{2*}} = O_p(1)$.

Lemma 4.4. (*Bhat, 2004*): Let $\{X_n\}$ be a sequence of random variables. Then $\sup_{n \geq 1} E|X_n| < \infty$ implies $X_n = O_p(1)$.

Hence, $T_2 = o_p(m^{-1})$ by (4.47). So, as a result $E(T_2) = R_2 = o(m^{-1})$.

Similarly, we can show that $R_1 = o(m^{-1})$. Then, we have that following second order approximation to $MSE(\hat{\mu}_{ij})$:

$$MSE(\hat{\mu}_{ij}) = g_{1ij}(\delta) + g_{2ij}(\delta) + g_{3ij}(\delta).$$

4.9.3 Proof of $(n - p)^{-1} \mathbf{y}' A_1^1 \mathbf{y} = O_p(1)$

To show that $(n - p)^{-1} \mathbf{y}' A_1^1 \mathbf{y} - (n - p)^{-1} E(\mathbf{y}' A_1^1 \mathbf{y})$ is $O_p(m^{-1/2})$ where $E(\mathbf{y}' A_1^1 \mathbf{y}) = -tr(A_1 \tilde{J}_n) = O(m)$, we need to show that $Var[(n - p)^{-1} \mathbf{y}' A_1^1 \mathbf{y}]$ is $O(m^{-1})$. We have

$$Var[(n - p)^{-1} \mathbf{y}' A_1^1 \mathbf{y}] = (n - p)^{-2} [2tr(A_1^1 V)^2 + 4(X\beta)' A_1^1 V A_1^1 (X\beta)].$$

Further,

$$\begin{aligned} (X\beta)' A_1^1 V A_1^1 (X\beta) &= -\beta' X' A_1 \tilde{J}_n A_1 V A_1^1 X \beta \\ &= -\beta' X' [V^{-1} - V^{-1} X (X' V^{-1} X)^{-1} X' V^{-1}] \tilde{J}_n A_1 V A_1^1 X \beta = 0, \end{aligned}$$

and

$$\begin{aligned} tr(A_1^1 V)^2 &= tr(A_1 \tilde{J}_n A_1 V A_1 \tilde{J}_n A_1 V) = tr(\tilde{J}_n A_1 V A_1 \tilde{J}_n A_1 V A_1) = tr(\tilde{J}_n A_1 \tilde{J}_n A_1) \\ &= O(m), \text{ where } A_1 V A_1 = A_1. \text{ Hence, } Var[(n - p)^{-1} \mathbf{y}' A_1^1 \mathbf{y}] = O(m^{-1}) \text{ and } (n - p)^{-1} \mathbf{y}' A_1^1 \mathbf{y} = (n - p)^{-1} E(\mathbf{y}' A_1^1 \mathbf{y}) + O_p(m^{-1/2}) \text{ or } (n - p)^{-1} \mathbf{y}' A_1^1 \mathbf{y} = O_p(1). \end{aligned}$$

4.9.4 Proof of $(n-p)^{-1}\mathbf{y}'A_1^2\mathbf{y} = O_p(1)$

Similar as before, to show that $(n-p)^{-1}\mathbf{y}'A_1^2\mathbf{y} - (n-p)^{-1}E(\mathbf{y}'A_1^2\mathbf{y})$ is $O_p(m^{-1/2})$ where $E(\mathbf{y}'A_1^2\mathbf{y}) = -tr(A_1) = O(m)$, we need to show that $Var[(n-p)^{-1}\mathbf{y}'A_1^2\mathbf{y}] = O(m^{-1})$. We can write that

$$Var[(n-p)^{-1}\mathbf{y}'A_1^2\mathbf{y}] = (n-p)^{-2}[2tr(A_1^2V)^2 + 4(X\beta)'A_1^2VA_1^2(X\beta)],$$

where $(X\beta)'A_1^2VA_1^2(X\beta) = 0$ and $tr(A_1^2V)^2 = tr(A_1A_1) = O(m)$. Hence, $Var[(n-p)^{-1}\mathbf{y}'A_1^2\mathbf{y}] = O(m^{-1})$. Then the result is obtained.

4.9.5 Proof of $(m-p)^{-1}\bar{\mathbf{y}}'A_2^1\bar{\mathbf{y}} = O_p(1)$

Similarly, to show that $(m-p)^{-1}\bar{\mathbf{y}}'A_2^1\bar{\mathbf{y}} - (m-p)^{-1}E(\bar{\mathbf{y}}'A_2^1\bar{\mathbf{y}})$ is $O_p(m^{-1/2})$ where $E(\bar{\mathbf{y}}'A_2^1\bar{\mathbf{y}}) = -tr(A_2) = O(m)$, we need to show that $Var[(m-p)^{-1}\bar{\mathbf{y}}'A_2^1\bar{\mathbf{y}}] = O(m^{-1})$. We have that

$$Var[(m-p)^{-1}\bar{\mathbf{y}}'A_2^1\bar{\mathbf{y}}] = (m-p)^{-2}[2tr(A_2^1V^*)^2 + 4(\bar{X}\beta)'A_2^1V^*A_2^1(\bar{X}\beta)],$$

where $(\bar{X}\beta)'A_2^1V^*A_2^1(\bar{X}\beta) = 0$ and $tr(A_2^1V^*)^2 = tr(A_2A_2) = O(m)$. Hence, $Var[(m-p)^{-1}\bar{\mathbf{y}}'A_2^1\bar{\mathbf{y}}] = O(m^{-1})$. Then the result is established.

4.9.6 Proof of $(m-p)^{-1}\bar{\mathbf{y}}'A_2^2\bar{\mathbf{y}} = O_p(1)$

Similarly, to show that $(m-p)^{-1}\bar{\mathbf{y}}'A_2^2\bar{\mathbf{y}} - (m-p)^{-1}E(\bar{\mathbf{y}}'A_2^2\bar{\mathbf{y}})$ is $O_p(m^{-1/2})$ where $E(\bar{\mathbf{y}}'A_2^2\bar{\mathbf{y}}) = -tr(A_2\tilde{N}^{-1}) = O(m)$, we need to show that $Var[(m-p)^{-1}\bar{\mathbf{y}}'A_2^2\bar{\mathbf{y}}] =$

$O(m^{-1})$. We know that

$$\text{Var}[(m-p)^{-1}\bar{y}'A_2^2\bar{y}] = (m-p)^{-2}[2\text{tr}(A_2^2V^*)^2 + 4(\bar{X}\beta)'A_2^2V^*A_2^2(\bar{X}\beta)],$$

where $(\bar{X}\beta)'A_2^2V^*A_2^2(\bar{X}\beta) = 0$ and $\text{tr}(A_2^2V^*)^2 = \text{tr}(\tilde{N}^{-1}A_2\tilde{N}^{-1}A_2) = O(m)$. Hence, $\text{Var}[(m-p)^{-1}\bar{y}'A_2^2\bar{y}] = O(m^{-1})$. Then the result is obtained.

4.9.7 Proof that $E[mse(\hat{\mu}_{ij})] - MSE(\hat{\mu}_{ij}) = o(m^{-1})$

In this Section, we show that $mse(\hat{\mu}_{ij})$ is correct to $O_p(m^{-1})$. By Taylor expansion,

$$g_{1ij}(\hat{\delta}) = g_{1ij}(\delta) + (\hat{\delta} - \delta)' \nabla g_{1ij}(\delta) + \frac{1}{2}(\hat{\delta} - \delta)' \nabla^2 g_{1ij}(\delta) (\hat{\delta} - \delta) + o_p(m^{-1}).$$

Now taking expectation of the above equation,

$$E[g_{1ij}(\hat{\delta})] = g_{1ij}(\delta) + E[(\hat{\delta} - \delta)' \nabla g_{1ij}(\delta)] + \frac{1}{2}E[(\hat{\delta} - \delta)' \nabla^2 g_{1ij}(\delta) (\hat{\delta} - \delta)] + o(m^{-1}),$$

where

$$E[(\hat{\delta} - \delta)' \nabla g_{1ij}(\delta)] = E[(\hat{\sigma}_v^2 - \sigma_v^2) \frac{dg_{1ij}}{d\sigma_v^2} + (\hat{\sigma}_u^2 - \sigma_u^2) \frac{dg_{1ij}}{d\sigma_u^2}] = b(\hat{\sigma}_v^2) \frac{dg_{1ij}}{d\sigma_v^2} + b(\hat{\sigma}_u^2) \frac{dg_{1ij}}{d\sigma_u^2}.$$

In addition, we have

$$\begin{aligned} E[(\hat{\delta} - \delta)' \nabla^2 g_{1ij}(\delta) (\hat{\delta} - \delta)] &= \sum_{k=1}^2 \sum_{l=1}^2 \nabla^2 g_{1ijkl}(\delta) E[(\hat{\delta}_k - \delta_k)(\hat{\delta}_l - \delta_l)] \\ &= \sum_{k=1}^2 \sum_{l=1}^2 \nabla^2 g_{1ijkl}(\delta) E[(\hat{\delta}_k - \tilde{\delta}_k + \tilde{\delta}_k - \delta_k)(\hat{\delta}_l - \tilde{\delta}_l + \tilde{\delta}_l - \delta_l)] \\ &= \sum_{k=1}^2 \sum_{l=1}^2 \nabla^2 g_{1ijkl}(\delta) E[(\tilde{\delta}_k - \delta_k)(\tilde{\delta}_l - \delta_l)] + \sum_{k=1}^2 \sum_{l=1}^2 \nabla^2 g_{1ijkl}(\delta) E[(\hat{\delta}_k - \tilde{\delta}_k)(\hat{\delta}_l - \tilde{\delta}_l)] \\ &\quad + \sum_{k=1}^2 \sum_{l=1}^2 \nabla^2 g_{1ijkl}(\delta) E[(\hat{\delta}_k - \tilde{\delta}_k)(\tilde{\delta}_l - \delta_l)] + \sum_{k=1}^2 \sum_{l=1}^2 \nabla^2 g_{1ijkl}(\delta) E[(\tilde{\delta}_k - \delta_k)(\hat{\delta}_l - \tilde{\delta}_l)] \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^2 \sum_{l=1}^2 \nabla^2 g_{1ijkl}(\delta) E[(\tilde{\delta}_k - \delta_k)(\tilde{\delta}_l - \delta_l)] + o(m^{-1}) \\
&= \text{tr}[\text{Var}(\tilde{\delta}) \nabla^2 g_{1ij}(\delta)] + o(m^{-1}),
\end{aligned}$$

since by applying Holder's inequality

$$E[(\hat{\delta}_k - \tilde{\delta}_k)(\hat{\delta}_l - \tilde{\delta}_l)] \leq (E|\hat{\delta}_k - \tilde{\delta}_k|^2)^{1/2} \cdot (E|\hat{\delta}_l - \tilde{\delta}_l|^2)^{1/2} = o(m^{-1}),$$

and noting that $\nabla^2 g_{1ijkl}(\delta) = O(1)$, $E|\hat{\delta}_l - \tilde{\delta}_l|^2 = O(m^{-2})$ and $E|\hat{\delta}_k - \tilde{\delta}_k|^2 = O(m^{-2})$ by the following argument :

$$\begin{aligned}
E|\hat{\delta}_k - \tilde{\delta}_k|^2 &= E\{|\tilde{\delta}_k|^2 I(\tilde{\delta}_k \leq 0)\} \\
&\leq (E|\tilde{\delta}_k|^4)^{1/2} \cdot (E(I(\tilde{\delta}_k \leq 0))^2)^{1/2} \quad \text{by Holder's inequality} \\
&= O(m^{-2}),
\end{aligned}$$

since $(E(I(\tilde{\delta}_k \leq 0))^2)^{1/2} = [Pr(\tilde{\delta}_k \leq 0)]^{1/2} = O(m^{-2})$ and

$$\begin{aligned}
E|\tilde{\delta}_k - \delta_k + \delta_k|^4 &= E[(\tilde{\delta}_k - \delta_k)^2 + \delta_k^2 + 2\delta_k|\tilde{\delta}_k - \delta_k|]^2 \\
&\leq 2\{E[(\tilde{\delta}_k - \delta_k)^2 + \delta_k^2]^2 + 4E[\delta_k^2(\tilde{\delta}_k - \delta_k)^2]\} \\
&= 2\{E(\tilde{\delta}_k - \delta_k)^4 + \delta_k^4 + 2\delta_k^2 E(\tilde{\delta}_k - \delta_k)^2 + 4\delta_k^2 E(\tilde{\delta}_k - \delta_k)^2\} = O(1).
\end{aligned}$$

Similarly, $E[(\hat{\delta}_k - \tilde{\delta}_k)(\tilde{\delta}_l - \delta_l)] \leq (E|\hat{\delta}_k - \tilde{\delta}_k|^2)^{1/2} (E|\tilde{\delta}_l - \delta_l|^2)^{1/2} = o(m^{-1})$, by noting that $E|\tilde{\delta}_l - \delta_l|^2 = \text{Var}(\tilde{\delta}_l) = O(m^{-1})$ and $E|\hat{\delta}_k - \tilde{\delta}_k|^2 = o(m^{-1})$. By a similar argument, $E[(\tilde{\delta}_k - \delta_k)(\hat{\delta}_l - \tilde{\delta}_l)] = o(m^{-1})$. Combining above results, we have

$$E[g_{1ij}(\hat{\delta})] = g_{1ij}(\delta) + b(\hat{\sigma}_\nu^2) \frac{dg_{1ij}(\delta)}{d\sigma_\nu^2} + b(\hat{\sigma}_u^2) \frac{dg_{1ij}(\delta)}{d\sigma_u^2} + \frac{1}{2} \text{tr}[\text{Var}(\tilde{\delta}) \nabla^2 g_{1ij}(\delta)] + o(m^{-1}).$$

It is easy to show that $\frac{1}{2}tr[Var(\tilde{\delta})\nabla^2 g_{1ij}(\delta)] = -g_{3ij}(\delta)$.

Hence,

$$E[g_{1ij}(\hat{\delta})] = g_{1ij}(\delta) + b(\hat{\sigma}_\nu^2) \frac{dg_{1ij}}{d\sigma_\nu^2} + b(\hat{\sigma}_u^2) \frac{dg_{1ij}}{d\sigma_u^2} - g_{3ij}(\delta) + o(m^{-1}).$$

Now, since $g_{2ij}(\delta)$ and $g_{3ij}(\delta)$ are of order $O(m^{-1})$, it follows that

$$E[g_{2ij}(\hat{\delta})] = g_{2ij}(\delta) + o(m^{-1}),$$

and

$$E[g_{3ij}(\hat{\delta})] = g_{3ij}(\delta) + o(m^{-1}).$$

Therefore, $g_{2ij}(\hat{\delta})$ and $g_{3ij}(\hat{\delta})$ are the correct estimators of $g_{2ij}(\delta)$ and $g_{3ij}(\delta)$ respectively, each with bias $o(m^{-1})$. Combining above results, we have that

$$E[mse(\hat{\mu}_{ij})] = MSE(\hat{\mu}_{ij}) + o(m^{-1}).$$

4.9.8 Some Lemmas

Lemma 4.5. *Assume the regularity conditions (C1)-(C6) mentioned in Section 1.4.4 are satisfied. Then*

$$E[(\hat{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta}]^2 = E[(\tilde{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta}]^2 + o(m^{-1}).$$

Proof: We can write

$$\begin{aligned} E[(\hat{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta}]^2 &= E[\sum_{j=1}^2 (\tilde{\delta}_j - \delta_j) \frac{d\tilde{\mu}_{ij}}{d\delta_j} - \sum_{j=1}^2 \frac{d\tilde{\mu}_{ij}}{d\delta_j} \tilde{\delta}_j I(\tilde{\delta}_j \leq 0)]^2 \\ &= E[(\tilde{\delta} - \delta)' \frac{d\tilde{\mu}_{ij}}{d\delta}]^2 + E[\sum_{j=1}^2 \frac{d\tilde{\mu}_{ij}}{d\delta_j} \tilde{\delta}_j I(\tilde{\delta}_j \leq 0)]^2 \end{aligned}$$

$$-2 \sum_{k=1}^2 \sum_{l=1}^2 E \left[\frac{d\tilde{\mu}_{ij}}{d\delta_k} \frac{d\tilde{\mu}_{ij}}{d\delta_l} (\tilde{\delta}_k - \delta_k) \tilde{\delta}_l I(\tilde{\delta}_l \leq 0) \right].$$

To complete the proof, we need to show that the last two terms are orders $o(m^{-1})$.

First,

$$\begin{aligned} E \left[\sum_{j=1}^2 \frac{d\tilde{\mu}_{ij}}{d\delta_j} I(\tilde{\delta}_j \leq 0) \tilde{\delta}_j \right]^2 &\leq \left[\sum_{j=1}^2 (E \left(\frac{d\tilde{\mu}_{ij}}{d\delta_j} I(\tilde{\delta}_j \leq 0) \right)^2)^{1/2} \cdot (E(\tilde{\delta}_j^2))^{1/2} \right]^2 \\ &\leq \left[\sum_{j=1}^2 (E \left| \frac{d\tilde{\mu}_{ij}}{d\delta_j} \right|^4)^{1/4} \cdot (E(I(\tilde{\delta}_j \leq 0))^4)^{1/4} \cdot (E(\tilde{\delta}_j^2))^{1/2} \right]^2 \\ &= \left[\sum_{j=1}^2 [E \left| \frac{d\tilde{\mu}_{ij}}{d\delta_j} \right|^4 E(I(\tilde{\delta}_j \leq 0))^4]^{1/4} (E(\tilde{\delta}_j^2))^{1/2} \right]^2, \end{aligned}$$

where $E \left| \frac{d\tilde{\mu}_{ij}}{d\delta_j} \right|^4 = O(1)$ by Lemma 4.6 with $s = 2$.

In addition, noting that $E(I(\tilde{\delta}_j \leq 0))^4 = Pr(\tilde{\delta}_j \leq 0) = O(m^{-4})$ and $E(\tilde{\delta}_j)^2 = E(\tilde{\delta}_j - \delta_j + \delta_j)^2 \leq 2(E(\tilde{\delta}_j - \delta_j)^2 + E\delta_j^2) = O(1)$, we have $E[\sum_{j=1}^2 \frac{d\tilde{\mu}_{ij}}{d\delta_j} I(\tilde{\delta}_j \leq 0) \tilde{\delta}_j]^2 = o(m^{-1})$.

Next, we have

$$\begin{aligned} &E \left[\frac{d\tilde{\mu}_{ij}}{d\delta_k} \frac{d\tilde{\mu}_{ij}}{d\delta_l} (\tilde{\delta}_k - \delta_k) \tilde{\delta}_l I(\tilde{\delta}_l \leq 0) \right] \\ &\leq \left[\left(E \left(\frac{d\tilde{\mu}_{ij}}{d\delta_k} \right)^2 \left(\frac{d\tilde{\mu}_{ij}}{d\delta_l} \right)^2 (\tilde{\delta}_k - \delta_k)^2 \right)^{1/2} \left(E(\tilde{\delta}_l^2) I(\tilde{\delta}_l \leq 0) \right)^{1/2} \right] \\ &\leq \left[\left(E \left(\frac{d\tilde{\mu}_{ij}}{d\delta_k} \right)^4 \right)^{1/4} \left(\left(E \left(\frac{d\tilde{\mu}_{ij}}{d\delta_l} \right)^8 \right)^{1/2} (E(\tilde{\delta}_k - \delta_k)^8)^{1/2} \right)^{1/4} \right]^{1/2} \\ &\quad \cdot \left[(E(\tilde{\delta}_l^4))^{1/2} (E(I(\tilde{\delta}_l \leq 0))^4)^{1/2} \right]^{1/2} \\ &= [E \left(\frac{d\tilde{\mu}_{ij}}{d\delta_k} \right)^4 E \left(\frac{d\tilde{\mu}_{ij}}{d\delta_l} \right)^8 E(\tilde{\delta}_k - \delta_k)^8]^{1/8} [E(\tilde{\delta}_l^4) E(I(\tilde{\delta}_l \leq 0))]^{1/4} = o(m^{-1}), \end{aligned}$$

since $E \left| \frac{d\tilde{\mu}_{ij}}{d\delta_k} \right|^4 = O(1)$, $\left| \frac{d\tilde{\mu}_{ij}}{d\delta_l} \right| = O(1)$, $E(\tilde{\delta}_k - \delta_k)^8 = O(m^{-4})$, $E(\tilde{\delta}_l^4) = O(1)$ and $E(I(\tilde{\delta}_k \leq 0)) = Pr(\tilde{\delta}_k \leq 0) = O(m^{-4})$. Then the Lemma 4.5 is established.

Lemma 4.6. Assume the regularity conditions (C1)-(C6) mentioned in section 1.4.4 are satisfied. Then $E|\frac{d\tilde{\mu}_{ij}}{d\sigma_j^2}|^{2s} = O(1)$ for $j = 1, 2$ and $s \in \mathbb{N}, s > 1$.

Proof: We have shown that $\tilde{\mu}_{ij} = x'_{ij}\tilde{\beta} + \gamma_i(1 - \gamma_{ij})(\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) + \gamma_{ij}(y_{ij} - x'_{ij}\tilde{\beta})$.

We may write $\frac{d\tilde{\mu}_{ij}}{d\sigma_\nu^2}$ as

$$\frac{d\tilde{\mu}_{ij}}{d\sigma_\nu^2} = \frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2} + \frac{d\gamma_i}{d\sigma_\nu^2}(1 - \gamma_{ij})(\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) - \gamma_i(1 - \gamma_{ij})\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2} - \gamma_{ij}\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2},$$

then

$$\begin{aligned} \left|\frac{d\tilde{\mu}_{ij}}{d\sigma_\nu^2}\right| &\leq (1 - \gamma_*)(\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2} + \frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}) + c_1(1 - \gamma_*)(\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) \\ &\leq c_2\left[\left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2}\right| + \left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right| + |\bar{y}_{i\gamma}| + |\bar{x}'_{i\gamma}\tilde{\beta}|\right], \end{aligned}$$

then

$$E\left|\frac{d\tilde{\mu}_{ij}}{d\sigma_\nu^2}\right|^{2s} \leq c_2^{2s} \cdot 2^{4s-2} \left[E\left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s} + E\left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s} + E|\bar{y}_{i\gamma}|^{2s} + E|\bar{x}'_{i\gamma}\tilde{\beta}|^{2s} \right]. \quad (4.48)$$

Note that $|\frac{d\gamma_i}{d\sigma_\nu^2}| = |\frac{\gamma_i^2}{\gamma_i} \frac{\sigma_u^2}{\sigma_\nu^4}| \leq \frac{\sigma_u^2}{2\sigma_\nu^4\gamma_*} \equiv c_1$, $|\frac{d\bar{y}_{i\gamma}}{d\sigma_\nu^2}| = 0$, $|\frac{d\bar{x}'_{i\gamma}}{d\sigma_\nu^2}| = 0$, $\inf_i \inf_j \gamma_{ij} \geq \gamma_* = \sigma_u^2(\sigma_u^2 + \sigma_e^{2*})^{-1}$, $\sup_i \gamma_i \leq 1$, $\sup_i \sup_j \gamma_{ij} \leq 1$, $\sigma_e^{2*} = \max_{i,j} \sigma_{eij}^2$ such that σ_{eij}^2 are uniformly bounded and c_1 and c_2 are known. In fact, we used C_r inequality in above which is defined in the following Proposition:

Proposition 4.2. *C_r inequality:* For random variables x and y , we have $E|x + y|^r \leq C_r(E|x|^r + E|y|^r)$ where $C_r = 1$ if $r < 1$ and $C_r = 2^{r-1}$ if $r \geq 1$.

Then we need to show that the terms of right side of (4.48) are of order $O(1)$.

We have, by using the Cauchy-Schwarz inequality,

$$\begin{aligned} \left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2}\right| &= |x'_{ij}[-(X'V^{-1}X)^{-1}X'\dot{V}_\nu^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}y + (X'V^{-1}X)^{-1}X'\dot{V}_\nu^{-1}y]| \\ &= |x'_{ij}(X'V^{-1}X)^{-1}X'\dot{V}_\nu^{-1}(y - X\tilde{\beta})| \end{aligned}$$

$$\leq (mx'_{ij}(X'V^{-1}X)^{-1}x_{ij})^{1/2}[m^{-1}(y - X\tilde{\beta})'\dot{V}_\nu^{-2}V(y - X\tilde{\beta})]^{1/2},$$

where $\dot{V}_\nu^{-1} = \frac{dV^{-1}}{d\sigma_\nu^2}$ and $\dot{V}_\nu^{-2} = (\frac{dV^{-1}}{d\sigma_\nu^2})(\frac{dV^{-1}}{d\sigma_\nu^2})$. Therefore, we may write

$$E\left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s} \leq [mx'_{ij}(X'V^{-1}X)^{-1}x_{ij}]^s E[m^{-1}(y - X\tilde{\beta})'\dot{V}_\nu^{-2}V(y - X\tilde{\beta})]^s, \quad (4.49)$$

where $s \in \mathbb{N}, s > 1$. But, we can write

$$\begin{aligned} m^{-1}(y - X\tilde{\beta})'\dot{V}_\nu^{-2}V(y - X\tilde{\beta}) &= m^{-1} \sum_{i=1}^m (y_i - x_i\tilde{\beta})'\dot{V}_{\nu i}^{-2}V_i(y_i - x_i\tilde{\beta}) \\ &= m^{-1} \sum_{i=1}^m (y_i - x_i\beta)'\dot{V}_{\nu i}^{-2}V_i(y_i - x_i\beta) + O_p(m^{-1/2}), \end{aligned}$$

where $\dot{V}_{\nu i}^{-2} = (\frac{dV_i^{-1}}{d\sigma_\nu^2})(\frac{dV_i^{-1}}{d\sigma_\nu^2})$. However, we have $m^{-1} \sum_{i=1}^m (y_i - x_i\beta)'\dot{V}_{\nu i}^{-2}V_i(y_i - x_i\beta) = m^{-1} \sum_i \sum_j \sum_k z_{ij}z_{ik}l_{1ijk}$ where $z_{ij} = y_{ij} - x'_{ij}\beta$ and l_{1ijk} = the (j,k)th element of $[\dot{V}_{\nu i}^{-2}V_i]$. On the other hand, $E[z_{ij}\sqrt{l_{1ijj}}] = 0$ and

$$Cov(z_{ij}\sqrt{l_{1ijk}}, z_{ik}\sqrt{l_{1ijk}}) = \begin{cases} l_{1jj}(\sigma_\nu^2 + \sigma_u^2 + \sigma_{eij}^2) & \text{if } k = j \\ l_{1jk}(\sigma_\nu^2) & \text{if } k \neq j \end{cases}$$

We know that $\max_{i,j} \sigma_{eij}^2 = \sigma_{e*}^2 = O(1)$, then we need to show that $\max_{j,k} l_{1ijk} = O(1)$ as follows

$$\begin{aligned} l_{1ijk} &= \left(\frac{\sigma_u^4 \gamma_i^2}{\sigma_\nu^4 \gamma_i^2}\right)^2 \left[\sigma_\nu^2 w_{ik} \sum_{l=1}^{n_i} \sum_{p=1}^{n_i} w_{ip} w_{il}^2 + (\sigma_u^2 + \sigma_{eij}^2) w_{ij} w_{ik} \sum_l w_{il}^2 \right] \\ &\leq [\sigma_u^4 \sigma_\nu^{-4} n_l^{-2}]^2 \left\{ \sigma_\nu^2 n_u^2 (\sigma_u^2 + \sigma_{e*}^2)^{-4} + n_u (\sigma_u^2 + \sigma_{e*}^2)^{-3} \right\} \\ &\leq \left(\frac{\sigma_u^4}{\sigma_\nu^4}\right)^2 (\sigma_u^2 + \sigma_{e*}^2)^{-3} n_l^{-3} \left\{ \sigma_\nu^2 n_u (\sigma_u^2 + \sigma_{e*}^2)^{-1} + 1 \right\} = O(1), \end{aligned}$$

where $n_l = \min_i n_i$, $n_u = \max_i n_i$ and $\sigma_{e*}^2 = \min_{i,j} \sigma_{eij}^2$. Hence $E[m^{-1} \sum_i \sum_j \sum_k z_{ij}z_{ik}l_{1ijk}]^s = O(1)$ for $s \in \mathbb{N}, s > 1$.

On the other hand, since $(X'V^{-1}X)^{-1} = O(m^{-1})$ and x_{ij} is uniformly bounded, then $(mx'_{ij}(X'V^{-1}X)^{-1}x_{ij}) = O(1)$. Therefore, combining above results, we may write

$$E\left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s} = O(1).$$

We must then show that $E\left|\frac{d\tilde{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s}$, $E|\bar{y}_{i\gamma}|^{2s}$ and $E|\tilde{x}'_{i\gamma}\tilde{\beta}|^{2s}$ are $O(1)$. We may write

$$E\left|\frac{d\tilde{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s} \leq (m\tilde{x}'_{i\gamma}(X'V^{-1}X)^{-1}\tilde{x}_{i\gamma})^s E\left[m^{-1}(y - X\tilde{\beta})'\dot{V}_\nu^{-2}V(y - X\tilde{\beta})\right]^s = O(1), \quad (4.50)$$

noting that (4.50) is similar to (4.49). In addition, we can write

$$E|\bar{y}_{i\gamma}|^{2s} = E|\gamma_i^{-1} \sum_{j=1}^{n_i} \gamma_{ij} y_{ij}|^{2s} \leq E\left|\sum_{j=1}^{n_i} y_{ij}\right|^{2s} \leq 2^{(2s-1)n_i/2} \sum_{j=1}^{n_i} E|y_{ij}|^{2s} < \infty,$$

since $|y_{ij}| < \infty$. Also,

$$E|\tilde{x}'_{i\gamma}\tilde{\beta}|^{2s} = E|\gamma_i^{-1} \sum_{j=1}^{n_i} \gamma_{ij} x'_{ij}\tilde{\beta}|^{2s} \leq E\left|\sum_{j=1}^{n_i} x'_{ij}\tilde{\beta}\right|^{2s} \leq 2^{(2s-1)n_i/2} \sum_{j=1}^{n_i} E|x'_{ij}\tilde{\beta}|^{2s} < \infty,$$

where

$$\begin{aligned} E|x'_{ij}\tilde{\beta}| &= |x'_{ij}(X'V^{-1}X)^{-1}| E\left|\sum_{i=1}^m \sum_{j=1}^{n_i} \sum_{k=1}^{n_i} x_{ij} V_{ijk}^{-1} y_{ik}\right| \\ &\leq |x'_{ij}(X'V^{-1}X)| \sum_i \sum_j \sum_k E|x_{ij} V_{ijk}^{-1} y_{ik}|, \end{aligned}$$

then

$$E|x'_{ij}\tilde{\beta}|^{2s} \leq |mx'_{ij}(X'V^{-1}X)^{-1}|^{2s} \left[m^{-1} \sum_i \sum_j \sum_k E|x_{ij} V_{ijk}^{-1} y_{ik}|\right]^{2s} = O(1),$$

since under the assumption of normality of the random errors, the first moment of the y_{ij} 's exists and is finite. Note that x_{ij} and V_{ijk}^{-1} are uniformly bounded where

V_{ijk}^{-1} = the (j,k) element of V_i^{-1} . Hence, combining above results, we have $E|\frac{d\tilde{\mu}_{ij}}{d\sigma_u^2}|^{2s} = O(1)$.

Similarly, we can write

$$\begin{aligned} \frac{d\tilde{\mu}_{ij}}{d\sigma_u^2} &= \frac{dx'_{ij}\tilde{\beta}}{d\sigma_u^2} + \frac{d\gamma_i}{d\sigma_u^2}(1 - \gamma_{ij})(\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) + \gamma_i(1 - \gamma_{ij})\left(\frac{d\bar{y}_{i\gamma}}{d\sigma_u^2} - \frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2} - \frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}\right) \\ &\quad - \frac{d\gamma_{ij}}{d\sigma_u^2}\gamma_i(\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) + \frac{d\gamma_{ij}}{d\sigma_u^2}(y_{ij} - x'_{ij}\tilde{\beta}) - \gamma_{ij}\frac{dx'_{ij}\tilde{\beta}}{d\sigma_u^2}, \end{aligned}$$

then, we have that

$$\begin{aligned} \left|\frac{d\tilde{\mu}_{ij}}{d\sigma_u^2}\right| &\leq 2\left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_u^2}\right| + c_5(1 - \gamma_*)[|\bar{y}_{i\gamma}| + |\bar{x}'_{i\gamma}\tilde{\beta}|] + (1 - \gamma_*)\left[c_3|\bar{y}_{i\gamma}| + \left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}\right| + c_4|\bar{x}'_{i\gamma}\tilde{\beta}|\right] \\ &\quad + (\sigma_u^2 + \sigma_{e*}^2)[|\bar{y}_{i\gamma}| + |\bar{x}'_{i\gamma}\tilde{\beta}| + |y_{ij}| + |x'_{ij}\tilde{\beta}|] \\ &\leq c_6\left[\left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_u^2}\right| + \left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}\right| + |\bar{y}_{i\gamma}| + |\bar{x}'_{i\gamma}\tilde{\beta}| + |y_{ij}| + |x'_{ij}\tilde{\beta}|\right], \end{aligned}$$

then

$$\begin{aligned} E\left|\frac{d\tilde{\mu}_{ij}}{d\sigma_u^2}\right|^{2s} &\leq c_6^{2s} \cdot 2^{6s-3} \left[E\left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_u^2}\right|^{2s} + E\left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}\right|^{2s} \right. \\ &\quad \left. + E|\bar{y}_{i\gamma}|^{2s} + E|\bar{x}'_{i\gamma}\tilde{\beta}|^{2s} + E|y_{ij}|^{2s} + E|x'_{ij}\tilde{\beta}|^{2s} \right]. \end{aligned} \quad (4.51)$$

Note that $|\frac{d\bar{y}_{i\gamma}}{d\sigma_u^2}| = |\frac{1}{\gamma_i \cdot \sigma_u^2}[\bar{y}_{i\gamma} \sum_{j=1}^{n_i} \gamma_{ij}^2 - \sum_{j=1}^{n_i} y_{ij} \gamma_{ij}^2]| \leq \frac{1}{\gamma_* \sigma_u^2} |n_u \bar{y}_{i\gamma}| = c_3 |\bar{y}_{i\gamma}|$, $|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}| \leq \frac{1}{\gamma_* \sigma_u^2} |n_u \bar{x}'_{i\gamma}| = c_4 |\bar{x}'_{i\gamma}|$, $|\frac{d\gamma_i}{d\sigma_u^2}| = |\frac{-\gamma_i^2}{\gamma_i^2 \sigma_u^2} \sum_{j=1}^{n_i} \gamma_{ij}^2| \leq \frac{n_u}{4\sigma_u^2 \gamma_*^2} \equiv c_5$, where $c_3, \dots, c_6$ are known.

We then need to show that $E|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_u^2}|^{2s}$, $E|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}|^{2s}$ are of order $O(1)$. We have, by Cauchy-Schwarz inequality,

$$\begin{aligned} \left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_u^2}\right| &= |x'_{ij}[-(X'V^{-1}X)^{-1}X'\dot{V}_u^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}y + (X'V^{-1}X)^{-1}X'\dot{V}_u^{-1}y]| \\ &= |x'_{ij}(X'V^{-1}X)^{-1}X'\dot{V}_u^{-1}(y - X\tilde{\beta})| \end{aligned}$$

$$\leq (mx'_{ij}(X'V^{-1}X)^{-1}x_{ij})^{1/2}[m^{-1}(y - X\tilde{\beta})'\dot{V}_u^{-2}V(y - X\tilde{\beta})]^{1/2},$$

where $\dot{V}_u^{-1} = \frac{dV^{-1}}{d\sigma_u^2}$ and $\dot{V}_u^{-2} = (\frac{dV^{-1}}{d\sigma_u^2})(\frac{dV^{-1}}{d\sigma_u^2})$. Therefore, we may write

$$E|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_u^2}|^{2s} \leq [mx'_{ij}(X'V^{-1}X)^{-1}x_{ij}]^s E[m^{-1}(y - X\tilde{\beta})'\dot{V}_u^{-2}V(y - X\tilde{\beta})]^s, \quad (4.52)$$

with $s \in \mathbb{N}, s > 1$. As before, we may write

$$m^{-1}(y - X\tilde{\beta})'\dot{V}_u^{-2}V(y - X\tilde{\beta}) = m^{-1} \sum_i (y_i - x_i\beta)'\dot{V}_{ui}^{-2}V_i(y_i - x_i\beta) + O_p(m^{-1/2}),$$

where $\dot{V}_{ui}^{-2} = (\frac{dV_i^{-1}}{d\sigma_u^2})(\frac{dV_i^{-1}}{d\sigma_u^2})$. However, we may write

$$m^{-1} \sum_i (y_i - x_i\beta)'\dot{V}_{ui}^{-2}V_i(y_i - x_i\beta) = m^{-1} \sum_{i=1}^m \sum_{j=1}^{n_i} \sum_{k=1}^{n_i} z_{ij}z_{ik}l_{2ijk},$$

where $z_{ij} = y_{ij} - x_{ij}\beta$ and l_{2ijk} =the (j,k) element of $[\dot{V}_{ui}^{-2}V_i]$. In addition, $z_{ij}\sqrt{l_{2ijj}}$ has zero mean and

$$Cov(z_{ij}\sqrt{l_{2ijj}}, z_{ik}\sqrt{l_{2ijk}}) = \begin{cases} l_{2ijj}(\sigma_\nu^2 + \sigma_u^2 + \sigma_{eij}^2) & \text{if } k = j \\ l_{2ijk}(\sigma_\nu^2) & \text{if } k \neq j \end{cases}$$

We then show that $\max_{j,k} l_{2ijk} = O(1)$ as follows

$$\begin{aligned} \dot{V}_{uilj}^{-1} &= \text{the } (l,j)\text{element of } [\dot{V}_{ui}^{-1}] = -(\sigma_u^2 + \sigma_{eij}^2)^{-2}\delta_{lj} - (\sigma_u^2 + \sigma_{eil}^2)^{-1}(\sigma_u^2 + \sigma_{eij}^2)^{-1} \\ &\cdot \left[\frac{\gamma_i}{\gamma_i^3} \frac{\sum_j \gamma_{ij}^2}{\sigma_\nu^2} (\gamma_i \sigma_\nu^2 - \gamma_i \sigma_u^2) - \frac{\gamma_i \sigma_u^2}{\gamma_i} (\sigma_u^2 + \sigma_{eil}^2)^{-1} (\sigma_u^2 + \sigma_{eij}^2)^{-1} \right], \end{aligned}$$

where

$$\delta_{lj} = \begin{cases} 1 & \text{if } l = j \\ 0 & \text{if } l \neq j \end{cases}$$

Then

$$l_{2ijk} = \sigma_\nu^2 \left\{ -n_i(\sigma_u^2 + \sigma_{eik}^2)^{-2} + (\sigma_u^2 + \sigma_{eik}^2)^{-1} \sum_{p=1}^{n_i} \sum_{m=1}^{n_i} (\sigma_u^2 + \sigma_{eip}^2)^{-1} (\sigma_u^2 + \sigma_{eim}^2)^{-2} \right\}$$

$$\begin{aligned}
& \cdot \left[\frac{\gamma_i}{\gamma_i^3} \frac{\sum_j \gamma_{ij}^2}{\sigma_\nu^2} (\gamma_i \sigma_\nu^2 - \gamma_i \sigma_u^2) - \frac{\gamma_i \sigma_u^2}{\gamma_i} (\sigma_u^2 + \sigma_{eip}^2)^{-1} (\sigma_u^2 + \sigma_{eim}^2)^{-1} \right] \\
& \cdot \left[\frac{\gamma_i}{\gamma_i^3} \frac{\sum_j \gamma_{ij}^2}{\sigma_\nu^2} (\gamma_i \sigma_\nu^2 - \gamma_i \sigma_u^2) - \frac{\gamma_i \sigma_u^2}{\gamma_i} (\sigma_u^2 + \sigma_{eim}^2)^{-1} (\sigma_u^2 + \sigma_{eik}^2)^{-1} \right] \Big\} \\
& - (\sigma_u^2 + \sigma_{eij}^2)^{-1} \left\{ (\sigma_u^2 + \sigma_{eik}^2)^{-2} - (\sigma_u^2 + \sigma_{eij}^2)^{-1} (\sigma_u^2 + \sigma_{eik}^2)^{-1} \sum_m (\sigma_u^2 + \sigma_{eim}^2)^{-2} \right. \\
& \cdot \left[\frac{\gamma_i}{\gamma_i^3} \frac{\sum_j \gamma_{ij}^2}{\sigma_\nu^2} (\gamma_i \sigma_\nu^2 - \gamma_i \sigma_u^2) - \gamma_i \gamma_i^{-1} \sigma_u^2 (\sigma_u^2 + \sigma_{eij}^2)^{-1} (\sigma_u^2 + \sigma_{eim}^2)^{-1} \right] \\
& \cdot \left. \left[\frac{\gamma_i}{\gamma_i^3} \frac{\sum_j \gamma_{ij}^2}{\sigma_\nu^2} (\gamma_i \sigma_\nu^2 - \gamma_i \sigma_u^2) - \gamma_i \gamma_i^{-1} \sigma_u^2 (\sigma_u^2 + \sigma_{eim}^2)^{-1} (\sigma_u^2 + \sigma_{eik}^2)^{-1} \right] \right\} \\
& \leq n_u \sigma_\nu^2 (\sigma_u^2 + \sigma_{e*}^2)^{-2} + (\sigma_u^2 + \sigma_{e*}^2)^{-3} + \left[n_u^2 \sigma_\nu^2 (\sigma_u^2 + \sigma_{e*}^2)^{-4} + n_u (\sigma_u^2 + \sigma_{e*}^2)^{-5} \right] \\
& \cdot \left[n_l^{-2} \sigma_\nu^{-2} (n_u \sigma_\nu^2 - \sigma_u^2) - n_l^{-1} \sigma_u^2 (\sigma_u^2 + \sigma_{e*}^2)^{-2} \right]^2 = O(1),
\end{aligned}$$

where $\dot{V}_{ui}^{-1} = \frac{dV_{ui}^{-1}}{d\sigma_u^2} = -diag_j (\sigma_u^2 + \sigma_{eij}^2)^{-2} - \frac{\gamma_i}{\gamma_i^3} \frac{\sum_{j=1}^{n_i} \gamma_{ij}^2}{\sigma_\nu^2} (\gamma_i \sigma_\nu^2 - \gamma_i \sigma_u^2) w_i w_i' + \frac{\sigma_u^2 \gamma_i}{\gamma_i} u_i u_i'$,

with $u_i = (u_{i1}, \dots, u_{in_i})'$, $u_{ij} = (\sigma_u^2 + \sigma_{eij}^2)^{-2}$, $\dot{V}_u^{-1} = diag_i (\dot{V}_{ui}^{-1})$. Then,

$E[m^{-1} \sum_i \sum_j \sum_k z_{ij} z_{ik} l_{2ijk}]^s = O(1)$ and as before $(m x'_{ij} (X' V^{-1} X)^{-1} x_{ij})^s = O(1)$.

Combining above results, we have

$$E \left| \frac{dx'_{ij} \tilde{\beta}}{d\sigma_u^2} \right|^{2s} = O(1).$$

In addition, we have

$$\left| \frac{d\bar{x}'_{i\gamma} \tilde{\beta}}{d\sigma_u^2} \right|^{2s} \leq (m \bar{x}'_{i\gamma} (X' V^{-1} X)^{-1} \bar{x}_{i\gamma})^s [m^{-1} (y - X \tilde{\beta})' \dot{V}_u^{-2} V (y - X \tilde{\beta})]^s,$$

then

$$E \left| \frac{d\bar{x}'_{i\gamma} \tilde{\beta}}{d\sigma_u^2} \right|^{2s} = O(1),$$

which is similar to (4.52). Then, by (4.51), we have

$$E \left| \frac{d\tilde{\mu}_{ij}}{d\sigma_u^2} \right|^{2s} = O(1),$$

since we have shown that $E|\bar{y}_{i\gamma}|^{2s}$, $E|\bar{x}'_{i\gamma}\tilde{\beta}|^{2s}$ and $E|x'_{ij}\tilde{\beta}|^{2s}$ are of order $O(1)$ and $E|y_{ij}|^{2s} = O(1)$, since $|y_{ij}| < \infty$. Therefore, the result is obtained.

Lemma 4.7. *Assume the regularity conditions (C1)-(C6) mentioned in section 1.4.4 are satisfied. Then $E|\frac{d^2\tilde{\mu}_{ij}}{d\sigma_\nu^2 d\sigma_u^2}|^{2s} = O(1)$ for $s \in \mathbb{N}, s > 1$.*

Proof: We may write

$$\begin{aligned} \frac{d^2\tilde{\mu}_{ij}}{d\sigma_\nu^2 d\sigma_u^2} &= \frac{d^2x'_{ij}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2} + \frac{d^2\gamma_i}{d\sigma_\nu^2 d\sigma_u^2}(1 - \gamma_{ij})(\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) + \frac{d\gamma_i}{d\sigma_\nu^2}(1 - \gamma_{ij})\left(\frac{d\bar{y}_{i\gamma}}{d\sigma_u^2} \right. \\ &\quad \left. - \frac{d\bar{x}'_{i\gamma}}{d\sigma_u^2}\tilde{\beta} - \frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}\right) - \frac{d\gamma_{ij}}{d\sigma_u^2}\frac{d\gamma_i}{d\sigma_\nu^2}(\bar{y}_{i\gamma} - \bar{x}'_{i\gamma}\tilde{\beta}) - (1 - \gamma_{ij})\frac{d\gamma_i}{d\sigma_u^2}\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2} \\ &\quad - \gamma_i(1 - \gamma_{ij})\left(\frac{d\bar{x}'_{i\gamma}}{d\sigma_u^2}\frac{d\tilde{\beta}}{d\sigma_\nu^2} + \frac{d^2\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right) + \gamma_i\frac{d\gamma_{ij}}{d\sigma_u^2}\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2} - \frac{d\gamma_{ij}}{d\sigma_u^2}\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2} - \gamma_{ij}\frac{d^2x'_{ij}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}. \end{aligned}$$

By using Lemma 4.6, we may write

$$\begin{aligned} \left|\frac{d^2\gamma_i}{d\sigma_\nu^2 d\sigma_u^2}\right| &= \left|\frac{\gamma_i \cdot \sigma_\nu^4(\gamma_i^2 - 2\gamma_i^3\sigma_u^2\gamma_i^{-2}\sigma_\nu^{-2}\sum_{j=1}^{n_i}\gamma_{ij}^2) - \gamma_i^2\sigma_\nu^4(\gamma_i - \sum_{j=1}^{n_i}\gamma_{ij}^2)}{(\gamma_i \cdot \sigma_\nu^4)^2}\right| \\ &\leq \left|\frac{n_u\sigma_\nu^4 + 2\sigma_u^2\sigma_\nu^{-2}n_u(2\gamma_*)^{-2}n_u\sigma_\nu^4 + 2\sigma_u^4n_u}{(\sigma_\nu^4)^2(2\gamma_*)^2}\right| \equiv c_7. \end{aligned}$$

Also, $\left|\frac{d\gamma_{ij}}{d\sigma_u^2}\right| = \left|\frac{\sigma_{eij}^2}{(\sigma_u^2 + \sigma_{eij}^2)^2}\right| = |\gamma_{ij}(1 - \gamma_{ij})/\sigma_u^2| \leq \frac{1-\gamma_*}{\sigma_u^2} \equiv c_8$. Then we have

$$\begin{aligned} \left|\frac{d^2\tilde{\mu}_{ij}}{d\sigma_\nu^2 d\sigma_u^2}\right| &\leq 2\left|\frac{d^2x'_{ij}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right| + c_7(1 - \gamma_*)[|\bar{y}_{i\gamma}| + |\bar{x}'_{i\gamma}\tilde{\beta}|] + c_1(1 - \gamma_*) \\ &\quad \cdot [c_3|\bar{y}_{i\gamma}| + c_4|\bar{x}'_{i\gamma}\tilde{\beta}| + \left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right|] + c_1 \cdot c_8(|\bar{y}_{i\gamma}| + |\bar{x}'_{i\gamma}\tilde{\beta}|) + (1 - \gamma_*)c_2\left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right| \\ &\quad + (1 - \gamma_*)[c_4\left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right| + \left|\frac{d^2\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right|] + (\sigma_u^2 + \sigma_{e*}^2)^{-1}\left[\left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right| + \left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2}\right|\right] \\ &\leq c_9\left[\left|\frac{d^2x'_{ij}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right| + |\bar{y}_{i\gamma}| + |\bar{x}'_{i\gamma}\tilde{\beta}| + \left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}\right| + \left|\frac{d\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right| + \left|\frac{d^2\bar{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right| + \left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2}\right|\right], \end{aligned}$$

then

$$E\left|\frac{d^2\tilde{\mu}_{ij}}{d\sigma_\nu^2 d\sigma_u^2}\right|^{2s} \leq c_{10} \left[E\left|\frac{d^2x'_{ij}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right|^{2s} + E|\tilde{y}_{i\gamma}|^{2s} \right. \\ \left. + E|\tilde{x}'_{i\gamma}\tilde{\beta}|^{2s} + E\left|\frac{d\tilde{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}\right|^{2s} + E\left|\frac{d\tilde{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s} + E\left|\frac{d^2\tilde{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right|^{2s} + E\left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s} \right], \quad (4.53)$$

where $c_7, \dots, c_{10}$ are known. In addition,

$$\left|\frac{d^2x'_{ij}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right| = \left| x'_{ij} \left[(X'V^{-1}X)^{-1}X'\dot{V}_u^{-1}X(X'V^{-1}X)^{-1}X'\dot{V}_\nu^{-1}X\tilde{\beta} - (X'V^{-1}X)^{-1}X'\dot{V}_{\nu\nu}^{-1} \right. \right. \\ \left. \left. \cdot X\tilde{\beta} + (X'V^{-1}X)^{-1}X'\dot{V}_\nu^{-1}X(X'V^{-1}X)^{-1}X'\dot{V}_u^{-1}X\tilde{\beta} - (X'V^{-1}X)^{-1}X'\dot{V}_\nu^{-1}X(X'V^{-1} \right. \right. \\ \left. \left. \cdot X)^{-1}X'\dot{V}_u^{-1}y - (X'V^{-1}X)^{-1}X'\dot{V}_u^{-1}X(X'V^{-1}X)^{-1}X'\dot{V}_\nu^{-1}y + (X'V^{-1}X)^{-1}X'\dot{V}_{\nu\nu}^{-1}y \right] \right| \\ = \left| x'_{ij}(X'V^{-1}X)^{-1} \left[- (X'\dot{V}_u^{-1}X)(X'V^{-1}X)^{-1}X'\dot{V}_\nu^{-1} + X'\dot{V}_{\nu\nu}^{-1} \right. \right. \\ \left. \left. - (X'\dot{V}_\nu^{-1}X)(X'V^{-1}X)^{-1}X'\dot{V}_u^{-1} \right] (y - X\tilde{\beta}) \right| \\ \leq (mx'_{ij}(X'V^{-1}X)^{-1}x_{ij})^{1/2} \left\{ 2[m^{-1}(y - X\tilde{\beta})'\dot{V}_\nu^{-2}\dot{V}_u^{-2}V^3(y - X\tilde{\beta})]^{1/2} \right. \\ \left. + [m^{-1}(y - X\tilde{\beta})'\dot{V}_{\nu\nu}^{-2}V(y - X\tilde{\beta})]^{1/2} \right\},$$

where $\dot{V}_{\nu\nu}^{-1} = \frac{d\dot{V}_\nu^{-1}}{d\sigma_\nu^2} = \sigma_\nu^{-8} \text{diag}_i \left\{ [2\sigma_u^{-2}w_iw'_i - 2\gamma_i\gamma_i^{-2}\sigma_\nu^{-2}\sum_{j=1}^{n_i}\gamma_{ij}^2w_iw'_i - u_iu'_i] \right. \\ \left. \cdot \gamma_i^2\sigma_u^4\sigma_\nu^4\gamma_i^{-2} - 2\sigma_\nu^4\sigma_u^2\gamma_i^2\gamma_i^{-4}w_iw'_i(\gamma_i - \sum_{j=1}^{n_i}\gamma_{ij}^2) \right\}$. But, we can write

$$m^{-1}(y - X\tilde{\beta})'\dot{V}_\nu^{-2}\dot{V}_u^{-2}V^3(y - X\tilde{\beta}) = m^{-1} \sum_i (y_i - x_i\beta)'\dot{V}_{\nu i}^{-2}\dot{V}_{u i}^{-2}V_i^3(y_i - x_i\beta)$$

$+O_p(m^{-1/2})$ and

$$m^{-1}(y - X\tilde{\beta})'\dot{V}_{\nu\nu}^{-2}V(y - X\tilde{\beta}) = m^{-1} \sum_i (y_i - x_i\beta)'\dot{V}_{\nu\nu i}^{-2}V_i(y_i - x_i\beta) + O_p(m^{-1/2}).$$

By using very similar approach as Lemma 4.6, we can show that

$$E[m^{-1}(y - X\tilde{\beta})'\dot{V}_\nu^{-2}\dot{V}_u^{-2}V^3(y - X\tilde{\beta})]^s = O(1),$$

and

$$E[m^{-1}(y - X\tilde{\beta})'\dot{V}_{\nu\nu}^{-2}V(y - X\tilde{\beta})]^s = O(1),$$

for $s \in \mathbb{N}, s > 1$. Then

$$\begin{aligned} E\left|\frac{d^2 x'_{ij}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right|^{2s} &\leq (mx'_{ij}(X'V^{-1}X)^{-1}x_{ij})^s \cdot 2^{2s-1} \left\{ 2^{2s} E\left[m^{-1}(y - X\tilde{\beta})'\dot{V}_\nu^{-2}\dot{V}_u^{-2} \right. \right. \\ &\quad \left. \left. \cdot V^3(y - X\tilde{\beta})\right]^s + E\left[m^{-1}(y - X\tilde{\beta})'\dot{V}_{\nu\nu}^{-2}V(y - X\tilde{\beta})\right]^s \right\} = O(1). \end{aligned} \quad (4.54)$$

Note that $E\left|\frac{d^2 \tilde{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2 d\sigma_u^2}\right|^{2s} = O(1)$ which is similar to (4.54). Moreover, we have shown in Lemma 4.6 that $E|\tilde{y}_{i\gamma}|^{2s}$, $E|\tilde{x}'_{i\gamma}\tilde{\beta}|^{2s}$, $E\left|\frac{d\tilde{x}'_{i\gamma}\tilde{\beta}}{d\sigma_u^2}\right|^{2s}$, $E\left|\frac{d\tilde{x}'_{i\gamma}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s}$ and $E\left|\frac{dx'_{ij}\tilde{\beta}}{d\sigma_\nu^2}\right|^{2s}$ are of order $O(1)$. Hence, we have, by (4.53),

$$E\left|\frac{d^2 \tilde{\mu}_{ij}}{d\sigma_\nu^2 d\sigma_u^2}\right|^{2s} = O(1).$$

Therefore, the result is established.

Lemma 4.8. *Assume the regularity conditions (C1)-(C6) mentioned in section 1.4.4 are satisfied. Then $E|\tilde{\delta}_j - \delta_j|^{2s} = O(m^{-s})$ for $j = 1, 2$ and $s \in \mathbb{N}, s > 1$.*

Proof: We may write, by (4.25),

$$\begin{aligned} \tilde{\sigma}_\nu^2 - \sigma_\nu^2 &= -[tr(\tilde{J}_n V^{-1})tr(\tilde{N}^{-1}V^{*-1}) - tr(V^{-1})tr(V^{*-1})]^{-1} \\ &\quad \cdot [tr(\tilde{N}^{-1}V^{*-1})(n - p - y'A_1 y) - tr(V^{-1})(m - p - \bar{y}'A_2 \bar{y})] + O_p(m^{-1}). \end{aligned}$$

Then

$$E(\tilde{\sigma}_\nu^2 - \sigma_\nu^2)^2 \leq 2 \left[tr(\tilde{J}_n V^{-1})tr(\tilde{N}^{-1}V^{*-1}) - tr(V^{-1})tr(V^{*-1}) \right]^{-2}$$

$\cdot \left[(tr(\tilde{N}^{-1}V^{*-1}))^2 E(n-p-y'A_1y)^2 + (tr(V^{-1}))^2 E(m-p-\bar{y}'A_2\bar{y})^2 \right] = O(m^{-1})$,
since $tr(\tilde{J}_n V^{-1})$, $tr(\tilde{N}^{-1}V^{*-1})$, $tr(V^{-1})$ and $tr(V^{*-1})$ are order $O(m)$, and $E(n-p-y'A_1y)^2 = Var(y'A_1y) = 2(n-p)$, noting that $y'A_1y \sim \chi^2(n-p)$ and $E(m-p-\bar{y}'A_2\bar{y})^2 = Var(\bar{y}'A_2\bar{y}) = 2(m-p)$, noting that $\bar{y}'A_2\bar{y} \sim \chi^2(m-p)$.

Similarly,

$$E(\tilde{\sigma}_v^2 - \sigma_v^2)^{2s} \leq 2^{2s-1} \left[tr(\tilde{J}_n V^{-1}) tr(\tilde{N}^{-1}V^{*-1}) - tr(V^{-1}) tr(V^{*-1}) \right]^{-2s} \\
\cdot \left[(tr(\tilde{N}^{-1}V^{*-1}))^{2s} E(n-p-y'A_1y)^{2s} + (tr(V^{-1}))^{2s} E(m-p-\bar{y}'A_2\bar{y})^{2s} \right] = O(m^{-s}),$$

by using C_r inequality and noting that $E(n-p-y'A_1y)^{2s} = O(m^s)$, since $E(n-p-y'A_1y) = 0$ and similarly $E(m-p-\bar{y}'A_2\bar{y})^{2s} = O(m^s)$, since $E(m-p-\bar{y}'A_2\bar{y}) = 0$.

Similarly, we may write, by (4.25),

$$\tilde{\sigma}_u^2 - \sigma_u^2 = - \left[tr(\tilde{J}_n V^{-1}) tr(\tilde{N}^{-1}V^{*-1}) - tr(V^{-1}) tr(V^{*-1}) \right]^{-1} \\
\cdot \left[tr(\tilde{J}_n V^{-1})(m-p-\bar{y}'A_2\bar{y}) - tr(V^{*-1})(n-p-y'A_1y) \right] + O_p(m^{-1}).$$

Then

$$E(\tilde{\sigma}_u^2 - \sigma_u^2)^2 \leq 2 \left[tr(\tilde{J}_n V^{-1}) tr(\tilde{N}^{-1}V^{*-1}) - tr(V^{-1}) tr(V^{*-1}) \right]^{-2} \\
\cdot \left[(tr(\tilde{J}_n V^{-1}))^2 E(m-p-\bar{y}'A_2\bar{y})^2 + (tr(V^{*-1}))^2 E(n-p-y'A_1y)^2 \right] = O(m^{-1}).$$

Similarly,

$$E(\tilde{\sigma}_u^2 - \sigma_u^2)^{2s} \leq 2^{2s-1} \left[tr(\tilde{J}_n V^{-1}) tr(\tilde{N}^{-1}V^{*-1}) - tr(V^{-1}) tr(V^{*-1}) \right]^{-2s} \\
\cdot \left[(tr(\tilde{J}_n V^{-1}))^{2s} E(m-p-\bar{y}'A_2\bar{y})^{2s} + (tr(V^{*-1}))^{2s} E(n-p-y'A_1y)^{2s} \right] = O(m^{-s}).$$

Then, the result is obtained.

σ_ν^2	σ_u^2	40	100	140	200	400
40		55.73	82.95	95.88	110.65	139.32
100		67.42	92.05	103.73	117.07	143.03
140		71.58	95.47	106.76	119.64	144.66
200		75.58	98.87	109.84	122.31	146.45
400		81.67	104.30	114.89	126.87	149.81

Table 4.1: Average mean squared error of sub-area EBLUP

σ_ν^2	σ_u^2	40	100	140	200	400
40		30.36	33.11	34.61	36.53	41.43
100		47.00	53.63	57.47	62.54	75.33
140		52.85	61.19	66.11	72.76	90.23
200		58.48	68.72	74.91	83.41	106.51
400		67.07	80.76	89.36	101.56	137.02

Table 4.2: Average mean squared error of area EBLUP

σ_ν^2	σ_u^2	40	100	140	200	400
40		(-0.8,6.1)	(0.8,1.5)	(2.1,1.1)	(4.4,0.8)	(13.5,0.5)
100		(-0.2,6.1)	(-0.1,1.5)	(-0.1,1.1)	(-0.1,0.8)	(0.7,0.5)
140		(0.1,6.1)	(0.2,1.5)	(0.2,1.1)	(0.3,0.8)	(0.4,0.5)
200		(0.4,6.1)	(0.5,1.5)	(0.5,1.1)	(0.5,0.8)	(0.6,0.5)
400		(0.7,6.1)	(0.7,1.5)	(0.8,1.1)	(0.8,0.8)	(0.9,0.5)

Table 4.3: Percent relative bias of variance component estimators ($\hat{\sigma}_\nu^2, \hat{\sigma}_u^2$)

σ_ν^2	σ_u^2	40	100	140	200	400
40		-8.3	-5.5	-4.2	-3.0	-1.4
100		-5.7	-4.6	-3.8	-3.0	-1.7
140		-4.8	-4.0	-3.3	-2.7	-1.6
200		-3.9	-3.5	-2.9	-2.3	-1.5
400		-2.9	-2.7	-2.3	-1.8	-1.2

Table 4.4: Percent average relative bias of sub-area naive MSE estimator

σ_ν^2	σ_u^2	40	100	140	200	400
40		-17.6	-18.8	-19.0	-18.8	-16.4
100		-7.1	-9.6	-10.8	-12.5	-16.2
140		-4.3	-6.2	-7.2	-8.6	-12.4
200		-2.2	-3.7	-4.4	-5.4	-8.5
400		0.3	-0.7	-1.1	-1.6	-3.4

Table 4.5: Percent average relative bias of area naive MSE estimator

σ_ν^2	σ_u^2	40	100	140	200	400
40		3.0	0.8	0.7	0.6	0.6
100		2.4	0.4	0.2	0.1	0.1
140		2.5	0.5	0.3	0.2	0.0
200		2.5	0.6	0.4	0.3	0.1
400		2.6	0.8	0.6	0.4	0.2

Table 4.6: Percent average relative bias of sub-area MSE estimator

σ_ν^2	σ_u^2	40	100	140	200	400
40		2.1	4.9	7.4	11.5	26.0
100		1.9	1.2	1.1	1.3	3.4
140		2.6	2.0	1.8	1.8	2.2
200		3.2	2.6	2.4	2.3	2.3
400		4.0	3.4	3.3	3.1	2.8

Table 4.7: Percent average relative bias of area MSE estimator

σ_ν^2	σ_u^2	40	100	140	200	400
40		(0.05,0.07)	(0.09,0.00)	(0.11,0.00)	(0.15,0.00)	(0.23,0.00)
100		(0.00,0.07)	(0.00,0.00)	(0.00,0.00)	(0.01,0.00)	(0.04,0.00)
140		(0.00,0.07)	(0.00,0.00)	(0.00,0.00)	(0.00,0.00)	(0.01,0.00)
200		(0.00,0.07)	(0.00,0.00)	(0.00,0.00)	(0.00,0.00)	(0.00,0.00)
400		(0.00,0.07)	(0.00,0.00)	(0.00,0.00)	(0.00,0.00)	(0.00,0.00)

Table 4.8: Monte carlo probabilities ($pr(\tilde{\sigma}_\nu^2 \leq 0), pr(\tilde{\sigma}_u^2 \leq 0)$)

σ_ν^2	σ_u^2	40	100	140	200	400
40		189.7	155.5	144.6	134.6	120.0
100		172.5	147.6	139.1	130.9	118.4
140		167.4	144.9	137.1	129.5	117.8
200		162.9	142.4	135.1	128.1	117.0
400		156.7	138.7	132.1	125.7	115.7

Table 4.9: Percent average relative efficiency of sub-area EBLUP over direct estimator

σ_ν^2	σ_u^2	40	100	140	200	400
40		161.5	172.9	180.1	190.2	218.9
100		129.8	135.9	139.8	145.4	162.3
140		122.4	127.2	130.3	134.8	148.3
200		116.4	120.0	122.4	125.9	136.5
400		108.7	110.7	112.1	114.1	120.4

Table 4.10: Percent average relative efficiency of area EBLUP over direct estimator

σ_ν^2	σ_u^2	40	100	140	200	400
40		108.2	103.4	102.1	101.2	100.3
100		126.3	112.4	108.3	105.0	101.5
140		137.7	118.8	112.9	107.9	102.5
200		153.5	128.1	119.7	112.6	104.1
400		197.6	156.1	141.3	127.7	110.2

Table 4.11: Percent average relative efficiency of sub-area EBLUP over the Fay-Herriot estimator

Chapter 5

Estimation of Conditional MSE in Small Area Estimation

Most of the research on small area estimation has focused on unconditional MSE estimation under an assumed small area model. Datta, Kubokawa, and Rao (2004) studied conditional MSE for the Fay-Herriot area level model. In this Chapter, we follow their approach to study conditional MSE estimation for the nested error linear regression model. We derive a second order approximation to the conditional MSE given the area-specific response data. We then derive a nearly unbiased estimator of the conditional MSE. A data analysis on the performance of our approach is also provided.

5.1 Introduction

A nested error linear regression model can be written as

$$y_{ij} = x'_{ij}\beta + v_i + e_{ij}; j = 1, \dots, N_i; i = 1, \dots, m \quad (5.1)$$

where y_{ij} is the variable of interest, $x_{ij} = (x_{ij1}, \dots, x_{ijp})'$ are the unit-specific auxiliary variables for each population element j in each small area i . The random effects v_i and e_{ij} are mutually independent and i.i.d. normal variables with zero means and variances σ_v^2 and σ_e^2 respectively, and N_i and m denote the population size in area i and the number of areas respectively. We further assume that a sample of size n_i is taken from the N_i units in the i -th area ($i=1, \dots, m$) and that the sample values $\{y_{ij}, x_{ij}; j = 1, \dots, n_i; i = 1, \dots, m\}$ also obey the assumed model (5.1). The mean of the i -th area is approximately equal to $\mu_i = \bar{X}_i' \beta + v_i$ for large N_i , where $\bar{X}_i$ is the known population mean of i -th area. We take μ_i as the parameter of interest.

The best (or Bayes) estimator of μ_i under squared loss is then given by

$$\tilde{\mu}_i^B = \bar{X}_i' \beta + (1 - B_i)(\bar{y}_i - \bar{x}_i' \beta),$$

where $B_i(\delta) = B_i = \sigma_e^2 / (\sigma_e^2 + n_i \sigma_v^2)$ with $\delta = (\sigma_v^2, \sigma_e^2)'$. The BLUP estimator of μ_i (without normality assumption) is

$$\tilde{\mu}_i(\delta) = \tilde{\mu}_i = \bar{X}_i' \tilde{\beta} + (1 - B_i)(\bar{y}_i - \bar{x}_i' \tilde{\beta}),$$

where $\tilde{\beta}(\delta) = \tilde{\beta} = (\sum_i X_i' V_i^{-1} X_i)^{-1} (\sum_i X_i' V_i^{-1} y_i)$, X_i is a $n_i \times p$ matrix with i -th row x_{ij}' , and $y_i = (y_{i1}, \dots, y_{in_i})'$ is a $n_i \times 1$ vector. We then use the transformation method to estimate the variance components σ_v^2 and σ_e^2 , as in Chapter 3. Therefore, the EBLUP (or EB) estimator of μ_i is given by

$$\hat{\mu}_i = \tilde{\mu}_i(\hat{\delta}) = \bar{X}_i' \hat{\beta} + (1 - \hat{B}_i)(\bar{y}_i - \bar{x}_i' \hat{\beta}),$$

where $\hat{\beta} = \tilde{\beta}(\hat{\delta})$, $\hat{B}_i = B_i(\hat{\delta})$ and $\hat{\delta}$ is the estimator of δ .

Prasad and Rao (1990) have shown that the unconditional MSE of $\hat{\mu}_i$ can be

approximated as

$$MSE(\hat{\mu}_i) = E(\hat{\mu}_i - \mu_i)^2 = g_{1i}(\delta) + g_{2i}(\delta) + g_{3i}(\delta) + o(m^{-1}),$$

where $g_{1i}(\delta) = B_i\sigma_\nu^2$, $g_{2i}(\delta) = (\bar{X}_i - (1 - B_i)\bar{x}_i)'(\sum_i X_i'V_i^{-1}X_i)^{-1}(\bar{X}_i - (1 - B_i)\bar{x}_i)$ and $g_{3i}(\delta) = B_i^2(1 - B_i)\sigma_e^{-4}h(\delta)$ with $h(\delta) = \sigma_\nu^4Var(\hat{\sigma}_e^2) + \sigma_e^4Var(\hat{\sigma}_\nu^2) - 2\sigma_\nu^2\sigma_e^2Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)$.

A second order unbiased estimator of the MSE of $\hat{\mu}_i$ is given by

$$mse(\hat{\mu}_i) = g_{1i}(\hat{\delta}) + g_{2i}(\hat{\delta}) + 2g_{3i}(\hat{\delta}).$$

In Section 5.2, we derive an approximation to the conditional MSE of $\hat{\mu}_i$ given the area-specific response data $y_i = (y_{i1}, \dots, y_{in_i})'$.

5.2 Second Order Approximation to Conditional MSE

It may be more appealing to practitioners to consider the estimation of the conditional MSE, $MSE^c(\hat{\mu}_i) = E[(\hat{\mu}_i - \mu_i)^2|y_i]$, where the expectation is conditional on the observations y_i in the i -th area. The conditional MSE of $\hat{\mu}_i$ given y_i is defined by

$$\begin{aligned} MSE^c[\hat{\mu}_i|y_i] &= E[(\hat{\mu}_i - \mu_i)^2|y_i] = E[(\hat{\mu}_i - \tilde{\mu}_i)^2|y_i] + E[(\tilde{\mu}_i - \mu_i)^2|y_i] \\ &\quad + 2E[(\hat{\mu}_i - \tilde{\mu}_i)(\tilde{\mu}_i - \mu_i)|y_i], \end{aligned} \tag{5.2}$$

where

$$E[(\tilde{\mu}_i - \mu_i)^2|y_i] = E[(\tilde{\mu}_i - \tilde{\mu}_i^B)^2|y_i] + E[(\tilde{\mu}_i^B - \mu_i)^2|y_i] + 2E[(\tilde{\mu}_i - \tilde{\mu}_i^B)(\tilde{\mu}_i^B - \mu_i)|y_i].$$

Since $E(\mu_i|y_i) = \tilde{\mu}_i^B$, we have that

$$E[(\tilde{\mu}_i - \tilde{\mu}_i^B)(\tilde{\mu}_i^B - \mu_i)|y_i] = 0.$$

Moreover,

$$E[(\tilde{\mu}_i^B - \mu_i)^2|y_i] = \text{Var}(\mu_i|y_i) = g_{1i}(\delta) = B_i\sigma_\nu^2.$$

Also,

$$\begin{aligned} E[(\tilde{\mu}_i - \tilde{\mu}_i^B)^2|y_i] &= E\left\{\left[(\bar{X}_i - (1 - B_i)\bar{x}_i)'(\tilde{\beta} - \beta)\right]^2|y_i\right\} \\ &= E\left\{\left[(\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}X'V^{-1}(y - X\beta)\right]^2|y_i\right\} \\ &= E\left\{(\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}X'V^{-1}(y - X\beta)(y - X\beta)'V^{-1}X(X'V^{-1}X)^{-1}\right. \\ &\quad \left. \cdot (\bar{X}_i - (1 - B_i)\bar{x}_i)|y_i\right\}. \end{aligned} \quad (5.3)$$

Now, using $E(y|y_i) = X\beta + a_i(y_i - X_i\beta)$ and $\text{Var}(y|y_i) = V - a_iV_ia_i'$, we have

$$\begin{aligned} E[(y - X\beta)(y - X\beta)'|y_i] &= \text{Var}(y - X\beta|y_i) + [E((y - X\beta)|y_i)][E((y - X\beta)'|y_i)] \\ &= V - a_iV_ia_i' + a_i(y_i - X_i\beta)(y_i - X_i\beta)'a_i' \\ &= V + a_i[(y_i - X_i\beta)(y_i - X_i\beta)' - V_i]a_i', \end{aligned}$$

where $V = \text{Var}(y)$, $a_i = (0, 0, \dots, I_{n_i}, \dots, 0, 0)'$ of dimension $n \times n_i$ with I_{n_i} as an $n_i \times n_i$ identity matrix.

Then, (5.3) equals to

$$\begin{aligned}
& (\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}X'V^{-1}\left\{V + a_i[(y_i - X_i\beta)(y_i - X_i\beta)' - V_i]a_i'\right\} \\
& \cdot V^{-1}X(X'V^{-1}X)^{-1}(\bar{X}_i - (1 - B_i)\bar{x}_i) \\
= & (\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}(\bar{X}_i - (1 - B_i)\bar{x}_i) + (\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}X'V^{-1} \\
& \cdot a_i[(y_i - X_i\beta)(y_i - X_i\beta)' - V_i]a_i'V^{-1}X(X'V^{-1}X)^{-1}(\bar{X}_i - (1 - B_i)\bar{x}_i) \\
= & (\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}(\bar{X}_i - (1 - B_i)\bar{x}_i) + (\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}X_i' \\
& \cdot [V_i^{-1}(y_i - X_i\beta)(y_i - X_i\beta)'V_i^{-1} - V_i^{-1}]X_i(X'V^{-1}X)^{-1}(\bar{X}_i - (1 - B_i)\bar{x}_i) \\
= & (\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}(\bar{X}_i - (1 - B_i)\bar{x}_i) + o_p(m^{-1}) \\
\equiv & g_{2i}(\delta) + o_p(m^{-1}),
\end{aligned}$$

where $(X'V^{-1}X)^{-1} = O(m^{-1})$.

In addition, to derive $E[(\hat{\mu}_i - \tilde{\mu}_i)^2|y_i]$, we need the following Taylor expansions:

$$\hat{B}_i = B_i + (\partial B_i / \partial \sigma_\nu^2)(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + (\partial B_i / \partial \sigma_e^2)(\hat{\sigma}_e^2 - \sigma_e^2) + O_p(m^{-1}),$$

$$\hat{\beta} = \tilde{\beta} + (\partial \tilde{\beta} / \partial \sigma_\nu^2)(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + (\partial \tilde{\beta} / \partial \sigma_e^2)(\hat{\sigma}_e^2 - \sigma_e^2) + o_p(m^{-1}),$$

since $\hat{\sigma}_\nu^2 - \sigma_\nu^2 = O_p(m^{-1/2})$, $\hat{\sigma}_e^2 - \sigma_e^2 = O_p(m^{-1/2})$, $\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} = O_p(m^{-1/2})$ and $\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = O_p(m^{-1/2})$ (see Sections 5.7.1 and 5.7.2).

Then we have

$$\begin{aligned}
E[(\hat{\mu}_i - \tilde{\mu}_i)^2|y_i] &= E\left\{\left[(\bar{X}_i - \bar{x}_i)'(\hat{\beta} - \tilde{\beta}) - \hat{B}_i(\bar{y}_i - \bar{x}_i'\hat{\beta}) + B_i(\bar{y}_i - \bar{x}_i'\tilde{\beta})\right]^2\middle|y_i\right\} \\
= & E\left\{\left[(\bar{X}_i - \bar{x}_i)'\left[\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2}(\hat{\sigma}_e^2 - \sigma_e^2)\right] - (B_i + \frac{\partial B_i}{\partial \sigma_\nu^2}(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial B_i}{\partial \sigma_e^2}(\hat{\sigma}_e^2 - \sigma_e^2))\right.\right. \\
& \left.\left. \cdot (\bar{y}_i - \bar{x}_i'(\tilde{\beta} + \frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2}(\hat{\sigma}_e^2 - \sigma_e^2))) + B_i(\bar{y}_i - \bar{x}_i'\tilde{\beta})\right]^2\middle|y_i\right\} + o_p(m^{-1})
\end{aligned}$$

$$\begin{aligned}
&= E \left\{ \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)' \left[\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right] - \left[\frac{\partial B_i}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial B_i}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right] \right. \right. \\
&\quad \left. \left. \cdot \left[\bar{y}_i - \bar{x}_i' \tilde{\beta} - \bar{x}_i' \left[\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right] \right] \right]^2 \middle| y_i \right\} + o_p(m^{-1}) \\
&= E \left\{ \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)' \left[\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right] - \left[\frac{\partial B_i}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial B_i}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right] \right. \right. \\
&\quad \left. \left. \cdot (\bar{y}_i - \bar{x}_i' \tilde{\beta}) \right]^2 \middle| y_i \right\} + o_p(m^{-1}).
\end{aligned}$$

Further,

$$\begin{aligned}
E[(\hat{\mu}_i - \tilde{\mu}_i)^2 | y_i] &= E \left\{ \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)' \frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} - \frac{\partial B_i}{\partial \sigma_\nu^2} (\bar{y}_i - \bar{x}_i' \tilde{\beta}) \right]^2 (\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 \middle| y_i \right\} \\
&\quad + E \left\{ \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)' \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} - \frac{\partial B_i}{\partial \sigma_e^2} (\bar{y}_i - \bar{x}_i' \tilde{\beta}) \right]^2 (\hat{\sigma}_e^2 - \sigma_e^2)^2 \middle| y_i \right\} \\
&\quad + 2E \left\{ \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)' \frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} - \frac{\partial B_i}{\partial \sigma_\nu^2} (\bar{y}_i - \bar{x}_i' \tilde{\beta}) \right] \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)' \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} - \frac{\partial B_i}{\partial \sigma_e^2} (\bar{y}_i - \bar{x}_i' \tilde{\beta}) \right] \right. \\
&\quad \left. \cdot (\hat{\sigma}_\nu^2 - \sigma_\nu^2)(\hat{\sigma}_e^2 - \sigma_e^2) \middle| y_i \right\} + o_p(m^{-1}) \\
&= \left(\frac{\partial B_i}{\partial \sigma_\nu^2} \right)^2 (\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] + \left(\frac{\partial B_i}{\partial \sigma_e^2} \right)^2 (\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] \\
&\quad + 2 \frac{\partial B_i}{\partial \sigma_\nu^2} \frac{\partial B_i}{\partial \sigma_e^2} (\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)(\hat{\sigma}_e^2 - \sigma_e^2) | y_i] + o_p(m^{-1}) \\
&= (\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 \left[\frac{n_i^2 \sigma_e^4}{(\sigma_e^2 + n_i \sigma_\nu^2)^4} \text{Var}(\hat{\sigma}_\nu^2) + \frac{n_i^2 \sigma_\nu^4}{(\sigma_e^2 + n_i \sigma_\nu^2)^4} \text{Var}(\hat{\sigma}_e^2) \right. \\
&\quad \left. - \frac{2n_i^2 \sigma_\nu^2 \sigma_e^2}{(\sigma_e^2 + n_i \sigma_\nu^2)^4} \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \right] + o_p(m^{-1}) \equiv g_{3i}^c(\theta),
\end{aligned}$$

where $\theta = (\beta', \delta')'$.

The above result is based on the following facts:

$$E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] = \text{Var}(\hat{\sigma}_\nu^2) + o_p(m^{-1}), \quad E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] = \text{Var}(\hat{\sigma}_e^2) + o_p(m^{-1}),$$

$E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)(\hat{\sigma}_e^2 - \sigma_e^2)|y_i] = Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) + o_p(m^{-1})$ (see Section 5.7.3) and noting that $g_{3i}^c(\theta) = \frac{(\bar{y}_i - \bar{x}_i'\beta)^2}{(\sigma_\nu^2 + \sigma_e^2/n_i)} g_{3i}(\delta)$ where $g_{3i}(\delta) = n_i^{-2}(\sigma_\nu^2 + \sigma_e^2/n_i)^{-3} [\sigma_e^4 Var(\hat{\sigma}_\nu^2) + \sigma_\nu^4 Var(\hat{\sigma}_e^2) - 2\sigma_\nu^2\sigma_e^2 Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)]$. Thus, we have

$$E[(\hat{\mu}_i - \tilde{\mu}_i)^2|y_i] = \frac{(\bar{y}_i - \bar{x}_i'\beta)^2}{\sigma_\nu^2 + \sigma_e^2/n_i} g_{3i}(\delta).$$

To calculate the cross-product term (5.2), we know that $E(\mu_i|y_i) = \tilde{\mu}_i^B$. Hence

$$E[(\tilde{\mu}_i^B - \mu_i)(\hat{\mu}_i - \tilde{\mu}_i)|y_i] = 0.$$

Therefore, we calculate the reduced cross-product term defined by (5.2) as follows

$$\begin{aligned} E[(\tilde{\mu}_i - \tilde{\mu}_i^B)(\hat{\mu}_i - \tilde{\mu}_i)|y_i] &= E\left\{ [\bar{X}_i - (1 - B_i)\bar{x}_i]'(\tilde{\beta} - \beta) [\bar{X}_i'(\hat{\beta} - \tilde{\beta}) + (1 - \hat{B}_i)(\bar{y}_i - \bar{x}_i'\hat{\beta}) \right. \\ &\quad \left. - (1 - B_i)(\bar{y}_i - \bar{x}_i'\tilde{\beta})] \middle| y_i \right\} \\ &= E\left\{ [\bar{X}_i - (1 - B_i)\bar{x}_i]'(\tilde{\beta} - \beta) \left[\bar{X}_i' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right) \right. \right. \\ &\quad \left. \left. + \left[(1 - B_i) + \left(1 - \frac{\partial B_i}{\partial \sigma_\nu^2} \right) (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \left(1 - \frac{\partial B_i}{\partial \sigma_e^2} \right) (\hat{\sigma}_e^2 - \sigma_e^2) \right] (\bar{y}_i - \bar{x}_i'\tilde{\beta}) - \bar{x}_i' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \right. \right. \right. \\ &\quad \left. \left. \left. + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right) \right] - (1 - B_i)(\bar{y}_i - \bar{x}_i'\tilde{\beta}) \right] \middle| y_i \right\} + o_p(m^{-1}) \\ &= E\left\{ [\bar{X}_i - (1 - B_i)\bar{x}_i]'(\tilde{\beta} - \beta) \left[\bar{X}_i' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right) - (1 - B_i)\bar{x}_i' \right. \right. \\ &\quad \left. \left. \cdot \left[\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right] + \left[\left(1 - \frac{\partial B_i}{\partial \sigma_\nu^2} \right) (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \left(1 - \frac{\partial B_i}{\partial \sigma_e^2} \right) (\hat{\sigma}_e^2 - \sigma_e^2) \right] (\bar{y}_i - \bar{x}_i'\tilde{\beta}) \right. \right. \\ &\quad \left. \left. \left. - \bar{x}_i' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right) \right] \right] \middle| y_i \right\} + o_p(m^{-1}) \end{aligned}$$

$$\begin{aligned}
&= E \left\{ [\bar{X}_i - (1 - B_i)\bar{x}_i]'(\tilde{\beta} - \beta) \left[\left[(\bar{X}_i - (1 - B_i)\bar{x}_i)' \frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} + \left(1 - \frac{\partial B_i}{\partial \sigma_\nu^2}\right)(\bar{y}_i - \bar{x}_i'\tilde{\beta}) \right] (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \right. \right. \\
&\quad \left. \left. + \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)' \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} + \left(1 - \frac{\partial B_i}{\partial \sigma_e^2}\right)(\bar{y}_i - \bar{x}_i'\tilde{\beta}) \right] (\hat{\sigma}_e^2 - \sigma_e^2) \right] \middle| y_i \right\} + o_p(m^{-1}) \\
&= E \left\{ [\bar{X}_i - (1 - B_i)\bar{x}_i]'(\tilde{\beta} - \beta)(\bar{y}_i - \bar{x}_i'\beta) \left[\left(1 - \frac{\partial B_i}{\partial \sigma_\nu^2}\right)(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \left(1 - \frac{\partial B_i}{\partial \sigma_e^2}\right)(\hat{\sigma}_e^2 - \sigma_e^2) \right] \middle| y_i \right\} \\
&\quad + o_p(m^{-1}) \tag{5.4}
\end{aligned}$$

We assume that $\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2 = O_p(m^{-1})$ and $\tilde{\beta}_{-i} - \beta = O_p(m^{-1})$, where $\hat{\sigma}_{\nu,-i}^2$ excludes observations of i -th area to calculate $\hat{\sigma}_\nu^2$ and similarly $\tilde{\beta}_{-i} = (\sum_i X_i' V_i^{-1} X_i)^{-1} \cdot (\sum_{j \neq i} X_j' V_j^{-1} y_j)$. Then (5.4) equals to

$$\begin{aligned}
&(\bar{y}_i - \bar{x}_i'\beta)[\bar{X}_i - (1 - B_i)\bar{x}_i]' \left\{ \left(1 - \frac{\partial B_i}{\partial \sigma_\nu^2}\right) E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2) | y_i] + \left(1 - \frac{\partial B_i}{\partial \sigma_e^2}\right) \right. \\
&\quad \left. \cdot E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{e,-i}^2 - \sigma_e^2) | y_i] \right\} + o_p(m^{-1}) \\
&= (\bar{y}_i - \bar{x}_i'\beta)[\bar{X}_i - (1 - B_i)\bar{x}_i]' \left\{ \left(1 - \frac{\partial B_i}{\partial \sigma_\nu^2}\right) E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2)] \right. \\
&\quad \left. + \left(1 - \frac{\partial B_i}{\partial \sigma_e^2}\right) E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{e,-i}^2 - \sigma_e^2)] \right\} + o_p(m^{-1}) = o_p(m^{-1}),
\end{aligned}$$

where $E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2)] = 0$ and $E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{e,-i}^2 - \sigma_e^2)] = 0$, since it is third moment and due to normality. Therefore,

$$E[(\tilde{\mu}_i - \mu_i)(\hat{\mu}_i - \tilde{\mu}_i) | y_i] = o_p(m^{-1}).$$

Theorem 5.1. *The conditional mean squared error of the EBLUP $\hat{\mu}_i$ given y_i , is obtained as*

$$MSE^c(\hat{\mu}_i | y_i) = g_{1i}(\delta) + g_{2i}(\delta) + g_{3i}^c(\theta) + o_p(m^{-1}).$$

5.3 Estimation of Conditional MSE

We now derive a second order unbiased estimator of the conditional MSE of $\hat{\mu}_i$ given y_i . By Taylor expansion, we may write

$$\begin{aligned} E[g_{1i}(\hat{\delta})|y_i] &= g_{1i}(\delta) + \frac{\partial g_{1i}(\delta)}{\partial \sigma_\nu^2} E(\hat{\sigma}_\nu^2 - \sigma_\nu^2|y_i) + \frac{\partial g_{1i}(\delta)}{\partial \sigma_e^2} E(\hat{\sigma}_e^2 - \sigma_e^2|y_i) + \frac{1}{2} E \left\{ \left[\frac{\partial^2 g_{1i}(\delta)}{\partial \sigma_\nu^4} \right. \right. \\ &\quad \left. \left. + \frac{\partial^2 g_{1i}(\delta)}{\partial \sigma_e^4} (\hat{\sigma}_e^2 - \sigma_e^2)^2 + 2 \frac{\partial^2 g_{1i}(\delta)}{\partial \sigma_\nu^2 \partial \sigma_e^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2)(\hat{\sigma}_e^2 - \sigma_e^2) \right] \middle| y_i \right\} + o_p(m^{-1}) \\ &= g_{1i}(\delta) + B_i^2 b(\hat{\sigma}_\nu^2|y_i) + \frac{(1 - B_i)^2}{n_i} b(\hat{\sigma}_e^2|y_i) - g_{3i}(\delta) + o_p(m^{-1}). \end{aligned}$$

It will be shown in Sections 5.7.4 and 5.7.5 that $g_{2i}(\hat{\delta})$ and $g_{3i}^c(\hat{\theta})$ are correct to order $O_p(m^{-1})$, in the sense that $E[g_{2i}(\hat{\delta})|y_i] = g_{2i}(\delta) + o_p(m^{-1})$ and $E[g_{3i}^c(\hat{\theta})|y_i] = g_{3i}^c(\theta) + o_p(m^{-1})$.

Hence, we have the following theorem.

Theorem 5.2. *A second order unbiased estimator of $MSE^c(\hat{\mu}_i|y_i)$ is given by*

$$mse^c(\hat{\mu}_i|y_i) = g_{1i}(\hat{\delta}) + g_{2i}(\hat{\delta}) + g_{3i}(\hat{\delta}) + g_{3i}^c(\hat{\theta}) - \hat{B}_i^2 \hat{b}(\hat{\sigma}_\nu^2|y_i) - \frac{(1 - \hat{B}_i)^2}{n_i} \hat{b}(\hat{\sigma}_e^2|y_i), \quad (5.5)$$

i.e.,

$$E[mse^c(\hat{\mu}_i|y_i)] + o_p(m^{-1}) = MSE^c(\hat{\mu}_i|y_i).$$

In Section 5.4, we find the conditional bias of $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$.

5.4 Conditional Bias of $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$

The fitting-of-constants estimators $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$ can be written as $\hat{\sigma}_e^2 = SSE_1/\nu$ and $\hat{\sigma}_\nu^2 = \eta^{-1}[SSE_2 - (n - p)\hat{\sigma}_e^2]$ where $\eta = n - \sum_i n_i^2 \bar{x}_i' (\sum_i X_i' X_i)^{-1} \bar{x}_i$, $\nu = n - m - p_1$, p_1 is the number of non-zero $x_{ij}^* (= x_{ij} - \bar{x}_i)$, SSE_1 is the residual sum of squares by regressing through the origin $y_{ij}^* (= y_{ij} - \bar{y}_i)$ on nonzero x_{ij}^* for those areas with $n_i > 1$. In addition, $SSE_2 = \sum_i \sum_j (y_{ij} - x_{ij}' \tilde{\beta}_{ols})^2$ where $\tilde{\beta}_{ols} = (\sum_i X_i' X_i)^{-1} (\sum_i X_i' y_i)$.

In fact, we can write (5.1) in matrix notation by $y = X\beta + D\nu + e$, where y is a $n \times 1$ vector with elements $y_{ij} (j = 1, \dots, n_i; i = 1, \dots, m)$, X is a $n \times p$ matrix with elements X_i ,

$$D = \begin{pmatrix} 1_{n_1} & 0 & 0 & \dots & 0 \\ 0 & 1_{n_2} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & 0 & \dots & 1_{n_m} \end{pmatrix},$$

with dimension $n \times m$, 1_{n_i} is a $n_i \times 1$ vector of one's, ν is a $m \times 1$ vector with elements ν_i , e is a $n \times 1$ vector with elements e_{ij} . We may then write $SSE_1 = y^{*'} [I_n - X^* (X^{*'} X^*)^{-1} X^{*'}] y^*$ where $y^* = [I_n - D(D'D)^{-1} D'] y$ is a $n \times 1$ vector with elements $y_{ij}^* (j = 1, \dots, n_i; i = 1, \dots, m)$, I_n is an $n \times n$ identity matrix and $X^* = [I_n - D(D'D)^{-1} D'] X$ is a $n \times p$ matrix with elements $x_{ij}^* (j = 1, \dots, n_i; i = 1, \dots, m)$. Then we may write $E(y^* | y_i) = X^* \beta + a_i (y_i^* - X_i^* \beta)$ and $Var(y^* | y_i) = V^* - a_i V_i^* a_i'$ where $V^* = Var(y^*) = diag(V_1^*, \dots, V_m^*)$ where $V_i^* = [I_{n_i} - D_i(D'D)^{-1} D_i'] V_i [I_{n_i} - D_i(D'D)^{-1} D_i']$ and $a_i = (0, 0, \dots, I_{n_i}, \dots, 0, 0)'$ with dimension $n \times n_i$. Therefore, we have that

$$E(SSE_1 | y_i) = E \left\{ y^{*'} [I_n - X^* (X^{*'} X^*)^{-1} X^{*'}] y^* | y_i \right\}$$

$$\begin{aligned}
&= \text{tr} \left\{ \left[I_n - X^*(X^{*'}X^*)^{-1}X^{*'} \right] V^* \right\} \\
&\quad - \text{tr} \left\{ a_i' \left[I_n - X^*(X^{*'}X^*)^{-1}X^{*'} \right] a_i V_i^* \right\} \\
&\quad + \text{tr} \left\{ a_i' \left[I_n - X^*(X^{*'}X^*)^{-1}X^{*'} \right] a_i (y_i^* - X_i^* \beta) (y_i^* - X_i^* \beta)' \right\} \\
&= \sum_i (n_i - 1) \sigma_e^2 - p_1 \sigma_e^2 + \text{tr} \left\{ a_i' \left[I_n - X^*(X^{*'}X^*)^{-1}X^{*'} \right] a_i \right. \\
&\quad \cdot \left. \left[(y_i^* - X_i^* \beta) (y_i^* - X_i^* \beta)' - V_i^* \right] \right\} \\
&= \sigma_e^2 [n - m - p_1] + \sum_{j=1}^{n_i} \left\{ \left[1 - x_{ij}^{*'} (X^{*'}X^*)^{-1} x_{ij}^* \right] \right. \\
&\quad \cdot \left. \left[(y_{ij}^* - x_{ij}^{*'} \beta)^2 - \frac{n_i - 1}{n_i} \sigma_e^2 \right] \right\}.
\end{aligned}$$

Then, we may express the conditional bias of $\hat{\sigma}_e^2$ given y_i as

$$b(\hat{\sigma}_e^2 | y_i) = \frac{\sum_{j=1}^{n_i} (y_{ij}^* - x_{ij}^{*'} \beta)^2 - (n_i - 1) \sigma_e^2}{(n - m - p_1)} + O_p(m^{-2}). \quad (5.6)$$

We now turn to the conditional bias of $\hat{\sigma}_v^2$ given y_i . We may write $SSE_2 = y'[I_n - X(X'X)^{-1}X']y$ where y is a $n \times 1$ vector and X is a $n \times p$ matrix such that model (5.1) can be written as matrix notation by $y = X\beta + v + e$, then we have that $E(y|y_i) = X\beta + a_i(y_i - X_i\beta)$ and $V(y|y_i) = V - a_i V_i a_i'$, where $V = \text{Var}(y) = \text{diag}_{1 \leq i \leq m}(V_i)$ with $V_i = \sigma_v^2 J_{n_i} + \sigma_e^2 I_{n_i}$.

Therefore, we have

$$\begin{aligned}
E(SSE_2 | y_i) &= E \left\{ y' \left[I_n - X(X'X)^{-1}X' \right] y \middle| y_i \right\} \\
&= \text{tr} \left\{ \left[I_n - X(X'X)^{-1}X' \right] V \right\} - \text{tr} \left\{ \left[I_n - X(X'X)^{-1}X' \right] a_i V_i a_i' \right\}
\end{aligned}$$

$$\begin{aligned}
& + (y_i - X_i\beta)' a_i' [I_n - X(X'X)^{-1}X'] a_i (y_i - X_i\beta) \\
= & \sigma_\nu^2 \sum_i \left[n_i - n_i^2 \bar{x}_i' (X'X)^{-1} \bar{x}_i \right] + n\sigma_e^2 - \sigma_e^2 \sum_i \text{tr} \left[X_i (X'X)^{-1} X_i' \right] \\
& + \text{tr} \left\{ a_i' [I_n - X(X'X)^{-1}X'] a_i \left[(y_i - X_i\beta)(y_i - X_i\beta)' - V_i \right] \right\} \\
= & \sigma_\nu^2 \sum_i \left[n_i - n_i^2 \bar{x}_i' (X'X)^{-1} \bar{x}_i \right] + \sigma_e^2 \left[n - \sum_i \text{tr} (X_i (X'X)^{-1} X_i') \right] \\
& + \sum_{j=1}^{n_i} \left\{ \left[1 - x_{ij}' (X'X)^{-1} x_{ij} \right] \left[(y_{ij} - x_{ij}'\beta)^2 - (\sigma_\nu^2 + \sigma_e^2) \right] \right\}. \tag{5.7}
\end{aligned}$$

Then, based on (5.6) and (5.7), we may write

$$\begin{aligned}
b(\hat{\sigma}_\nu^2 | y_i) &= E(\hat{\sigma}_\nu^2 - \sigma_\nu^2 | y_i) = \eta^{-1} E \left\{ \left[SSE_2 - (n-p)\hat{\sigma}_e^2 \right] | y_i \right\} - \sigma_\nu^2 \\
&= \eta^{-1} E(SSE_2 | y_i) - \frac{n-p}{\eta} E(\hat{\sigma}_e^2 | y_i) - \sigma_\nu^2 \\
&= \sigma_\nu^2 + \frac{n-p}{\eta} \sigma_e^2 + \eta^{-1} \sum_{j=1}^{n_i} \left\{ \left[1 - x_{ij}' (X'X)^{-1} x_{ij} \right] \left[(y_{ij} - x_{ij}'\beta)^2 - (\sigma_\nu^2 + \sigma_e^2) \right] \right\} \\
&\quad - \frac{n-p}{\eta} E(\hat{\sigma}_e^2 | y_i) - \sigma_\nu^2.
\end{aligned}$$

Then the conditional bias of $\hat{\sigma}_\nu^2$ given y_i is obtained by

$$b(\hat{\sigma}_\nu^2 | y_i) = \frac{\sum_{j=1}^{n_i} (y_{ij} - x_{ij}'\beta)^2 - n_i(\sigma_\nu^2 + \sigma_e^2)}{\eta} - \frac{n-p}{\eta} b(\hat{\sigma}_e^2 | y_i) + O_p(m^{-2}). \tag{5.8}$$

5.5 Data Analysis

To illustrate our method, we employ the real data set given by Battese, Harter and Fuller (1988). We have already used this data set in Chapter 3. We consider the estimation of mean hectares of corn per segment for 12 counties (area) in north-central Iowa. The total sample size, $n = \sum_i n_i$, for the 12 counties is 36, and the

sample size n_i within each county ranged from 1 to 5. The population size N_i within each county ranged from 402 to 965. The nested error regression model is

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \nu_i + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m$$

where y_{ij} is the number of hectares of corn in the j -th segment of the i -th county, x_{1ij} and x_{2ij} are the number of pixels classified as corn and soybeans respectively, in the j -th segment of the i -th county. The small area mean $\mu_i = \beta_0 + \beta_1 \bar{X}_{1i} + \beta_2 \bar{X}_{2i} + \nu_i$, where $\bar{X}_{1i}$ and $\bar{X}_{2i}$ are the known population mean number of pixels classified as corn and soybeans per segment respectively in the i -th county, by using LANDSAT satellite readings. We estimated the variance components based on fitting-of-constants method and regression parameters by generalized least squared method. We found $\hat{\sigma}_\nu^2 = 139.7$, $\hat{\sigma}_e^2 = 149.6$ and $\hat{\beta} = (51.05, 0.33, -0.13)'$. We use the formula (5.5) to evaluate the conditional MSE estimation.

Booth and Hobert (1998) studied the conditional MSE estimation based on ML approach. Booth and Hobert (1998) used the parametric bootstrap to get the bias term needed in MSE estimation. Booth and Hobert (1998) also applied their approach to the data set given by Battese, Harter and Fuller (1988). To compare our approach with the competitor, we report the results in Table 5.1. As we can see from Table 5.1, our leading term, g_{1i} , is larger than Booth and Hobert (1998), since we used the fitting-of-constants method to estimate variance components ($\hat{\sigma}_\nu^2 = 139.7$, $\hat{\sigma}_e^2 = 149.6$) in contrast with ML estimation ($\hat{\sigma}_\nu^2 = 121.1$, $\hat{\sigma}_e^2 = 137.3$) used by Booth and Hobert (1998); note that ML estimators of σ_ν^2 and σ_e^2 may be significantly biased here because $m = 12$ is not large enough. However, our g_{2i} and g_{3i}^c are close in most counties to the ones used in Booth and Hobert (1998). Moreover, our bias term is significantly smaller than the bias term presented by Booth and Hobert

(1998). For example, for the county Worth in Table 5.1, $Bias_{TR} = 2.42$ compared to $Bias_{BH} = 11.93$. Our second order conditional unbiased estimators of MSE are smaller than the corresponding MSE estimation used by Booth and Hobert (1998) for most counties; for example, for the county Worth in Table 5.1, $mse_{T-R} = 92.58$ compared to $mse_{B-H} = 100.03$.

5.6 Summary

In this Chapter, we studied the conditional mean squared error of area mean given the area-specific response data under a nested error linear regression model. We derived a second order approximation to the conditional MSE of the EBLUP estimator of a small area mean. We then derived an estimator of the conditional MSE that is conditionally unbiased to second order. We justified the accuracy of approximation to the conditional MSE by showing that any neglected terms are of order $o(m^{-1})$, for large m . A data analysis on the performance of our approach is also provided.

5.7 Proofs

5.7.1 Proof of $\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} = O_p(m^{-1/2})$

We may write $\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}$ as

$$\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} = -\left(\sum_i X_i' V_i^{-1} X_i\right)^{-1} \frac{\partial}{\partial \sigma_\nu^2} \left[\left(\sum_i X_i' V_i^{-1} X_i\right) \right] \left(\sum_i X_i' V_i^{-1} X_i\right)^{-1} \left(\sum_i X_i' V_i^{-1} y_i\right)$$

$$+(\sum_i X_i' V_i^{-1} X_i)^{-1} \frac{\partial}{\partial \sigma_\nu^2} (\sum_i X_i' V_i^{-1} y_i),$$

where $V_i^{-1} = \sigma_e^{-2} [I_{n_i} - \frac{(1-B_i)}{n_i} J_{n_i}]$. Then $\frac{\partial V_i^{-1}}{\partial \sigma_e^2} = -B_i^2 J_{n_i} / \sigma_e^4$, so $\frac{\partial}{\partial \sigma_e^2} (\sum_i X_i' V_i^{-1} X_i) = -\sigma_e^{-4} [\sum_i B_i^2 X_i' J_{n_i} X_i]$ and $\frac{\partial}{\partial \sigma_e^2} (\sum_i X_i' V_i^{-1} y_i) = -\sigma_e^{-4} [\sum_i B_i^2 X_i' J_{n_i} y_i]$. Therefore, we have

$$\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} = -(\sum_i X_i' V_i^{-1} X_i)^{-1} \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} (y_i - X_i \tilde{\beta}). \quad (5.9)$$

To complete the proof, we must show that $Var(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}) = O(m^{-1})$, noting that $E(\partial \tilde{\beta} / \partial \sigma_\nu^2) = 0$. Hence, we may write

$$Var(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}) = (\sum_i X_i' V_i^{-1} X_i)^{-1} Var \left[\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} (y_i - X_i \tilde{\beta}) \right] \\ \cdot (\sum_i X_i' V_i^{-1} X_i)^{-1},$$

where $Var \left[\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} (y_i - X_i \tilde{\beta}) \right] = \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-4} X_i' J_{n_i} \\ \cdot [V_i - X_i (\sum_i X_i' V_i^{-1} X_i)^{-1} X_i'] J_{n_i} X_i$.

Then, we have that

$$Var(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}) = (\sum_i X_i' V_i^{-1} X_i)^{-1} \left[\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-4} (J_{n_i} X_i)' V_i (J_{n_i} X_i) - \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-4} \right. \\ \left. \cdot (X_i' J_{n_i} X_i)' (\sum_i X_i' V_i^{-1} X_i)^{-1} (X_i' J_{n_i} X_i) \right] (\sum_i X_i' V_i^{-1} X_i)^{-1} = O(m^{-1}). \quad (5.10)$$

As a result, $\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} = O_p(m^{-1/2})$.

5.7.2 Proof of $\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = O_p(m^{-1/2})$

We may write $\frac{\partial \tilde{\beta}}{\partial \sigma_e^2}$ as

$$\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = -(\sum_i X_i' V_i^{-1} X_i)^{-1} \frac{\partial}{\partial \sigma_e^2} \left[(\sum_i X_i' V_i^{-1} X_i) \right] (\sum_i X_i' V_i^{-1} X_i)^{-1} (\sum_i X_i' V_i^{-1} y_i)$$

$+(\sum_i X_i' V_i^{-1} X_i)^{-1} \frac{\partial}{\partial \sigma_e^2} (\sum_i X_i' V_i^{-1} y_i)$
 where $\frac{\partial V_i^{-1}}{\partial \sigma_e^2} = -\sigma_e^{-2} V_i^{-1} + n_i^{-1} \sigma_e^{-4} B_i(1 - B_i) J_{n_i}$, then $\frac{\partial}{\partial \sigma_e^2} (\sum_i X_i' V_i^{-1} X_i)$
 $= -\sigma_e^{-2} (\sum_i X_i' V_i^{-1} X_i) + \sigma_e^{-4} \sum_i \frac{B_i(1-B_i)}{n_i} X_i' J_{n_i} X_i$ and $\frac{\partial}{\partial \sigma_e^2} (\sum_i X_i' V_i^{-1} y_i) = -\sigma_e^{-2}$
 $\cdot (\sum_i X_i' V_i^{-1} y_i) + \sigma_e^{-4} \sum_i \frac{B_i(1-B_i)}{n_i} X_i' J_{n_i} y_i$. Therefore, we have

$$\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = \sigma_e^{-4} (\sum_i X_i' V_i^{-1} X_i)^{-1} \sum_i \frac{B_i(1-B_i)}{n_i} X_i' J_{n_i} (y_i - X_i \tilde{\beta}). \quad (5.11)$$

To complete the proof, we must show that $Var(\frac{\partial \tilde{\beta}}{\partial \sigma_e^2}) = O(m^{-1})$, noting that
 $E(\partial \tilde{\beta} / \partial \sigma_e^2) = 0$. However, comparing (5.9) to (5.11), we have $\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = \frac{-\sigma_e^2}{\sigma_e^2} \frac{\partial \tilde{\beta}}{\partial \sigma_e^2}$, so we
 can write by (5.10) that

$$Var(\frac{\partial \tilde{\beta}}{\partial \sigma_e^2}) = (\frac{\sigma_e^2}{\sigma_e^2})^2 Var(\frac{\partial \tilde{\beta}}{\partial \sigma_e^2}) = O(m^{-1}).$$

As a result, $\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = O_p(m^{-1/2})$, then the result is desired.

5.7.3 Proof of $E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] = E(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 + o_p(m^{-1})$

We assume $\hat{\sigma}_{\nu,-i}^2 - \hat{\sigma}_\nu^2 = O_p(m^{-1})$ where $\hat{\sigma}_{\nu,-i}^2$ is $\hat{\sigma}_\nu^2$ without the information of the
i-th area. Therefore, we may write

$$\begin{aligned}
 E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] &= E[(\hat{\sigma}_\nu^2 - \hat{\sigma}_{\nu,-i}^2)^2 + (\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2)^2 + 2(\hat{\sigma}_\nu^2 - \hat{\sigma}_{\nu,-i}^2)(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2) | y_i] \\
 &= E(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2)^2 + o_p(m^{-1}) = E(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 + o_p(m^{-1}),
 \end{aligned}$$

then the result is established.

The proof of $E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] = E(\hat{\sigma}_e^2 - \sigma_e^2)^2 + o_p(m^{-1})$ is very similar as $E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] = E(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 + o_p(m^{-1})$.

5.7.4 Proof of $E[g_{2i}(\hat{\delta})|\mathbf{y}_i] = g_{2i}(\delta) + o_p(m^{-1})$

We may write $g_{2i}(\delta)$ by Taylor expansion around δ , as follows

$$\begin{aligned} g_{2i}(\hat{\delta}) - g_{2i}(\delta) &= (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \frac{\partial}{\partial \sigma_\nu^2} \left[(\bar{X}'_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}(\bar{X}'_i - (1 - B_i)\bar{x}_i) \right] \Big|_{\delta=\delta^*} \\ &\quad + (\hat{\sigma}_e^2 - \sigma_e^2) \frac{\partial}{\partial \sigma_e^2} \left[(\bar{X}'_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}(\bar{X}'_i - (1 - B_i)\bar{x}_i) \right] \Big|_{\delta=\delta^*} \end{aligned}$$

and we know that

$$\begin{aligned} \frac{\partial V^{-1}}{\partial \sigma_\nu^2} &= -\sigma_\nu^{-2}\sigma_e^{-2} \text{diag}_i \left[\frac{B_i(1-B_i)}{n_i} J_{n_i} \right], \quad \frac{\partial V^{-1}}{\partial \sigma_e^2} = -\sigma_e^{-4} \text{diag}_i \left[I_{n_i} - \frac{(1+B_i)(1-B_i)}{n_i} J_{n_i} \right]. \text{ Then} \\ \frac{\partial}{\partial \sigma_\nu^2} \left[(\bar{X}'_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}(\bar{X}'_i - (1 - B_i)\bar{x}_i) \right] &= \frac{-2B_i(1 - B_i)}{\sigma_\nu^2} \bar{x}'_i (X'V^{-1}X)^{-1} \\ &\quad \cdot (\bar{X}_i - (1 - B_i)\bar{x}_i) + \sigma_\nu^{-2}\sigma_e^{-2} (\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}X' \\ &\quad \cdot \text{diag}_i \left[\frac{B_i(1 - B_i)}{n_i} J_{n_i} \right] X (X'V^{-1}X)^{-1} (\bar{X}_i - (1 - B_i)\bar{x}_i), \end{aligned}$$

and

$$\begin{aligned} \frac{\partial}{\partial \sigma_e^2} \left[(\bar{X}'_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}(\bar{X}'_i - (1 - B_i)\bar{x}_i) \right] &= \frac{2B_i(1 - B_i)}{\sigma_e^2} \bar{x}'_i (X'V^{-1}X)^{-1} \\ &\quad \cdot (\bar{X}_i - (1 - B_i)\bar{x}_i) + \sigma_e^{-4} (\bar{X}_i - (1 - B_i)\bar{x}_i)'(X'V^{-1}X)^{-1}X' \text{diag}_i \left[I_{n_i} - \frac{(1 + B_i)(1 - B_i)}{n_i} J_{n_i} \right] \\ &\quad \cdot X (X'V^{-1}X)^{-1} (\bar{X}_i - (1 - B_i)\bar{x}_i), \end{aligned}$$

$$\text{then } |g_{2i}(\hat{\delta}) - g_{2i}(\delta)| < k_0 \left| \left[\bar{x}'_i + (\bar{X}_i - (1 - B_i^*)\bar{x}_i)'(X'V_*^{-1}X)^{-1}X'X + (\bar{X}_i - (1 - B_i^*)\bar{x}_i)' \right] (X'V_*^{-1}X)^{-1}(\bar{X}_i - (1 - B_i^*)\bar{x}_i) \right| \left[|\hat{\sigma}_\nu^2 - \sigma_\nu^2| + |\hat{\sigma}_e^2 - \sigma_e^2| \right],$$

where k_0 is a generic constant, $B_i^* = B_i|_{\delta=\delta^*}$ and $V_*^{-1} = V^{-1}|_{\delta=\delta^*}$. Hence,

$$\begin{aligned} E \left[|g_{2i}(\hat{\delta}) - g_{2i}(\delta)| | \mathbf{y}_i \right] &\leq \max_{1 \leq i \leq m} \left\{ \left[\bar{x}'_i + (\bar{X}_i - (1 - B_i)\bar{x}_i)'(I + (X'V^{-1}X)^{-1}(X'X)) \right] \right. \\ &\quad \left. \cdot (X'V^{-1}X)^{-1}(\bar{X}_i - (1 - B_i)\bar{x}_i) \right\} E \left\{ \left[|\hat{\sigma}_\nu^2 - \sigma_\nu^2| + |\hat{\sigma}_e^2 - \sigma_e^2| \right] | \mathbf{y}_i \right\}, \end{aligned}$$

$$\begin{aligned}
E \left[\left| g_{2i}(\hat{\delta}) - g_{2i}(\delta) \right| | y_i \right] &\leq \max_{1 \leq i \leq m} \left\{ \left[\bar{x}'_i + (\bar{X}_i - (1 - B_i)\bar{x}_i)'(I + (X'V^{-1}X)^{-1}(X'X)) \right] \right. \\
&\quad \left. \cdot (X'V^{-1}X)^{-1}(\bar{X}_i - (1 - B_i)\bar{x}_i) \right\} \left\{ \left(E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] \right)^{1/2} + \left(E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] \right)^{1/2} \right\} \\
&= o_p(m^{-1}), \text{ by Holder's inequality and } E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] = \text{Var}(\hat{\sigma}_\nu^2) + o_p(m^{-1}) = \\
&O_p(m^{-1}) \text{ by Section 5.7.3 and similarly } E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] = O_p(m^{-1}). \text{ In addition,}
\end{aligned}$$

$$\begin{aligned}
\max_{1 \leq i \leq m} \left\{ \left[\bar{x}'_i + (\bar{X}_i - (1 - B_i)\bar{x}_i)'(I + (X'V^{-1}X)^{-1}(X'X)) \right] (X'V^{-1}X)^{-1} \right. \\
\left. \cdot (\bar{X}_i - (1 - B_i)\bar{x}_i) \right\} = O(m^{-1}),
\end{aligned}$$

so we have that $E[g_{2i}(\hat{\delta}) | y_i] = g_{2i}(\delta) + o_p(m^{-1})$, then the result is desired.

5.7.5 Proof of $E[g_{3i}^c(\hat{\theta}) | y_i] = g_{3i}^c(\theta) + o_p(m^{-1})$

We may write

$$\begin{aligned}
g_{3i}^c(\hat{\theta}) - g_{3i}^c(\theta) &= \frac{(\bar{y}_i - \bar{x}'_i \hat{\beta})^2}{\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2/n_i} \left[\hat{\sigma}_e^4 \widehat{\text{Var}}(\hat{\sigma}_\nu^2) + \hat{\sigma}_\nu^4 \widehat{\text{Var}}(\hat{\sigma}_e^2) - 2\hat{\sigma}_\nu^2 \hat{\sigma}_e^2 \widehat{\text{Cov}}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \right] \\
&\quad - \frac{(\bar{y}_i - \bar{x}'_i \beta)^2}{\sigma_\nu^2 + \sigma_e^2/n_i} \left[\sigma_e^4 \text{Var}(\hat{\sigma}_\nu^2) + \sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2) - 2\sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \right],
\end{aligned}$$

where $\text{Var}(\hat{\sigma}_\nu^2) = 2\eta_1^{-2}[\nu_1^{-1}(n - p - \nu_1)(n - p)\sigma_e^4 + \eta_2\sigma_\nu^4 + 2\eta_1\sigma_\nu^2\sigma_e^2]$, $\text{Var}(\hat{\sigma}_e^2) = 2\nu_1^{-1}\sigma_e^4$, $\text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) = -2\eta_1^{-1}\nu_1^{-1}(n - p - \nu_1)\sigma_e^4$ with $\nu_1 = n - m - p + 1$, $\eta_1 = n - \sum_i n_i^2 \bar{x}'_i (X'X)^{-1} \bar{x}_i$, $\eta_2 = \sum_i n_i^2 (1 - n_i \bar{x}'_i (X'X)^{-1} \bar{x}_i) + \text{tr} \left[(X'X)^{-1} \sum_i n_i^2 \bar{x}_i \bar{x}'_i \right]^2$.

Hence,

$$g_{3i}^c(\hat{\theta}) - g_{3i}^c(\theta) = \frac{(\bar{y}_i - \bar{x}'_i \hat{\beta})^2}{\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2/n_i} \left[\hat{\sigma}_e^4 \widehat{\text{Var}}(\hat{\sigma}_\nu^2) - \sigma_e^4 \text{Var}(\hat{\sigma}_\nu^2) + \hat{\sigma}_\nu^4 \widehat{\text{Var}}(\hat{\sigma}_e^2) - \sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2) \right]$$

$$\begin{aligned}
& -2\hat{\sigma}_\nu^2\hat{\sigma}_e^2\widehat{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) + 2\sigma_\nu^2\sigma_e^2Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \Big] + \Big[\sigma_e^4Var(\hat{\sigma}_\nu^2) + \sigma_\nu^4Var(\hat{\sigma}_e^2) \\
& -2\sigma_\nu^2\sigma_e^2Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \Big] \Big[\frac{(\bar{y}_i - \bar{x}'_i\hat{\beta})^2}{\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2/n_i} - \frac{(\bar{y}_i - \bar{x}'_i\beta)^2}{\sigma_\nu^2 + \sigma_e^2/n_i} \Big],
\end{aligned}$$

then we may write

$$\begin{aligned}
& \left| g_{3i}^c(\hat{\theta}) - g_{3i}^c(\theta) \right| \leq \left| (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \frac{\partial}{\partial \sigma_\nu^2} [\sigma_e^4 Var(\hat{\sigma}_\nu^2)] \Big|_{\delta=\delta^*} \right| + \left| (\hat{\sigma}_e^2 - \sigma_e^2) \frac{\partial}{\partial \sigma_e^2} [\sigma_e^4 Var(\hat{\sigma}_\nu^2)] \Big|_{\delta=\delta^*} \right| \\
& + \left| (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \frac{\partial}{\partial \sigma_\nu^2} [\sigma_\nu^4 Var(\hat{\sigma}_e^2)] \Big|_{\delta=\delta^*} \right| + \left| (\hat{\sigma}_e^2 - \sigma_e^2) \frac{\partial}{\partial \sigma_e^2} [\sigma_\nu^4 Var(\hat{\sigma}_e^2)] \Big|_{\delta=\delta^*} \right| \\
& + 2 \left| (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \frac{\partial}{\partial \sigma_\nu^2} [\sigma_\nu^2 \sigma_e^2 Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)] \Big|_{\delta=\delta^*} \right| + 2 \left| (\hat{\sigma}_e^2 - \sigma_e^2) \frac{\partial}{\partial \sigma_e^2} [\sigma_\nu^2 \sigma_e^2 Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)] \Big|_{\delta=\delta^*} \right| \\
& + \left| \left[\sigma_e^4 Var(\hat{\sigma}_\nu^2) + \sigma_\nu^4 Var(\hat{\sigma}_e^2) - 2\sigma_\nu^2 \sigma_e^2 Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \right] \frac{n_i(\bar{y}_i - \bar{x}'_i\beta^*)^2}{(n_i\sigma_\nu^{*2} + \sigma_e^{*2})^2} \right. \\
& \quad \left. \cdot \left[n_i(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + (\hat{\sigma}_e^2 - \sigma_e^2) + \frac{2(n_i\sigma_\nu^{*2} + \sigma_e^{*2})}{(\bar{y}_i - \bar{x}'_i\beta^*)} \bar{x}'_i(\hat{\beta} - \beta) \right] \right|. \tag{5.12}
\end{aligned}$$

Note that

$$\begin{aligned}
& \frac{(\bar{y}_i - \bar{x}'_i\hat{\beta})^2}{\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2/n_i} - \frac{(\bar{y}_i - \bar{x}'_i\beta)^2}{\sigma_\nu^2 + \sigma_e^2/n_i} = -n_i(\bar{y}_i - \bar{x}'_i\beta^*)^2 \left[\frac{n_i(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + (\hat{\sigma}_e^2 - \sigma_e^2)}{(n_i\sigma_\nu^{*2} + \sigma_e^{*2})^2} \right] \\
& + \frac{2(\bar{y}_i - \bar{x}'_i\beta^*)}{\sigma_\nu^{*2} + \sigma_e^{*2}/n_i} \bar{x}'_i(\hat{\beta} - \beta),
\end{aligned}$$

which is Taylor expansion around θ such that $\theta = \theta^*$ where $|\theta^* - \theta| < |\hat{\theta} - \theta|$.

We may express the terms of (5.12) as follows

$$\begin{aligned}
& \left| \frac{\partial}{\partial \sigma_\nu^2} (\sigma_e^4 Var(\hat{\sigma}_\nu^2)) \Big|_{\delta=\delta^*} \right| = \left| 4\sigma_e^{*4} \eta_1^{-2} (\eta_2 \sigma_\nu^{*2} + \eta_1 \sigma_e^{*2}) \right| \leq \left| 4\hat{\sigma}_e^4 \eta_1^{-2} (\eta_2 \hat{\sigma}_\nu^2 + \eta_1 \hat{\sigma}_e^2) \right| \\
& \leq k_1 n^{-1} \left| \hat{\sigma}_e^4 (\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2) \right|, \tag{5.13}
\end{aligned}$$

and

$$\begin{aligned}
& \left| \frac{\partial}{\partial \sigma_e^2} (\sigma_e^4 \text{Var}(\hat{\sigma}_\nu^2)) \right|_{\delta=\delta^*} \\
&= \left| 2\sigma_e^2 \text{Var}(\hat{\sigma}_\nu^2) + 2\sigma_e^4 \eta_1^{-2} (2\nu_1^{-1}(n-p)(n-p-\nu_1)\sigma_e^2 + 2\eta_1 \sigma_\nu^2) \right|_{\delta=\delta^*} \\
&= \left| 4\sigma_e^{*2} \eta_1^{-2} (\nu_1^{-1}(n-p)(n-p-\nu_1)\sigma_e^{*4} + \eta_2 \sigma_\nu^{*4} + 2\eta_1 \sigma_e^{*2} \sigma_\nu^{*2}) \right. \\
&\quad \left. + 2\sigma_e^{*4} \eta_1^{-2} (2\nu_1^{-1}(n-p)(n-p-\nu_1)\sigma_e^{*2} + 2\eta_1 \sigma_\nu^{*2}) \right| \\
&\leq k_2 n^{-1} \left| \sigma_e^{*2} (\sigma_e^{*4} + \sigma_\nu^{*4} + \sigma_\nu^{*2} \sigma_e^{*2}) \right| \leq k_2 n^{-1} \left| \hat{\sigma}_e^2 (\hat{\sigma}_e^4 + \hat{\sigma}_\nu^4 + \hat{\sigma}_e^2 \hat{\sigma}_\nu^2) \right|. \tag{5.14}
\end{aligned}$$

Also,

$$\left| \frac{\partial}{\partial \sigma_\nu^2} (\sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2)) \right|_{\delta=\delta^*} = \left| 2\sigma_\nu^2 \text{Var}(\hat{\sigma}_e^2) \right|_{\delta=\delta^*} = \left| 4\sigma_\nu^{*2} \nu_1^{-1} \sigma_e^{*4} \right| \leq k_3 n^{-1} \left| \hat{\sigma}_\nu^2 \hat{\sigma}_e^4 \right|. \tag{5.15}$$

In addition,

$$\left| \frac{\partial}{\partial \sigma_e^2} (\sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2)) \right|_{\delta=\delta^*} = \left| 4\nu_1^{-1} \sigma_e^2 \sigma_\nu^4 \right|_{\delta=\delta^*} \leq k_4 n^{-1} \left| \hat{\sigma}_e^2 \hat{\sigma}_\nu^4 \right|. \tag{5.16}$$

Moreover,

$$\left| \frac{\partial}{\partial \sigma_\nu^2} [\sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)] \right|_{\delta=\delta^*} = \left| \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \right|_{\delta=\delta^*} \leq k_5 n^{-1} \left| \hat{\sigma}_e^6 \right|, \tag{5.17}$$

and

$$\begin{aligned}
\left| \frac{\partial}{\partial \sigma_e^2} [\sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)] \right|_{\delta=\delta^*} &= \left| \sigma_\nu^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) - 4\sigma_\nu^2 \sigma_e^4 \eta_1^{-1} \nu_1^{-1} (n-p-\nu_1) \right|_{\delta=\delta^*} \\
&\leq k_6 n^{-1} \left| \hat{\sigma}_\nu^2 \hat{\sigma}_e^4 \right|, \tag{5.18}
\end{aligned}$$

where $k_1, \dots, k_6$ are known. Hence, by (5.13), we may write

$$\begin{aligned}
& E \left| (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \frac{\partial}{\partial \sigma_\nu^2} (\sigma_e^4 \text{Var}(\hat{\sigma}_\nu^2)) \right|_{\delta=\delta^*} |y_i| \leq k_1 n^{-1} E \left| (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \hat{\sigma}_e^4 (\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2) \right| |y_i| \\
&\leq k_1 n^{-1} \left\{ E \left[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 |y_i| \right] E \left[(\hat{\sigma}_e^4 (\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2))^2 |y_i| \right] \right\}^{1/2} = o_p(m^{-1}), \tag{5.19}
\end{aligned}$$

by Holder's inequality and noting that $E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2|y_i] = O_p(m^{-1})$ and

$$E[(\hat{\sigma}_e^4(\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2))^2|y_i] = O_p(1), \quad (5.20)$$

since $E[(\hat{\sigma}_e^4(\hat{\sigma}_\nu^2 + \hat{\sigma}_e^2))^2|y_i] \leq 2\sqrt{2} \left\{ E(\hat{\sigma}_e^{16}|y_i) [E(\hat{\sigma}_\nu^8|y_i) + E(\hat{\sigma}_e^8|y_i)] \right\}^{1/2} = O_p(1)$.
Because, we have that

$$\begin{aligned} E(\hat{\sigma}_e^{2s}|y_i) &= E[(\hat{\sigma}_e^2 - \sigma_e^2 + \sigma_e^2)^{2s}|y_i] \\ &\leq 2^{2s-1} \left\{ E[(\hat{\sigma}_e^2 - \sigma_e^2)^{2s}|y_i] + \sigma_e^{4s} \right\} = O_p(1), \end{aligned}$$

since $E[(\hat{\sigma}_e^2 - \sigma_e^2)^{2s}|y_i] = O_p(m^{-s})$ for $s \in \mathbb{N}, s > 1$.

Also, similar as $E(\hat{\sigma}_e^{2s}|y_i)$, we have $E(\hat{\sigma}_\nu^{2s}|y_i) = O_p(1)$, then $E(\hat{\sigma}_e^{16}|y_i)$, $E(\hat{\sigma}_e^8|y_i)$ and $E(\hat{\sigma}_\nu^8|y_i)$ are of order $O_p(1)$, then (5.20) is desired.

Also, by (5.14),

$$\begin{aligned} E \left| (\hat{\sigma}_e^2 - \sigma_e^2) \frac{\partial}{\partial \sigma_e^2} [\sigma_e^4 \text{Var}(\hat{\sigma}_\nu^2)] \Big|_{\delta=\delta^*} \Big| y_i \right| &\leq n^{-1} k_2 E \left| (\hat{\sigma}_e^2 - \sigma_e^2) \hat{\sigma}_e^2 (\hat{\sigma}_e^4 + \hat{\sigma}_\nu^4 + \hat{\sigma}_e^2 \hat{\sigma}_\nu^2) \Big| y_i \right| \\ &\leq k_2 n^{-1} \left\{ E[(\hat{\sigma}_e^2 - \sigma_e^2)^2|y_i] E[(\hat{\sigma}_e^2(\hat{\sigma}_e^4 + \hat{\sigma}_\nu^4 + \hat{\sigma}_e^2 \hat{\sigma}_\nu^2))^2|y_i] \right\}^{1/2} = o_p(m^{-1}), \quad (5.21) \end{aligned}$$

since $E[(\hat{\sigma}_e^2 - \sigma_e^2)^2|y_i] = O_p(m^{-1})$ and the second term of right hand side of (5.21) is of order $O_p(1)$, because of the same argument of pervious case. In addition, by (5.15), we have that

$$\begin{aligned} E \left| (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \frac{\partial}{\partial \sigma_\nu^2} (\sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2)) \Big|_{\delta=\delta^*} \Big| y_i \right| &\leq n^{-1} k_3 E \left| (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \hat{\sigma}_\nu^2 \hat{\sigma}_e^4 \Big| y_i \right| \\ &\leq n^{-1} k_3 \left\{ E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2|y_i] E(\hat{\sigma}_\nu^4 \hat{\sigma}_e^8|y_i) \right\}^{1/2} \end{aligned}$$

$$\leq n^{-1}k_3 \left\{ E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] \left[E(\hat{\sigma}_\nu^8 | y_i) E(\hat{\sigma}_e^{16} | y_i) \right]^{1/2} \right\}^{1/2} = o_p(m^{-1}). \quad (5.22)$$

Moreover, by (5.16) we may write

$$\begin{aligned} E \left| (\hat{\sigma}_e^2 - \sigma_e^2) \frac{\partial}{\partial \sigma_e^2} (\sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2)) \right|_{\delta=\delta^*} | y_i | &\leq k_4 n^{-1} E \left| (\hat{\sigma}_e^2 - \sigma_e^2) \hat{\sigma}_e^2 \hat{\sigma}_\nu^4 | y_i \right| \\ &\leq n^{-1} k_4 \left\{ E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] E(\hat{\sigma}_e^4 \hat{\sigma}_\nu^8 | y_i) \right\}^{1/2} \\ &\leq k_4 n^{-1} \left\{ E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] \left[E(\hat{\sigma}_e^8 | y_i) E(\hat{\sigma}_\nu^{16} | y_i) \right]^{1/2} \right\}^{1/2} = o_p(m^{-1}), \end{aligned} \quad (5.23)$$

and by (5.17), we have that

$$\begin{aligned} E \left| (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \frac{\partial}{\partial \sigma_\nu^2} (\sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)) \right|_{\delta=\delta^*} | y_i | &\leq k_5 n^{-1} E \left| (\hat{\sigma}_\nu^2 - \sigma_\nu^2) \hat{\sigma}_e^6 | y_i \right| \\ &\leq k_5 n^{-1} \left\{ E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] E(\hat{\sigma}_e^{12} | y_i) \right\}^{1/2} = o_p(m^{-1}), \end{aligned} \quad (5.24)$$

and also by (5.18), we may have

$$\begin{aligned} E \left| (\hat{\sigma}_e^2 - \sigma_e^2) \frac{\partial}{\partial \sigma_e^2} (\sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)) \right|_{\delta=\delta^*} | y_i | &\leq k_6 n^{-1} E \left[|\hat{\sigma}_e^2 - \sigma_e^2| \hat{\sigma}_\nu^2 \hat{\sigma}_e^4 | y_i \right] \\ &\leq k_6 n^{-1} \left\{ E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] E(\hat{\sigma}_\nu^4 \hat{\sigma}_e^8 | y_i) \right\}^{1/2} \\ &\leq k_6 n^{-1} \left\{ E[(\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] \left[E(\hat{\sigma}_\nu^8 | y_i) E(\hat{\sigma}_e^{16} | y_i) \right]^{1/2} \right\}^{1/2} = o_p(m^{-1}), \end{aligned} \quad (5.25)$$

and we may write the last term of (5.12) as

$$\begin{aligned} E \left\{ \left[\sigma_e^4 \text{Var}(\hat{\sigma}_\nu^2) + \sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2) - 2\sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \right] \frac{n_i(\bar{y}_i - \bar{x}_i' \beta^*)^2}{(n_i \sigma_\nu^{*2} + \sigma_e^{*2})^2} \right. \\ \left. \cdot \left[n_i(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + (\hat{\sigma}_e^2 - \sigma_e^2) + \frac{2(n_i \sigma_\nu^{*2} + \sigma_e^{*2})}{(\bar{y}_i - \bar{x}_i' \beta^*)} \bar{x}_i'(\hat{\beta} - \beta) \right] \right| | y_i \right\} \end{aligned}$$

$$\leq k_7 \left| \left[\sigma_e^4 \text{Var}(\hat{\sigma}_\nu^2) + \sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2) - 2\sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \right] \right. \\ \left. \cdot \left\{ E(\hat{\sigma}_\nu^2 - \sigma_\nu^2 | y_i) + E(\hat{\sigma}_e^2 - \sigma_e^2 | y_i) + E[\bar{x}'_i(\hat{\beta} - \beta) | y_i] \right\} \right| = O_p(m^{-2}), \quad (5.26)$$

where k_7 is known and noting that $\sigma_e^4 \text{Var}(\hat{\sigma}_\nu^2) + \sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2) - 2\sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) = O(m^{-1})$, $E(\hat{\sigma}_\nu^2 - \sigma_\nu^2 | y_i) = O_p(m^{-1})$ and $E(\hat{\sigma}_e^2 - \sigma_e^2 | y_i) = O_p(m^{-1})$. In addition, we have $E[\bar{x}'_i(\hat{\beta} - \beta) | y_i] = O_p(m^{-1})$, since we can write

$$E[\bar{x}'_i(\hat{\beta} - \beta) | y_i] = E[\bar{x}'_i(\tilde{\beta} - \beta) | y_i] + E[\bar{x}'_i(\hat{\beta} - \tilde{\beta}) | y_i],$$

where

$$\begin{aligned} E[\bar{x}'_i(\tilde{\beta} - \beta) | y_i] &= E\left[\bar{x}'_i(X'V^{-1}X)^{-1}X'V^{-1}(y - X\beta) | y_i\right] \\ &= \bar{x}'_i(X'V^{-1}X)^{-1}X'V^{-1}a_i(y_i - X_i\beta) \\ &= \bar{x}'_i(X'V^{-1}X)^{-1}X'_iV_i^{-1}(y_i - X_i\beta) = O_p(m^{-1}). \end{aligned}$$

Also,

$$E[\bar{x}'_i(\hat{\beta} - \tilde{\beta}) | y_i] = E\left[\bar{x}'_i\left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2}(\hat{\sigma}_e^2 - \sigma_e^2)\right) | y_i\right] + o_p(m^{-1}),$$

where

$$\begin{aligned} E\left[\bar{x}'_i \frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}(\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i\right] &= -\bar{x}'_i \left(\sum_i X'_i V_i^{-1} X_i \right)^{-1} \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X'_i J_{n_i} \\ &\cdot \left[(y_i - X_i \beta) b(\hat{\sigma}_\nu^2 | y_i) - X_i E[(\tilde{\beta} - \beta)(\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i] \right] = O_p(m^{-1}), \end{aligned}$$

where $b(\hat{\sigma}_\nu^2 | y_i) = O_p(m^{-1})$ and $E[(\tilde{\beta} - \beta)(\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i] = o_p(m^{-1})$.

Similarly, we can show that $E\left[\bar{x}'_i \frac{\partial \tilde{\beta}}{\partial \sigma_e^2}(\hat{\sigma}_e^2 - \sigma_e^2) | y_i\right] = O_p(m^{-1})$.

As a result, by (5.19)-(5.26), we have $E\left[g_{3i}^c(\hat{\theta}) - g_{3i}^c(\theta) | y_i\right] = o_p(m^{-1})$, i.e., $E(g_{3i}^c(\hat{\theta}) | y_i) = g_{3i}^c(\theta) + o_p(m^{-1})$. Then the result is established.

County	T-R method					B-H method				
	g_{1i}	g_{2i}	g_{3i}^c	Bias	mse_i	g_{1i}	g_{2i}	g_{3i}^c	Bias	mse_i
Cerro	72.22	9.99	0.02	8.59	90.82	64.34	8.82	0.02	17.35	90.53
Ham.	72.22	7.94	0.74	8.58	89.48	64.34	7.14	1.08	17.28	89.84
Worth	72.22	5.01	12.93	2.42	92.58	64.34	4.62	19.14	11.93	100.03
Humb.	48.70	9.09	3.03	0.37	61.19	43.81	8.14	4.62	8.71	65.28
Frank.	36.74	1.34	4.57	0.69	43.34	33.21	1.27	7.15	4.57	46.20
Pocah.	36.74	2.00	2.25	3.47	44.46	33.21	1.86	3.49	7.26	45.82
Winn.	36.74	1.82	2.01	1.56	42.13	33.21	1.67	3.14	6.52	44.54
Wrigh	36.74	3.04	0.07	3.76	43.61	33.21	2.79	0.11	9.16	45.27
Webs.	29.49	0.85	1.95	3.79	36.08	26.75	0.81	3.06	6.06	36.68
Hanc.	24.64	1.69	0.11	2.46	28.90	22.39	1.56	0.18	5.85	29.98
Koss.	24.64	0.72	2.33	2.81	30.50	22.39	0.68	3.73	4.57	31.37
Hard.	24.64	4.60	0.79	2.11	32.14	22.39	4.18	1.24	4.57	32.38

Table 5.1: The components of conditional MSE estimation of predicted hectares of corn by Battese et al. data set (1988) for our approach and Booth and Hobert approach (1998)

Chapter 6

Estimation of MSE of Weighted Predictors

For the nested error linear regression model, the EBLUP estimator is a weighted combination of the regression synthetic estimator and the survey regression estimator. The EBLUP estimator runs into difficulties when the estimated variance of the random effect, $\hat{\sigma}_\nu^2$, is zero. In this case, it reduces to the regression synthetic estimator regardless of the size of the sampling variance of the direct estimator. One way to get around this problem is to use a weighted combination, $\hat{\mu}_{iw}$, of the synthetic and survey regression estimators with fixed weights w_i and $1 - w_i$ respectively. In this Chapter, we derive second order approximations of the MSE and the conditional MSE given $y_i = (y_{i1}, \dots, y_{in_i})'$ of $\hat{\mu}_{iw}$ under the nested error linear regression model. Moreover, we derive the estimators of MSE and conditional MSE of $\hat{\mu}_{iw}$ that are nearly unbiased.

6.1 Introduction

In Chapter 5, we studied the nested error linear regression model

$$y_{ij} = x'_{ij}\beta + v_i + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m$$

where y_{ij} is the variable of interest, $x_{ij} = (x_{ij1}, \dots, x_{ijp})'$ are the unit-specific auxiliary variable data for sample element j in small area i and β is the regression parameter. The random effects v_i and e_{ij} are i.i.d. normal with zero means and variances σ_v^2 and σ_e^2 respectively. Further, n_i and m denote the sample size in area i and the number of areas respectively. The mean of i -th area is given by $\mu_i = \bar{X}'_i\beta + v_i$ where $\bar{X}_i$ is the known population mean of x_{ij} 's. The BLUP of μ_i can be expressed as $\tilde{\mu}_i(\delta) = \bar{X}'_i\tilde{\beta} + (1 - B_i)(\bar{y}_i - \bar{x}'_i\tilde{\beta})$, where $\delta = (\sigma_v^2, \sigma_e^2)'$, $B_i(\delta) = \sigma_e^2/(\sigma_e^2 + n_i\sigma_v^2)$ and $\tilde{\beta}(\delta) = (\sum_i X'_i V_i^{-1} X_i)^{-1}(\sum_i X'_i V_i^{-1} y_i)$ with X_i is a $n_i \times p$ matrix, y_i is a $n_i \times 1$ vector and $V_i = \sigma_v^2 J_{n_i} + \sigma_e^2 I_{n_i}$ with J_{n_i} as a $n_i \times n_i$ matrix of one's and I_{n_i} as an $n_i \times n_i$ identity matrix. We can use the fitting-of-constants method to estimate variance components σ_v^2 and σ_e^2 , as in Chapter 3. The weighted estimator of μ_i with fixed weights w_i and $1 - w_i$ is given by

$$\hat{\mu}_{iw} = w_i \bar{X}'_i \hat{\beta} + (1 - w_i) \left[\bar{y}_i + (\bar{X}_i - \bar{x}_i)' \hat{\beta} \right],$$

where $\hat{\beta} = \tilde{\beta}(\hat{\delta})$ and $\hat{\delta} = (\hat{\sigma}_v^2, \hat{\sigma}_e^2)'$ is the estimator of $\delta = (\sigma_v^2, \sigma_e^2)'$. Prior guesses of σ_v^2 and σ_e^2 , say σ_{v0}^2 and σ_{e0}^2 , available from past studies, may be used to construct the weight w_i as $w_i = \sigma_{e0}^2 / (n_i \sigma_{v0}^2 + \sigma_{e0}^2)$. Also, the synthetic estimator $\bar{X}'_i \hat{\beta}$ is a special case of the weighted estimator $\hat{\mu}_{iw}$ with $w_i = 1$. Note that $\hat{\mu}_{iw}$ remains model-unbiased for μ_i for any fixed weight w_i , provided the linking model holds, that is, $E(\mu_i) = \bar{X}'_i \beta$.

In Section 6.2, we derive a second order MSE approximation of the weighted estimator $\hat{\mu}_{iw}$.

6.2 Second Order Approximation of the MSE of Weighted Predictors

Following Prasad and Rao (1990), we derive a second order approximation to the MSE of the weighted predictor $\hat{\mu}_{iw}$. We have

$$MSE(\hat{\mu}_{iw}) = E(\hat{\mu}_{iw} - \mu_i)^2 = E(\hat{\mu}_{iw} - \tilde{\mu}_i)^2 + E(\tilde{\mu}_i - \mu_i)^2,$$

where

$$\begin{aligned} E(\hat{\mu}_{iw} - \tilde{\mu}_i)(\tilde{\mu}_i - \mu_i) &= E(\hat{\mu}_{iw} - \tilde{\mu}_i)(\tilde{\mu}_i - \tilde{\mu}_i^B) \\ &= E\left[(\hat{\mu}_{iw} - \tilde{\mu}_i)(\bar{X}_i - (1 - B_i)\bar{x}_i)'(\tilde{\beta} - \beta)\right] \\ &= E\left[(\hat{\mu}_{iw} - \tilde{\mu}_i)(\bar{X}_i - (1 - B_i)\bar{x}_i)'E(\tilde{\beta} - \beta|A'y)\right] \\ &= 0, \end{aligned}$$

noting that $E(\tilde{\beta} - \beta|A'y) = E(\tilde{\beta} - \beta) = 0$ because of the condition $A'X = 0$. Also,

$$E(\tilde{\mu}_i - \mu_i)^2 = g_{1i}(\delta) + g_{2i}(\delta)$$

with

$$g_{1i}(\delta) = B_i\sigma_\nu^2, \tag{6.1}$$

and

$$g_{2i}(\delta) = (\bar{X}_i - (1 - B_i)\bar{x}_i)' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} (\bar{X}_i - (1 - B_i)\bar{x}_i). \tag{6.2}$$

Also, we may write

$$\begin{aligned} E(\hat{\mu}_{iw} - \tilde{\mu}_i)^2 &= E\left[(\bar{X}_i - \bar{x}_i)'(\hat{\beta} - \tilde{\beta}) - w_i(\bar{y}_i - \bar{x}_i'\hat{\beta}) + B_i(\bar{y}_i - \bar{x}_i'\tilde{\beta})\right]^2 \\ &= E\left[(\bar{X}_i - \bar{x}_i)'(\hat{\beta} - \tilde{\beta}) + (B_i - w_i)(\bar{y}_i - \bar{x}_i'\tilde{\beta}) + w_i\bar{x}_i'(\hat{\beta} - \tilde{\beta})\right]^2 \end{aligned}$$

$$\begin{aligned}
&= E \left[((\bar{X}_i - \bar{x}_i) + w_i \bar{x}_i)' (\hat{\beta} - \tilde{\beta}) + (B_i - w_i)(\bar{y}_i - \bar{x}_i' \tilde{\beta}) \right]^2 \\
&= (B_i - w_i)^2 E(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 + (\bar{X}_i - (1 - w_i) \bar{x}_i)' E \left[(\hat{\beta} - \tilde{\beta})(\hat{\beta} - \tilde{\beta})' \right] (\bar{X}_i - (1 - w_i) \bar{x}_i) \\
&\quad + 2(B_i - w_i) E \left\{ \left[\bar{X}_i - (1 - w_i) \bar{x}_i \right]' (\hat{\beta} - \tilde{\beta})(\bar{y}_i - \bar{x}_i' \tilde{\beta}) \right\}.
\end{aligned}$$

Note that $E(\hat{\beta} - \tilde{\beta})(\hat{\beta} - \tilde{\beta})' = [o(m^{-1})]_{p \times p}$, since by using Taylor expansion about $\hat{\beta}$, we may write

$$\hat{\beta} = \tilde{\beta} + \frac{\partial \tilde{\beta}}{\partial \delta'} (\hat{\delta} - \delta) + [o_p(m^{-1})]_{p \times 1}.$$

In addition,

$$\begin{aligned}
E(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 &= E \left[(\bar{y}_i - \bar{x}_i' \beta) - \bar{x}_i' (\tilde{\beta} - \beta) \right]^2 \\
&= \sigma_\nu^2 + \sigma_e^2/n_i - \bar{x}_i' \left(\sum_i X_i V_i^{-1} X_i \right)^{-1} \bar{x}_i.
\end{aligned}$$

Hence,

$$\begin{aligned}
E(\hat{\mu}_{iw} - \tilde{\mu}_i)^2 &= (B_i - w_i)^2 \left[\sigma_\nu^2 + \sigma_e^2/n_i - \bar{x}_i' \left(\sum_i X_i V_i^{-1} X_i \right)^{-1} \bar{x}_i \right] \\
&\quad + 2(B_i - w_i) E \left[(\bar{y}_i - \bar{x}_i' \tilde{\beta}) [\bar{X}_i - (1 - w_i) \bar{x}_i]' \frac{\partial \tilde{\beta}}{\partial \delta'} (\hat{\delta} - \delta) \right] + o(m^{-1}).
\end{aligned}$$

To evaluate the above expectation, we need the following Lemma due to Stein (1981).

Lemma 6.1. *Let $Z \sim N(\mu, \sigma^2)$, then*

$$E[h(Z)(Z - \mu)] = \sigma^2 E \left[\frac{\partial h(Z)}{\partial Z} \right],$$

provided the expectations exist and $\lim_{z \rightarrow \pm\infty} h(z) \exp[-\frac{1}{2\sigma^2}(z - \mu)^2] = 0$.

We have

$$E \left[(\bar{y}_i - \bar{x}_i' \tilde{\beta}) [\bar{X}_i - (1 - w_i) \bar{x}_i]' \frac{\partial \tilde{\beta}}{\partial \delta'} (\hat{\delta} - \delta) \right]$$

$$\begin{aligned}
&= E\left[(\bar{y}_i - \bar{x}_i'\beta)[\bar{X}_i - (1 - w_i)\bar{x}_i]'\frac{\partial\tilde{\beta}}{\partial\delta'}(\hat{\delta} - \delta)\right] + o(m^{-1}) \\
&= (\sigma_\nu^2 + \sigma_e^2/n_i)E\left\{\frac{\partial}{\partial\bar{y}_i}\left[(\bar{X}_i - (1 - w_i)\bar{x}_i)'\frac{\partial\tilde{\beta}}{\partial\delta'}(\hat{\delta} - \delta)\right]\right\} + o(m^{-1}),
\end{aligned}$$

by Lemma 6.1. In addition, we can write

$$\begin{aligned}
\frac{\partial}{\partial\bar{y}_i}\left[(\bar{X}_i - (1 - w_i)\bar{x}_i)'\frac{\partial\tilde{\beta}}{\partial\delta'}(\hat{\delta} - \delta)\right] &= \frac{\partial}{\partial\bar{y}_i}\left[(\bar{X}_i - (1 - w_i)\bar{x}_i)'\frac{\partial\tilde{\beta}}{\partial\sigma_\nu^2}\right](\hat{\sigma}_\nu^2 - \sigma_\nu^2) \\
&+ \frac{\partial}{\partial\bar{y}_i}\left[(\bar{X}_i - (1 - w_i)\bar{x}_i)'\frac{\partial\tilde{\beta}}{\partial\sigma_e^2}\right](\hat{\sigma}_e^2 - \sigma_e^2) + (\bar{X}_i - (1 - w_i)\bar{x}_i)'\frac{\partial\tilde{\beta}}{\partial\sigma_\nu^2}\frac{\partial\hat{\sigma}_\nu^2}{\partial\bar{y}_i} \\
&+ (\bar{X}_i - (1 - w_i)\bar{x}_i)'\frac{\partial\tilde{\beta}}{\partial\sigma_e^2}\frac{\partial\hat{\sigma}_e^2}{\partial\bar{y}_i},
\end{aligned}$$

with

$$\frac{\partial}{\partial\bar{y}_i}\left[(\bar{X}_i - (1 - w_i)\bar{x}_i)'\frac{\partial\tilde{\beta}}{\partial\sigma_\nu^2}\right] = \frac{\partial}{\partial\sigma_\nu^2}(\bar{X}_i - (1 - w_i)\bar{x}_i)'\frac{\partial\tilde{\beta}}{\partial\bar{y}_i}.$$

Also, we may write

$$\frac{\partial\tilde{\beta}}{\partial\bar{y}_i} = \frac{\partial}{\partial\bar{y}_i}\left\{(\sum_i X_i'V_i^{-1}X_i)^{-1}(\sum_i X_i'V_i^{-1}y_i)\right\} = \frac{-n_i\sigma_\nu^2\sigma_e^{-2}(\sum_i X_i'V_i^{-1}X_i)^{-1}\bar{x}_i}{(\sigma_\nu^2 + \sigma_e^2/n_i)}.$$

We then have

$$\begin{aligned}
&\frac{\partial}{\partial\sigma_\nu^2}\left[\frac{(-n_i\sigma_\nu^2\sigma_e^{-2})(\bar{X}_i - (1 - w_i)\bar{x}_i)'(\sum_i X_i'V_i^{-1}X_i)^{-1}\bar{x}_i}{\sigma_\nu^2 + \sigma_e^2/n_i}\right] \\
&= (\bar{X}_i - (1 - w_i)\bar{x}_i)'\left\{\frac{-n_i\sigma_e^{-2}(\sum_i X_i'V_i^{-1}X_i)^{-1}\bar{x}_i}{\sigma_\nu^2 + \sigma_e^2/n_i} - n_i\sigma_\nu^2\sigma_e^{-2}\left[\sigma_e^{-4}(\sum_i X_i'V_i^{-1}X_i)^{-1}\right.\right. \\
&\quad \cdot (\sum_i B_i^2 X_i'J_{n_i}X_i)(\sum_i X_i'V_i^{-1}X_i)^{-1}\bar{x}_i(\sigma_\nu^2 + \sigma_e^2/n_i) - (\sum_i X_i'V_i^{-1}X_i)^{-1}\bar{x}_i\left.\right] \\
&\quad \left.\cdot (\sigma_\nu^2 + \sigma_e^2/n_i)^{-2}\right\} = O(m^{-1}).
\end{aligned}$$

Similarly, we have

$$\frac{\partial}{\partial\sigma_e^2}\left[\frac{(-n_i\sigma_\nu^2\sigma_e^{-2})(\bar{X}_i - (1 - w_i)\bar{x}_i)'(\sum_i X_i'V_i^{-1}X_i)^{-1}\bar{x}_i}{\sigma_\nu^2 + \sigma_e^2/n_i}\right]$$

$$\begin{aligned}
&= (\bar{X}_i - (1 - w_i)\bar{x}_i)' \left\{ \frac{n_i \sigma_\nu^2 \sigma_e^{-4} (\sum_i X_i' V_i^{-1} X_i)^{-1} \bar{x}_i}{\sigma_\nu^2 + \sigma_e^2/n_i} - n_i \sigma_\nu^2 \sigma_e^{-2} \left[-\sigma_e^2 (\sum_i X_i' V_i^{-1} X_i)^{-1} \bar{x}_i \right. \right. \\
&\quad \left. \left. + \sigma_e^{-4} (\sum_i X_i' V_i^{-1} X_i)^{-1} (\sum_i B_i (1 - B_i) n_i^{-1} X_i' J_{n_i} X_i) (\sum_i X_i' V_i^{-1} X_i)^{-1} \bar{x}_i (\sigma_\nu^2 + \sigma_e^2/n_i) \right. \right. \\
&\quad \left. \left. - n_i^{-1} (\sum_i X_i' V_i^{-1} X_i)^{-1} \bar{x}_i \right] / (\sigma_\nu^2 + \sigma_e^2/n_i)^2 \right\} = O(m^{-1}).
\end{aligned}$$

In addition, we may write

$$\begin{aligned}
\frac{\partial \hat{\sigma}_e^2}{\partial \bar{y}_i} &= \frac{\partial}{\partial \bar{y}_i} (SSE_1/\nu) \\
&= \nu^{-1} \frac{\partial}{\partial \bar{y}_i} \left\{ \sum_i y_i^* \left[I_{n_i} - X_i^* (X^{*'} X^*)^{-1} X_i^{*'} \right] y_i^* \right\} \\
&= \nu^{-1} \frac{\partial}{\partial \bar{y}_i} \left\{ \sum_i (y_i - \bar{y}_i 1_i)' \left[I_{n_i} - X_i^* (X^{*'} X^*)^{-1} X_i^{*'} \right] (y_i - \bar{y}_i 1_i) \right\} \\
&= -2\nu^{-1} 1_i' \left\{ I_{n_i} - X_i^* (X^{*'} X^*)^{-1} X_i^{*'} \right\} y_i^* \\
&= O_p(m^{-1}), \tag{6.3}
\end{aligned}$$

where $\nu = n - m - p_1$, p_1 is the number of non-zero x_{ij}^* , $x_{ij}^* = x_{ij} - \bar{x}_i$, y_i^* is a $n_i \times 1$ vector with elements $y_{ij}^* = y_{ij} - \bar{y}_i$, I_{n_i} is an $n_i \times n_i$ identity matrix, 1_i is a $n_i \times 1$ vector of one's and X^* is a $n \times p$ matrix with elements $x_{ij}^* (j = 1, \dots, n_i; i = 1, \dots, m)$.

Moreover, we may write

$$\begin{aligned}
\frac{\partial \hat{\sigma}_\nu^2}{\partial \bar{y}_i} &= \frac{\partial}{\partial \bar{y}_i} \left\{ \eta^{-1} [SSE_2 - (n - p) \hat{\sigma}_e^2] \right\} = \eta^{-1} \frac{\partial}{\partial \bar{y}_i} \left[\sum_i \sum_j (y_{ij} - x_{ij}' \tilde{\beta}_{ols})^2 \right] - \frac{n - p}{\eta} \frac{\partial \hat{\sigma}_e^2}{\partial \bar{y}_i} \\
&= \eta^{-1} \frac{\partial}{\partial \bar{y}_i} \left\{ \sum_i y_i' \left[I_{n_i} - X_i (X' X)^{-1} X_i' \right] y_i \right\} - \frac{n - p}{\eta} \frac{\partial \hat{\sigma}_e^2}{\partial \bar{y}_i} \\
&= -\frac{n - p}{\eta} \frac{\partial \hat{\sigma}_e^2}{\partial \bar{y}_i} = O_p(m^{-1}) \quad \text{by (6.3).}
\end{aligned}$$

We have also shown in Chapter 5 that $\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = O_p(m^{-1/2})$ and $\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = O_p(m^{-1/2})$. Note that $\hat{\sigma}_\nu^2 - \sigma_\nu^2 = O_p(m^{-1/2})$ and $\hat{\sigma}_e^2 - \sigma_e^2 = O_p(m^{-1/2})$. As a result, $E \left\{ \frac{\partial}{\partial \bar{y}_i} \left[(\bar{X}_i - (1 - \right.$

$w_i)\bar{x}_i)' \frac{\partial \tilde{\beta}}{\partial \delta'}(\hat{\delta} - \delta) \Big] \Big\} = o(m^{-1})$. Therefore, we have

$$E \left[(\bar{y}_i - \bar{x}_i' \tilde{\beta})(\bar{X}_i - (1 - w_i)\bar{x}_i)' \frac{\partial \tilde{\beta}}{\partial \delta'}(\hat{\delta} - \delta) \right] = o(m^{-1}),$$

and

$$E(\hat{\mu}_{iw} - \tilde{\mu}_i)^2 = (B_i - w_i)^2 \left[\sigma_\nu^2 + \sigma_e^2/n_i - \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i \right] + o(m^{-1}).$$

Hence, we have

$$MSE(\hat{\mu}_{iw}) = g_{1i}(\delta) + g_{2i}(\delta) + g_{wi}(\delta) + o(m^{-1}),$$

where $g_{wi}(\delta) = (B_i - w_i)^2 \left[\sigma_\nu^2 + \sigma_e^2/n_i - \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i \right]$. Note that $g_{wi}(\delta) = 0$ when $w_i = B_i$.

Theorem 6.1. *The second order approximation of the MSE of the weighted predictor $\hat{\mu}_{iw}$ is given by*

$$MSE(\hat{\mu}_{iw}) \approx g_{1i}(\delta) + g_{2i}(\delta) + g_{wi}(\delta).$$

In Section 6.3, we derive the second order unbiased estimator of MSE of $\hat{\mu}_{iw}$.

6.3 MSE Estimation for Weighted Predictors

We rewrite $MSE(\hat{\mu}_{iw})$ to derive a second order unbiased estimator of the MSE of the weighted predictor $\hat{\mu}_{iw}$:

$$\begin{aligned} MSE(\hat{\mu}_{iw}) &= \left[g_{1i}(\delta) + (B_i - w_i)^2 (\sigma_\nu^2 + \sigma_e^2/n_i) \right] \\ &+ \left[g_{2i}(\delta) - (B_i - w_i)^2 \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i \right] + o(m^{-1}) \end{aligned}$$

$$= A_{1iw}(\delta) + A_{2iw}(\delta) + o(m^{-1}),$$

where

$$\begin{aligned} A_{1iw}(\delta) &= \sigma_\nu^2 B_i + (B_i - w_i)^2 (\sigma_\nu^2 + \sigma_e^2/n_i) \\ &= \sigma_\nu^2 B_i + B_i \sigma_e^2/n_i + w_i^2 (\sigma_\nu^2 + \sigma_e^2/n_i) - 2w_i \sigma_e^2/n_i \\ &= (1 - w_i)^2 \sigma_e^2/n_i + w_i^2 \sigma_\nu^2. \end{aligned}$$

Then,

$$E[A_{1iw}(\hat{\delta})] = (1 - w_i)^2 n_i^{-1} \sigma_e^2 + w_i^2 \sigma_\nu^2 = A_{1iw}(\delta),$$

since our estimators, $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$, (based on the fitting-of-constants method) are unbiased. Also,

$$E[A_{2iw}(\hat{\delta})] = E\left[g_{2i}(\hat{\delta}) - (\hat{B}_i - w_i)^2 \bar{x}_i' \left(\sum_i X_i' \hat{V}_i^{-1} X_i\right)^{-1} \bar{x}_i\right] = A_{2iw}(\delta) + o(m^{-1}).$$

Theorem 6.2. *A second order unbiased estimator of the MSE of the weighted predictor $\hat{\mu}_{iw}$ is given by*

$$mse[\hat{\mu}_{iw}(\hat{\delta})] = g_{1i}(\hat{\delta}) + g_{2i}(\hat{\delta}) + g_{wi}(\hat{\delta}),$$

namely, $E[mse(\hat{\mu}_{iw}(\hat{\delta}))] = MSE[\hat{\mu}_{iw}(\hat{\delta})] + o(m^{-1})$.

In Section 6.4, we study the conditional MSE of the weighted predictor $\hat{\mu}_{iw}$ given the area-specific response data.

6.4 Conditional MSE Approximation of Weighted Predictors

We derive a second order approximation of the conditional MSE of the weighted predictor $\hat{\mu}_{iw}$ given $y_i = (y_{i1}, \dots, y_{in_i})'$. This conditional MSE can be written as

$$MSE^c(\hat{\mu}_{iw}|y_i) = E[(\hat{\mu}_{iw} - \mu_i)^2|y_i]$$

$$= E[(\hat{\mu}_{iw} - \tilde{\mu}_i)^2 | y_i] + E[(\tilde{\mu}_i - \mu_i)^2 | y_i] + 2E[(\hat{\mu}_{iw} - \tilde{\mu}_i)(\tilde{\mu}_i - \mu_i) | y_i]. \quad (6.4)$$

We have shown in Chapter 5 that

$$E[(\tilde{\mu}_i - \mu_i)^2 | y_i] = g_{1i}(\delta) + g_{2i}(\delta) + o_p(m^{-1}). \quad (6.5)$$

In addition, we can evaluate $E[(\hat{\mu}_{iw} - \tilde{\mu}_i)^2 | y_i]$ as follows :

$$\begin{aligned} E[(\hat{\mu}_{iw} - \tilde{\mu}_i)^2 | y_i] &= E \left\{ \left[(\bar{X}_i - \bar{x}_i)'(\hat{\beta} - \tilde{\beta}) - w_i(\bar{y}_i - \bar{x}_i' \hat{\beta}) + B_i(\bar{y}_i - \bar{x}_i' \tilde{\beta}) \right]^2 | y_i \right\} \\ &= E \left\{ \left[(\bar{X}_i - (1 - w_i)\bar{x}_i)'(\hat{\beta} - \tilde{\beta}) + (B_i - w_i)(\bar{y}_i - \bar{x}_i' \tilde{\beta}) \right]^2 | y_i \right\} \\ &= (B_i - w_i)^2 E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 | y_i] + E \left\{ \left[(\bar{X}_i - (1 - w_i)\bar{x}_i)'(\hat{\beta} - \tilde{\beta}) \right]^2 | y_i \right\} \\ &\quad + 2(B_i - w_i) E \left\{ (\bar{X}_i - (1 - w_i)\bar{x}_i)'(\hat{\beta} - \tilde{\beta})(\bar{y}_i - \bar{x}_i' \tilde{\beta}) | y_i \right\}, \end{aligned} \quad (6.6)$$

where the first expectation on the right side of (6.6) can be written as

$$\begin{aligned} E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 | y_i] &= E \left\{ \left[(\bar{y}_i - \bar{x}_i' \beta) - \bar{x}_i'(\tilde{\beta} - \beta) \right]^2 | y_i \right\} \\ &= (\bar{y}_i - \bar{x}_i' \beta)^2 + \bar{x}_i' E[(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)' | y_i] \bar{x}_i - 2(\bar{y}_i - \bar{x}_i' \beta) E[\bar{x}_i'(\tilde{\beta} - \beta) | y_i] \\ &= (\bar{y}_i - \bar{x}_i' \beta)^2 + \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i - 2(\bar{y}_i - \bar{x}_i' \beta) E \left\{ \bar{x}_i' (X' V^{-1} X)^{-1} \right. \\ &\quad \left. \cdot X' V^{-1} (y - X\beta) | y_i \right\} + o_p(m^{-1}), \end{aligned}$$

with

$$\begin{aligned} &E \left[\bar{x}_i' (X' V^{-1} X)^{-1} X' V^{-1} (y - X\beta) | y_i \right] \\ &= \bar{x}_i' (X' V^{-1} X)^{-1} X' V^{-1} E[(y - X\beta) | y_i] \\ &= \bar{x}_i' (X' V^{-1} X)^{-1} X' V^{-1} a_i(y_i - X_i \beta) \\ &= \bar{x}_i' (X' V^{-1} X)^{-1} X_i' V_i^{-1} (y_i - X_i \beta) \\ &= \sigma_e^{-2} \bar{x}_i' (X' V^{-1} X)^{-1} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \beta) - n_i \bar{x}_i (1 - B_i) (\bar{y}_i - \bar{x}_i' \beta) \right]. \end{aligned}$$

Then,

$$E[(\bar{y}_i - \bar{x}'_i \tilde{\beta})^2 | y_i] = (\bar{y}_i - \bar{x}'_i \beta)^2 + \bar{x}'_i \left(\sum_i X'_i V_i^{-1} X_i \right)^{-1} \bar{x}_i - 2\sigma_e^{-2} (\bar{y}_i - \bar{x}'_i \beta) \\ \cdot \left\{ \bar{x}'_i \left(\sum_i X'_i V_i^{-1} X_i \right)^{-1} \left[\sum_j x_{ij} (y_{ij} - x'_{ij} \beta) - n_i (1 - B_i) \bar{x}_i (\bar{y}_i - \bar{x}'_i \beta) \right] \right\} + o_p(m^{-1}). \quad (6.7)$$

Also, the second expectation on the right side of (6.6) is obtained by

$$E \left\{ \left[(\bar{X}_i - (1 - w_i) \bar{x}_i)' (\hat{\beta} - \tilde{\beta}) \right]^2 | y_i \right\} \\ = E \left\{ \left[(\bar{X}_i - (1 - w_i) \bar{x}_i)' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right) \right]^2 | y_i \right\} + o_p(m^{-1}) = o_p(m^{-1}). \quad (6.8)$$

Finally, the third expectation on the right side of (6.6) may be expressed as

$$E \left[(\bar{X}_i - (1 - w_i) \bar{x}_i)' (\hat{\beta} - \tilde{\beta}) (\bar{y}_i - \bar{x}'_i \tilde{\beta}) | y_i \right] \\ = (\bar{y}_i - \bar{x}'_i \beta) E \left[(\bar{X}_i - (1 - w_i) \bar{x}_i)' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right) | y_i \right] + o_p(m^{-1}). \quad (6.9)$$

But, we have that

$$E \left[(\bar{X}_i - (1 - w_i) \bar{x}_i)' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right) | y_i \right] \\ = (\bar{X}_i - (1 - w_i) \bar{x}_i)' \left(\sum_i X_i V_i^{-1} X_i \right)^{-1} \left[\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X'_i J_{n_i} X_i \right] \\ \cdot E \left\{ \left[\left(\sum_i X'_i V_i^{-1} X_i \right)^{-1} \sum_i X'_i V_i^{-1} (y_i - X_i \beta) - \left[\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X'_i J_{n_i} X_i \right]^{-1} \right. \right. \\ \left. \left. \cdot \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X'_i J_{n_i} (y_i - X_i \beta) \right] \left[(\hat{\sigma}_\nu^2 - \sigma_\nu^2) - \sigma_\nu^2 \sigma_e^{-2} (\hat{\sigma}_e^2 - \sigma_e^2) \right] | y_i \right\},$$

where $(\partial \tilde{\beta} / \partial \sigma_\nu^2)$ can be written as

$$\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} = - \left(\sum_i X'_i V_i^{-1} X_i \right)^{-1} \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X'_i J_{n_i} (y_i - X_i \tilde{\beta}) \quad (\text{by (5.9) in Chapter 5})$$

$$\begin{aligned}
&= \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} X_i \left\{ \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \sum_i X_i' V_i^{-1} (y_i - X_i \beta) \right. \\
&\quad \left. - \left[\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} X_i \right]^{-1} \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} (y_i - X_i \beta) \right\}.
\end{aligned}$$

Further, we know from Chapter 5 that $\frac{\partial \tilde{\beta}}{\partial \sigma_e^2} = \frac{-\sigma_\nu^2}{\sigma_e^2} \frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}$. Hence

$$\begin{aligned}
&E \left[\left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \sum_i X_i' V_i^{-1} (y_i - X_i \beta) (\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i \right] \\
&= E[(\tilde{\beta} - \beta)(\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i] = E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2) | y_i] + o_p(m^{-1}) \\
&= E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2)] + o_p(m^{-1}) = o_p(m^{-1}), \tag{6.10}
\end{aligned}$$

where $\tilde{\beta}_{-i} - \tilde{\beta} = O_p(m^{-1})$, $\hat{\sigma}_{\nu,-i}^2 - \hat{\sigma}_\nu^2 = O_p(m^{-1})$ and $E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2)] = 0$, since third order moments are all zero due to normality.

Similarly, we have that

$$\begin{aligned}
&E \left[\left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \sum_i X_i' V_i^{-1} (y_i - X_i \beta) (\hat{\sigma}_e^2 - \sigma_e^2) | y_i \right] \\
&= E[(\tilde{\beta} - \beta)(\hat{\sigma}_e^2 - \sigma_e^2) | y_i] = E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{e,-i}^2 - \sigma_e^2) | y_i] + o_p(m^{-1}) \\
&= E[(\tilde{\beta}_{-i} - \beta)(\hat{\sigma}_{e,-i}^2 - \sigma_e^2)] + o_p(m^{-1}) = o_p(m^{-1}), \tag{6.11}
\end{aligned}$$

where $\hat{\sigma}_{e,-i}^2 - \hat{\sigma}_e^2 = O_p(m^{-1})$.

On the other hand, we may write

$$\begin{aligned}
&E \left\{ \left[\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} X_i \right]^{-1} \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} (y_i - X_i \beta) (\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i \right\} \\
&= E[(\tilde{\beta}' - \beta)(\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i] \\
&= E[(\tilde{\beta}'_{-i} - \beta)(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2) | y_i] + o_p(m^{-1}) \\
&= E[(\tilde{\beta}'_{-i} - \beta)(\hat{\sigma}_{\nu,-i}^2 - \sigma_\nu^2)] + o_p(m^{-1}) \\
&= o_p(m^{-1}), \tag{6.12}
\end{aligned}$$

where $\tilde{\beta}' = [\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} X_i]^{-1} [\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} y_i]$, $\tilde{\beta}'_{-i} - \tilde{\beta}' = O_p(m^{-1})$ and $E(\tilde{\beta}') = \beta$.

Also, we may write

$$\begin{aligned}
E \left\{ \left[\sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} X_i \right]^{-1} \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} (y_i - X_i \beta) (\hat{\sigma}_e^2 - \sigma_e^2) \middle| y_i \right\} \\
= E[(\tilde{\beta}' - \beta)(\hat{\sigma}_e^2 - \sigma_e^2) | y_i] \\
= E[(\tilde{\beta}'_{-i} - \beta)(\hat{\sigma}_{e,-i}^2 - \sigma_e^2) | y_i] + o_p(m^{-1}) \\
= E[(\tilde{\beta}'_{-i} - \beta)(\hat{\sigma}_{e,-i}^2 - \sigma_e^2)] + o_p(m^{-1}) \\
= o_p(m^{-1}).
\end{aligned}$$

Then by (6.9)- (6.12), we have that

$$E \left[(\bar{X}_i - (1 - w_i) \bar{x}_i)' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right) \middle| y_i \right] = o_p(m^{-1}).$$

Hence, by (6.9), we have

$$E \left[(\bar{X}_i - (1 - w_i) \bar{x}_i)' (\hat{\beta} - \tilde{\beta}) (\bar{y}_i - \bar{x}_i' \tilde{\beta}) \middle| y_i \right] = o_p(m^{-1}). \quad (6.13)$$

Therefore, we may write by (6.7), (6.8) and (6.13) as

$$\begin{aligned}
E[(\hat{\mu}_{iw} - \tilde{\mu}_i)^2 | y_i] &= (B_i - w_i)^2 \left\{ (\bar{y}_i - \bar{x}_i' \beta)^2 + \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i - 2 \sigma_e^{-2} (\bar{y}_i - \bar{x}_i' \beta) \right. \\
&\quad \left. \cdot \bar{x}_i \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \beta) - n_i \bar{x}_i (1 - B_i) (\bar{y}_i - \bar{x}_i' \beta) \right] \right\} + o_p(m^{-1}) \\
&= (B_i - w_i)^2 \left[(\bar{y}_i - \bar{x}_i' \beta)^2 - \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i \right] + 2(B_i - w_i)^2 \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \\
&\quad \cdot \left\{ \bar{x}_i - (\bar{y}_i - \bar{x}_i' \beta) \sigma_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \beta) - n_i (1 - B_i) \bar{x}_i (\bar{y}_i - \bar{x}_i' \beta) \right] \right\} + o_p(m^{-1})
\end{aligned}$$

$$\equiv g_{wi1}^c(\delta) + o_p(m^{-1}). \quad (6.14)$$

In addition, since $E(\mu_i|y_i) = \tilde{\mu}_i^B$, we have

$$E[(\hat{\mu}_{iw} - \tilde{\mu}_i)(\tilde{\mu}_i^B - \mu_i)|y_i] = 0. \quad (6.15)$$

Therefore, we may write

$$\begin{aligned} & E[(\hat{\mu}_{iw} - \tilde{\mu}_i)(\tilde{\mu}_i - \tilde{\mu}_i^B)|y_i] \\ &= E \left\{ \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)'(\tilde{\beta} - \beta) \right] \left[(\bar{X}_i - (1 - w_i)\bar{x}_i)'(\hat{\beta} - \tilde{\beta}) + (B_i - w_i)(\bar{y}_i - \bar{x}_i'\tilde{\beta}) \right] | y_i \right\} \\ &= E \left\{ \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)'(\tilde{\beta} - \beta) \right] \left[(\bar{X}_i - (1 - w_i)\bar{x}_i)' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2}(\hat{\sigma}_e^2 - \sigma_e^2) \right) \right. \right. \\ & \quad \left. \left. + (B_i - w_i)(\bar{y}_i - \bar{x}_i'\tilde{\beta}) \right] | y_i \right\} + o_p(m^{-1}) \\ &= E \left\{ (\bar{X}_i - (1 - B_i)\bar{x}_i)'(\tilde{\beta} - \beta)(\bar{X}_i - (1 - w_i)\bar{x}_i)' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2}(\hat{\sigma}_e^2 - \sigma_e^2) \right) | y_i \right\} \\ & \quad + (B_i - w_i)E \left\{ (\bar{X}_i - (1 - B_i)\bar{x}_i)'(\tilde{\beta} - \beta)(\bar{y}_i - \bar{x}_i'\tilde{\beta}) | y_i \right\} + o_p(m^{-1}), \end{aligned}$$

where

$$\begin{aligned} & E \left\{ (\bar{X}_i - (1 - B_i)\bar{x}_i)'(\tilde{\beta} - \beta)(\bar{X}_i - (1 - w_i)\bar{x}_i)' \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2}(\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2}(\hat{\sigma}_e^2 - \sigma_e^2) \right) | y_i \right\} \\ &= o_p(m^{-1}), \end{aligned}$$

and

$$\begin{aligned} & E \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)'(\tilde{\beta} - \beta)(\bar{y}_i - \bar{x}_i'\tilde{\beta}) | y_i \right] \\ &= (\bar{y}_i - \bar{x}_i'\beta)(\bar{X}_i - (1 - B_i)\bar{x}_i)'E[\tilde{\beta} - \beta|y_i] - E \left[(\bar{X}_i - (1 - B_i)\bar{x}_i)'(\tilde{\beta} - \beta)\bar{x}_i'(\tilde{\beta} - \beta) | y_i \right] \\ &= (\bar{y}_i - \bar{x}_i'\beta)(\bar{X}_i - (1 - B_i)\bar{x}_i)' \left(\sum_i X_i'V_i^{-1}X_i \right)^{-1} \sigma_e^{-2} \left[\sum_j x_{ij}(y_{ij} - x_{ij}'\beta) \right] \end{aligned}$$

$$\begin{aligned}
& -n_i(1-B_i)\bar{x}_i(\bar{y}_i - \bar{x}'_i\beta) \Big] - (\bar{X}_i - (1-B_i)\bar{x}_i)' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i + o_p(m^{-1}) \\
& = (\bar{X}_i - (1-B_i)\bar{x}_i)' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \left\{ \sigma_e^{-2}(\bar{y}_i - \bar{x}'_i\beta) \left[\sum_j x_{ij}(y_{ij} - x'_{ij}\beta) - n_i(1-B_i)\bar{x}_i \right. \right. \\
& \quad \left. \left. \cdot (\bar{y}_i - \bar{x}'_i\beta) \right] - \bar{x}_i \right\} + o_p(m^{-1}). \tag{6.16}
\end{aligned}$$

Therefore, we may write the cross-product term of (6.4), by using (6.15) and (6.16), as

$$\begin{aligned}
E[(\hat{\mu}_{iw} - \tilde{\mu}_i)(\tilde{\mu}_i - \mu_i)|y_i] &= (B_i - w_i)(\bar{X}_i - (1-B_i)\bar{x}_i)' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \\
& \cdot \left\{ \sigma_e^{-2}(\bar{y}_i - \bar{x}'_i\beta) \left[\sum_j x_{ij}(y_{ij} - x'_{ij}\beta) - n_i(1-B_i)\bar{x}_i(\bar{y}_i - \bar{x}'_i\beta) \right] - \bar{x}_i \right\} + o_p(m^{-1}) \\
& \equiv g_{wi2}^c(\delta) + o_p(m^{-1}). \tag{6.17}
\end{aligned}$$

Theorem 6.3. *A second order approximation of the conditional MSE of the weighted predictor $\hat{\mu}_{iw}(\hat{\delta})$ is obtained by*

$$MSE^c[\hat{\mu}_{iw}(\hat{\delta})|y_i] \approx g_{1i}(\delta) + g_{2i}(\delta) + g_{wi1}^c(\delta) + 2g_{wi2}^c(\delta),$$

where $g_{1i}(\delta)$, $g_{2i}(\delta)$, $g_{wi1}^c(\delta)$ and $g_{wi2}^c(\delta)$ are given by (6.1), (6.2), (6.14) and (6.17) respectively.

6.5 Conditional MSE Estimation of Weighted Predictors

We now derive a second order unbiased estimator of the conditional MSE of the weighted predictor $\hat{\mu}_{iw}(\hat{\delta})$. We may write $g_{1i}(\delta)$, using Taylor expansion, as

$$g_{1i}(\delta) = E \left\{ g_{1i}(\hat{\delta}) - \frac{\partial g_{1i}}{\partial \sigma_\nu^2} \Big|_{\delta=\hat{\delta}} \hat{b}(\hat{\sigma}_\nu^2|y_i) - \frac{\partial g_{1i}}{\partial \sigma_e^2} \Big|_{\delta=\hat{\delta}} \hat{b}(\hat{\sigma}_e^2|y_i) - \frac{1}{2} \frac{\partial^2 g_{1i}}{\partial \sigma_\nu^4} \Big|_{\delta=\hat{\delta}} \hat{V}(\hat{\sigma}_\nu^2|y_i) \right.$$

$$-\frac{1}{2} \frac{\partial^2 g_{1i}}{\partial \sigma_e^4} \Big|_{\delta=\hat{\delta}} \hat{V}(\hat{\sigma}_e^2 | y_i) - \frac{\partial^2 g_{1i}}{\partial \sigma_\nu^2 \partial \sigma_e^2} \Big|_{\delta=\hat{\delta}} \widehat{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2 | y_i) \Big\} + o(m^{-1}).$$

Then, we have

$$g_{1i}(\delta) = E \left[g_{1i}(\hat{\delta}) - \hat{B}_i^2 \hat{b}(\hat{\sigma}_\nu^2 | y_i) - (1 - \hat{B}_i)^2 n_i^{-1} \hat{b}(\hat{\sigma}_e^2 | y_i) + g_{3i}(\hat{\delta}) \right] + o(m^{-1}).$$

Also, we have shown in Chapter 5 that

$$E[g_{2i}(\hat{\delta}) | y_i] = g_{2i}(\delta) + o_p(m^{-1}).$$

In the next step, we find second order conditional unbiased estimators for $g_{wi1}^c(\delta) + 2g_{wi2}^c(\delta)$. We may rewrite $g_{wi1}^c(\delta) + 2g_{wi2}^c(\delta)$ as follows

$$\begin{aligned} g_{wi1}^c(\delta) + 2g_{wi2}^c(\delta) &= (B_i - w_i)^2 \left\{ (\bar{y}_i - \bar{x}_i' \beta)^2 + \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i - 2(\bar{y}_i - \bar{x}_i' \beta) \bar{x}_i' \right. \\ &\quad \cdot \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \sigma_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \beta) - n_i (1 - B_i) \bar{x}_i (\bar{y}_i - \bar{x}_i' \beta) \right] \Big\} + 2(B_i - w_i) \\ &\quad \cdot (\bar{X}_i - (1 - B_i) \bar{x}_i)' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \left\{ (\bar{y}_i - \bar{x}_i' \beta) \sigma_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \beta) \right. \right. \\ &\quad \left. \left. - n_i (1 - B_i) \bar{x}_i (\bar{y}_i - \bar{x}_i' \beta) \right] - \bar{x}_i \right\} \\ &= (B_i - w_i)^2 \left[(\bar{y}_i - \bar{x}_i' \beta)^2 - \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i \right] + 2(B_i - w_i) (\bar{X}_i - (1 - w_i) \bar{x}_i)' \\ &\quad \cdot \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \left\{ (\bar{y}_i - \bar{x}_i' \beta) \sigma_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \beta) - n_i (1 - B_i) \bar{x}_i (\bar{y}_i - \bar{x}_i' \beta) \right] - \bar{x}_i \right\}. \end{aligned}$$

We have

$$\begin{aligned} &E \left[(\hat{B}_i - w_i)^2 \bar{x}_i' \left(\sum_i X_i' \hat{V}_i^{-1} X_i \right)^{-1} \bar{x}_i | y_i \right] \\ &= (B_i - w_i)^2 \bar{x}_i' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i + o_p(m^{-1}). \end{aligned} \tag{6.18}$$

Moreover, we may write

$$\begin{aligned}
& E \left\{ (\hat{B}_i - w_i)(\bar{X}_i - (1 - w_i)\bar{x}_i)' \left(\sum_i X_i' \hat{V}_i^{-1} X_i \right)^{-1} \left[(\bar{y}_i - \bar{x}_i' \hat{\beta}) \hat{\sigma}_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \hat{\beta}) \right. \right. \right. \\
& \quad \left. \left. \left. - n_i \bar{x}_i (1 - \hat{B}_i) (\bar{y}_i - \bar{x}_i' \hat{\beta}) \right] - \bar{x}_i \right] | y_i \right\} \\
& = (B_i - w_i)(\bar{X}_i - (1 - w_i)\bar{x}_i)' \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \left\{ (\bar{y}_i - \bar{x}_i' \beta) \sigma_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \beta) \right. \right. \\
& \quad \left. \left. - n_i (1 - B_i) \bar{x}_i (\bar{y}_i - \bar{x}_i' \beta) \right] - \bar{x}_i \right\} + o_p(m^{-1}). \tag{6.19}
\end{aligned}$$

In addition, by using Taylor expansion, we may write

$$\begin{aligned}
& E[(\hat{B}_i - w_i)^2 (\bar{y}_i - \bar{x}_i' \hat{\beta})^2 | y_i] = (B_i - w_i)^2 E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 | y_i] \\
& \quad + 2 \frac{\partial B_i}{\partial \sigma_\nu^2} (B_i - w_i) E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 (\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i] \\
& \quad + 2 \frac{\partial B_i}{\partial \sigma_e^2} (B_i - w_i) E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 (\hat{\sigma}_e^2 - \sigma_e^2) | y_i] - 2(B_i - w_i)^2 \left\{ E \left[(\bar{y}_i - \bar{x}_i' \tilde{\beta}) \bar{x}_i' \frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i \right] \right. \\
& \quad \left. + E \left[(\bar{y}_i - \bar{x}_i' \tilde{\beta}) \bar{x}_i' \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) | y_i \right] \right\} + \left(\left(\frac{\partial B_i}{\partial \sigma_\nu^2} \right)^2 + \frac{\partial^2 B_i}{\partial \sigma_\nu^4} (B_i - w_i) \right) \\
& \quad \cdot E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 (\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i] + \left(\left(\frac{\partial B_i}{\partial \sigma_e^2} \right)^2 + \frac{\partial^2 B_i}{\partial \sigma_e^4} (B_i - w_i) \right) E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 (\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i] \\
& \quad + \left[\frac{\partial B_i}{\partial \sigma_\nu^2} \frac{\partial B_i}{\partial \sigma_e^2} + \frac{\partial^2 B_i}{\partial \sigma_\nu^2 \partial \sigma_e^2} (B_i - w_i) \right] E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 (\hat{\sigma}_\nu^2 - \sigma_\nu^2) (\hat{\sigma}_e^2 - \sigma_e^2) | y_i] + o_p(m^{-1}),
\end{aligned}$$

where $E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 | y_i]$ is given by (6.7).

We can also write

$$\begin{aligned}
& E[(\bar{y}_i - \bar{x}_i' \tilde{\beta})^2 (\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i] \\
& = (\bar{y}_i - \bar{x}_i' \beta)^2 E[(\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i] + E[(\bar{x}_i' (\tilde{\beta} - \beta))^2 (\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i] \\
& \quad - 2(\bar{y}_i - \bar{x}_i' \beta) E[\bar{x}_i' (\tilde{\beta} - \beta) (\hat{\sigma}_\nu^2 - \sigma_\nu^2) | y_i]
\end{aligned}$$

$$= (\bar{y}_i - \bar{x}'_i \beta)^2 b(\hat{\sigma}_\nu^2 | y_i) + o_p(m^{-1}).$$

Similarly,

$$E \left[(\bar{y}_i - \bar{x}'_i \tilde{\beta})^2 (\hat{\sigma}_e^2 - \sigma_e^2) | y_i \right] = (\bar{y}_i - \bar{x}'_i \beta)^2 b(\hat{\sigma}_e^2 | y_i) + o_p(m^{-1}).$$

Also, we have

$$\begin{aligned} & E \left\{ (\bar{y}_i - \bar{x}'_i \tilde{\beta}) \bar{x}'_i \left(\frac{\partial \tilde{\beta}}{\partial \sigma_\nu^2} (\hat{\sigma}_\nu^2 - \sigma_\nu^2) + \frac{\partial \tilde{\beta}}{\partial \sigma_e^2} (\hat{\sigma}_e^2 - \sigma_e^2) \right) | y_i \right\} \\ &= -(\bar{y}_i - \bar{x}'_i \beta) \bar{x}'_i \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} (y_i - X_i \beta) \\ & \quad \cdot \left[b(\hat{\sigma}_\nu^2 | y_i) - \sigma_\nu^2 \sigma_e^{-2} b(\hat{\sigma}_e^2 | y_i) \right] + o_p(m^{-1}), \end{aligned}$$

and

$$\begin{aligned} E \left[(\bar{y}_i - \bar{x}'_i \tilde{\beta})^2 (\hat{\sigma}_\nu^2 - \sigma_\nu^2)^2 | y_i \right] &= (\bar{y}_i - \bar{x}'_i \beta)^2 Var(\hat{\sigma}_\nu^2) + o_p(m^{-1}), \\ E \left[(\bar{y}_i - \bar{x}'_i \tilde{\beta})^2 (\hat{\sigma}_e^2 - \sigma_e^2)^2 | y_i \right] &= (\bar{y}_i - \bar{x}'_i \beta)^2 Var(\hat{\sigma}_e^2) + o_p(m^{-1}), \\ E \left[(\bar{y}_i - \bar{x}'_i \tilde{\beta})^2 (\hat{\sigma}_\nu^2 - \sigma_\nu^2)(\hat{\sigma}_e^2 - \sigma_e^2) | y_i \right] &= (\bar{y}_i - \bar{x}'_i \beta)^2 Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) + o_p(m^{-1}). \end{aligned}$$

Finally, we may write

$$\begin{aligned} E \left[(\hat{B}_i - w_i)^2 (\bar{y}_i - \bar{x}'_i \hat{\beta})^2 | y_i \right] &= (B_i - w_i)^2 (\bar{y}_i - \bar{x}'_i \beta)^2 + (B_i - w_i)^2 \left\{ \bar{x}'_i \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \bar{x}_i \right. \\ & \quad \left. - 2(\bar{y}_i - \bar{x}'_i \beta) \bar{x}'_i \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \sigma_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x'_{ij} \beta) - n_i (1 - B_i) \bar{x}_i (\bar{y}_i - \bar{x}'_i \beta) \right] \right\} \\ & \quad + 2 \frac{\partial B_i}{\partial \sigma_\nu^2} (B_i - w_i) (\bar{y}_i - \bar{x}'_i \beta)^2 b(\hat{\sigma}_\nu^2 | y_i) + 2 \frac{\partial B_i}{\partial \sigma_e^2} (B_i - w_i) (\bar{y}_i - \bar{x}'_i \beta)^2 b(\hat{\sigma}_e^2 | y_i) + 2 (B_i - w_i)^2 (\bar{y}_i \\ & \quad - \bar{x}'_i \beta) \bar{x}'_i \left(\sum_i X_i' V_i^{-1} X_i \right)^{-1} \sum_i (n_i \sigma_\nu^2 + \sigma_e^2)^{-2} X_i' J_{n_i} (y_i - X_i \beta) \left[b(\hat{\sigma}_\nu^2 | y_i) - \sigma_\nu^2 \sigma_e^{-2} b(\hat{\sigma}_e^2 | y_i) \right] \\ & \quad + \left(\left(\frac{\partial B_i}{\partial \sigma_\nu^2} \right)^2 + \frac{\partial^2 B_i}{\partial \sigma_\nu^4} (B_i - w_i) \right) (\bar{y}_i - \bar{x}'_i \beta)^2 V(\hat{\sigma}_\nu^2) + \left(\left(\frac{\partial B_i}{\partial \sigma_e^2} \right)^2 + \frac{\partial^2 B_i}{\partial \sigma_e^4} (B_i - w_i) \right) (\bar{y}_i - \bar{x}'_i \beta)^2 \end{aligned}$$

$$.V(\hat{\sigma}_e^2) + \left(\frac{\partial B_i}{\partial \sigma_\nu^2} \frac{\partial B_i}{\partial \sigma_e^2} + \frac{\partial^2 B_i}{\partial \sigma_\nu^2 \partial \sigma_e^2} (B_i - w_i) \right) (\bar{y}_i - \bar{x}'_i \beta)^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) + o_p(m^{-1}). \quad (6.20)$$

Then, by (6.20), we may write

$$\begin{aligned} (B_i - w_i)^2 (\bar{y}_i - \bar{x}'_i \beta)^2 &= E \left\{ (\hat{B}_i - w_i)^2 \left[(\bar{y}_i - \bar{x}'_i \hat{\beta})^2 - \bar{x}'_i \left(\sum_i X'_i \hat{V}_i^{-1} X_i \right)^{-1} \bar{x}_i \right. \right. \\ &\quad \left. \left. + 2(\bar{y}_i - \bar{x}'_i \hat{\beta}) \bar{x}'_i \left(\sum_i X'_i \hat{V}_i^{-1} X_i \right) \hat{\sigma}_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x'_{ij} \hat{\beta}) - n_i (1 - \hat{B}_i) \bar{x}_i (\bar{y}_i - \bar{x}'_i \hat{\beta}) \right] \right] \right. \\ &\quad \left. - 2 \frac{\partial B_i}{\partial \sigma_\nu^2} \Big|_{\delta=\hat{\delta}} (\hat{B}_i - w_i) (\bar{y}_i - \bar{x}'_i \hat{\beta})^2 \hat{b}(\hat{\sigma}_\nu^2 | y_i) - 2 \frac{\partial B_i}{\partial \sigma_e^2} \Big|_{\delta=\hat{\delta}} (\hat{B}_i - w_i) (\bar{y}_i - \bar{x}'_i \hat{\beta})^2 \hat{b}(\hat{\sigma}_e^2 | y_i) \right. \\ &\quad \left. - 2(\hat{B}_i - w_i)^2 (\bar{y}_i - \bar{x}'_i \hat{\beta}) \bar{x}'_i \left(\sum_i X'_i \hat{V}_i^{-1} X_i \right)^{-1} \sum_i (n_i \hat{\sigma}_\nu^2 + \hat{\sigma}_e^2)^{-2} X'_i J_{n_i} (y_i - X_i \hat{\beta}) \right. \\ &\quad \cdot \left[\hat{b}(\hat{\sigma}_\nu^2 | y_i) - \hat{\sigma}_\nu^2 \hat{\sigma}_e^2 \hat{b}(\hat{\sigma}_e^2 | y_i) \right] - \left(\left(\frac{\partial B_i}{\partial \sigma_\nu^2} \right)^2 \Big|_{\delta=\hat{\delta}} + \frac{\partial^2 B_i}{\partial \sigma_\nu^4} \Big|_{\delta=\hat{\delta}} (\hat{B}_i - w_i) \right) (\bar{y}_i - \bar{x}'_i \hat{\beta})^2 \hat{V}(\hat{\sigma}_\nu^2) \\ &\quad - \left(\left(\frac{\partial B_i}{\partial \sigma_e^2} \right)^2 \Big|_{\delta=\hat{\delta}} + \frac{\partial^2 B_i}{\partial \sigma_e^4} \Big|_{\delta=\hat{\delta}} (\hat{B}_i - w_i) \right) (\bar{y}_i - \bar{x}'_i \hat{\beta})^2 \hat{V}(\hat{\sigma}_e^2) - \left(\frac{\partial B_i}{\partial \sigma_\nu^2} \frac{\partial B_i}{\partial \sigma_e^2} \right. \\ &\quad \left. + \frac{\partial^2 B_i}{\partial \sigma_\nu^2 \partial \sigma_e^2} (\hat{B}_i - w_i) \right) \Big|_{\delta=\hat{\delta}} (\bar{y}_i - \bar{x}'_i \hat{\beta})^2 \widehat{\text{Cov}}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \Big\} + o_p(m^{-1}). \quad (6.21) \end{aligned}$$

Hence, by (6.18), (6.19) and (6.21), we may write

$$\begin{aligned} g_{wi}^c(\hat{\delta}) &\equiv (\hat{B}_i - w_i)^2 \left\{ (\bar{y}_i - \bar{x}'_i \hat{\beta})^2 - \bar{x}'_i \left(\sum_i X'_i \hat{V}_i^{-1} X_i \right)^{-1} \bar{x}_i + 2(\bar{y}_i - \bar{x}'_i \hat{\beta}) \bar{x}'_i \right. \\ &\quad \cdot \left(\sum_i X'_i \hat{V}_i^{-1} X_i \right)^{-1} \hat{\sigma}_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x'_{ij} \hat{\beta}) - n_i \bar{x}_i (1 - \hat{B}_i) (\bar{y}_i - \bar{x}'_i \hat{\beta}) \right] \Big\} \\ &\quad - 2(\bar{y}_i - \bar{x}'_i \hat{\beta})^2 (\hat{B}_i - w_i) \left(\frac{\partial B_i}{\partial \sigma_\nu^2} \Big|_{\delta=\hat{\delta}} \hat{b}(\hat{\sigma}_\nu^2 | y_i) + \frac{\partial B_i}{\partial \sigma_e^2} \Big|_{\delta=\hat{\delta}} \hat{b}(\hat{\sigma}_e^2 | y_i) \right) - 2(\hat{B}_i - w_i) (\bar{y}_i - \bar{x}'_i \hat{\beta}) \bar{x}'_i \\ &\quad \cdot \left(\sum_i X'_i \hat{V}_i^{-1} X_i \right)^{-1} \sum_i (n_i \hat{\sigma}_\nu^2 + \hat{\sigma}_e^2)^{-2} X'_i J_{n_i} (y_i - X_i \hat{\beta}) \left[\hat{b}(\hat{\sigma}_\nu^2 | y_i) - \hat{\sigma}_\nu^2 \hat{\sigma}_e^{-2} \hat{b}(\hat{\sigma}_e^2 | y_i) \right] \\ &\quad - (\bar{y}_i - \bar{x}'_i \hat{\beta})^2 \left\{ \left(\left(\frac{\partial B_i}{\partial \sigma_\nu^2} \right)^2 \Big|_{\delta=\hat{\delta}} + \frac{\partial^2 B_i}{\partial \sigma_\nu^4} \Big|_{\delta=\hat{\delta}} (\hat{B}_i - w_i) \right) \hat{V}(\hat{\sigma}_\nu^2) + \left(\left(\frac{\partial B_i}{\partial \sigma_e^2} \right)^2 \Big|_{\delta=\hat{\delta}} + \frac{\partial^2 B_i}{\partial \sigma_e^4} \Big|_{\delta=\hat{\delta}} \right. \right. \\ &\quad \left. \cdot (\hat{B}_i - w_i) \right) \hat{V}(\hat{\sigma}_e^2) \Big\} - \left(\frac{\partial B_i}{\partial \sigma_\nu^2} \frac{\partial B_i}{\partial \sigma_e^2} + \frac{\partial^2 B_i}{\partial \sigma_\nu^2 \partial \sigma_e^2} (\hat{B}_i - w_i) \right) \Big|_{\delta=\hat{\delta}} (\bar{y}_i - \bar{x}'_i \hat{\beta})^2 \widehat{\text{Cov}}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \end{aligned}$$

$$\begin{aligned}
& -(\hat{B}_i - w_i)^2 \bar{x}_i' (\sum_i X_i' \hat{V}_i^{-1} X_i)^{-1} \bar{x}_i + 2(\hat{B}_i - w_i) (\bar{X}_i - (1 - w_i) \bar{x}_i)' (\sum_i X_i' \hat{V}_i^{-1} X_i)^{-1} \\
& \cdot \left\{ (\bar{y}_i - \bar{x}_i' \hat{\beta}) \hat{\sigma}_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \hat{\beta}) - n_i \bar{x}_i (1 - \hat{B}_i) (\bar{y}_i - \bar{x}_i' \hat{\beta}) \right] - \bar{x}_i \right\} \\
& = (\bar{y}_i - \bar{x}_i' \hat{\beta})^2 \left\{ (\hat{B}_i - w_i)^2 - 2(\hat{B}_i - w_i) \left(\frac{\partial B_i}{\partial \sigma_\nu^2} \hat{b}(\hat{\sigma}_\nu^2 | y_i) + \frac{\partial B_i}{\partial \sigma_e^2} \hat{b}(\hat{\sigma}_e^2 | y_i) \right) \right\}_{\delta=\hat{\delta}} - [(\hat{B}_i - w_i) \\
& \cdot \left(\frac{\partial^2 B_i}{\partial \sigma_\nu^4} \right)_{\delta=\hat{\delta}} + \left(\frac{\partial^2 B_i}{\partial \sigma_\nu^2} \right)_{\delta=\hat{\delta}}] \hat{V}(\hat{\sigma}_\nu^2) - \left[\left(\frac{\partial^2 B_i}{\partial \sigma_e^2} \right)_{\delta=\hat{\delta}} + (\hat{B}_i - w_i) \frac{\partial^2 B_i}{\partial \sigma_e^4} \right]_{\delta=\hat{\delta}}] \hat{V}(\hat{\sigma}_e^2) \\
& - \left(\frac{\partial B_i}{\partial \sigma_\nu^2} \frac{\partial B_i}{\partial \sigma_e^2} \right)_{\delta=\hat{\delta}} + \frac{\partial^2 B_i}{\partial \sigma_\nu^2 \partial \sigma_e^2} \Big|_{\delta=\hat{\delta}} (\hat{B}_i - w_i) \widehat{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \Big\} - 2(\hat{B}_i - w_i) (\bar{y}_i - \bar{x}_i' \hat{\beta}) \bar{x}_i' \\
& \cdot \left(\sum_i X_i' \hat{V}_i^{-1} X_i \right)^{-1} \sum_i (n_i \hat{\sigma}_\nu^2 + \hat{\sigma}_e^2)^{-2} X_i' J_{n_i} (y_i - X_i \hat{\beta}) \left[\hat{b}(\hat{\sigma}_\nu^2 | y_i) - \hat{\sigma}_\nu^2 \hat{\sigma}_e^{-2} \hat{b}(\hat{\sigma}_e^2 | y_i) \right] \\
& + 2(\hat{B}_i - w_i) (\bar{X}_i - (1 - w_i) \bar{x}_i)' (\sum_i X_i' \hat{V}_i^{-1} X_i)^{-1} \left\{ (\bar{y}_i - \bar{x}_i' \hat{\beta}) \hat{\sigma}_e^{-2} \right. \\
& \cdot \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \hat{\beta}) - n_i (1 - \hat{B}_i) \bar{x}_i (\bar{y}_i - \bar{x}_i' \hat{\beta}) \right] - \bar{x}_i \Big\} + 2(\hat{B}_i - w_i)^2 \bar{x}_i' (\sum_i X_i' \hat{V}_i^{-1} X_i)^{-1} \\
& \cdot \left\{ (\bar{y}_i - \bar{x}_i' \hat{\beta}) \hat{\sigma}_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \hat{\beta}) - n_i (1 - \hat{B}_i) \bar{x}_i (\bar{y}_i - \bar{x}_i' \hat{\beta}) \right] - \bar{x}_i \right\}.
\end{aligned}$$

Therefore, we may write the second order unbiased estimator of $g_{wi1}^c(\delta) + 2g_{wi2}^c(\delta)$ as

$$\begin{aligned}
g_{wi}^c(\hat{\delta}) & = (\bar{y}_i - \bar{x}_i' \hat{\beta})^2 \left\{ (\hat{B}_i - w_i)^2 - 2(\hat{B}_i - w_i) \left(\frac{\partial B_i}{\partial \sigma_\nu^2} \Big|_{\delta=\hat{\delta}} \hat{b}(\hat{\sigma}_\nu^2 | y_i) + \frac{\partial B_i}{\partial \sigma_e^2} \Big|_{\delta=\hat{\delta}} \hat{b}(\hat{\sigma}_e^2 | y_i) \right) \right. \\
& - \left[\left(\frac{\partial B_i}{\partial \sigma_\nu^2} \right)_{\delta=\hat{\delta}}^2 + \frac{\partial^2 B_i}{\partial \sigma_\nu^4} \Big|_{\delta=\hat{\delta}} (\hat{B}_i - w_i) \right] \hat{V}(\hat{\sigma}_\nu^2) - \left[\left(\frac{\partial B_i}{\partial \sigma_e^2} \right)_{\delta=\hat{\delta}}^2 + \frac{\partial^2 B_i}{\partial \sigma_e^4} \Big|_{\delta=\hat{\delta}} (\hat{B}_i - w_i) \right] \\
& \cdot \hat{V}(\hat{\sigma}_e^2) - \left(\frac{\partial B_i}{\partial \sigma_\nu^2} \frac{\partial B_i}{\partial \sigma_e^2} + \frac{\partial^2 B_i}{\partial \sigma_\nu^2 \partial \sigma_e^2} (\hat{B}_i - w_i) \right) \Big|_{\delta=\hat{\delta}} \widehat{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \Big\} + 2(\hat{B}_i - w_i) \\
& \cdot (\bar{X}_i - (1 - \hat{B}_i) \bar{x}_i)' (\sum_i X_i' \hat{V}_i^{-1} X_i)^{-1} \left\{ (\bar{y}_i - \bar{x}_i' \hat{\beta}) \hat{\sigma}_e^{-2} \left[\sum_j x_{ij} (y_{ij} - x_{ij}' \hat{\beta}) \right. \right. \\
& \left. \left. - n_i \bar{x}_i (1 - \hat{B}_i) (\bar{y}_i - \bar{x}_i' \hat{\beta}) \right] - \bar{x}_i \right\} - 2(\hat{B}_i - w_i) (\bar{y}_i - \bar{x}_i' \hat{\beta}) \bar{x}_i' (\sum_i X_i' \hat{V}_i^{-1} X_i)^{-1} \sum_i (n_i \hat{\sigma}_\nu^2 + \hat{\sigma}_e^2)^{-2}
\end{aligned}$$

$$X_i' J_{n_i} (y_i - X_i \hat{\beta}) \left[\hat{b}(\hat{\sigma}_\nu^2 | y_i) - \hat{\sigma}_\nu^2 \hat{\sigma}_e^{-2} \hat{b}(\hat{\sigma}_e^2 | y_i) \right].$$

That means,

$$E[g_{wi}^c(\hat{\delta}) | y_i] = g_{wi1}^c(\delta) + 2g_{wi2}^c(\delta) + o_p(m^{-1}).$$

Theorem 6.4. *A second order unbiased estimator of $MSE^c[\hat{\mu}_{iw}(\hat{\delta}) | y_i]$ is given by*

$$mse^c[\hat{\mu}_{iw}(\hat{\delta}) | y_i] = g_{1i}(\hat{\delta}) - \hat{B}_i^2 \hat{b}(\hat{\sigma}_\nu^2 | y_i) - (1 - \hat{B}_i)^2 n_i^{-1} \hat{b}(\hat{\sigma}_e^2 | y_i) + g_{2i}(\hat{\delta}) + g_{3i}(\hat{\delta}) + g_{wi}^c(\hat{\delta}),$$

$$\text{namely, } E\left\{mse^c[\hat{\mu}_{iw}(\hat{\delta}) | y_i]\right\} + o_p(m^{-1}) = MSE^c[\hat{\mu}_{iw}(\hat{\delta}) | y_i].$$

6.6 Summary

In this Chapter, we studied the unconditional and conditional MSE of a weighed predictor of a small area mean with fixed weights under a nested error linear regression model. We derived a second order approximation to the unconditional and conditional MSE of the weighted predictor and then obtained estimators of MSE that are approximately unbiased.

Chapter 7

Structural Measurement Error Models Under Nested Error Linear Regression Model (EB Approach)

In this Chapter, we study the nested error linear regression model when the covariates are subject to measurement error. We study the effect of structural measurement error in the context of small area estimation. Specifically, we consider the nested error linear regression model and find the Bayes estimator when the sampling fractions are non-negligible in each area and covariate terms are subject to measurement error. We then estimate the variance components as well as the regression parameters. In addition, we derive a second order approximation to mean squared error (MSE) of empirical Bayes (EB) estimator in order to measure the precision of our estimator. Moreover, we derive an estimator of MSE which is nearly unbiased. We then show that the EB estimators are asymptotically optimal. Finally, we undertake

a simulation study to show the behavior of our estimator when the covariate terms are subject to measurement error and compare with the EB estimator when there is no error in the covariate terms. We study how much efficiency is lost if we have measurement error in our covariate terms.

7.1 Introduction

A nested error linear regression model may be written as:

$$y_{ij} = b_0 + b_1 x_{ij} + \nu_i + e_{ij}; j = 1, \dots, n_i, i = 1, \dots, m \quad (7.1)$$

In addition, $\nu_i \stackrel{i.i.d.}{\sim} N(0, \sigma_\nu^2)$ and $e_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma_e^2)$. Under measurement error on x , we have $X_{ij} = x_{ij} + u_{ij}$ as observed values, where $u_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma_u^2)$, (Fuller, 1987). In the structural case, the x_{ij} are assumed to be i.i.d. $N(\mu_x, \sigma_x^2)$. Moreover, we assume σ_u^2 is known and x_{ij}, ν_i, e_{ij} and u_{ij} are mutually independent. The goal is to estimate the i -th area mean $\gamma_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij}$ from the available data $\{(y_{ij}, X_{ij}), j = 1, \dots, n_i; i = 1, \dots, m\}$.

Letting y_i and y_i^* respectively denote sampled and nonsampled vectors of y , we have

$$[(y_i, y_i^*) | \theta] \sim N \left[(b_0 + b_1 \mu_x)(1_{n_i}, 1_{N_i - n_i}), \begin{pmatrix} \sigma_\nu^2 J_{n_i} + (b_1^2 \sigma_x^2 + \sigma_e^2) I_{n_i} & \sigma_\nu^2 1_{n_i} 1'_{N_i - n_i} \\ \sigma_\nu^2 1_{N_i - n_i} 1'_{n_i} & \sigma_\nu^2 J_{N_i - n_i} + (b_1^2 \sigma_x^2 + \sigma_e^2) I_{N_i - n_i} \end{pmatrix} \right],$$

with $\theta = (b_0, b_1, \mu_x, \sigma_x^2, \sigma_e^2, \sigma_\nu^2)'$. The Bayes (or best) predictor of y_i^* given y_i and θ can be then written as:

$$\begin{aligned}
E(y_i^*|y_i, \theta) &= (b_0 + b_1\mu_x)1_{N_i-n_i} + \sigma_\nu^2 1_{N_i-n_i} 1'_{n_i} \left[(\sigma_e^2 + b_1^2\sigma_x^2)^{-1} (I_{n_i} - \frac{1}{n_i} J_{n_i}) + \frac{1}{n_i} J_{n_i} \right. \\
&\quad \left. \cdot (b_1^2\sigma_x^2 + \sigma_e^2 + n_i\sigma_\nu^2)^{-1} \right] \left(y_i - (b_0 + b_1\mu_x)1_{n_i} \right) \\
&= (b_0 + b_1\mu_x)1_{N_i-n_i} + \sigma_\nu^2 1_{N_i-n_i} 1'_{n_i} (b_1^2\sigma_x^2 + \sigma_e^2 + n_i\sigma_\nu^2)^{-1} \left(y_i - (b_0 + b_1\mu_x)1_{n_i} \right) \\
&= \left[(b_0 + b_1\mu_x) + n_i\sigma_\nu^2 (b_1^2\sigma_x^2 + \sigma_e^2 + n_i\sigma_\nu^2)^{-1} (\bar{y}_i - b_0 - b_1\mu_x) \right] 1_{N_i-n_i} \\
&= \left[(1 - B_i)\bar{y}_i + B_i(b_0 + b_1\mu_x) \right] 1_{N_i-n_i},
\end{aligned}$$

where $B_i = (b_1^2\sigma_x^2 + \sigma_e^2)/(b_1^2\sigma_x^2 + \sigma_e^2 + n_i\sigma_\nu^2)$. It follows that the Bayes (or best) predictor of γ_i as:

$$\hat{\gamma}_i^B = E(\gamma_i|y_i) = (1 - f_i B_i)\bar{y}_i + f_i B_i(b_0 + b_1\mu_x),$$

where $f_i = 1 - n_i/N_i$.

7.2 Empirical Bayes

In this Section, we estimate θ to get the empirical Bayes (EB) estimator of γ_i . We estimate the model parameters based on the method of moment. In fact, we can write the mean and variance-covariance terms for the random variables X and y as follows: $\sigma_y^2 = b_1^2\sigma_x^2 + \sigma_\nu^2 + \sigma_e^2$; $\sigma_X^2 = \sigma_x^2 + \sigma_u^2$; $\mu_y = b_0 + b_1\mu_x$; $\mu_X = \mu_x$ and $\sigma_{Xy} = b_1\sigma_x^2$. Then the corresponding estimators are given by $m_{yy} = \hat{b}_1^2\hat{\sigma}_x^2 + \hat{\sigma}_e^2$, $m_{XX} = \hat{\sigma}_x^2 + \hat{\sigma}_u^2$; $\bar{y} = \hat{b}_0 + \hat{b}_1\hat{\mu}_x$; $\bar{X} = \hat{\mu}_x$ and $m_{Xy} = \hat{b}_1\hat{\sigma}_x^2$, where $\bar{y} = n^{-1} \sum_i n_i \bar{y}_i$, $\bar{X} = n^{-1} \sum_i n_i \bar{X}_i$, $m_{XX} = (n - m)^{-1} \sum_{i=1}^m \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$ and similar expression for m_{yy} and m_{Xy} . Therefore, we may have:

$$\begin{aligned}
\hat{\sigma}_e^2 &= m_{yy} - \hat{b}_1 m_{Xy}, \\
\hat{\sigma}_x^2 &= m_{XX} - \hat{\sigma}_u^2, \\
\hat{b}_0 &= \bar{y} - \hat{b}_1 \bar{X},
\end{aligned} \tag{7.2}$$

$$\hat{\mu}_x = \bar{X},$$

and

$$\hat{b}_1 = (m_{XX} - \sigma_u^2)^{-1} m_{Xy}. \quad (7.3)$$

In addition, we need to estimate σ_ν^2 . Writing m_{yy} as $m_{yy} = (n - m)^{-1} \sum_{i=1}^m \sum_{j=1}^{n_i} [b_1(x_{ij} - \bar{x}_i) + e_{ij} - \bar{e}_i]^2$, we have

$$E(m_{yy}) = (n - m)^{-1} \sum_i \sum_j \left[\left(1 - \frac{1}{n_i}\right) (b_1^2 \sigma_x^2 + \sigma_e^2) \right] = b_1^2 \sigma_x^2 + \sigma_e^2. \quad (7.4)$$

In addition, letting $M_{yy} = (m - 1)^{-1} \sum_{i=1}^m n_i (\bar{y}_i - \bar{y})^2$, we have

$$E(M_{yy}) = b_1^2 \sigma_x^2 + \sigma_e^2 + \sigma_\nu^2 g_m / (m - 1), \quad (7.5)$$

noting that $\bar{y}_i - \bar{y} = b_1(\bar{x}_i - \bar{x}) + \bar{e}_i - \bar{e} + \nu_i - \bar{\nu}$, where $\bar{x} = n^{-1} \sum_i n_i \bar{x}_i$, $\bar{e} = n^{-1} \sum_i n_i \bar{e}_i$ and $\bar{\nu} = n^{-1} \sum_i n_i \nu_i$ and $g_m = n - \sum_i n_i^2 / n$. Hence, by (7.4) and (7.5), an unbiased estimator of σ_ν^2 is given by

$$\hat{\sigma}_\nu^2 = \frac{m - 1}{g_m} (M_{yy} - m_{yy}).$$

To have proper estimators in the parameter space, $\hat{\sigma}_x^2$, $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$ must be nonnegative.

The estimators of σ_x^2 , σ_ν^2 and σ_e^2 will be positive if and only if

$$m_{yy}(m_{XX} - \sigma_u^2) - m_{Xy}^2 > 0 \quad (7.6)$$

for σ_x^2 and σ_e^2 , and

$$M_{yy} - m_{yy} > 0 \quad (7.7)$$

for σ_ν^2 . Then, estimates for samples in which (7.6) and (7.7) are violated are taken as $\hat{\sigma}_e^2 = 0$, $\hat{\sigma}_x^2 = m_{yy}^{-1} m_{Xy}^2$, $\hat{b}_1 = m_{Xy}^{-1} m_{yy}$, $\hat{b}_0 = \bar{y} - \hat{b}_1 \bar{X}$ and $\hat{\sigma}_\nu^2 = 0$.

The empirical Bayes (EB) predictor of γ_i is now given by

$$\hat{\gamma}_i^{EB} = (1 - f_i \hat{B}_i) \bar{y}_i + f_i \hat{B}_i (\hat{b}_0 + \hat{b}_1 \bar{X}), \quad (7.8)$$

where $\hat{B}_i = B_i|_{\delta=\hat{\delta}}$ with $\delta = (b_1, \sigma_x^2, \sigma_\nu^2, \sigma_e^2)'$.

7.3 MSE Approximation

In this Section, we derive a second order approximation to the mean squared error of the EB estimator $\hat{\gamma}_i^{EB}$ given by (7.8). We may write $MSE(\hat{\gamma}_i^{EB})$ as follows:

$$\begin{aligned} MSE(\hat{\gamma}_i^{EB}) &= E(\hat{\gamma}_i^{EB} - \gamma_i)^2 \\ &= E(\hat{\gamma}_i^B - \gamma_i)^2 + E(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B)^2 + 2E(\hat{\gamma}_i^B - \gamma_i)(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B), \end{aligned}$$

where $E(\hat{\gamma}_i^B - \gamma_i)^2 \equiv g_{1i}(\delta)$; $E(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B)^2 \equiv g_i(\delta) + o(m^{-1})$ and $E(\hat{\gamma}_i^B - \gamma_i)(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B) = E[(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B)E(\hat{\gamma}_i^B - \gamma_i|y_i)] = 0$. Then

$$MSE(\hat{\gamma}_i^{EB}) = g_{1i}(\delta) + g_i(\delta) + o(m^{-1}).$$

Therefore we need to find $g_{1i}(\delta)$ and $g_i(\delta)$.

We may write $g_{1i}(\delta) = E(\hat{\gamma}_i^B - \gamma_i)^2 = Var(\gamma_i|y_i) = N_i^{-2}Var(1'_{N_i-n_i}y_i^*|y_i)$, where

$$Var(y_i^*|y_i) = \sigma_\nu^2 J_{N_i-n_i} + (b_1^2 \sigma_x^2 + \sigma_e^2) I_{N_i-n_i} - \sigma_\nu^2 1_{N_i-n_i} 1'_{n_i} \left[\sigma_\nu^2 J_{n_i} + (b_1^2 \sigma_x^2 + \sigma_e^2) I_{n_i} \right]^{-1}$$

$$\cdot \sigma_\nu^2 1_{n_i} 1'_{N_i-n_i} = \sigma_\nu^2 J_{N_i-n_i} + (b_1^2 \sigma_x^2 + \sigma_e^2) I_{N_i-n_i} - \sigma_\nu^4 1_{N_i-n_i} 1'_{n_i} A^{-1} 1_{n_i} 1'_{N_i-n_i}$$

with $A^{-1} = (\sigma_e^2 + b_1^2 \sigma_x^2)^{-1} (I_{n_i} - \frac{1}{n_i} J_{n_i}) + \frac{1}{n_i} J_{n_i} (b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2)^{-1}$. Then

$$Var(y_i^*|y_i) = \sigma_\nu^2 J_{N_i-n_i} + (b_1^2 \sigma_x^2 + \sigma_e^2) I_{N_i-n_i} - \sigma_\nu^{-4} 1_{N_i-n_i} 1'_{n_i}$$

$$\begin{aligned}
& \cdot \left[(\sigma_e^2 + b_1^2 \sigma_x^2)^{-1} (I_{n_i} - \frac{1}{n_i} J_{n_i}) + \frac{1}{n_i} J_{n_i} (b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2)^{-1} \right] 1_{n_i} 1'_{N_i - n_i} \\
& = \sigma_\nu^2 J_{N_i - n_i} + (b_1^2 \sigma_x^2 + \sigma_e^2) I_{N_i - n_i} - \sigma_\nu^4 1_{N_i - n_i} 1'_{n_i} (b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2)^{-1} 1_{n_i} 1'_{N_i - n_i} \\
& = (b_1^2 \sigma_x^2 + \sigma_e^2) I_{N_i - n_i} + \frac{(b_1^2 \sigma_x^2 + \sigma_e^2) \sigma_\nu^2}{b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2} J_{N_i - n_i}.
\end{aligned}$$

Hence,

$$\begin{aligned}
g_{1i}(\delta) & = N_i^{-2} 1'_{N_i - n_i} \left[(b_1^2 \sigma_x^2 + \sigma_e^2) I_{N_i - n_i} + \frac{(b_1^2 \sigma_x^2 + \sigma_e^2) \sigma_\nu^2}{b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2} J_{N_i - n_i} \right] 1_{N_i - n_i} \\
& = N_i^{-2} \left[(N_i - n_i) (b_1^2 \sigma_x^2 + \sigma_e^2) + \sigma_\nu^2 (N_i - n_i)^2 \frac{b_1^2 \sigma_x^2 + \sigma_e^2}{b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2} \right] \\
& = (b_1^2 \sigma_x^2 + \sigma_e^2) f_i \left[N_i^{-1} + \frac{f_i \sigma_\nu^2}{b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2} \right]. \tag{7.9}
\end{aligned}$$

In order to find $g_i(\delta)$, we need the following terms:

$$\begin{aligned}
\partial \hat{\gamma}_i^B / \partial b_0 & = f_i B_i; \quad \partial \hat{\gamma}_i^B / \partial \mu_x = f_i B_i b_1; \\
\partial \hat{\gamma}_i^B / \partial b_1 & = f_i B_i \mu_x + \frac{2 f_i b_1 \sigma_x^2 (1 - B_i) (b_0 + b_1 \mu_x - \bar{y}_i)}{b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2}; \quad \partial B_i / \partial \sigma_x^2 = \frac{n_i b_1^2 \sigma_\nu^2}{(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)^2} \\
\partial B_i / \partial \sigma_\nu^2 & = -\frac{n_i (\sigma_e^2 + b_1^2 \sigma_x^2)}{(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)^2}; \quad \partial B_i / \partial \sigma_e^2 = \frac{n_i \sigma_\nu^2}{(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)^2}; \\
\partial \hat{\gamma}_i^B / \partial B_i & = f_i (b_0 + b_1 \mu_x - \bar{y}_i).
\end{aligned}$$

Hence,

$$\begin{aligned}
\partial \hat{\gamma}_i^B / \partial \sigma_x^2 & = \frac{f_i n_i b_1^2 \sigma_\nu^2 (b_0 + b_1 \mu_x - \bar{y}_i)}{(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)^2}; \quad \partial \hat{\gamma}_i^B / \partial \sigma_\nu^2 = -\frac{n_i f_i (\sigma_e^2 + b_1^2 \sigma_x^2) (b_0 + b_1 \mu_x - \bar{y}_i)}{(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)^2}; \\
\partial \hat{\gamma}_i^B / \partial \sigma_e^2 & = \frac{n_i f_i \sigma_\nu^2 (b_0 + b_1 \mu_x - \bar{y}_i)}{(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)^2}.
\end{aligned}$$

We may write:

$$\begin{aligned}
g_i(\delta) & = (f_i B_i)^2 \text{Var}(\hat{b}_0) + E[f_i B_i \mu_x + \frac{2 f_i b_1 \sigma_x^2 (1 - B_i) (b_0 + b_1 \mu_x - \bar{y}_i)}{b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2}]^2 \text{Var}(\hat{b}_1) \\
& + 2 f_i^2 B_i^2 \mu_x \text{Cov}(\hat{b}_0, \hat{b}_1) + (f_i B_i b_1)^2 \text{Var}(\hat{\mu}_x) + 2 f_i^2 B_i^2 b_1 \text{Cov}(\hat{\mu}_x, \hat{b}_0) + (1 - B_i)^2
\end{aligned}$$

$$\begin{aligned}
& \cdot \frac{f_i^2 b_1^4}{n_i(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)} \text{Var}(\hat{\sigma}_x^2) + \frac{n_i f_i^2 B_i^2}{\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2} \text{Var}(\hat{\sigma}_\nu^2) + \frac{f_i^2 (1 - B_i)^2}{n_i(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)} \\
& \cdot \text{Var}(\hat{\sigma}_e^2) + \frac{2f_i^2 (1 - B_i)^2 b_1^2}{n_i(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)} \text{Cov}(\hat{\sigma}_x^2, \hat{\sigma}_e^2) - \frac{2B_i (1 - B_i) b_1^2 f_i^2}{\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2} \text{Cov}(\hat{\sigma}_x^2, \hat{\sigma}_\nu^2) \\
& - \frac{2f_i^2 B_i (1 - B_i)}{\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2} \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) + \frac{4n_i f_i^2 b_1^3 \sigma_\nu^4 \sigma_x^2}{(b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2)^3} \text{Cov}(\hat{\sigma}_x^2, \hat{b}_1) \\
& - \frac{4n_i f_i^2 b_1 \sigma_x^2 \sigma_\nu^2 (\sigma_e^2 + b_1^2 \sigma_x^2)}{(b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2)^3} \text{Cov}(\hat{\sigma}_\nu^2, \hat{b}_1) + \frac{4n_i f_i^2 \sigma_x^2 \sigma_\nu^4 b_1}{(\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2)^3} \text{Cov}(\hat{\sigma}_e^2, \hat{b}_1) + o(m^{-1}) \\
& = n^{-1} f_i^2 B_i^2 \left[\left(\sum_i n_i^2 / n \right) \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2 \right] + \frac{4f_i^2 b_1^2 \sigma_x^4 (1 - B_i)^2}{n_i (b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2)} \text{Var}(\hat{b}_1) \\
& + \frac{f_i^2}{\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2} \left[\frac{(1 - B_i)^2}{n_i} \left(b_1^4 \text{Var}(\hat{\sigma}_x^2) + \text{Var}(\hat{\sigma}_e^2) + 2b_1^2 \text{Cov}(\hat{\sigma}_x^2, \hat{\sigma}_e^2) \right) \right. \\
& + B_i \left(n_i B_i \text{Var}(\hat{\sigma}_\nu^2) - 2(1 - B_i) b_1^2 \text{Cov}(\hat{\sigma}_x^2, \hat{\sigma}_\nu^2) - 2(1 - B_i) \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \right) + 4n_i^{-1} \sigma_x^2 b_1 (1 - B_i) \\
& \left. \cdot \left[b_1^2 (1 - B_i) \text{Cov}(\hat{\sigma}_x^2, \hat{b}_1) - n_i B_i \text{Cov}(\hat{\sigma}_\nu^2, \hat{b}_1) + (1 - B_i) \text{Cov}(\hat{\sigma}_e^2, \hat{b}_1) \right] \right] + o(m^{-1}). \quad (7.10)
\end{aligned}$$

Note that $E(b_0 + b_1 \mu_x - \bar{y}_i)^2 = \text{Var}(\bar{y}_i) = n_i^{-1} (b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2)$, $\text{Cov}(\hat{\mu}_x, \hat{b}_1) = 0$, $\text{Cov}(\hat{\mu}_x, \hat{b}_0) = -b_1 \sigma_u^2 / n$ and $\text{Var}(\hat{b}_1), \text{Var}(\hat{\sigma}_x^2), \text{Var}(\hat{\sigma}_e^2), \text{Var}(\hat{\sigma}_\nu^2), \text{Cov}(\hat{\sigma}_x^2, \hat{\sigma}_e^2), \text{Cov}(\hat{\sigma}_x^2, \hat{\sigma}_\nu^2), \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2), \text{Cov}(\hat{\sigma}_x^2, \hat{b}_1), \text{Cov}(\hat{\sigma}_\nu^2, \hat{b}_1), \text{Cov}(\hat{\sigma}_e^2, \hat{b}_1)$ are given in Section 7.4. In fact, to evaluate (7.10), we used the results $\text{Var}(\hat{b}_0), \text{Var}(\hat{b}_1)$ and $\text{Cov}(\hat{b}_0, \hat{b}_1)$.

Theorem 7.1. *The mean squared error of the EB estimator $\hat{\gamma}_i^{EB}$ is approximated by*

$$MSE(\hat{\gamma}_i^{EB}) = E(\hat{\gamma}_i^{EB} - \gamma_i)^2 \simeq g_{1i}(\delta) + g_i(\delta),$$

where $g_{1i}(\delta)$ and $g_i(\delta)$ are given by (7.9) and (7.10) and neglected terms are $o(m^{-1})$.

We derive the variance-covariance matrix of model parameter estimators in Section 7.4.

7.4 Variances and Covariances of Model Parameter Estimators

In this Section, we derive the variances and covariances of estimators of regression parameters and variance components. To find the variance-covariance matrix of estimators of regression parameters, we have the following theorem.

Theorem 7.2. *Let model (7.1) hold with σ_u^2 known, $\sigma_x^2 > 0$ and $\sigma_v^2 > 0$. Then $m^{1/2}[(\hat{b}_0 - b_0), (\hat{b}_1 - b_1)]$ where $(\hat{b}_0, \hat{b}_1)$ are defined by (7.2) and (7.3), is a normal vector variable with zero mean and covariance matrix with elements*

$$\begin{aligned} Var[m^{1/2}(\hat{b}_0 - b_0)] &= \frac{m}{n} \left[\left(\sum_i n_i^2 / n \right) \sigma_v^2 + \sigma_e^2 + b_1^2 \sigma_u^2 \right] + \frac{m}{n-m} \mu_x^2 \sigma_x^{-4} \left[(\sigma_x^2 + \sigma_u^2)(\sigma_v^2 + \sigma_e^2) \right. \\ &\quad \left. + b_1^2 \sigma_u^2 \right] + \frac{n-2m + \sum_i n_i^{-1}}{n-m} b_1^2 \sigma_u^4 + \sigma_v^2 (\sigma_x^2 + \sigma_u^2) + \sigma_v^2 \mu_x^2 (n-m)^{-1} \left(\sum_i n_i^2 - n \right) \Big] + O(m^{-1}), \end{aligned}$$

$$\begin{aligned} Var[m^{1/2}(\hat{b}_1 - b_1)] &= \frac{m}{n-m} \sigma_x^{-4} \left[(\sigma_x^2 + \sigma_u^2)(\sigma_v^2 + \sigma_e^2 + b_1^2 \sigma_u^2) + \frac{n-2m + \sum_i n_i^{-1}}{n-m} b_1^2 \sigma_u^4 \right. \\ &\quad \left. + \sigma_v^2 (\sigma_x^2 + \sigma_u^2) + \sigma_v^2 \mu_x^2 (n-m)^{-1} \left(\sum_i n_i^2 - n \right) \right] + O(m^{-1}) \end{aligned}$$

and

$$Cov[m^{1/2}(\hat{b}_0 - b_0), m^{1/2}(\hat{b}_1 - b_1)] = -\mu_x Var[m^{1/2}(\hat{b}_1 - b_1)] + O(m^{-1}).$$

Proof: Since the sample moments converge in probability to the population moments, we can expand $\hat{b}_1$ in a Taylor series about the population values to obtain:

$$\hat{b}_1 = (\sigma_X^2 - \sigma_u^2)^{-1} \sigma_{Xy} + (\sigma_X^2 - \sigma_u^2)^{-1} (m_{Xy} - \sigma_{Xy}) - (\sigma_X^2 - \sigma_u^2)^{-2} \sigma_{Xy} (m_{XX} - \sigma_X^2) + O_p(m^{-1})$$

where $\sigma_X^2 = \sigma_x^2 + \sigma_u^2$; $\sigma_{Xy} = b_1\sigma_x^2$. Then,

$$m^{1/2}(\hat{b}_1 - b_1) = m^{1/2}\sigma_x^{-2}(m_{Xy} - b_1(m_{XX} - \sigma_u^2)) + O_p(m^{-1/2}).$$

Therefore, we may write

$$m^{1/2}(\hat{b}_1 - b_1) = m^{1/2}\sigma_x^{-2}(m_{Xw} - \sigma_{Xw}) + O_p(m^{-1/2}), \quad (7.11)$$

where $w_{ij} = \nu_i + e_{ij} - b_1 u_{ij}$ and $m_{Xw} = (n - m)^{-1} \sum_i \sum_j (X_{ij} - \bar{X}_i) w_{ij}$. Note that

$$\begin{aligned} m_{Xw} - \sigma_{Xw} &= m_{Xw} + b_1\sigma_u^2 = (n - m)^{-1} \sum_i \sum_j (X_{ij} - \bar{X}_i) w_{ij} + b_1\sigma_u^2 \\ &= (n - m)^{-1} \sum_i \sum_j (X_{ij} - \bar{X}_i)(y_{ij} - b_0 - b_1 X_{ij}) + b_1\sigma_u^2 \\ &= (n - m)^{-1} \left[\sum_i \sum_j (X_{ij} - \bar{X}_i) y_{ij} - b_1 \sum_i \sum_j (X_{ij} - \bar{X}_i) X_{ij} \right] + b_1\sigma_u^2 \\ &= m_{Xy} - b_1 m_{XX} + b_1\sigma_u^2. \end{aligned}$$

Then, by (7.11), we may write:

$$\begin{aligned} \text{Var}[m^{1/2}(\hat{b}_1 - b_1)] &= \frac{m}{n - m} \sigma_x^{-4} \left[(\sigma_x^2 + \sigma_u^2)(\sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_u^2) + \frac{n - 2m + \sum_i n_i^{-1}}{n - m} b_1^2 \sigma_u^4 \right. \\ &\quad \left. + \sigma_\nu^2(\sigma_x^2 + \sigma_u^2) + \sigma_\nu^2 \mu_x^2 (n - m)^{-1} \left(\sum_i n_i^2 - n \right) \right] + O(m^{-1}). \end{aligned}$$

For $\hat{b}_0$, we may write:

$$\begin{aligned} \hat{b}_0 &= \bar{y} - \hat{b}_1 \bar{X} = b_0 + b_1 \bar{x} + \bar{\nu} + \bar{e} - \hat{b}_1(\bar{x} + \bar{u}) \\ &= b_0 - (\hat{b}_1 - b_1)\mu_x + \bar{w} + O_p(m^{-1}), \end{aligned}$$

where $\bar{w} = \bar{\nu} + \bar{e} - b_1 \bar{u}$. Then

$$m^{1/2}(\hat{b}_0 - b_0) = m^{1/2}(\bar{w} - (\hat{b}_1 - b_1)\mu_x) + O_p(m^{-1/2}). \quad (7.12)$$

Therefore, we may write

$$\begin{aligned}
Var[m^{1/2}(\hat{b}_0 - b_0)] &= m[(\sum_i n_i^2/n^2)\sigma_\nu^2 + \frac{\sigma_e^2}{n} + b_1^2 \frac{\sigma_u^2}{n}] + \mu_x^2 Var[m^{1/2}(\hat{b}_1 - b_1)] \\
&\quad - 2\mu_x Cov[m^{1/2}(\hat{b}_1 - b_1), m^{1/2}\bar{w}] + O(m^{-1}) \\
&= (m/n)[(\sum_i n_i^2/n)\sigma_\nu^2 + b_1^2\sigma_u^2 + \sigma_e^2] + \mu_x^2 Var[m^{1/2}(\hat{b}_1 - b_1)] + O(m^{-1}),
\end{aligned}$$

since

$$\begin{aligned}
Cov(m_{Xw}, \bar{w}) &= Cov[(n-m)^{-1} \sum_i \sum_j (X_{ij} - \bar{X}_i)w_{ij}, \bar{w}] \\
&= (n-m)^{-1} [\sum_i \sum_j Cov(X_{ij}w_{ij}, \bar{w}) - \sum_i n_i Cov(\bar{X}_i \bar{w}_i, \bar{w})] = 0.
\end{aligned}$$

Hence,

$$\begin{aligned}
Var[m^{1/2}(\hat{b}_0 - b_0)] &= (m/n)[(\sum_i n_i^2/n)\sigma_\nu^2 + \sigma_e^2 + b_1^2\sigma_u^2] + \frac{m}{n-m}\mu_x^2\sigma_x^{-4}[(\sigma_x^2 + \sigma_u^2) \\
&\quad \cdot (\sigma_\nu^2 + \sigma_e^2 + b_1^2\sigma_u^2) + \frac{n-2m + \sum_i n_i^{-1}}{n-m}b_1^2\sigma_u^4 + \sigma_\nu^2(\sigma_x^2 + \sigma_u^2) + \sigma_\nu^2\mu_x^2(n-m)^{-1}(\sum_i n_i^2 - n)] \\
&\quad + O(m^{-1}).
\end{aligned}$$

Finally,

$$Cov[m^{1/2}(\hat{b}_0 - b_0), m^{1/2}(\hat{b}_1 - b_1)] = -\mu_x Var[m^{1/2}(\hat{b}_1 - b_1)] + O(m^{-1}),$$

by (7.11) and (7.12). Hence, the results of Theorem 7.2 are obtained.

In addition, we have

$$Var(\hat{\mu}_x) = Var(\bar{X}) = \frac{\sigma_x^2 + \sigma_u^2}{n}.$$

Also,

$$Var(\hat{\sigma}_x^2) = Var(m_{XX}) = 2(n-m)^{-2} \left[(n-2m + \sum_i n_i^{-1})(\sigma_x^2 + \sigma_u^2)^2 \right]$$

$$+(\sigma_x^4 + \sigma_u^4) \sum_i (1 - 1/n_i)(3 - 4/n_i) \Big]. \quad (7.13)$$

Moreover, $Var(\hat{\sigma}_\nu^2) = (\frac{m-1}{g_m})^2 [Var(M_{yy}) + Var(m_{yy})]$, where

$$\begin{aligned} Var(M_{yy}) = 2(m-1)^{-2} \Big\{ \sum_i n_i^2 \Big[(b_1^2 \sigma_x^2 + \sigma_e^2)^2 / n_i^2 + \sigma_\nu^4 \\ + 2(b_1^2 \sigma_x^2 + \sigma_e^2) \sigma_\nu^2 / n_i \Big] \Big\} + o(m^{-1}), \end{aligned} \quad (7.14)$$

and

$$\begin{aligned} Var(m_{yy}) = 2(n-m)^{-2} \Big[(n-2m + \sum_i n_i^{-1})(b_1^2 \sigma_x^2 + \sigma_e^2)^2 \\ + (b_1^4 \sigma_x^4 + \sigma_e^4) \sum_i (1 - 1/n_i)(3 - 4/n_i) \Big]. \end{aligned} \quad (7.15)$$

Note that $Cov(M_{yy}, m_{yy}) = 0$. Then, by (7.14) and (7.15), we have $Var(\hat{\sigma}_\nu^2)$.

In addition, $Var(\hat{\sigma}_e^2) = Var(m_{yy}) + Var(\hat{b}_1 m_{Xy}) - 2Cov(m_{yy}, \hat{b}_1 m_{Xy})$, where $Var(\hat{b}_1 m_{Xy}) = Var[(m_{XX} - \sigma_u^2)^{-1} m_{Xy}^2]$. We may write the following term based on Taylor expansion:

$$\hat{b}_1 m_{Xy} = b_1^2 \sigma_x^2 + 2b_1(m_{Xy} - b_1 \sigma_x^2) - b_1^2(m_{XX} - \sigma_x^2 - \sigma_u^2) + O_p(m^{-1}). \quad (7.16)$$

Then, we may write

$$Var[\hat{b}_1 m_{Xy}] = b_1^2 \Big[4Var(m_{Xy}) + b_1^2 Var(m_{XX}) - 4b_1 Cov(m_{XX}, m_{Xy}) \Big] + o(m^{-1}),$$

where

$$\begin{aligned} Var(m_{Xy}) = (n-m)^{-2} \Big[(n-2m + \sum_i n_i^{-1}) \Big[(\sigma_x^2 + \sigma_u^2)(b_1^2 \sigma_x^2 + \sigma_e^2) + b_1^2 \sigma_x^4 \Big] \\ + 2b_1^2 \sigma_x^4 \sum_i (1 - 1/n_i)(3 - 4/n_i) \Big], \end{aligned} \quad (7.17)$$

and

$$\begin{aligned} Cov(m_{XX}, m_{Xy}) &= 2b_1\sigma_x^2(n-m)^{-2} \left[(n-2m + \sum_i n_i^{-1})(\sigma_x^2 + \sigma_u^2) \right. \\ &\quad \left. + \sigma_x^2 \sum_i (1 - 1/n_i)(3 - 4/n_i) \right]. \end{aligned} \quad (7.18)$$

In addition, by (7.16), we have

$$Cov(m_{yy}, \hat{b}_1 m_{Xy}) = -b_1^2 Cov(m_{yy}, m_{XX}) + 2b_1 Cov(m_{yy}, m_{Xy}) + o(m^{-1}), \quad (7.19)$$

where

$$Cov(m_{yy}, m_{XX}) = 2b_1^2\sigma_x^4(n-m)^{-2} \left[(n-2m + \sum_i n_i^{-1}) + \sum_i (1 - 1/n_i)(3 - 4/n_i) \right], \quad (7.20)$$

and

$$\begin{aligned} Cov(m_{yy}, m_{Xy}) &= 2b_1\sigma_x^2(n-m)^{-2} \left[(n-2m + \sum_i n_i^{-1})(b_1^2\sigma_x^2 + \sigma_e^2) \right. \\ &\quad \left. + b_1^2\sigma_x^2 \sum_i (1 - 1/n_i)(3 - 4/n_i) \right]. \end{aligned} \quad (7.21)$$

Therefore, combining above results, $Var(\hat{\sigma}_e^2)$ is obtained.

In addition, we derive the covariance terms as follows:

$$\begin{aligned} Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) &= -\frac{m-1}{g_m} Var(m_{yy}) + \frac{m-1}{g_m} Cov(m_{yy}, \hat{b}_1 m_{Xy}) + \frac{m-1}{g_m} Cov(M_{yy}, m_{yy}) \\ &\quad - \frac{m-1}{g_m} Cov(M_{yy}, \hat{b}_1 m_{Xy}) \\ &= \frac{(m-1)}{g_m} \left[Cov(m_{yy}, \hat{b}_1 m_{Xy}) - Var(m_{yy}) \right] + o(m^{-1}), \end{aligned}$$

since $Cov(M_{yy}, m_{yy}) = 0$ and $Cov(M_{yy}, \hat{b}_1 m_{Xy}) = 0$ (see the Appendix for details).

Therefore, by (7.15) and (7.19), $Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)$ is obtained.

Moreover, we may write

$$\begin{aligned} Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_x^2) &= \frac{m-1}{g_m} \left[Cov(M_{yy}, m_{XX}) - Cov(m_{yy}, m_{XX}) \right] \\ &= -2b_1^2 \sigma_x^4 (m-1)(n-m)^{-2} g_m^{-1} \left[(n-2m + \sum_i n_i^{-1}) + \sum_i (1-1/n_i)(3-4/n_i) \right] + o(m^{-1}), \end{aligned}$$

since $Cov(M_{yy}, m_{XX}) = 0$ (see the Appendix).

In addition, we have

$$Cov(\hat{\sigma}_e^2, \hat{\sigma}_x^2) = Cov(m_{yy}, m_{XX}) - Cov(m_{XX}, \hat{b}_1 m_{Xy}),$$

where $Cov(m_{yy}, m_{XX})$ is given by (7.20), and by (7.16),

$$Cov(m_{XX}, \hat{b}_1 m_{Xy}) = -b_1^2 Var(m_{XX}) + 2b_1 Cov(m_{XX}, m_{Xy}) + o(m^{-1}), \quad (7.22)$$

where $Var(m_{XX})$ and $Cov(m_{XX}, m_{Xy})$ are given by (7.13) and (7.18) respectively.

Therefore, combining above results, $Cov(\hat{\sigma}_e^2, \hat{\sigma}_x^2)$ is obtained.

Moreover, we may write

$$Cov(\hat{\sigma}_x^2, \hat{b}_1) = \sigma_x^{-2} \left[Cov(m_{XX}, m_{Xy}) - b_1 Var(m_{XX}) \right] + o(m^{-1}).$$

Also, we have that

$$Cov(\hat{\sigma}_\nu^2, \hat{b}_1) = -\frac{(m-1)}{\sigma_x^2 g_m} \left[Cov(m_{yy}, m_{Xy}) - b_1 Cov(m_{yy}, m_{XX}) \right] + o(m^{-1}). \quad (7.23)$$

Then by (7.20) and (7.21), we obtain $Cov(\hat{\sigma}_\nu^2, \hat{b}_1)$.

In addition, we may write

$$\begin{aligned} Cov(\hat{\sigma}_e^2, \hat{b}_1) &= Cov(m_{yy}, \hat{b}_1) - Cov(\hat{b}_1 m_{Xy}, \hat{b}_1) \\ &= -b_1 \sigma_x^{-2} (Cov(m_{yy}, m_{XX}) - Cov(\hat{b}_1 m_{Xy}, m_{XX})) \\ &+ \sigma_x^{-2} (Cov(m_{yy}, m_{Xy}) - Cov(\hat{b}_1 m_{Xy}, m_{Xy})) + o(m^{-1}), \end{aligned} \quad (7.24)$$

where $Cov(m_{yy}, m_{XX}), Cov(\hat{b}_1 m_{Xy}, m_{XX}), Cov(m_{yy}, m_{Xy})$ are given by (7.20), (7.22) and (7.21) respectively. Also, using (7.16),

$$Cov(\hat{b}_1 m_{Xy}, m_{Xy}) = -b_1^2 Cov(m_{XX}, m_{Xy}) + 2b_1 Var(m_{Xy}) + o(m^{-1}),$$

where $Cov(m_{XX}, m_{Xy})$ and $Var(m_{Xy})$ are given by (7.18) and (7.17) respectively. Hence, $Cov(\hat{b}_1 m_{Xy}, m_{Xy})$ is obtained. Finally, $Cov(\hat{\sigma}_e^2, \hat{b}_1)$ is obtained from (7.24) using above results.

7.5 MSE Estimation

We derive a second order unbiased estimator of mean squared error of $\hat{\gamma}_i^{EB}$ in the sense that $E[mse(\hat{\gamma}_i^{EB})] - MSE(\hat{\gamma}_i^{EB}) = o(m^{-1})$. To evaluate the bias of $g_{1i}(\hat{\delta})$ as an estimator of $g_{1i}(\delta)$, we use a Taylor expansion of $g_{1i}(\hat{\delta})$ around δ :

$$E(g_{1i}(\hat{\delta})) = g_{1i}(\delta) + \frac{\partial g_{1i}}{\partial \delta'} E(\hat{\delta} - \delta) + \frac{1}{2} tr \left\{ \frac{\partial^2 g_{1i}}{\partial \delta \partial \delta'} E(\hat{\delta} - \delta)(\hat{\delta} - \delta)' \right\} + o(m^{-1}). \quad (7.25)$$

To evaluate the bias of $\hat{\delta}$, we may write

$$E(\hat{\sigma}_x^2) = E(m_{XX}) - \sigma_u^2 = \sigma_x^2 + \sigma_u^2 - \sigma_u^2 = \sigma_x^2.$$

That means, $\hat{\sigma}_x^2$ is an unbiased estimator of σ_x^2 . Also, we show that $\hat{\sigma}_\nu^2$ is an unbiased estimator of σ_ν^2 as follows:

$$E(\hat{\sigma}_\nu^2) = \frac{m-1}{g_m} [E(M_{yy}) - E(m_{yy})] = \frac{m-1}{g_m} \left[\frac{g_m}{m-1} \sigma_\nu^2 + b_1^2 \sigma_x^2 + \sigma_e^2 - b_1^2 \sigma_x^2 - \sigma_e^2 \right] = \sigma_\nu^2.$$

Moreover, to find the bias of $\hat{\sigma}_e^2$, we may write the following Taylor expansion

$$\hat{b}_1 m_{Xy} = b_1^2 \sigma_x^2 + 2b_1(m_{Xy} - b_1 \sigma_x^2) - b_1^2(m_{XX} - \sigma_x^2 - \sigma_u^2) + \sigma_x^{-2}(m_{Xy} - b_1 \sigma_x^2)^2$$

$$+b_1^2\sigma_x^{-2}(m_{XX}-\sigma_x^2-\sigma_u^2)^2-2b_1\sigma_x^{-2}(m_{Xy}-b_1\sigma_x^2)(m_{XX}-\sigma_x^2-\sigma_u^2)+o_p(m^{-1}),$$

then,

$$E(\hat{b}_1 m_{Xy}) = b_1^2\sigma_x^2 + \sigma_x^{-2} \left[Var(m_{Xy}) + b_1^2 Var(m_{XX}) - 2b_1 Cov(m_{Xy}, m_{XX}) \right] + o(m^{-1}),$$

where $Var(m_{XX})$, $Cov(m_{Xy}, m_{XX})$ and $Var(m_{Xy})$ are given by (7.13), (7.18) and (7.17) respectively. Hence, we can write

$$E(\hat{\sigma}_e^2 - \sigma_e^2) = -\sigma_x^{-2} \left[Var(m_{Xy}) + b_1^2 Var(m_{XX}) - 2b_1 Cov(m_{Xy}, m_{XX}) \right] + o(m^{-1}).$$

In addition, to find the bias of $\hat{b}_1$, we may write

$$\begin{aligned} \hat{b}_1 &= b_1 - b_1\sigma_x^{-2}(m_{XX}-\sigma_X^2) + \sigma_x^{-2}(m_{Xy}-\sigma_{Xy}) + b_1\sigma_x^{-4}(m_{XX}-\sigma_X^2)^2 - \sigma_x^{-4}(m_{XX}-\sigma_X^2) \\ &\quad \cdot (m_{Xy}-\sigma_{Xy}) + o_p(m^{-1}). \end{aligned}$$

Then,

$$E(\hat{b}_1 - b_1) = \sigma_x^{-4} \left[b_1 Var(m_{XX}) - Cov(m_{XX}, m_{Xy}) \right] + o(m^{-1}).$$

To evaluate (7.25), we need to have the following terms:

$$\begin{aligned} \partial g_{1i}/\partial b_1 &= 2b_1 f_i \sigma_x^2 [N_i^{-1} + n_i^{-1} f_i (1 - B_i)^2], \\ \partial g_{1i}/\partial \sigma_e^2 &= f_i \left[N_i^{-1} + f_i (1 - B_i)^2 / n_i \right], \\ \partial^2 g_{1i}/\partial b_1^2 &= 2f_i \sigma_x^2 [N_i^{-1} + f_i n_i^{-1} (1 - B_i)^2] - \frac{8b_1^2 f_i^2 \sigma_x^4 (1 - B_i)^2}{n_i (n_i \sigma_\nu^2 + b_1^2 \sigma_x^2 + \sigma_e^2)}, \\ \partial^2 g_{1i}/\partial \sigma_\nu^4 &= -2n_i f_i^2 B_i^2 / (n_i \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2), \\ \partial^2 g_{1i}/\partial \sigma_x^4 &= -2b_1^4 f_i^2 (1 - B_i)^2 / n_i (n_i \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2), \\ \partial^2 g_{1i}/\partial \sigma_e^4 &= -2n_i^{-1} f_i^2 (1 - B_i)^2 / (n_i \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2), \\ \partial^2 g_{1i}/\partial b_1 \partial \sigma_\nu^2 &= 4b_1 f_i^2 \sigma_x^2 B_i (1 - B_i) / (n_i \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2), \end{aligned}$$

$$\begin{aligned}
\partial^2 g_{1i}/\partial b_1 \partial \sigma_x^2 &= 2b_1 f_i [N_i^{-1} + f_i n_i^{-1} (1 - B_i)^2] - 4b_1^3 f_i^2 \sigma_x^2 (1 - B_i)^2 / n_i (n_i \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2), \\
\partial^2 g_{1i}/\partial b_1 \partial \sigma_e^2 &= -4b_1 f_i^2 \sigma_x^2 (1 - B_i)^2 / n_i (n_i \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2), \\
\partial^2 g_{1i}/\partial \sigma_\nu^2 \partial \sigma_x^2 &= 2b_1^2 f_i^2 B_i (1 - B_i) / (n_i \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2), \\
\partial^2 g_{1i}/\partial \sigma_\nu^2 \partial \sigma_e^2 &= 2f_i^2 B_i (1 - B_i) / (n_i \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2), \\
\partial^2 g_{1i}/\partial \sigma_x^2 \partial \sigma_e^2 &= -2b_1^2 f_i^2 (1 - B_i)^2 / n_i (n_i \sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_x^2).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\frac{1}{2} \text{tr} \left\{ \frac{\partial^2 g_{1i}}{\partial \delta \partial \delta'} E(\hat{\delta} - \delta)(\hat{\delta} - \delta)' \right\} &= \frac{1}{2} \left[(\partial^2 g_{1i}/\partial b_1^2) \text{Var}(\hat{b}_1) + (\partial^2 g_{1i}/\partial \sigma_\nu^4) \text{Var}(\hat{\sigma}_\nu^2) \right. \\
&\quad \left. + (\partial^2 g_{1i}/\partial \sigma_x^4) \text{Var}(\hat{\sigma}_x^2) + (\partial^2 g_{1i}/\partial \sigma_e^4) \text{Var}(\hat{\sigma}_e^2) \right] + (\partial^2 g_{1i}/\partial b_1 \partial \sigma_\nu^2) \text{Cov}(\hat{b}_1, \hat{\sigma}_\nu^2) \\
&\quad + (\partial^2 g_{1i}/\partial b_1 \partial \sigma_x^2) \text{Cov}(\hat{b}_1, \hat{\sigma}_x^2) + (\partial^2 g_{1i}/\partial b_1 \partial \sigma_e^2) \text{Cov}(\hat{b}_1, \hat{\sigma}_e^2) + (\partial^2 g_{1i}/\partial \sigma_\nu^2 \partial \sigma_x^2) \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_x^2) \\
&\quad + (\partial^2 g_{1i}/\partial \sigma_\nu^2 \partial \sigma_e^2) \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) + (\partial^2 g_{1i}/\partial \sigma_x^2 \partial \sigma_e^2) \text{Cov}(\hat{\sigma}_x^2, \hat{\sigma}_e^2) \equiv -g'_i(\delta).
\end{aligned}$$

On the other hand, we have $E[g_i(\hat{\delta})] - g_i(\delta) = o(m^{-1})$.

Theorem 7.3. *A second order unbiased estimator of $MSE(\hat{\gamma}_i^{EB})$ is given by*

$$mse(\hat{\gamma}_i^{EB}) = g_{1i}(\hat{\delta}) - \frac{\partial g_{1i}}{\partial \sigma_e^2} E(\hat{\sigma}_e^2 - \sigma_e^2)|_{\delta=\hat{\delta}} - \frac{\partial g_{1i}}{\partial b_1} E(\hat{b}_1 - b_1)|_{\delta=\hat{\delta}} + g_i(\hat{\delta}) + g'_i(\hat{\delta}), \quad (7.26)$$

in the sense that $E[mse(\hat{\gamma}_i^{EB})] - MSE(\hat{\gamma}_i^{EB}) = o(m^{-1})$.

7.6 Asymptotic Result on $m^{-1} \sum_{i=1}^m E(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B)^2$

In this Section, we show that $\hat{\gamma}_i^{EB}$ is asymptotically optimal in the sense that the Bayes risk

$$m^{-1} \sum_{i=1}^m E(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B)^2 \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

We know that $\hat{\gamma}_i^B = (1 - f_i B_i) \bar{y}_i + f_i B_i (b_0 + b_1 \mu_x)$ and $\hat{\gamma}_i^{EB} = (1 - f_i \hat{B}_i) \bar{y}_i + f_i \hat{B}_i \bar{y}$, where $\bar{y}$ is sample mean of y_{ij} for the whole sample. To evaluate the Bayes risk, we write

$$\begin{aligned} \hat{\gamma}_i^{EB} - \hat{\gamma}_i^B &= (1 - f_i \hat{B}_i) \bar{y}_i + f_i \hat{B}_i \bar{y} - (1 - f_i B_i) \bar{y}_i - f_i B_i (b_0 + b_1 \mu_x) \\ &= f_i \left[(B_i - \hat{B}_i) (\bar{y}_i - \bar{y}) + B_i (\bar{y} - b_0 - b_1 \mu_x) \right]. \end{aligned}$$

Then,

$$m^{-1} \sum_{i=1}^m E(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B)^2 \leq 2m^{-1} \sum_{i=1}^m E \left[(B_i - \hat{B}_i)^2 (\bar{y}_i - \bar{y})^2 + B_i^2 (\bar{y} - b_0 - b_1 \mu_x)^2 \right],$$

by C_r inequality. However, we have that

$$\begin{aligned} E(\bar{y} - b_0 - b_1 \mu_x)^2 &= Var(\bar{y}) = Var(n^{-1} \sum_i n_i \bar{y}_i) = n^{-2} \sum_i n_i^2 (\sigma_\nu^2 + (b_1^2 \sigma_x^2 + \sigma_e^2)/n_i) \\ &= n^{-2} [nb_1^2 \sigma_x^2 + \sigma_\nu^2 (\sum_i n_i^2) + n\sigma_e^2] \\ &\leq n^{-1} [b_1^2 \sigma_x^2 + \sigma_e^2 + k_0 \sigma_\nu^2] \rightarrow 0 \quad \text{as } m \rightarrow \infty, \end{aligned}$$

where $\max_{1 \leq i \leq m} n_i \leq k_0 < \infty$. Also, we may write

$$\begin{aligned} \max_{1 \leq i \leq m} |\hat{B}_i - B_i| &= \max_{1 \leq i \leq m} \left| \frac{\hat{\sigma}_e^2 + \hat{b}_1^2 \hat{\sigma}_x^2}{\hat{\sigma}_e^2 + \hat{b}_1^2 \hat{\sigma}_x^2 + n_i \hat{\sigma}_\nu^2} - \frac{\sigma_e^2 + b_1^2 \sigma_x^2}{\sigma_e^2 + b_1^2 \sigma_x^2 + n_i \sigma_\nu^2} \right| \\ &= \max_{1 \leq i \leq m} \left| \frac{1}{1 + n_i \hat{M}} - \frac{1}{1 + n_i M} \right| = \max_{1 \leq i \leq m} \frac{n_i |M - \hat{M}|}{(1 + n_i \hat{M})(1 + n_i M)} \leq n_i |M - \hat{M}| \end{aligned}$$

where $M = \sigma_\nu^2 / (\sigma_e^2 + b_1^2 \sigma_x^2)$. Note that $\hat{M} \rightarrow M$ in probability and also $m^{-1} \sum_i n_i (\bar{y}_i - \bar{y})^2 = \frac{m-1}{m} M_{yy}$ such that

$$M_{yy} \rightarrow b_1^2 \sigma_x^2 + \sigma_e^2 + c \sigma_\nu^2 \quad \text{in probability,} \quad (7.27)$$

where $\frac{cm}{m-1} \rightarrow c(>0)$ as $m \rightarrow \infty$. It is easy to show (7.27) using a similar proof the one used in Ghosh, Sinha, and Kim (2006), Lemma 1. Hence, by combining above

results, we have

$$m^{-1} \sum_{i=1}^m (\hat{B}_i - B_i)^2 (\bar{y}_i - \bar{y})^2 \rightarrow 0 \quad \text{in probability.}$$

In addition, $m^{-1} \sum_i (\hat{B}_i - B_i)^2 (\bar{y}_i - \bar{y})^2 \leq m^{-1} \sum_i n_i (\bar{y}_i - \bar{y})^2$. Since we have

$$\begin{aligned} E \left[m^{-1} \sum_i n_i (\bar{y}_i - \bar{y})^2 \right]^2 &= \text{Var} \left[m^{-1} \sum_i n_i (\bar{y}_i - \bar{y})^2 \right] + \left[E \left[m^{-1} \sum_i n_i (\bar{y}_i - \bar{y})^2 \right] \right]^2 \\ &= 2m^{-2} \sum_i (b_1^2 \sigma_x^2 + \sigma_e^2 + n_i \sigma_\nu^2)^2 + \left(\frac{m-1}{m} \right)^2 \left(b_1^2 \sigma_x^2 + \sigma_e^2 + (m-1)^{-1} g_m \sigma_\nu^2 \right)^2 + o(m^{-1}) \\ &= O(1), \end{aligned}$$

it follows that $\sup_{m \geq 1} m^{-1} \sum_{i=1}^m (\hat{B}_i - B_i)^2 (\bar{y}_i - \bar{y})^2$ is uniformly integrable in m and $E \left[m^{-1} \sum_i (\hat{B}_i - B_i)^2 (\bar{y}_i - \bar{y})^2 \right] \rightarrow 0$ as $m \rightarrow \infty$ (see Resnick, 1999, p.191). Therefore, by combining above results, we may write

$$m^{-1} \sum_{i=1}^m E(\hat{\gamma}_i^{EB} - \hat{\gamma}_i^B)^2 \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

7.7 Simulation Study

We conducted a simulation study to compare the performance of the proposed method with the naive method when the covariates are subject to measurement errors. To this end, we used the data in Battese et al. (1988). We have already used this data set in Chapters 3 and 5. Battese et al. (1988) considered the estimation of mean hectares of soybeans per segment for 12 counties (area) in north-central Iowa. The total sample size, $n = \sum_i n_i$, for the 12 counties is 36, and the sample size n_i within each county ranged from 1 to 5. The population size N_i within each county

ranged from 402 to 965. The nested error regression model is

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \nu_i + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m$$

where y_{ij} is the number of hectares of soybeans in the j -th segment of the i -th county, x_{1ij} and x_{2ij} are the number of pixels classified as corn and soybeans respectively, in the j -th segment of the i -th county. Battese et. al (1988) observed that only the coefficient of soybean pixels is significantly different from 0 for the soybean function. Therefore, we consider prediction of soybean data using soybean pixels only as co-variates in our simulation study. Hence, the nested error regression model reduces to

$$y_{ij} = b_0 + b_1 x_{ij} + \nu_i + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m$$

However, the x_{ij} , soybean pixels, are not observable, instead we have $X_{ij} = x_{ij} + u_{ij}$ as observed values. If we fit the above regression model to data set given by Battese et al. (1988), we have $\hat{b}_0 = -3.12$, $\hat{b}_1 = 0.472$, $\hat{\sigma}_\nu^2 = 259.93$ and $\hat{\sigma}_e^2 = 191.14$. We assume $\sigma_u^2 = 500$ as known. We also used the population area means of x_{ij} to estimate σ_x^2 as $\hat{\sigma}_x^2 = (m-1)^{-1} \sum_i^m (\bar{X}_{ip} - \bar{X}_p)^2 = 527.88$, where $\bar{X}_{ip}$ is the i -th population area mean of x_{ij} and $\bar{X}_p = m^{-1} \sum_i^m \bar{X}_{ip}$. We duplicated x_{ij} in our simulation study based on the number of areas. We treated $\hat{b}_0, \hat{b}_1, \hat{\sigma}_\nu^2, \hat{\sigma}_e^2$ and $\hat{\sigma}_x^2$ as true parameters in our simulation study.

We generated $B = 5000$ independent sets of normally distributed random variables $\{\nu_i; i = 1, \dots, m\}, \{e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m\}$ having mean zero and specified variances σ_ν^2 and σ_e^2 . We also generated $\{x_{ij}; j = 1, \dots, n_i; i = 1, \dots, m\}$ with mean $\bar{X}_p$ and variance σ_x^2 . From these generated data sets, observations $\{(y_{ij}, X_{ij}); j = 1, \dots, n_i; i = 1, \dots, m\}$ are obtained using the model $y_{ij} = b_0 + b_1 x_{ij} + \nu_i + e_{ij}$ and

$X_{ij} = x_{ij} + u_{ij}$. By using the generated samples $(y_{ij}^{(b)}, X_{ij}^{(b)})$, $b=1, \dots, B=5000$, we calculated $\hat{b}_0^{(b)}, \hat{b}_1^{(b)}, \hat{\sigma}_x^{2(b)}, \hat{\sigma}_\nu^{2(b)}$ and $\hat{\sigma}_e^{2(b)}$ by the method of moments mentioned in Section 7.2. For each generated sample, we calculated

$$ARB_p = B^{-1} \sum_{b=1}^B (est_p^{(b)} / par) - 1$$

and

$$ARE_p = B^{-1} \sum_{b=1}^B |(est_p^{(b)} - par) / par|,$$

as average relative bias and absolute relative error of the proposed model parameters where par and $est_p^{(b)}$ denote the model parameters and the estimation of model parameters based on the proposed method of Section 7.2. We also calculated the standard error of EB estimators based on (7.26) over the number of areas as $\sqrt{m^{-1} \sum_{i=1}^m mse_{i.p.}}$.

On the other hand, we used the naive approach to show the effect of measurement error in the estimation of model parameters. In this approach, using the generated samples $(y_{ij}^{(b)}, X_{ij}^{(b)})$, $b=1, \dots, B=5000$, we calculated $\hat{b}_0^{(b)}, \hat{b}_1^{(b)}, \hat{\sigma}_\nu^{2(b)}$ and $\hat{\sigma}_e^{2(b)}$ by the generalized least squared method for regression parameters and the fitting-of-constants method for variance components σ_ν^2 and σ_e^2 as mentioned in Chapter 3. For each generated sample, we calculated

$$ARB_n = B^{-1} \sum_{b=1}^B (est_n^{(b)} / par) - 1$$

and

$$ARE_n = B^{-1} \sum_{b=1}^B |(est_n^{(b)} - par) / par|,$$

as average relative bias and absolute relative error of naive model parameters where $est_n^{(b)}$ is the estimation of model parameters based on the naive method. We also

calculated the standard error of EB estimators as described in Chapter 1, as MSE estimation of nested error regression model, over the number of areas as

$$\sqrt{m^{-1} \sum_{i=1}^m mse_{i.n.}}$$

In the third scenario, we used the generated samples without measurement error $(y_{ij}^{(b)}, x_{ij}^{(b)})$, $b=1, \dots, B=5000$, and followed the procedure described in naive method to calculate ARB_t and ARE_t as

$$ARB_t = B^{-1} \sum_{b=1}^B (est_t^{(b)} / par) - 1$$

and

$$ARE_t = B^{-1} \sum_{b=1}^B |(est_t^{(b)} - par) / par|,$$

as average relative bias and absolute relative error of true (without measurement error) model and the standard error of EB estimators over the number of areas as

$$\sqrt{m^{-1} \sum_{i=1}^m mse_{i.t.}}$$

Results on percent average relative bias, percent absolute relative error and the standard error of EB estimators are reported in Tables 7.1 and 7.2 respectively. In fact, we calculated the average relative bias and absolute relative error for different values of m but, to save the space, we only present the percent average relative bias and percent absolute relative error for $m = 768$; but results are similar for other values of m .

However, for the standard error of EB estimators, we report the results for different values of number of areas m . As shown in Table 7.1, the absolute value of ARB_p for b_0, b_1 and σ_e^2 are significantly smaller than the corresponding values in ARB_n . However, ARB_p for σ_v^2 is larger than the corresponding ARB_n . As a result, if the

covariates are subject to measurement errors and we ignore them, we will lose some efficiency in the sense of estimation of model parameters. For example, ARB for σ_e^2 for the proposed method is -0.19% compared to 29.97% for the naive method. We have similar results for ARE_p for the proposed method compared to ARE_n for the naive method. Moreover, we want to check the lose of efficiency in the sense of ARB and ARE if the covariates are subject to measurement errors. To this end, we compare ARB_p and ARE_p with corresponding ARB_t and ARE_t . As expected, ARB and ARE for the true method (without measurement error) are smaller than ARB and ARE for the proposed method. However, the lose in efficiency in the later case is much smaller than previous case where we use the naive method wrongly.

In addition, we report the standard error of EB estimators over m for the proposed (s_p), naive (s_n) and true (s_t) method. Table 7.2 shows that with increasing m , s_p , s_n and s_t are decreasing. In addition, the rate of decrease for s_p is faster than the rate for the other two methods. However, even though s_p is decreasing faster than s_n and s_t with increasing m , the standard error of EB estimators for the proposed method is still larger than for the naive method, but the difference between s_p and s_n is not very significant for large m . As expected, the standard error of EB estimators for the true method is smaller than the proposed and the naive methods.

7.8 Summary

In this Chapter, we studied a structural measurement error models with nested error regression models, when the covariates in the model are random and also subject to measurement errors. We developed empirical Bayes (EB) estimators of area means

and derived a second order approximation to the mean squared error of EB estimators. We then derived the estimators of MSE that are approximately unbiased. In addition, we showed that the EB estimators are asymptotically optimal. A simulation study on the performance of the proposed method is also provided.

7.9 Appendix

In this Section, we calculate $Cov(M_{yy}, m_{Xy})$ and $Cov(M_{yy}, m_{XX})$. We show that $Cov(M_{yy}, m_{Xy}) = Cov(M_{yy}, m_{XX}) = 0$ and consequently $Cov(M_{yy}, \hat{b}_1 m_{Xy}) = Cov(M_{yy}, \hat{b}_1 m_{XX}) = 0$.

To derive $Cov(M_{yy}, m_{Xy})$, we may write

$$\begin{aligned} Cov(M_{yy}, m_{Xy}) &= Cov\left[(m-1)^{-1} \sum_i n_i (\bar{y}_i - \bar{y})^2, (n-m)^{-1} \sum_i \sum_j (X_{ij} - \bar{X}_i)(y_{ij} - \bar{y}_i)\right] \\ &= (m-1)^{-1}(n-m)^{-1} Cov(\bar{y}'_m A \bar{y}_m, x'_n B y_n), \end{aligned}$$

where $\bar{y}_m = (\bar{y}_1, \dots, \bar{y}_m)'$; $x_n = (x_{11}, \dots, x_{mn_m})'$; $y_n = (y_{11}, \dots, y_{mn_m})'$; $A = D_m - n^{-1}d_m d'_m$; $B = I_n - n^{-1}1_n 1'_n - B_{nm}(D_m - n^{-1}d_m d'_m)B_{mn} = I_n - n^{-1}J_n - B_{nm}AB_{mn}$ with $D_m = diag(n_1, \dots, n_m)$; $d_m = (n_1, \dots, n_m)'$;

$$B_{mn} = \begin{pmatrix} 1'_{n_1}/n_1 & 0 & \dots & 0 \\ 0 & 1'_{n_2}/n_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1'_{n_m}/n_m \end{pmatrix}$$

and J_n is a $n \times n$ matrix of one's. Then we may write

$$Cov(M_{yy}, m_{Xy}) = [(m-1)(n-m)]^{-1} \left[tr(AC_{\bar{y}x}BC_{y\bar{y}} + AC_{\bar{y}y}B'C_{x\bar{y}}) \right]$$

$$+2\mu'_{\bar{y}}AC_{\bar{y}x}B\mu_y + 2\mu'_{\bar{y}}AC_{\bar{y}y}B'\mu_X\Big],$$

where $C_{\bar{y}x} = E(\bar{y}_m - \mu_{\bar{y}})(x_n - \mu_X)'$; $C_{y\bar{y}} = E(y_n - \mu_y)(\bar{y}_m - \mu_{\bar{y}})'$; $\mu_{\bar{y}} = (b_0 + b_1\mu_x)1_m$; $\mu_y = (b_0 + b_1\mu_x)1_n$; $\mu_X = \mu_x 1_n$.

Indeed, we used the following Lemma to get above equation.

Lemma 7.1. (Searle, 1971, p.66): Let x_1, x_2, x_3 and x_4 be vectors order n_1, n_2, n_3 and n_4 respectively and be normally distributed with respective means μ_1, μ_2, μ_3 and μ_4 and covariance matrices C_{ij} , of order $n_i \times n_j$, for $i, j = 1, 2, 3$ and 4:

$$C_{ij} = E(x_i - \mu_i)(x_j - \mu_j)' = (C_{ji})'.$$

Then, the covariance between two bilinear forms $x'_1 A_{12} x_2$ and $x'_3 A_{34} x_4$ are given by

$$\begin{aligned} Cov(x'_1 A_{12} x_2, x'_3 A_{34} x_4) &= tr(A_{12} C_{23} A_{34} C_{41} + A_{12} C_{24} A_{43} C_{31}) \\ &+ \mu'_1 A_{12} C_{23} A_{34} \mu_4 + \mu'_1 A_{12} C_{24} A_{43} \mu_3 + \mu'_2 A_{21} C_{13} A_{34} \mu_4 + \mu'_2 A_{21} C_{14} A_{43} \mu_3. \end{aligned}$$

Note that $B = B'$; $A = A'$. After some calculation, we get $BC_{y\bar{y}} = 0$; $C_{\bar{y}x}B = 0$. Therefore, combining above results, we have $Cov(M_{yy}, m_{Xy}) = 0$.

Similarly, we may write

$$\begin{aligned} Cov(M_{yy}, m_{XX}) &= Cov\left[(m-1)^{-1} \sum_i n_i (\bar{y}_i - \bar{y})^2, \right. \\ &\quad \left. (n-m)^{-1} \sum_i \sum_j (X_{ij} - \bar{X}_i)(X_{ij} - \bar{X}_i)\right] \\ &= (m-1)^{-1}(n-m)^{-1} \left[tr(AC_{\bar{y}x}BC_{x\bar{y}} + AC_{\bar{y}x}B'C_{x\bar{y}}) + 4\mu'_{\bar{y}}AC_{\bar{y}x}B\mu_X \right], \end{aligned}$$

by using the Lemma 7.1. After some calculation, we get $C_{\bar{y}x}B = 0$. Then, we have $Cov(M_{yy}, m_{XX}) = 0$.

Method	Quality measure	b_0	b_1	σ_ν^2	σ_e^2	σ_x^2
Proposed	ARB(%)	8.23	0.24	-0.12	-0.19	0.07
	ARE(%)	152.93	4.86	6.12	4.63	5.54
Naive	ARB(%)	-1521.49	-48.66	-0.03	29.97	-
	ARE(%)	1521.49	48.66	5.84	29.99	-
True	ARB(%)	-1.17	-0.05	0.01	0.06	-
	ARE(%)	77.02	2.41	5.50	2.93	-

Table 7.1: Percent average relative bias and percent absolute relative error of model parameters based on the proposed, naive and true methods for $m = 768$

Number of areas	s_p	s_n	s_t
36	10.57	8.84	8.12
60	10.00	8.73	8.02
84	9.83	8.71	7.98
108	9.66	8.69	7.96
132	9.58	8.68	7.94
156	9.52	8.66	7.93
192	9.48	8.66	7.92
384	9.35	8.64	7.90
768	9.30	8.63	7.89

Table 7.2: Standard error of EB estimators for the proposed, naive and true methods for different number of areas

Chapter 8

Functional Measurement Error Models Under Nested Error Linear Regression Model (EB Approach)

In this Chapter, we study small area estimation under the nested error linear regression model when the covariates are measured with error. We assume here that the unknown true covariates are non-stochastic. This is the so-called functional measurement error model. This is in contrast with the structural measurement error (Chapter 7) where the unknown true covariates are treated as random. In this Chapter, we obtain a pseudo-Bayes estimator of a small area mean when the sampling fractions are non-negligible in each area. We then estimate the variance components as well as the regression parameters to get a pseudo-empirical Bayes (PEB) estimator. Moreover, we derive a second order approximation to the mean squared error (MSE) of the pseudo-EB estimator in order to measure the precision of our estimator. In addition, we derive an estimator of MSE which is nearly unbiased. We then show that the pseudo-EB estimators are asymptotically optimal. A simulation study is also conducted to compare the proposed pseudo-EB estimator to the naive

EB estimator which ignores measurement errors in the covariates.

8.1 Introduction

A nested error linear regression model for the sample data, assuming no sample selection bias, may be written as:

$$y_{ij} = b_0 + b_1 x_{ij} + \nu_i + e_{ij}; j = 1, \dots, n_i, i = 1, \dots, m \quad (8.1)$$

where $\nu_i \stackrel{i.i.d.}{\sim} N(0, \sigma_\nu^2)$ and $e_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma_e^2)$. Under measurement error on x, we have $X_{ij} = x_{ij} + u_{ij}$ as observed values, where x_{ij} is fixed and unknown and $u_{ij} \stackrel{i.i.d.}{\sim} N(0, \sigma_u^2)$, (Fuller, 1987). Moreover, we assume that σ_u^2 is known and ν_i, e_{ij} and u_{ij} are mutually independent. The goal is to estimate the i -th area mean $\gamma_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij}$ from the available data $\{(y_{ij}, X_{ij}), j = 1, \dots, n_i; i = 1, \dots, m\}$.

Letting y_i and y_i^* respectively denote sampled and nonsampled vectors of y, we have

$$[(y_i, y_i^*) | \theta] \sim N \left[(b_0 1_{n_i} + b_1 x_i, b_0 1_{N_i - n_i} + b_1 x_i^*), \begin{pmatrix} \sigma_\nu^2 J_{n_i} + \sigma_e^2 I_{n_i} & \sigma_\nu^2 1_{n_i} 1'_{N_i - n_i} \\ \sigma_\nu^2 1_{N_i - n_i} 1'_{n_i} & \sigma_\nu^2 J_{N_i - n_i} + \sigma_e^2 I_{N_i - n_i} \end{pmatrix} \right],$$

with $\theta = (b_0, b_1, \sigma_e^2, \sigma_\nu^2)'$. The Bayes (or best) predictor of y_i^* given y_i and θ can then be written as:

$$\begin{aligned} E(y_i^* | y_i, \theta) &= b_0 1_{N_i - n_i} + b_1 x_i^* + \sigma_\nu^2 1_{N_i - n_i} 1'_{n_i} \left[\sigma_e^{-2} (I_{n_i} - \frac{1}{n_i} J_{n_i}) + \frac{1}{n_i} J_{n_i} \right. \\ &\quad \left. \cdot (\sigma_e^2 + n_i \sigma_\nu^2)^{-1} \right] (y_i - b_0 1_{n_i} - b_1 x_i) \\ &= b_0 1_{N_i - n_i} + b_1 x_i^* + \sigma_\nu^2 1_{N_i - n_i} 1'_{n_i} (\sigma_e^2 + n_i \sigma_\nu^2)^{-1} (y_i - b_0 1_{n_i} - b_1 x_i) \end{aligned}$$

$$\begin{aligned}
&= b_0 1_{N_i - n_i} + b_1 x_i^* + n_i \sigma_\nu^2 (\sigma_e^2 + n_i \sigma_\nu^2)^{-1} (\bar{y}_i - b_0 - b_1 \bar{x}_i) 1_{N_i - n_i} \\
&= \left[(1 - B_i) \bar{y}_i + B_i b_0 \right] 1_{N_i - n_i} + b_1 \left(x_i^* - (1 - B_i) \bar{x}_i 1_{N_i - n_i} \right),
\end{aligned}$$

where $B_i = \sigma_e^2 / (\sigma_e^2 + n_i \sigma_\nu^2)$ and (x_i, x_i^*) as sampled and non-sampled vectors of $\mathbf{x}$. It leads to the Bayes (or best) predictor of γ_i as:

$$\hat{\gamma}_i^B = E(\gamma_i | y_i) = (1 - f_i B_i) \bar{y}_i + f_i B_i \left[b_0 + \frac{b_1}{B_i} (\bar{x}_i^* - (1 - B_i) \bar{x}_i) \right],$$

where $f_i = 1 - n_i / N_i$ and $\bar{x}_i^* = \frac{1}{f_i} (\bar{x}_{ip} - \frac{n_i}{N_i} \bar{x}_i)$ with $\bar{x}_{ip}$ and $\bar{x}_i$ as the population mean and sample mean of x for the i -th area. We take $\bar{x}_{ip} = \bar{X}_{ip}$, the population mean of X for the i -th area, noting that $\bar{U}_{ip} = N_i^{-1} \sum_{j=1}^{N_i} u_{ij} \approx 0$ for large N_i .

However, the $\bar{x}_i$ are not observable. In fact, we observe instead $X_{ij} | x_{ij} \stackrel{i.i.d.}{\sim} N(x_{ij}, \sigma_u^2)$. Then a pseudo-Bayes estimator of γ_i is obtained by replacing $\bar{x}_i$ in $\hat{\gamma}_i^B$ by $\bar{X}_i$, i.e. we estimate γ_i by

$$\hat{\gamma}_i^{PB} = (1 - f_i B_i) \bar{y}_i + f_i B_i b_0 + b_1 \left[\bar{X}_{ip} - (1 - f_i B_i) \bar{X}_i \right],$$

where $\bar{X}_i = n_i^{-1} \sum_{j=1}^{n_i} X_{ij}$ is the sample mean of X for the i -th area.

In Section 8.2, we estimate variance components as well as the regression parameters.

8.2 Empirical Bayes

In this Section, we estimate θ in $\hat{\gamma}_i^{PB}$ to get a pseudo-EB estimator of γ_i . As in Chapter 7, $m_{XX} = (n - m)^{-1} \sum_{i=1}^m \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$, $\bar{X}$ = overall mean of X and similar expressions for y . We estimate the model parameters based on the method

of moment. We may write that

$$m_{yy} = (n - m)^{-1} \sum_i \sum_j (y_{ij} - \bar{y}_i)^2 = (n - m)^{-1} \sum_i \sum_j \left[b_1(x_{ij} - \bar{x}_i) + e_{ij} - \bar{e}_i \right]^2.$$

Then

$$E(m_{yy}) = b_1^2 m_{xx} + \sigma_e^2, \quad (8.2)$$

where $m_{xx} = (n - m)^{-1} \sum_i \sum_j (x_{ij} - \bar{x}_i)^2$. Also, we have that

$$E(m_{XX}) = (n - m)^{-1} \sum_i \sum_j E(x_{ij} - \bar{x}_i + u_{ij} - \bar{u}_i)^2 = m_{xx} + \sigma_u^2. \quad (8.3)$$

In addition, we define $M_{yy} = (m - 1)^{-1} \sum_i n_i (\bar{y}_i - \bar{y})^2$ where $\bar{y} = n^{-1} \sum_i n_i \bar{y}_i$.

Then

$$\begin{aligned} E(M_{yy}) &= (m - 1)^{-1} \left[b_1^2 \sum_i n_i (\bar{x}_i - \bar{x})^2 + \sigma_v^2 \left(n - \frac{\sum_i n_i^2}{n} \right) + (m - 1) \sigma_e^2 \right] \\ &= b_1^2 M_{xx} + \sigma_v^2 \frac{g_m}{m - 1} + \sigma_e^2, \end{aligned} \quad (8.4)$$

where $g_m = n - \sum_i n_i^2/n$ and $M_{xx} = (m - 1)^{-1} \sum_i n_i (\bar{x}_i - \bar{x})^2$. Moreover,

$$E(m_{Xy}) = (n - m)^{-1} \sum_i \sum_j E(X_{ij} - \bar{X}_i)(y_{ij} - \bar{y}_i) = b_1 m_{xx}. \quad (8.5)$$

As a result, we obtain $\hat{m}_{xx} = m_{XX} - \sigma_u^2$ as the estimator of m_{xx} , by (8.3). Also, by (8.5), we obtain

$$\hat{b}_1 = m_{Xy} (m_{XX} - \sigma_u^2)^{-1}, \quad (8.6)$$

$$\hat{b}_0 = \bar{y} - \hat{b}_1 \bar{X}, \quad (8.7)$$

as estimators of b_1 and b_0 . In addition, by (8.2) and (8.5), we have

$$\hat{\sigma}_e^2 = m_{yy} - \hat{b}_1 m_{Xy}, \quad (8.8)$$

as an estimator of σ_e^2 . Therefore, by (8.4) and (8.8), we have

$$\hat{\sigma}_\nu^2 = \frac{m-1}{g_m} [M_{yy} - \hat{\sigma}_e^2 - \hat{b}_1^2(M_{XX} - \sigma_u^2)], \quad (8.9)$$

as an estimator of σ_ν^2 .

To have proper estimators in the parameter space, $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$ must be nonnegative. The estimators of σ_ν^2 and σ_e^2 will be positive if and only if

$$m_{yy}(m_{XX} - \sigma_u^2) - m_{Xy}^2 > 0 \quad (8.10)$$

for σ_e^2 , and

$$(M_{yy} - m_{yy})(m_{XX} - \sigma_u^2)^2 + m_{Xy}^2(m_{XX} - M_{XX}) > 0 \quad (8.11)$$

for σ_ν^2 . Then, estimates for sample in which (8.10) and (8.11) are violated are taken as $\hat{\sigma}_e^2 = 0$, $\hat{b}_1 = m_{Xy}^{-1}m_{yy}$, $\hat{b}_0 = \bar{y} - \hat{b}_1\bar{X}$ and $\hat{\sigma}_\nu^2 = 0$.

The pseudo-EB of γ_i can be then written by

$$\hat{\gamma}_i^{PEB} = (1 - f_i\hat{B}_i)\bar{y}_i + f_i\hat{B}_i\hat{b}_0 + \hat{b}_1[\bar{X}_{ip} - (1 - f_i\hat{B}_i)\bar{X}_i], \quad (8.12)$$

where $\hat{B}_i = B_i|_{\delta_1=\hat{\delta}_1}$, with $\delta_1 = (\sigma_\nu^2, \sigma_e^2)'$.

8.3 MSE Approximation

In this Section, we derive a second order approximation to the mean squared error of the pseudo-EB estimator $\hat{\gamma}_i^{PEB}$ given by (8.12). We may write $MSE(\hat{\gamma}_i^{PEB})$ as

follows:

$$\begin{aligned}
MSE(\hat{\gamma}_i^{PEB}) &= E(\hat{\gamma}_i^{PEB} - \gamma_i)^2 \\
&= E(\hat{\gamma}_i^{PB} - \gamma_i)^2 + E(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})^2 \\
&\quad + 2E(\hat{\gamma}_i^{PB} - \gamma_i)(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})
\end{aligned} \tag{8.13}$$

with

$$E(\hat{\gamma}_i^{PB} - \gamma_i)^2 = E(\hat{\gamma}_i^B - \gamma_i)^2 + E(\hat{\gamma}_i^{PB} - \hat{\gamma}_i^B)^2 + 2E(\hat{\gamma}_i^B - \gamma_i)(\hat{\gamma}_i^{PB} - \hat{\gamma}_i^B),$$

where $E(\hat{\gamma}_i^B - \gamma_i)(\hat{\gamma}_i^{PB} - \hat{\gamma}_i^B) = 0$ since $E(\gamma_i - \hat{\gamma}_i^B | y_i) = 0$. Also,

$$E(\hat{\gamma}_i^B - \gamma_i)^2 = Var(\gamma_i | y_i) = N_i^{-2} Var(1'_{N_i-n_i} y_i^* | y_i),$$

where

$$\begin{aligned}
Var(y_i^* | y_i) &= (\sigma_\nu^2 J_{N_i-n_i} + \sigma_e^2 I_{N_i-n_i}) - \sigma_\nu^2 1_{N_i-n_i} 1'_{n_i} \left[\sigma_\nu^2 J_{n_i} + \sigma_e^2 I_{n_i} \right]^{-1} \sigma_\nu^2 1_{n_i} 1'_{N_i-n_i} \\
&= \sigma_\nu^2 J_{N_i-n_i} + \sigma_e^2 I_{N_i-n_i} - \sigma_\nu^4 1_{N_i-n_i} 1'_{n_i} \left[\sigma_e^{-2} (I_{n_i} - \frac{1}{n_i} J_{n_i}) + n_i^{-1} J_{n_i} (\sigma_e^2 + n_i \sigma_\nu^2)^{-1} \right] 1_{n_i} 1'_{N_i-n_i} \\
&= \sigma_\nu^2 J_{N_i-n_i} + \sigma_e^2 I_{N_i-n_i} - \sigma_\nu^4 1_{N_i-n_i} 1'_{n_i} (\sigma_e^2 + n_i \sigma_\nu^2)^{-1} 1_{n_i} 1'_{N_i-n_i} \\
&= \sigma_e^2 I_{N_i-n_i} + \frac{\sigma_e^2 \sigma_\nu^2}{\sigma_e^2 + n_i \sigma_\nu^2} J_{N_i-n_i}.
\end{aligned}$$

Then we may write

$$\begin{aligned}
E(\hat{\gamma}_i^B - \gamma_i)^2 &= N_i^{-2} \left[(N_i - n_i) \sigma_e^2 + \sigma_\nu^2 (N_i - n_i)^2 \frac{\sigma_e^2}{\sigma_e^2 + n_i \sigma_\nu^2} \right] \\
&= \sigma_e^2 f_i \left[N_i^{-1} + \frac{f_i \sigma_\nu^2}{\sigma_e^2 + n_i \sigma_\nu^2} \right].
\end{aligned} \tag{8.14}$$

Also,

$$E(\hat{\gamma}_i^{PB} - \hat{\gamma}_i^B)^2 = b_1^2 (1 - f_i B_i)^2 E(\bar{X}_i - \bar{x}_i)^2 = b_1^2 (1 - f_i B_i)^2 \sigma_u^2 / n_i. \tag{8.15}$$

Therefore, by (8.14) and (8.15), one may write

$$E(\hat{\gamma}_i^{PB} - \gamma_i)^2 = f_i \sigma_e^2 \left[N_i^{-1} + \frac{f_i \sigma_\nu^2}{\sigma_e^2 + n_i \sigma_\nu^2} \right] + b_1^2 (1 - f_i B_i)^2 \sigma_u^2 / n_i \equiv g_{1i}(\delta), \quad (8.16)$$

where $\delta = (b_1, \sigma_\nu^2, \sigma_e^2)'$. To find $E(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})^2$, we need the following terms:

$$\begin{aligned} \partial \hat{\gamma}_i^{PB} / \partial b_0 &= f_i B_i; \quad \partial \hat{\gamma}_i^{PB} / \partial b_1 = \bar{X}_{ip} - (1 - f_i B_i) \bar{X}_i; \\ \partial \hat{\gamma}_i^{PB} / \partial \sigma_\nu^2 &= -\frac{n_i f_i \sigma_e^2 (b_0 + b_1 \bar{X}_i - \bar{y}_i)}{(\sigma_e^2 + n_i \sigma_\nu^2)^2}; \quad \partial \hat{\gamma}_i^{PB} / \partial \sigma_e^2 = \frac{n_i f_i \sigma_\nu^2 (b_0 + b_1 \bar{X}_i - \bar{y}_i)}{(\sigma_e^2 + n_i \sigma_\nu^2)^2}. \end{aligned}$$

We may write

$$\begin{aligned} E(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})^2 &= (f_i B_i)^2 \text{Var}(\hat{b}_0) + \left[\bar{X}_{ip} - (1 - f_i B_i) \bar{x}_i \right]^2 + (1 - f_i B_i)^2 \sigma_u^2 / n_i \text{Var}(\hat{b}_1) \\ &+ 2 f_i B_i \left[\bar{X}_{ip} - (1 - f_i B_i) \bar{x}_i \right] \text{Cov}(\hat{b}_0, \hat{b}_1) + \frac{n_i f_i^2 (b_1^2 \sigma_u^2 + \sigma_e^2 + n_i \sigma_\nu^2)}{(\sigma_e^2 + n_i \sigma_\nu^2)^4} \left[\sigma_\nu^4 \text{Var}(\hat{\sigma}_e^2) + \sigma_e^4 \right. \\ &\cdot \text{Var}(\hat{\sigma}_\nu^2) - 2 \sigma_\nu^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \left. \right] - 2 \frac{(1 - f_i B_i) b_1 f_i \sigma_u^2}{(\sigma_e^2 + n_i \sigma_\nu^2)^2} \left[\sigma_\nu^2 \text{Cov}(\hat{b}_1, \hat{\sigma}_e^2) - \sigma_e^2 \text{Cov}(\hat{b}_1, \hat{\sigma}_\nu^2) \right] \\ &+ o(m^{-1}) \equiv g_i(\delta) + o(m^{-1}), \end{aligned} \quad (8.17)$$

where $\text{Var}(\hat{b}_0)$, $\text{Var}(\hat{b}_1)$, $\text{Var}(\hat{\sigma}_e^2)$, $\text{Var}(\hat{\sigma}_\nu^2)$, $\text{Cov}(\hat{b}_0, \hat{b}_1)$, $\text{Cov}(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2)$, $\text{Cov}(\hat{b}_1, \hat{\sigma}_\nu^2)$ and $\text{Cov}(\hat{b}_1, \hat{\sigma}_e^2)$ are given in Section 8.4. Technical details on the order of the reminder terms are omitted.

In addition, we need to evaluate the cross-term defined in (8.13). We have

$$\begin{aligned} C_i(\delta) &\equiv E(\hat{\gamma}_i^{PB} - \gamma_i)(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB}) = E\left[(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})E(\hat{\gamma}_i^{PB} - \gamma_i | y_i)\right] \\ &= E\left[(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})(\hat{\gamma}_i^{PB} - E(\gamma_i | y_i))\right] = E\left[(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})(\hat{\gamma}_i^{PB} - \hat{\gamma}_i^B)\right] \\ &= -b_1(1 - f_i B_i)E\left\{(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})\bar{u}_i\right\} = -b_1(1 - f_i B_i)E\left[E\left[(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})\bar{u}_i | \bar{u}_i\right]\right] \\ &= -b_1(1 - f_i B_i)E\left[\bar{u}_i \left[(\partial \hat{\gamma}_i^{PB} / \partial b_0) E(\hat{b}_0 - b_0 | \bar{u}_i) + (\partial \hat{\gamma}_i^{PB} / \partial b_1) E(\hat{b}_1 - b_1 | \bar{u}_i) \right] \right] \end{aligned}$$

$$\begin{aligned}
& +E((\partial\hat{\gamma}_i^{PB}/\partial\sigma_e^2)(\hat{\sigma}_e^2 - \sigma_e^2)|\bar{u}_i) + E((\partial\hat{\gamma}_i^{PB}/\partial\sigma_\nu^2)(\hat{\sigma}_\nu^2 - \sigma_\nu^2)|\bar{u}_i) + \frac{1}{2}E((\partial^2\hat{\gamma}_i^{PB}/\partial\sigma_\nu^4)(\hat{\sigma}_\nu^2 \\
& - \sigma_\nu^2)^2|\bar{u}_i) + \frac{1}{2}E((\partial^2\hat{\gamma}_i^{PB}/\partial\sigma_e^4)(\hat{\sigma}_e^2 - \sigma_e^2)^2|\bar{u}_i) + E((\partial^2\hat{\gamma}_i^{PB}/\partial b_0\partial\sigma_\nu^2)(\hat{b}_0 - b_0)(\hat{\sigma}_\nu^2 - \sigma_\nu^2)|\bar{u}_i) \\
& + E((\partial^2\hat{\gamma}_i^{PB}/\partial b_0\partial\sigma_e^2)(\hat{b}_0 - b_0)(\hat{\sigma}_e^2 - \sigma_e^2)|\bar{u}_i) + E((\partial^2\hat{\gamma}_i^{PB}/\partial b_1\partial\sigma_\nu^2)(\hat{b}_1 - b_1)(\hat{\sigma}_\nu^2 - \sigma_\nu^2)|\bar{u}_i) \\
& + E((\partial^2\hat{\gamma}_i^{PB}/\partial b_1\partial\sigma_e^2)(\hat{b}_1 - b_1)(\hat{\sigma}_e^2 - \sigma_e^2)|\bar{u}_i) + E((\partial^2\hat{\gamma}_i^{PB}/\partial\sigma_\nu^2\partial\sigma_e^2)(\hat{\sigma}_\nu^2 - \sigma_\nu^2)(\hat{\sigma}_e^2 - \sigma_e^2)|\bar{u}_i) \Big] \\
& + o(m^{-1}), \tag{8.18}
\end{aligned}$$

where we may evaluate (8.18) by the numerical methods.

Theorem 8.1. *The mean squared error of the pseudo-EB estimator $\hat{\gamma}_i^{PEB}$ is approximated by*

$$MSE(\hat{\gamma}_i^{PEB}) = E(\hat{\gamma}_i^{PEB} - \gamma_i)^2 \simeq g_{1i}(\delta) + g_i(\delta) + 2C_i(\delta),$$

where $g_{1i}(\delta)$, $g_i(\delta)$ and $C_i(\delta)$ are given by (8.16), (8.17) and (8.18).

We derive the variance-covariance matrix of model parameter estimators in Section 8.4.

8.4 Variances and Covariances of Model Parameter Estimators

In this Section, we derive the variances and covariances of estimators of regression parameters and variance components. To find the variance-covariance matrix of estimators of regression parameters, we have the following theorem.

Theorem 8.2. Let model (8.1) hold with σ_u^2 known, $\sigma_v^2 > 0$ and $\sigma_e^2 > 0$. Then $m^{1/2}[(\hat{b}_0 - b_0), (\hat{b}_1 - b_1)]$ where $(\hat{b}_0, \hat{b}_1)$ are defined by (8.7) and (8.6), is a normal vector variable with zero mean and covariance matrix with elements

$$\begin{aligned} Var[m^{1/2}(\hat{b}_0 - b_0)] &= \frac{m}{n} \left[\left(\sum_i n_i^2 / n \right) \sigma_v^2 + \sigma_e^2 + b_1^2 \sigma_u^2 \right] + \frac{m}{n-m} \bar{x}^2 m_{xx}^{-2} \\ &\cdot \left[m_{xx}(\sigma_v^2 + \sigma_e^2 + b_1^2 \sigma_u^2) + \sigma_u^2(\sigma_v^2 + \sigma_e^2) + 2b_1^2(n-m)^{-1}(n-2m + \sum_i n_i^{-1})\sigma_u^4 \right. \\ &\quad \left. - (n-m)^{-1} \sigma_v^2 \sum_i \sum_j x_{ij}(x_{ij} - \bar{x}_i) + \sigma_v^2 \sigma_u^2 \right] + O(m^{-1}), \\ Var[m^{1/2}(\hat{b}_1 - b_1)] &= \frac{m}{n-m} m_{xx}^{-2} \left[m_{xx}(\sigma_v^2 + \sigma_e^2 + b_1^2 \sigma_u^2) + \sigma_u^2(\sigma_v^2 + \sigma_e^2) \right. \\ &\quad \left. + 2b_1^2(n-m)^{-1}(n-2m + \sum_i n_i^{-1})\sigma_u^4 - (n-m)^{-1} \sigma_v^2 \sum_i \sum_j x_{ij}(x_{ij} - \bar{x}_i) + \sigma_v^2 \sigma_u^2 \right] + O(m^{-1}), \end{aligned}$$

and

$$Cov[m^{1/2}(\hat{b}_0 - b_0), m^{1/2}(\hat{b}_1 - b_1)] = -\bar{x} Var[m^{1/2}(\hat{b}_1 - b_1)] + O(m^{-1}).$$

Proof: Since the sample moments converge in probability to the population moments, we can expand $\hat{b}_1$ in a Taylor series about the population values to obtain:

$$\hat{b}_1 = b_1 + m_{xx}^{-1}(m_{Xy} - b_1 m_{xx}) - b_1 m_{xx}^{-1}(m_{XX} - \sigma_u^2 - m_{xx}) + O_p(m^{-1}), \quad (8.19)$$

then

$$\begin{aligned} m^{1/2}(\hat{b}_1 - b_1) &= m^{1/2} m_{xx}^{-1} \left[m_{Xy} - b_1(m_{XX} - \sigma_u^2) \right] + O_p(m^{-1/2}) \\ &= m^{1/2} m_{xx}^{-1} [m_{Xw} - \sigma_{Xw}] + O_p(m^{-1/2}), \end{aligned}$$

(see Structural model-Chapter 7). Therefore,

$$Var[m^{1/2}(\hat{b}_1 - b_1)] = \frac{m}{n-m} m_{xx}^{-2} \left[m_{xx}(\sigma_v^2 + \sigma_e^2 + b_1^2 \sigma_u^2) + \sigma_u^2(\sigma_v^2 + \sigma_e^2) + 2b_1^2(n-m)^{-1} \right]$$

$$.(n - 2m + \sum_i n_i^{-1})\sigma_u^4 - (n - m)^{-1}\sigma_\nu^2 \sum_i \sum_j x_{ij}(x_{ij} - \bar{x}_i) + \sigma_\nu^2 \sigma_u^2] + O(m^{-1}).$$

For $\hat{b}_0$, we may write:

$$\begin{aligned}\hat{b}_0 &= \bar{y} - \hat{b}_1 \bar{X} = b_0 + b_1 \bar{x} + \bar{\nu} + \bar{e} - \hat{b}_1(\bar{x} + \bar{u}) \\ &= b_0 - (\hat{b}_1 - b_1)\bar{x} + \bar{w} + O_p(m^{-1}),\end{aligned}$$

where $\bar{w} = \bar{\nu} + \bar{e} - b_1 \bar{u}$. Then

$$m^{1/2}(\hat{b}_0 - b_0) = m^{1/2}(\bar{w} - (\hat{b}_1 - b_1)\bar{x}) + O_p(m^{-1/2}). \quad (8.20)$$

Therefore, we may write

$$\begin{aligned}Var[m^{1/2}(\hat{b}_0 - b_0)] &= m[(\sum_i n_i^2/n^2)\sigma_\nu^2 + \frac{\sigma_e^2}{n} + b_1^2 \frac{\sigma_u^2}{n}] + \bar{x}^2 Var[m^{1/2}(\hat{b}_1 - b_1)] \\ &\quad - 2\bar{x} Cov[m^{1/2}(\hat{b}_1 - b_1), m^{1/2}\bar{w}] + O(m^{-1}) \\ &= (m/n)[(\sum_i n_i^2/n)\sigma_\nu^2 + b_1^2 \sigma_u^2 + \sigma_e^2] + \bar{x}^2 Var[m^{1/2}(\hat{b}_1 - b_1)] + O(m^{-1}),\end{aligned}$$

since $Cov(m_{Xw}, \bar{w}) = 0$. Hence,

$$\begin{aligned}Var[m^{1/2}(\hat{b}_0 - b_0)] &= (m/n)[(\sum_i n_i^2/n)\sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_u^2] + \frac{m}{n-m} \bar{x}^2 m_{xx}^{-2} \\ &\quad \cdot [m_{xx}(\sigma_\nu^2 + \sigma_e^2 + b_1^2 \sigma_u^2) + \sigma_u^2(\sigma_\nu^2 + \sigma_e^2) + 2b_1^2(n-m)^{-1}(n-2m + \sum_i n_i^{-1})\sigma_u^4 - (n-m)^{-1}\sigma_\nu^2 \\ &\quad \cdot \sum_i \sum_j x_{ij}(x_{ij} - \bar{x}_i) + \sigma_\nu^2 \sigma_u^2] + O(m^{-1}).\end{aligned}$$

Finally,

$$Cov[m^{1/2}(\hat{b}_0 - b_0), m^{1/2}(\hat{b}_1 - b_1)] = -\bar{x} Var[m^{1/2}(\hat{b}_1 - b_1)] + O(m^{-1}),$$

by (8.19) and (8.20). Hence, the results of Theorem 8.2 are obtained.

In addition,

$$Var(\hat{\sigma}_e^2) = Var(m_{yy}) + Var(\hat{b}_1 m_{Xy}) - 2Cov(m_{yy}, \hat{b}_1 m_{Xy}), \quad (8.21)$$

where

$$\begin{aligned} Var(m_{yy}) = & 2(n-m)^{-2} \sigma_e^2 \left[(n-2m + \sum_i n_i^{-1}) \sigma_e^2 + 2b_1^2 \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 (1 - 1/n_i) \right. \\ & \left. + \sigma_e^2 \sum_i (1 - 1/n_i)(3 - 4/n_i) + 2b_1^2 \sum_i n_i^{-1} \sum_j x_{ij}(x_{ij} - \bar{x}_i) \right]. \end{aligned}$$

We may write the following term based on Taylor expansion:

$$\hat{b}_1 m_{Xy} = b_1^2 m_{xx} + 2b_1(m_{Xy} - b_1 m_{xx}) - b_1^2(m_{XX} - \sigma_u^2 - m_{xx}) + O_p(m^{-1}),$$

then, we may write

$$Var[\hat{b}_1 m_{Xy}] = b_1^2 \left[4Var(m_{Xy}) + b_1^2 Var(m_{XX}) - 4b_1 Cov(m_{XX}, m_{Xy}) \right] + o(m^{-1}),$$

where

$$\begin{aligned} Var(m_{Xy}) = & (n-m)^{-2} \left[\sum_i \sum_j (x_{ij} - \bar{x}_i)^2 (1 - 1/n_i) (\sigma_e^2 + b_1^2 \sigma_u^2) + (n-2m + \sum_i n_i^{-1}) \right. \\ & \left. \cdot \sigma_u^2 \sigma_e^2 + (b_1^2 \sigma_u^2 + \sigma_e^2) \sum_i n_i^{-1} \sum_{j \neq k} (x_{ij} - \bar{x}_i)(x_{ik} - \bar{x}_i) + \sigma_u^2 \sigma_e^2 (m - \sum_i n_i^{-1}) \right], \\ Var(m_{XX}) = & 2(n-m)^{-2} \sigma_u^2 \left[(n-2m + \sum_i n_i^{-1}) \sigma_u^2 + 2 \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 (1 - 1/n_i) \right. \\ & \left. + \sigma_u^2 \sum_i (1 - 1/n_i)(3 - 4/n_i) + 2 \sum_i n_i^{-1} \sum_j x_{ij}(x_{ij} - \bar{x}_i) \right], \end{aligned}$$

and

$$Cov(m_{XX}, m_{Xy}) = 2(n-m)^{-2} b_1 \sigma_u^2 \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 (1 - 1/n_i).$$

In addition, to evaluate (8.21), we may write

$$Cov(m_{yy}, \hat{b}_1 m_{xy}) = 2b_1 Cov(m_{yy}, m_{xy}) + o(m^{-1}),$$

where

$$Cov(m_{yy}, m_{xy}) = 2b_1(n-m)^{-2}\sigma_e^2 \sum_i \sum_j (x_{ij} - \bar{x}_i)^2(1 - 1/n_i),$$

noting that $Cov(m_{yy}, m_{xx}) = 0$. Therefore, combining above results, $Var(\hat{\sigma}_e^2)$ is defined.

In addition, we may write

$$\begin{aligned} Var(\hat{\sigma}_e^2) = & \left(\frac{m-1}{g_m}\right)^2 \left[Var(M_{yy}) + Var(\hat{\sigma}_e^2) + Var(\hat{b}_1^2 M_{xx}) + \sigma_u^4 Var(\hat{b}_1^2) \right. \\ & \left. - 2Cov(M_{yy}, \hat{\sigma}_e^2) - 2Cov(M_{yy}, \hat{b}_1^2(M_{xx} - \sigma_u^2)) + 2Cov(\hat{\sigma}_e^2, \hat{b}_1^2(M_{xx} - \sigma_u^2)) \right], \quad (8.22) \end{aligned}$$

where

$$\begin{aligned} Var(M_{yy}) = & 2(m-1)^{-2} \left\{ \sigma_\nu^4 \sum_i n_i^2 + (m-2)\sigma_e^4 + 2b_1^2 \sigma_\nu^2 \sum_i n_i^2 (\bar{x}_i - \bar{x})^2 \right. \\ & \left. + 2b_1^2 \sigma_e^2 \sum_i n_i (\bar{x}_i - \bar{x})^2 + 2n\sigma_\nu^2 \sigma_e^2 \right\} + o(m^{-1}), \end{aligned}$$

and $Cov(M_{yy}, \hat{\sigma}_e^2) = Cov(M_{yy}, m_{yy} - \hat{b}_1 m_{xy}) = -Cov(M_{yy}, \hat{b}_1 m_{xy}) = 0$ (similar to Structural model-Chapter 7- Appendix). Moreover, to evaluate (8.22), we may write

$$Var(\hat{b}_1^2 M_{xx}) = b_1^4(1 + 2m_{xx}^{-1}\sigma_u^2)^2 Var(M_{xx}) + o(m^{-1}),$$

based on Taylor expansion

$$\hat{b}_1^2 = b_1^2 + 2b_1 m_{xx}^{-1}(m_{xy} - b_1 m_{xx}) - 2b_1^2 m_{xx}^{-1}(m_{xx} - \sigma_u^2 - m_{xx}) + O_p(m^{-1}),$$

where M_{XX} is independent of m_{XX} and m_{Xy} . Note that

$$Var(M_{XX}) = 2(m-1)^{-2}\sigma_u^2 \left[\sigma_u^2(m-2) + 2 \sum_i n_i (\bar{x}_i - \bar{x})^2 \right].$$

Also, to evaluate (8.22), we may write

$$Var(\hat{b}_1^2) = 4b_1^2 \left[m_{xx}^{-2} Var(m_{Xy}) + (b_1 m_{xx}^{-1})^2 Var(m_{XX}) + 2b_1 m_{xx}^{-2} Cov(m_{XX}, m_{Xy}) \right] + o(m^{-1}),$$

and $Cov(M_{yy}, \hat{b}_1^2 M_{XX}) = 0$, noting that $Cov(M_{yy}, M_{XX}) = 0$. Also, note that $Cov(M_{yy}, \hat{b}_1^2) = 0$ and $Cov(\hat{\sigma}_e^2, \hat{b}_1^2 M_{XX}) = 0$. In addition, to evaluate (8.22), we may write

$$Cov(\hat{\sigma}_e^2, \hat{b}_1^2) = Cov(m_{yy}, \hat{b}_1^2) - Cov(\hat{b}_1 m_{Xy}, \hat{b}_1^2), \quad (8.23)$$

where

$$Cov(m_{yy}, \hat{b}_1^2) = 2b_1 m_{xx}^{-1} Cov(m_{yy}, m_{Xy}) + o(m^{-1}),$$

and

$$Cov(\hat{b}_1 m_{Xy}, \hat{b}_1^2) = -b_1^2 Cov(m_{XX}, \hat{b}_1^2) + 2b_1 Cov(m_{Xy}, \hat{b}_1^2) + o(m^{-1}), \quad (8.24)$$

with

$$Cov(m_{XX}, \hat{b}_1^2) = -2b_1^2 m_{xx}^{-1} Var(m_{XX}) + 2b_1 m_{xx}^{-1} Cov(m_{XX}, m_{Xy}) + o(m^{-1}).$$

Also, to evaluate (8.24), we may write

$$Cov(m_{Xy}, \hat{b}_1^2) = -2b_1^2 m_{xx}^{-1} Cov(m_{Xy}, m_{XX}) + 2b_1 m_{xx}^{-1} Var(m_{Xy}) + o(m^{-1}).$$

Therefore, combining above results, by (8.24), $Cov(\hat{b}_1 m_{Xy}, \hat{b}_1^2)$ is defined. Hence, we have $Cov(\hat{\sigma}_e^2, \hat{b}_1^2)$ from (8.23). Thus, $Var(\hat{\sigma}_e^2)$ is now defined by (8.22).

In addition, we may write by (8.9),

$$Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) = \frac{m-1}{g_m} \left[\sigma_u^2 Cov(\hat{\sigma}_e^2, \hat{b}_1^2) - Var(\hat{\sigma}_e^2) \right] + o(m^{-1}),$$

since $Cov(M_{yy}, \hat{\sigma}_e^2) = 0$ and $Cov(\hat{\sigma}_e^2, \hat{b}_1^2 M_{XX}) = 0$.

Moreover, we may write by (8.8),

$$Cov(\hat{b}_1, \hat{\sigma}_e^2) = Cov(\hat{b}_1, m_{yy}) - Cov(\hat{b}_1, \hat{b}_1 m_{Xy}), \quad (8.25)$$

where

$$Cov(\hat{b}_1, m_{yy}) = m_{xx}^{-1} Cov(m_{Xy}, m_{yy}) + o(m^{-1}),$$

and

$$Cov(\hat{b}_1, \hat{b}_1 m_{Xy}) = -b_1^2 Cov(\hat{b}_1, m_{XX}) + 2b_1 Cov(\hat{b}_1, m_{Xy}) + o(m^{-1}),$$

where

$$Cov(\hat{b}_1, m_{XX}) = m_{xx}^{-1} Cov(m_{Xy}, m_{XX}) - b_1 m_{xx}^{-1} Var(m_{XX}) + o(m^{-1}),$$

and

$$Cov(\hat{b}_1, m_{Xy}) = m_{xx}^{-1} Var(m_{Xy}) - b_1 m_{xx}^{-1} Cov(m_{Xy}, m_{XX}) + o(m^{-1}).$$

Therefore, combining above results, $Cov(\hat{b}_1, \hat{\sigma}_e^2)$ is now defined by (8.25).

Finally, we derive $Cov(\hat{b}_1, \hat{\sigma}_\nu^2)$ by (8.25) as

$$Cov(\hat{b}_1, \hat{\sigma}_\nu^2) = \frac{m-1}{g_m} \left[\sigma_u^2 Cov(\hat{b}_1, \hat{b}_1^2) - Cov(\hat{b}_1, \hat{\sigma}_e^2) \right],$$

where

$$Cov(\hat{b}_1, \hat{b}_1^2) = -2b_1^2 m_{xx}^{-1} Cov(\hat{b}_1, m_{XX}) + 2b_1 m_{xx}^{-1} Cov(\hat{b}_1, m_{Xy}) + o(m^{-1}).$$

Note that $Cov(\hat{b}_1, M_{yy}) = 0$ and $Cov(\hat{b}_1, \hat{b}_1^2 M_{XX}) = 0$, since M_{yy} and M_{XX} are independent of m_{XX} and m_{Xy} .

8.5 MSE Estimation

We derive a second order unbiased estimator of mean squared error of $\hat{\gamma}_i^{PEB}$ in the sense that $E[mse(\hat{\gamma}_i^{PEB})] - MSE(\hat{\gamma}_i^{PEB}) = o(m^{-1})$. To evaluate the bias of $g_{1i}(\hat{\delta})$ as an estimator of $g_{1i}(\delta)$, we use a Taylor expansion of $g_{1i}(\hat{\delta})$ around δ :

$$E(g_{1i}(\hat{\delta})) = g_{1i}(\delta) + \frac{\partial g_{1i}}{\partial \delta'} E(\hat{\delta} - \delta) + \frac{1}{2} tr \left\{ \frac{\partial^2 g_{1i}}{\partial \delta \partial \delta'} E(\hat{\delta} - \delta)(\hat{\delta} - \delta)' \right\} + o(m^{-1}). \quad (8.26)$$

To evaluate the bias of $\hat{\delta}$, we may write

$$E(\hat{\sigma}_e^2) = E(m_{yy} - \hat{b}_1 m_{xy}) = \sigma_e^2 + b_1^2 m_{xx} - E(\hat{b}_1 m_{xy}),$$

where

$$E(\hat{b}_1 m_{xy}) = b_1^2 m_{xx} + m_{xx}^{-1} Var(m_{xy}) + b_1^2 m_{xx}^{-1} Var(m_{xx}) - 2b_1 m_{xx}^{-1} Cov(m_{xy}, m_{xx}) + o(m^{-1}),$$

then,

$$E(\hat{\sigma}_e^2 - \sigma_e^2) = -m_{xx}^{-1} Var(m_{xy}) - b_1^2 m_{xx}^{-1} Var(m_{xx}) + 2b_1 m_{xx}^{-1} Cov(m_{xy}, m_{xx}) + o(m^{-1}).$$

In addition,

$$E(\hat{\sigma}_\nu^2 - \sigma_\nu^2) = \frac{m-1}{g_m} \left[b_1^2 M_{xx} - E(\hat{\sigma}_e^2 - \sigma_e^2) - E(\hat{b}_1^2 (M_{xx} - \sigma_u^2)) \right],$$

where

$$\begin{aligned} E(\hat{b}_1^2 (M_{xx} - \sigma_u^2)) &= E(\hat{b}_1^2) E(M_{xx} - \sigma_u^2) \\ &= \left[b_1^2 + m_{xx}^{-2} Var(m_{xy}) + 3b_1^2 m_{xx}^{-2} Var(m_{xx}) - 4b_1 m_{xx}^{-2} Cov(m_{xy}, m_{xx}) \right] M_{xx} + o(m^{-1}). \end{aligned}$$

Then, we have

$$E(\hat{\sigma}_\nu^2 - \sigma_\nu^2) = \frac{m-1}{g_m m_{xx}} \left[(1 - M_{xx} m_{xx}^{-1}) Var(m_{xy}) + b_1^2 (1 - 3m_{xx}^{-1} M_{xx}) Var(m_{xx}) \right]$$

$$-2b_1(1 - 2m_{xx}^{-1}M_{xx})Cov(m_{Xy}, m_{XX})] + o(m^{-1}).$$

Moreover, we may write

$$E(\hat{b}_1 - b_1) = b_1 m_{xx}^{-2} Var(m_{XX}) - m_{xx}^{-2} Cov(m_{XX}, m_{Xy}) + o(m^{-1}).$$

To evaluate (8.26), we need to have the following terms:

$$\begin{aligned} \partial g_{1i}/\partial b_1 &= 2b_1 \sigma_u^2 (1 - f_i B_i)^2 / n_i, \\ \partial g_{1i}/\partial \sigma_\nu^2 &= f_i \sigma_e^2 [f_i \sigma_e^2 + 2b_1^2 \sigma_u^2 (1 - f_i B_i)] / (n_i \sigma_\nu^2 + \sigma_e^2)^2, \\ \partial g_{1i}/\partial \sigma_e^2 &= f_i [N_i^{-1} + f_i (1 - B_i)^2 / n_i - 2b_1^2 \sigma_u^2 (1 - B_i)(1 - f_i B_i) / n_i (n_i \sigma_\nu^2 + \sigma_e^2)], \\ \partial^2 g_{1i}/\partial b_1^2 &= 2\sigma_u^2 (1 - f_i B_i)^2 / n_i, \\ \partial^2 g_{1i}/\partial \sigma_\nu^4 &= \frac{f_i \sigma_e^2}{(n_i \sigma_\nu^2 + \sigma_e^2)^2} \left[-2n_i f_i B_i + \frac{2n_i b_1^2 \sigma_u^2 (3f_i B_i - 2)}{n_i \sigma_\nu^2 + \sigma_e^2} \right], \\ \partial^2 g_{1i}/\partial \sigma_e^4 &= -\frac{2f_i^2 (1 - B_i)^2}{n_i (n_i \sigma_\nu^2 + \sigma_e^2)} - \frac{2n_i b_1^2 f_i \sigma_u^2 (1 - B_i)(3f_i B_i - f_i - 2)}{n_i^2 (n_i \sigma_\nu^2 + \sigma_e^2)^2}, \\ \partial^2 g_{1i}/\partial b_1 \partial \sigma_\nu^2 &= 4b_1 f_i \sigma_u^2 B_i (1 - f_i B_i) / (n_i \sigma_\nu^2 + \sigma_e^2), \\ \partial^2 g_{1i}/\partial b_1 \partial \sigma_e^2 &= -4b_1 f_i \sigma_u^2 (1 - B_i)(1 - f_i B_i) / n_i (n_i \sigma_\nu^2 + \sigma_e^2), \\ \partial^2 g_{1i}/\partial \sigma_\nu^2 \partial \sigma_e^2 &= \frac{2f_i}{(n_i \sigma_\nu^2 + \sigma_e^2)^3} \left[b_1^2 \sigma_u^2 (1 - f_i B_i)(n_i \sigma_\nu^2 - \sigma_e^2) + f_i \sigma_e^2 [n_i \sigma_\nu^2 - b_1^2 \sigma_u^2 (1 - B_i)] \right]. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \frac{1}{2} tr \left\{ \frac{\partial^2 g_{1i}}{\partial \delta \partial \delta'} E(\hat{\delta} - \delta)(\hat{\delta} - \delta)' \right\} &= \frac{1}{2} \left[(\partial^2 g_{1i}/\partial b_1^2) Var(\hat{b}_1) + (\partial^2 g_{1i}/\partial \sigma_\nu^4) Var(\hat{\sigma}_\nu^2) \right. \\ &+ (\partial^2 g_{1i}/\partial \sigma_e^4) Var(\hat{\sigma}_e^2) \left. \right] + (\partial^2 g_{1i}/\partial b_1 \partial \sigma_\nu^2) Cov(\hat{b}_1, \hat{\sigma}_\nu^2) + (\partial^2 g_{1i}/\partial b_1 \partial \sigma_e^2) Cov(\hat{b}_1, \hat{\sigma}_e^2) \\ &+ (\partial^2 g_{1i}/\partial \sigma_\nu^2 \partial \sigma_e^2) Cov(\hat{\sigma}_\nu^2, \hat{\sigma}_e^2) \equiv -g'_i(\delta). \end{aligned}$$

On the other hand, we have $E[g_i(\hat{\delta})] - g_i(\delta) = o(m^{-1})$ and $E[C_i(\hat{\delta})] - C_i(\delta) = o(m^{-1})$. We should point out that in $g_{1i}(\delta)$, other unknown terms of function of x_{ij} are involved which are easy to estimate.

Theorem 8.3. *A second order unbiased estimator of $MSE(\hat{\gamma}_i^{PEB})$ is given by*

$$\begin{aligned} mse(\hat{\gamma}_i^{PEB}) = & g_{1i}(\hat{\delta}) - \frac{\partial g_{1i}}{\partial b_1} E(\hat{b}_1 - b_1)|_{\delta=\hat{\delta}} - \frac{\partial g_{1i}}{\partial \sigma_\nu^2} E(\hat{\sigma}_\nu^2 - \sigma_\nu^2)|_{\delta=\hat{\delta}} - \frac{\partial g_{1i}}{\partial \sigma_e^2} E(\hat{\sigma}_e^2 - \sigma_e^2)|_{\delta=\hat{\delta}} \\ & + g_i(\hat{\delta}) + g_i'(\hat{\delta}) + 2C_i(\hat{\delta}), \end{aligned}$$

in the sense that $E[mse(\hat{\gamma}_i^{PEB})] - MSE(\hat{\gamma}_i^{PEB}) = o(m^{-1})$.

8.6 Asymptotic Result on $m^{-1} \sum_{i=1}^m E(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})^2$

The pseudo-EB estimator $\hat{\gamma}_i^{PEB}$ is asymptotically optimal in the sense that

$$m^{-1} \sum_{i=1}^m E(\hat{\gamma}_i^{PEB} - \hat{\gamma}_i^{PB})^2 \rightarrow 0 \quad \text{as } m \rightarrow \infty,$$

similar to the result under the structural model (Chapter 7). In the case of functional model, the calculations are very similar to the ones used in Ghosh and Sinha (2006). We only need to replace x_i with $\bar{x}_i$ everywhere in Theorem 3.1 of Ghosh and Sinha (2006).

8.7 Simulation Study

To investigate the performance of the proposed method when the covariates are subject to measurement errors, a simulation study was undertaken. To this end,

we used the data in Battese et al. (1988). We have already used this data set in Chapters 3, 5 and 7. Battese et al. (1988) considered the estimation of mean hectares of soybeans per segment for 12 counties (area) in north-central Iowa. The total sample size, $n = \sum_i n_i$, for the 12 counties is 36, and the sample size n_i within each county ranged from 1 to 5. The population size N_i within each county ranged from 402 to 965. The nested error regression model is

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \nu_i + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m$$

where y_{ij} is the number of hectares of soybeans in the j -th segment of the i -th county, x_{1ij} and x_{2ij} are the number of pixels classified as corn and soybeans respectively, in the j -th segment of the i -th county. As in Chapter 7, we consider prediction of soybean data using soybean pixels only as covariates in our simulation study. Hence, the nested error regression model reduces to

$$y_{ij} = b_0 + b_1 x_{ij} + \nu_i + e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m$$

Also, we have $X_{ij} = x_{ij} + u_{ij}$ where X_{ij} is as observed values instead of x_{ij} . If we fit the above regression model to data set given by Battese et al. (1988), we have $\hat{b}_0 = -3.12, \hat{b}_1 = 0.472, \hat{\sigma}_\nu^2 = 259.93$ and $\hat{\sigma}_e^2 = 191.14$. We assume $\sigma_u^2 = 500$ as known. We duplicated x_{ij} in our simulation study based on the number of areas. We treated $\hat{b}_0, \hat{b}_1, \hat{\sigma}_\nu^2$ and $\hat{\sigma}_e^2$ as true parameters in our simulation study.

We generated $B = 5000$ independent sets of normally distributed random variables $\{\nu_i; i = 1, \dots, m\}, \{e_{ij}; j = 1, \dots, n_i; i = 1, \dots, m\}$ having mean zero and specified variances σ_ν^2 and σ_e^2 . From these generated data sets, observations $\{(y_{ij}, X_{ij}); j = 1, \dots, n_i; i = 1, \dots, m\}$ are obtained using the model $y_{ij} = b_0 + b_1 x_{ij} + \nu_i + e_{ij}$ and $X_{ij} = x_{ij} + u_{ij}$. By using the generated samples $(y_{ij}^{(b)}, X_{ij}^{(b)})$, $b=1, \dots, B=5000$, we cal-

culated $\hat{b}_0^{(b)}, \hat{b}_1^{(b)}, \hat{\sigma}_\nu^{2(b)}$ and $\hat{\sigma}_e^{2(b)}$ by the method of moments mentioned in Section 8.2.

For each generated sample, we calculated

$$ARB_p = B^{-1} \sum_{b=1}^B (est_p^{(b)} / par) - 1$$

and

$$ARE_p = B^{-1} \sum_{b=1}^B |(est_p^{(b)} - par) / par|,$$

as average relative bias and absolute relative error of the proposed model parameters where par and $est_p^{(b)}$ denote the model parameters and the estimation of model parameters based on the proposed method of Section 8.2.

In addition, we generated the i -th area mean as $\gamma_i^{(b)} = b_0 + b_1 \bar{X}_{ip} + \nu_i^{(b)}$, $b = 1, \dots, B = 5000$, where $\bar{X}_{ip}$ is i -th population area mean of x_{ij} . To get the empirical MSE (EMSE), we calculated

$$EMSE_{ip} = B^{-1} \sum_{b=1}^B (\hat{\gamma}_{ip}^{PEB(b)} - \gamma_i^{(b)})^2,$$

as EMSE of PEB estimators of i -th area mean for the proposed method, where $\hat{\gamma}_{ip}^{PEB(b)}$ is given by (8.12) for each generated sample.

On the other hand, we used the naive approach to show the effect of measurement errors in the estimation of model parameters. In this approach, using the generated samples $(y_{ij}^{(b)}, X_{ij}^{(b)})$, $b=1, \dots, B=5000$, we calculated $\hat{b}_0^{(b)}, \hat{b}_1^{(b)}, \hat{\sigma}_\nu^{2(b)}$ and $\hat{\sigma}_e^{2(b)}$ by the generalized least squared method for regression parameters and the fitting-of-constants method for variance components σ_ν^2 and σ_e^2 as mentioned in Chapter 3. For each generated sample, we calculated

$$ARB_n = B^{-1} \sum_{b=1}^B (est_n^{(b)} / par) - 1$$

and

$$ARE_n = B^{-1} \sum_{b=1}^B |(est_n^{(b)} - par)/par|,$$

as average relative bias and absolute relative error of naive model parameters where $est_n^{(b)}$ is the estimation of model parameters based on the naive method. We also calculated

$$EMSE_{in} = B^{-1} \sum_{b=1}^B (\hat{\gamma}_{in}^{EB(b)} - \gamma_i^{(b)})^2,$$

as EMSE of EB (or EBLUP) estimators of i -th area mean for the naive method, where $\hat{\gamma}_{in}^{EB(b)}$ is the EBLUP of i -th area mean discussed in Chapter 1.

In the third scenario, we used the generated samples without measurement error $(y_{ij}^{(b)}, x_{ij})$, $b=1, \dots, B=5000$, and followed the procedure described in naive method to calculate ARB_t and ARE_t as

$$ARB_t = B^{-1} \sum_{b=1}^B (est_t^{(b)} / par) - 1$$

and

$$ARE_t = B^{-1} \sum_{b=1}^B |(est_t^{(b)} - par)/par|,$$

as average relative bias and absolute relative error of true (without measurement error) model and the EMSE of EB estimators of i -th area mean for the true method as $EMSE_{it}$.

Results on average relative bias, absolute relative error and the empirical MSE of EB estimators are reported in Tables 8.1 and 8.2 respectively. In fact, we calculated the average relative bias and absolute relative error for different values of m but, to save the space, we only present the average relative bias and absolute relative error for $m = 768$; but results are similar for other values of m .

However, for the empirical MSE of PEB estimators, we report the results for different values of number of areas m . As shown in Table 8.1, ARB_p for b_0, b_1, σ_ν^2 and σ_e^2 are significantly smaller than the corresponding values in ARB_n . As a result, if the covariates are subject to measurement errors and we ignore them, we will loss some efficiency in the sense of estimation of model parameters. For example, ARB for σ_e^2 for the proposed method is -0.10% compared to 50.47% for the naive method. We have similar results for ARE_p for the proposed method compared to ARE_n for the naive method. However, ARE_p for σ_ν^2 is 10.97% compared to 6.18% for the ARE_n . Moreover, we want to check the loss of efficiency in the sense of ARB and ARE if the covariates are subject to measurement errors. To this end, we compare ARB_p and ARE_p with corresponding ARB_t and ARE_t . As expected, ARB and ARE for the true method (without measurement error) are smaller then ARB and ARE for the proposed method. However, the loss in efficiency in the later case is much smaller than pervious case where we use naive method wrongly.

In addition, we report the empirical MSE of PEB estimators for selected m for the proposed ($EMSE_p$), naive ($EMSE_n$) and true ($EMSE_t$) methods. Table 8.2 shows that with increasing m , $EMSE_p, EMSE_n$ and $EMSE_t$ are decreasing. In addition, the rate of decrease for the $EMSE_p$ is faster than the rate for the other two methods. However, even though $EMSE_p$ is decreasing faster than $EMSE_n$ with increasing m , the EMSE of PEB estimators for the proposed method is still larger than for the naive method, the difference between $EMSE_p$ and $EMSE_n$ is not very significant for large m . As expected, the EMSE of EB estimators for the true method is smaller than the proposed and the naive methods.

To further study the behavior of MSE and evaluate the effect of the estimation

of model parameters on the MSE, we calculated the terms in the following decomposition:

$$\begin{aligned}
EMSE_{ip} = & B^{-1} \sum_{b=1}^B (\hat{\gamma}_{ip}^{PB(b)} - \gamma_i^{(b)})^2 + B^{-1} \sum_{b=1}^B (\hat{\gamma}_{ip}^{PEB(b)} - \hat{\gamma}_{ip}^{PB(b)})^2 \\
& + 2B^{-1} \sum_{b=1}^B (\hat{\gamma}_{ip}^{PEB(b)} - \hat{\gamma}_{ip}^{PB(b)}) (\hat{\gamma}_{ip}^{PB(b)} - \gamma_i^{(b)}). \tag{8.27}
\end{aligned}$$

The first term on right hand side of (8.27) is the leading term and the other terms are 0 when the all parameters are known. The second term accounts for the variability of model parameters. In fact, we have shown in Section 8.6 that for large m , the second term averaged over areas i tends to 0, i.e., the pseudo-EB estimator $\hat{\gamma}_i^{PEB}$ is asymptotically optimal. We have shown this fact here using simulation. To have comparable results, we can write similar equations as (8.27) for the naive and the true methods. The results for selected values of m are considered. Table 8.3 gives the first term on right hand side of (8.27) averaged over areas i as the leading term for the proposed method ($EMSE_{1p}$) and similarly for the naive ($EMSE_{1n}$) and the true ($EMSE_{1t}$) methods. As expected, the leading terms for the proposed and the naive methods are the same across values of m . However, the leading term for the true method is significantly smaller than the corresponding values for the proposed and the naive methods.

In addition, we evaluated the effect on MSE due to estimating model parameters. To that end, we calculated the second term on the right side of (8.27) averaged over areas i for the proposed method ($EMSE_{2p}$) and similarly for the naive ($EMSE_{2n}$) and the true ($EMSE_{2t}$) methods. Results are reported in Table 8.4. As expected, with increasing m , $EMSE_{2p}$, $EMSE_{2n}$ and $EMSE_{2t}$ are decreasing. However, the rate of decrease for the proposed method is faster than the rate for the other two

methods. In addition, for large m , $EMSE_{2p}$ tends to 0 as we have shown in Section 8.6. However, $EMSE_{2n}$ does not go to 0 for large m and tends to a positive value. Obviously, values $EMSE_{2t}$ for the true method are less than those for the other methods.

Moreover, we evaluated the cross-product term given in (8.27). To that end, we calculated the cross-product term on right hand side of (8.27) averaged over areas i for the proposed method ($EMSE_{3p}$) and similarly for the naive ($EMSE_{3n}$) and the true ($EMSE_{3t}$) methods. Results are given in Table 8.5. Similar to Table 8.4, $EMSE_{3p}$ tends to 0 for large m while the cross-product term for the naive method does not go to 0 and tends to a negative value for large m . As in Table 8.4, $EMSE_{3t}$ tends to 0 for large m and has smaller value for specified m comparing with the proposed and the naive methods.

8.8 Summary

In this Chapter, we studied a functional measurement error model with the nested error regression model when the covariates in the model are subject to measurement errors. In particular, we developed the pseudo-empirical Bayes (PEB) estimators of area means. We derived a second order approximation to the mean squared error of the PEB estimators, and then obtained an estimator of MSE that is approximately unbiased. In addition, we derived an asymptotic expression for the total MSE of the pseudo-EB estimators. Empirical results on the performance of the proposed method are also presented.

Method	Quality measure	b_0	b_1	σ_ν^2	σ_e^2
Proposed	ARB(%)	4.62	0.02	-0.19	-0.10
	ARE(%)	49.07	1.47	10.97	4.62
Naive	ARB(%)	-360.33	-11.96	1.24	50.47
	ARE(%)	360.33	11.96	6.18	50.47
True	ARB(%)	0.49	0.00	0.01	0.06
	ARE(%)	33.13	0.94	5.50	2.93

Table 8.1: Percent average relative bias and percent absolute relative error of model parameters based on the proposed, naive and true methods for $m = 768$

Number of areas	$EMSE_p$	$EMSE_n$	$EMSE_t$
60	95.62	85.92	63.68
84	92.40	85.21	63.69
108	91.01	84.65	63.14
132	90.22	84.73	63.11
156	89.58	84.34	62.81
192	89.12	84.15	62.63
384	88.18	83.66	62.44
768	87.97	83.33	62.16

Table 8.2: Empirical MSE of PEB estimators for the proposed method and similar estimators for the naive and true methods

Number of areas	$EMSE_{1p}$	$EMSE_{1n}$	$EMSE_{1t}$
60	87.46	87.46	61.72
84	87.65	87.65	62.27
108	87.33	87.33	62.04
132	87.86	87.86	62.20
156	87.71	87.71	62.07
192	87.75	87.75	62.01
384	87.69	87.69	62.13
768	87.69	87.69	62.01

Table 8.3: Leading terms of EMSE of PEB, naive and true methods

Number of areas	$EMSE_{2p}$	$EMSE_{2n}$	$EMSE_{2t}$
60	11.97	9.80	1.98
84	7.03	8.88	1.42
108	5.50	8.50	1.10
132	3.89	8.30	0.89
156	3.09	8.01	0.75
192	2.19	7.77	0.60
384	0.95	7.36	0.31
768	0.46	7.09	0.16

Table 8.4: Empirical values of $EMSE_{2p}$, $EMSE_{2n}$ and $EMSE_{2t}$

Number of areas	$EMSE_{3p}$	$EMSE_{3n}$	$EMSE_{3t}$
60	-3.80	-11.34	-0.02
84	-2.28	-11.31	-0.01
108	-1.82	-11.18	0.00
132	-1.53	-11.43	0.02
156	-1.21	-11.38	-0.01
192	-0.82	-11.37	0.02
384	-0.46	-11.39	-0.01
768	-0.17	-11.45	0.00

Table 8.5: Empirical values of the cross-product terms $EMSE_{3p}$, $EMSE_{3n}$ and $EMSE_{3t}$

Chapter 9

Future Research

In this thesis, we have studied various small area models. We have applied the EBLUP (EB) approach and have obtained the MSE of EBLUP for small area means under these models. We have also considered MSE estimation. For future research, we propose the following directions:

In Chapter 2, we studied a new generalized regression (GREG) estimator of a small area mean under a two-level model. We derived the MSE of the new GREG estimator under the model, assuming the model parameters are known. We can, however, estimate the model parameters to find the MSE approximation. Specifically, we can estimate the regression parameters by the generalized least squared method and estimate the variance-covariance matrix Σ_v by the restricted iterative generalized least squares (RIGLS) introduced by Goldstein (1989) and used by Moura and Holt (1999). The RIGLS is equivalent to REML under normality of errors. We can then derive the MSE estimation of EBLUP by the jackknife method introduced by Jiang, Lahiri, and Wan (2002).

In Chapter 3, we studied the pseudo-EBLUP estimator proposed by You and

Rao (2002). To obtain the pseudo-EBLUP, we used the REML to estimate the variance components σ_ν^2 and σ_e^2 . In fact, we used this approach with only a 2×2 matrix involved to calculate the cross-product term of MSE approximation. We can, however, estimate the variance components σ_ν^2 and σ_e^2 by ML or fitting-of-constants method. However, in this case, we will have a $(p + 2) \times (p + 2)$ matrix to find the cross-product in MSE approximation. We can then derive the MSE estimation and compare the results with the proposed MSE estimation presented in Chapter 3.

In Chapter 4, we introduced a sub-area level small area model. We used an iterative method of estimating the model variances based on weighted residual sum of squares to obtain EBLUP estimators of area and sub-area means. We can, however, estimate the variance components σ_ν^2 and σ_u^2 by REML, ML or fitting-of-constants method to find EBLUP estimators of area and sub-area means. We can then derive second order approximations to the mean squared error (MSE) of the model-based small area estimators and then derive MSE estimators unbiased to second order. We can then compare the results with the results presented in Chapter 4. Another direction to sub-area is to derive the MSE estimation of EBLUP by using the jackknife method introduced by Jiang, Lahiri, and Wan (2002).

In Chapter 5, we studied the conditional MSE estimation for the nested error linear regression model. We derived a second order approximation to the conditional MSE given the area-specific response data and derived a nearly unbiased estimator of the conditional MSE. Booth and Hobert (1998) studied the conditional MSE estimation in a very broad type of generalized linear mixed models. Booth and Hobert (1998) estimated the variance components based on ML approach. We can extend our approach to the two-level model and estimate the regression parameters

by the generalized least squared method and the variance-covariance matrix Σ_ν by the RIGLS (or REML) to find the EBLUP estimators of small area means. We can derive the conditional MSE estimation by the jackknife method introduced by Jiang, Lahiri, and Wan (2002). We can then compare the results with the Booth and Hobert (1998) method.

In Chapter 6, we derived second order approximations to the MSE and the conditional MSE, given $y_i = (y_{i1}, \dots, y_{in_i})'$, of $\hat{\mu}_{iw}$ under the nested error linear regression model. Moreover, we derived estimators of MSE and conditional MSE of $\hat{\mu}_{iw}$ that are nearly unbiased. Similar to Chapter 5, we can extend our approach to the two-level model and estimate the variance-covariance matrix Σ_ν by the RIGLS (or REML) to find the EBLUP estimators of $\hat{\mu}_{iw}$. Using jackknife method, we can derive the MSE and the conditional MSE estimators of $\hat{\mu}_{iw}$.

In Chapters 7 and 8, we studied the nested error linear regression model when the covariates are subject to measurement error (structural and functional model). We used the method of moments to estimate the model parameters to obtain EB (or pseudo-EB) estimators of small area means. We also derived estimators of MSE which are nearly unbiased. We can, however, estimate the model parameters by the generalized least squared method for regression parameters and the REML or ML method for variance components to obtain EB (or PEB) estimators of small area means. We can then derive second order unbiased estimators of MSE by the jackknife method. We can then compare the results with the proposed methods in Chapters 7 and 8.

Bibliography

- [1] Battese, G.E., Harter, R.H., and Fuller, W.A. (1988) An Error Components Model for Prediction of County Crop Areas Using Survey and Satellite Data. *Journal of the American Statistical Association*, **83**, 28-36.
- [2] Bhat, B.R. (2004) *Modern Probability Theory*, New Age International Pvt Ltd.
- [3] Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1975) *Discrete Multivariate Analysis: Theory and Practice*, Cambridge, Mass., MIT P.
- [4] Booth, J.G. and Hobert, J.P. (1998) Standard Errors of Prediction in Generalized Linear Mixed Models. *Journal of the American Statistical Association*, **93**, 262-272.
- [5] Carlin, B.P. and Louis, T.A. (2000) *Bayes and Empirical Bayes Methods for Data Analysis* (2nd ed.), London: Chapman and Hall.
- [6] Cressie, N. (1991) Small-Area Prediction of Undercount Using the General Linear Model, *Proceedings of Statistics Symposium 90: Measurement and Improvement of Data Quality*, Ottawa: Statistics Canada, pp. 93-105.

- [7] Das, K., Jiang, J., and Rao, J.N.K. (2004) Mean Squared Error of Empirical Predictor. *The Annals of Statistics*, **32**, 818-840.
- [8] Datta, G.S., Ghosh, M., Huang, E.T., Isaki, C.T., Schultz, L.K., and Tsay, J.H. (1992) Hierarchical and Empirical Bayes Methods for Adjustment of Census Undercount: The 1998 Missouri Dress Rehearsal Data, *Survey Methodology*, **18**, 95-108.
- [9] Datta, G.S., Kubokawa, T., and Rao, J.N.K. (2004) Estimation of MSE in Small Area Estimation. *Technical Report*, Department of Statistics, University of Georgia, Athens.
- [10] Datta, G.S. and Lahiri, P. (2000) A Unified Measure of Uncertainty of Estimated Best Linear Unbiased Predictors in Small Area Estimation Problems. *Statistica Sinica*, **10**, 613-627.
- [11] Datta, G.S., Rao, J.N.K., and Smith, D.D. (2005) On Measuring the Variability of Small Area Estimators Under a Basic Area Level Model. *Biometrika*, **92**, 183-196.
- [12] Fay, R.E. and Herriot, R.A. (1979) Estimation of Income from Small Places: An Application of James-Stein Procedures to Census Data. *Journal of the American Statistical Association*, **74**, 269-277.
- [13] Fuller, W.A. (1987) *Measurement Error Models*, John Wiley and Sons.
- [14] Fuller, W.A. and Goyeneche, J.J. (1998) Estimation of the State Variance Component. *A draft paper*.

- [15] Ghosh, M. and Sinha, K. (2006) Empirical Bayes Estimation in Finite Population Sampling Under Functional Measurement Error Models. *Journal of Statistical Planning and Inference*, In press.
- [16] Ghosh, M., Sinha, K., and Kim, D. (2006) Empirical and Hierarchical Bayesian Estimation in Finite Population Sampling Under Structural Measurement Error Models. *The Scandinavian Journal of Statistics*, **33**, 591-608.
- [17] Goldstein, H. (1989) Restricted Unbiased Iterative Generalized Least Squares Estimation. *Biometrika*, **76**, 622-623.
- [18] Henderson, C.R. (1953) Estimation of Variance and Covariance Components. *Biometrics*, **9**, 226-252.
- [19] Henderson, C.R. (1975) Best Linear Unbiased Estimation and Prediction Under a Selection Model. *Biometrics*, **31**, 423-447.
- [20] Jiang, J. and Lahiri, P. (2006) Estimation of Finite Population Domain Means: A Model-Assisted Empirical Best Prediction Approach. *Journal of the American Statistical Association*, **101**, 301-311.
- [21] Jiang, J., Lahiri, P., and Wan, S-M. (2002) A Unified Jackknife Theory for Empirical Best Prediction with M-Estimation. *The Annals of Statistics*, **30**, 1782-1810.
- [22] Kackar, R.N. and Harville, D.A. (1984) Approximations for Standard Errors of Estimators of Fixed and Random Effects in Mixed Linear Models. *Journal of the American Statistical Association*, **79**, 853-862.

- [23] Kleffe, J. and Rao, J.N.K. (1992) Estimation of Mean Square Error of Empirical Best Linear Unbiased Predictors Under a Random Error Variance Linear Model. *Journal of Multivariate Analysis*, **43**, 1-15.
- [24] Lehtonen, R. and Veijanen, A. (1999) Domain Estimation with Logistic Generalized Regression and Related Estimators. *Conference Proceedings, IASS Satellite Conference on Small Area Estimation*, 121-128, Riga, Latvia.
- [25] Martiz, J.S. and Lwin, T. (1989) *Empirical Bayes Methods* (2nd ed.), London: Chapman and Hall.
- [26] Moura, F.A.S. and Holt, D. (1999) Small Area Estimation Using Multi-level Models. *Survey Methodology*, **25**, 73-80.
- [27] Prasad, N.G.N. and Rao, J.N.K. (1990) The Estimation of the Mean Squared Error of Small-Area Estimators. *Journal of the American Statistical Association*, **85**, 163-171.
- [28] Rao, C.R. (2001) *Linear Statistical Inference and Its Applications*, New York: John Wiley.
- [29] Rao, J.N.K. (2003) *Small Area Estimation*, New York: John Wiley.
- [30] Rao, J.N.K. and Yu, M. (1992) Small Area Estimation by Combining Time Series and Cross-Sectional Data. *Proceedings of the Section on Survey Research Method, American Statistical Association*, pp. 1-9.

- [31] Rao, J.N.K. and Yu, M. (1994) Small Area Estimation by Combining Time Series and Cross-Sectional Data. *The Canadian Journal of Statistics*, **22**, 511-528.
- [32] Resnick, S.I. (1999) *A Probability Path*, Birkhäuser, Boston.
- [33] Särndal, C.E., Swensson, B., and Wretman, J.H. (1992) *Model Assisted Survey Sampling*, New York: Springer-Verlag.
- [34] Searle, S.R. (1971) *Linear Models*. New York: John Wiley.
- [35] Serfling, R.J. (2001) *Approximation Theorems of Mathematical Statistics*, John Wiley and Sons.
- [36] Stein, C. (1981) Estimation of the Mean of a Multivariate Normal Distribution. *The Annals of Statistics*, **9**, 1135-1151.
- [37] Stukel, D.M. (1991) Small Area Estimation Under One and Two-fold Nested Error Regression Models, *Unpublished Ph.D. Thesis*, Carleton University, Ottawa, Canada.
- [38] Stukel, D.M. and Rao, J.N.K. (1999) Small-Area Estimation Under Two-fold Nested Errors Regression Models. *Journal of Statistical Planning and Inference*, **78**, 131-147.
- [39] You, Y. and Rao, J.N.K. (2002) A Pseudo-Empirical Best Linear Unbiased Prediction Approach to Small Area Estimation Using Survey Weights. *The Canadian Journal of Statistics*, **30**, 431-439.