IMPROVED CONFIDENCE INTERVALS FOR A SMALL AREA MEAN UNDER THE FAY-HERRIOT MODEL

Yegnanew Alem Shiferaw

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the Degree of Doctor of Philosophy

Supervisor: Professor Jacky Galpin

Johannesburg, August 2016

Declaration

I declare that this Thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other university.

____

(Signature of candidate)

———————- day of ———————— 2016 in Johannesburg.

Abstract

There is a growing demand for small area estimates for policy and decision making, local planning and fund distribution. Surveys are generally designed to give representative estimates at national or regional level, but estimates of variables of interest are often also needed at the small area levels. These cannot be reliably obtained from the survey data as the sample sizes at these levels are too small. This problem is addressed by using small area estimation techniques. The main aim of this thesis is to develop confidence intervals (CIs) which are accurate to terms $O(m^{-3/2})$ under the FH model using the Taylor series expansion. Rao (2003a), among others, notes that there is a situation in mixed model estimation that the estimates of the variance component of the random effect, A, can take negative values. In this case, Prasad and Rao (1990) consider $\hat{A} = 0$. Under this situation, the contribution of the mean squared error (MSE) estimate, assuming all parameters are known, becomes zero. As a solution, Rao (2003a) among others proposed a weighted estimator with fixed weights (i.e., $w_i = \frac{1}{2}$). In addition, if the MSE estimate is negative, we cannot construct CIs based on the empirical best linear unbiased predictor (EBLUP) estimates. Datta, Kubokawa, Molina and Rao (2011) derived the MSE estimator for the weighted estimator with fixed weights which is always positive. We use their MSE estimator to derive CIs based on this estimator to overcome the above difficulties. The other criticism of the MSE estimator is that it is not area-specific since it does not involve the direct estimator in its expression. Following Rao (2001), we propose area specific MSE estimators and use them to construct CIs. The performance of the proposed CIs are investigated via simulation studies and compared with the Cox (1975) and Prasad and Rao (1990) methods. Our simulation results show that the proposed CIs have higher coverage probabilities. These methods are applied to standard poverty and percentage of food expenditure measures estimated from the 2010/11 Household Consumption Expenditure survey and the 2007 census data sets.

Keywords: Small area estimation, Weighted estimator with fixed weights, EBLUP, FH model, MSE, CI, Poverty, percentage of food expenditure

Papers Submitted to Journals, Conference Presentations and Prize Awards

Paper Accepted for Publication

- Yegnanew A. Shiferaw and Jacky S. Galpin (2015). Area specific confidence intervals for a small area mean under the Fay-Herriot model. *Journal of Iranian Statistical Society* YY, ZZ ZZ (Part of chapter 4)

Conference Presentations

- Small area estimation of poverty in Ethiopia at the 15th African Economic Conference at Democratic Republic of Congo, Kinshasa, 02-04 November, 2015.
- Confidence Intervals for a small area mean under the Fay-Herrot model, Sub-Saharan Conference on Spatial Statistics, School of Public Health, University of the Witwatersrand, 17-24 November 2014.
- Improved Confidence Intervals under the Fay-Herriot Model: Application to Poverty Incidence in Ethiopia, the 56th annual South African Statistical Association conference, Rhodes University in Grahamstown, 27-30 October 2014.
- Estimating Poverty in the Ethiopian Zones Using Small Area Estimation Models, the 4th ISIbalo Young African Statisticians Conference: Pretoria, South Africa, 31 July-2 August 2014.

Seminar

- Area specific confidence intervals under the Fay-Herriot model: application to percentage of food expenditure data, School of Statistics and Actuarial Science, University of the Witwatersrand weekly seminar on 08 October, 2015.

Symposium

- Confidence Interval Estimation of Small Area Parameters under the Linear Mixed Model: Application to Poverty and Percentage of Food Expenditure Data, University of the Witwatersrand, 7th Cross Faculty Graduate Symposium, 01-02 March 2016.

Prize Awards

- **Third Prize Winner**: South African Statistical Association Postgraduate Student Paper Competition, Sponsored by Stats SA, 2014.
- **Best Presenter Award**: the 4th ISIbalo Young African Statisticians Conference: Pretoria, South Africa, 31July-2 August 2014

Acknowledgements

First of all I would like to thank God for helping me, and enabling me to carry out this study. Secondly I wish to thank my supervisor, Professor Jacky Galpin, who has been supportive since the days I began working with her, and whose advice on the research has been priceless. Professor Jacky Galpin has supported me not only by providing a research assistantship, but also searching funding for my accommodation. I would like to thank all the staff of the School of Statistics and Actuarial Science and the University of the Witwatersrand for the use of their computers, libraries and other supports. I would also like to thank Central Statistics Agency (CSA) of Ethiopia for delivering the data for my application study. Special thanks goes to my beloved daughter Eliana, my wife Abeba and all of my families for their encouragement. I would also like to thank all of my friends who supported me in the strive towards my goal.

Acronyms

AL	Average Length
ANOVA	Analysis of Variance
BLUP	Best Linear Unbiased Predictor
CSA	Central Statistics Agency
CP	Coverage Probability
CV	Coefficient of Variation
CI	Confidence Interval
EB	Empirical Bayes
EBLUP	Empirical Best Linear Unbiased Predictor
FH	Fay-Herriot
GVF	Generalized Variance Function
HCE	Household Consumption Expenditure
MDG	Millennium Development Goals
ML	Maximum Likelihood
MoFED	Ministry of Finance and Economic Development
MSE	Mean Squared Error
NGO	Non-Governmental Organisation
OLS	Ordinary Least Squares

PR	Prasad-Rao
RB	Relative Bias
RR	Relative Risk
REML	Restricted Maximum Likelihood
SNNPRS	Southern Nations Nationalities and Peoples Regional State
WLS	Weighted Least Squares

Contents

Declaration **ii**

Abstract **iii**

Conference Presentations and Prize Award **iv**

Acknowledgements **vi**

Acronyms **vii**

List of Figures **xv**

List of Tables **xvii**

1 Introduction **1**

1.1 Small Area Estimation 1

1.2 General Overview of Confidence Interval Estimation 3

1.3 Statement of the Problem 6

1.4 Objectives of the Study 7

1.5 The Fay-Herriot Model 8

1.6 Estimation of the Small Area Mean in the FH Model 9

1.6.1 Best Linear Unbiased Predictor (BLUP) Estimator 9

1.6.2 Empirical Best Linear Unbiased Predictor (EBLUP) Estimator 10

1.6.3 Weighted Estimator with Fixed Weights 10

1.6.4 Empirical Bayes Estimator 11

1.7 Thesis Chapter Headings 11

2 Literature Review 13

2.1 Review of Literature Concerning Estimation of Variance Components 13

2.1.1 Estimation of Sampling Variance 13

2.1.2 Estimation of Model Variance 14

2.1.3 Asymptotic Variance Estimators 15

2.2 Review of Literature Concerning MSE 16

2.2.1 Approximation of the MSE of the EBLUP Estimator 16

2.2.2 Approximation of the MSE of the Weighted Estimator with Fixed Weights . . . 17

2.2.3 Estimation of the MSE of the EBLUP Estimator 18

2.2.4 Estimation of the MSE of the Weighted Estimator with Fixed Weights 20

2.2.5 MSE Estimation of the Non-sampled Areas 20

2.3 Review of Literature Concerning Confidence Interval Estimation of a Small Area Mean . 21

2.3.1 Confidence Intervals Based on the Taylor Series Expansion Methods 22

2.3.2 Confidence Intervals Based on Resampling Approaches 23

3 BLUP and Assessing Uncertainty of Point Estimators 25

3.1 BLUP Estimator 25

3.2 MSE Estimator 27

3.2.1 Approximation of MSE($\hat{\theta}_i^{EB}$) 28

3.2.2 Estimation of MSE($\hat{\theta}_i^{EB}$) 32

3.2.3 Area Specific MSE Estimation for EBLUP 32

3.2.3.1 Area Specific MSE Estimation Based on the ML or REML Estimators 33

3.2.3.2 Area Specific MSE Estimation Based on Method of Moments 33

3.2.4 MSE of the Weighted Estimator with Fixed Weights 34

3.2.4.1 MSE Approximation for the Weighted Estimator with Fixed Weights . 34

3.2.4.2 MSE Estimation for the Weighted Estimator with Fixed Weights . . . 34

4 Confidence Intervals with Improved Coverage Probability 36

4.1 Confidence Intervals for a Small Area Mean 36

4.1.1 Confidence Intervals for a Small Area Mean Based on the Weighted Estimator with Fixed Weights 36

4.1.2 Confidence Intervals Based on the EBLUP 43

4.2 Confidence Intervals for the Difference between Two Small Area Means 44

4.2.1 Confidence Intervals for $\theta_{\mathrm{i}} - \theta_{\mathrm{j}}$ based on the Weighted Estimator with Fixed Weights 45

4.2.2 Confidence Intervals for $\theta_{\mathrm{i}} - \theta_{\mathrm{j}}$ based on EBLUP 50

4.3 Confidence Intervals Based on EBLUP vs Weighted Estimator with Fixed Weights . . . 51

4.3.1 Confidence Intervals for a Small Area Parameter Based on EBLUP vs Weighted Estimator with Fixed Weights 51

4.3.2 Confidence Intervals for the Difference between Two Small Area Means Based on EBLUP vs Weighted Estimator with Fixed Weights 52

5 Finite Sample Performance of MSEs and CIs 53

5.1 Comparison of the MSE Estimators 53

5.1.1 Simulation Study I 53

5.1.2 Design of Simulation Study 54

5.1.3 Simulation Results without Covariates 56

5.1.4 Simulation Results with Covariates 60

5.1.5 Discussion of Simulation Results 65

5.2 Comparison of Confidence Intervals: Simulation Study II 66

5.2.1 Comparison of Confidence Intervals Based on the Weighted Estimator with Fixed Weights 66

5.2.2 Discussion of Simulation Results 67

5.2.3 Comparison of Confidence Intervals Based on EBLUP 74

5.2.4 Discussion of Simulation Results 75

5.3 Comparison of Confidence Intervals for the Difference of Two Small Area Means: Simulation Study III 80

5.3.1 Comparison of Confidence Intervals Based on the Weighted Estimator with Fixed weights 80

5.3.2 Discussion of Simulation Results 81

5.3.3 Comparison of Confidence Intervals Based on EBLUP 86

5.3.4 Discussion of Simulation Results 87

5.4 Comparison of the Proposed Methods with the Existing Methods 91

5.4.1 Comparison of Confidence Intervals for a Small Area Mean 91

5.4.2 Comparison of Confidence Intervals for the Difference of Two Small Area Means 93

6 Application to Real Data **96**

6.1 Data Sources 96

6.1.1 Household Consumption Expenditure Survey 96

6.1.2 The 2007 Population and Housing Census of Ethiopia 97

6.2 Small Area Estimation of Percentage of Food Expenditure in Ethiopia, 2011 98

6.3 Small Area Estimation of Poverty in Ethiopia, 2011 98

6.3.1 Method of Measuring Poverty in Small Areas 98

6.4 Specification of Variables 100

6.5 Procedures for Obtaining EBLUP Estimates in the Fay-Herriot Model 101

6.6 Generalized Variance Function 103

6.7 Percentage of Food Expenditure Estimates 105

6.8 Head Count Ratio Estimates 107

6.8.1 Total Head Count Ratio Estimates 107

6.8.2 Food Head Count Ratio Estimates 107

6.9 Model Diagnostics 109

6.10 The Links between Poverty and Spending on Food 115

7 Conclusions and Areas for Further Research **116**

7.1 Conclusions 116

7.2 Areas for Further Research 118

Appendices **119**

A Proof of Theorem 4.1.2 **119**

B Proof of Theorem 4.2.1 **125**

C Proof of Theorem 4.2.3 **127**

D Tables **132**

E Figures **142**

Bibliography **142**

List of Tables

5.1 Simulated values of $\widehat{\mathrm{MSE}}(\hat{\theta}_{\mathrm{i}}^{\mathrm{EB}})$ multiplied by 100 for m = 15, A = 1, ψ_{i} patterns (I), (II) and (III) and for the normal random effects distribution . . . 57

5.2 Simulated values of the RB of $\widehat{\mathrm{MSE}}(\hat{\theta}_{\mathrm{i}}^{\mathrm{EB}})$ for A = 1, m = 15, ψ_{i} pattern (I), (II), (III) and for the normal and the Laplace random effects distributions . . . 58

5.3 Simulated values of the RR of $\widehat{\mathrm{MSE}}(\hat{\theta}_{\mathrm{i}}^{\mathrm{EB}})$ for A = 1, m = 15, ψ_{i} pattern (I), (II), (III) and for the normal and the Laplace random effects distributions . . . 59

5.4 Simulated values of the RB of $\widehat{\mathrm{MSE}}(\hat{\theta}_{\mathrm{i}}^{\mathrm{EB}})$ for A = 1, m = 15, ψ_{i} pattern (I), (II), (III) and for the normal random effects distribution with covariates . . . 60

5.5 Simulated values of the RR of $\widehat{\mathrm{MSE}}(\hat{\theta}_{\mathrm{i}}^{\mathrm{EB}})$ for A = 1, m = 15, ψ_{i} pattern (I), (II), (III) and for the normal random effects distribution with covariates . . . 61

5.6 CPs and ALs of a small area mean for nominal 95% CIs for ψ_{i} pattern (I), (II) and (III). Results obtained for the normal random effects distributions and m = 20. . . . 72

5.7 CPs and ALs of a small area mean for nominal 95% CIs for ψ_{i} pattern (I), (II) and (III). Results obtained for the Laplace (double exponential) random effects distributions and m = 20. . . . 73

5.8 CPs of a small area mean for nominal 95% CIs for ψ_{i} pattern (II). Results obtained for the normal and the Laplace (double exponential) random effects distributions and m = 20. 76

5.9 ALs of a small area mean for nominal 95% CIs for ψ_{i} pattern (II). Results obtained for the normal and the Laplace (double exponential) random effects distributions and m = 20. 77

5.10 CPs and ALs of the difference between two small area means for nominal 95% CIs for ψ_{i} pattern (I), (II) and (III). Results obtained for the normal random effects distributions and m = 20. . . . 85

5.11 CPs of the difference between two small area means for nominal 95% CIs for ψ_i pattern (II). Results obtained for the normal and the Laplace (double exponential) random effects distributions and m = 20. 89

5.12 ALs of the difference between two small area means for nominal 95% CIs for ψ_i pattern (II). Results obtained for the normal and the Laplace (double exponential) random effects distributions and m = 20. 90

5.13 CPs and ALs of a small area mean for nominal 95% CIs for ψ_i pattern (IV). Results obtained for the normal random effects distributions and m = 15. Note that the first value in the cell is the proposed method and the value in the parenthesis is the naive method. . 92

5.14 CPs and ALs of a small area mean for nominal 95% CIs for ψ_i pattern (IV). Results obtained for the chi-square random effects distributions and m = 15. Note that the first value in the cell is the proposed method and the value in the parenthesis is the naive method. 93

5.15 CPs and ALs of the difference between two small area means for nominal 95% CIs for ψ_i pattern (I). Results obtained for the normal random effects distributions and m = 15. Note that the first value in the cell is the proposed method and the value in the parenthesis is the naive method. 94

D.1 Synthetic estimates of total poverty, food poverty and percentage of food expenditure measures and associated standard errors and normal theory 95% CIs in 2011: FH results 132

D.2 Direct and EBLUP percentage of food expenditure estimates and associated standard errors, CVs and 95% CIs in 2011: FH results 133

D.3 Direct and EBLUP food poverty head count ratio estimates and associated standard errors and CVs in 2011: FH results 136

D.4 Direct and EBLUP total poverty head count ratio and associated standard errors and CVs in 2011: FH results 139

List of Figures

5.1 Simulated values of the average of the absolute value of the RB of $\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB})$ for A = 1 and ψ_i pattern (I) and for the normal random effects distribution. The points are connected with lines for visibility purposes. 62

5.2 Simulated values of the average of the absolute value of the RB of $\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB})$ for A = 1 and ψ_i pattern (II) and for the normal random effects distribution. The points are connected with lines for visibility purposes. 63

5.3 Simulated values of the average of the absolute value of the RB of $\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB})$ for A = 1 and ψ_i pattern (III) and for the normal random effects distribution. The points are connected with lines for visibility purposes. 64

5.4 Simulated values of the CPs of Method IW, Method IIW and Method IIIW for nominal 95% CIs for ψ_i pattern (I). The points are connected with lines for visibility purposes. . . 69

5.5 Simulated values of the CPs of Method IW, Method IIW and Method IIIW for nominal 95% CIs for ψ_i pattern (II). The points are connected with lines for visibility purposes. . 70

5.6 Simulated values of the CPs of Method IW, Method IIW and Method IIIW for nominal 95% CIs for ψ_i pattern (III). The points are connected with lines for visibility purposes. . 71

5.7 Simulated values of the CPs of Method I, Method II, Method III, CIRao, CIJY and CIJY1 for nominal 95% CIs for ψ_i pattern (II). The points are connected with lines for visibility purposes. 78

5.8 Simulated values of the ALs of Method I, Method II, Method III, CIRao, CIJY and CIJY1 for nominal 95% CIs for ψ_i pattern (II). The points are connected with lines for visibility purposes. 79

5.9 Simulated values of the CPs of Method IWD, Method IIWD and Method IIIWD for nominal 95% CIs for ψ_i pattern I. The points are connected with lines for visibility purposes. 82

5.10 Simulated values of the CPs of Method IWD, Method IIWD and Method IIIWD for nominal 95% CIs for ψ_i pattern II. The points are connected with lines for visibility purposes. 83

5.11 Simulated values of the CPs of Method IWD, Method IIWD and Method IIIWD for nominal 95% CIs for ψ_i pattern III. The points are connected with lines for visibility purposes. 84

5.12 Simulated values of the CPs of Method ID, Method IID, Method IIID, CIRaoD, CIJYD and CIJY1D for nominal 95% CIs for ψ_i pattern II. The points are connected with lines for visibility purposes. 88

6.1 Dispersion plots for GVF in 2010/11 using the HCE survey 104

6.2 A plot illustrating the direct and EBLUP CIs and the width of CIs of percentage of food expenditure estimates versus Ethiopian zones in 2010/11 106

6.3 A plot illustrating the standard errors (bottom) and CVs (top) along the Ethiopian zones in 2010/11 112

6.4 Standardized residuals (top); residuals (bottom) versus fitted values 113

6.5 EBLUP estimates of food poverty measures (top), total poverty measures (middle) and percentage of food expenditure (bottom) against direct estimates for each zone 114

6.6 The scatter plot between poverty and spending on food. The points are connected with lines for visibility purposes 115

E.1 Index plots of residuals (top) and residuals against sampling weights (bottom) 142

E.2 Index plots of residuals for total poverty (left), food poverty (middle) and percentage of food expenditure (right). 143

E.3 A map of Ethiopia showing the food costs as a percentage of expenditure 144

E.4 A map of Ethiopia based on the food head count ratio 145

E.5 A map of Ethiopia based on the total head count ratio 146

Chapter 1

Introduction

1.1 Small Area Estimation

There is a growing need for population estimates at small geographic area levels, for use in regional planning and policy and decision making (Nandram, 1999, Li, 2006). An example is poverty mapping for a country or region, for which demographic information is needed, as well as information about household income and poverty levels. The resulting poverty maps can be used as a basis for targeting disadvantaged areas as well as the evaluation of geographically oriented actions and programmes involved in the national poverty strategy (Goh and Virola, 2005; World Bank, 2010).

The census data gives demographic information excluding income and expenditure information, because the questionnaire has to be kept short due to time constraints also, collecting such data for an entire country would be expensive (World Bank, 1998; Tarozzi and Deaton, 2007). Detailed poverty related information is obtained using a separate survey, for which the sample design aims to provide reliable data for large areas such as national or regional levels and pays little or no attention to the smaller areas of interest such as sub-regional area (Battese, Harter and Fuller, 1988; Chatterjee and Lahiri, 2002; Isidro, 2010). For example, in Ethiopia, sub-national administrative levels include zones within each region which are in turn composed of districts (called woredas). Districts are usually composed of villages (called kebeles). Those villages are the primary sampling units in the survey design for national surveys such as Household Consumption Expenditure (HCE) (CSA, 2012).

Furthermore, the 2010/11 Ethiopian HCE survey covered all rural and urban areas of the country except the non-sedentary zones (three zones in Afar and six zones in Somali regional states). In this survey, a total of 28032 households (10,368 in rural and 17,664 in urban areas) were sampled at a country level using a stratified random sampling design (CSA, 2012). Details of the HCE survey will be given in Chapter 6. Estimates from the HCE survey can only be computed for all zones which contain respondents to the sample survey. For example, we cannot compute an estimate for the three zones in Afar and six zones in Somali as there are no sample units selected from these zones. In addition to this, some of the sampled zones and a substantial number of woredas have small sample size and the estimators will thus have low precision. This results in the estimators having unacceptably large standard errors, thus making the estimates statistically unreliable (Prasad and Rao, 1990; Datta and Ghosh, 2012).

We cannot conduct a new survey due to cost and operational considerations. However, we need sufficient information under these restricted resources, in order to increase the effective sample size, thus decreasing the standard error to achieve acceptable statistical precision (Schenker and Raghunathan, 2007; Chandra and Sud, 2012). Moreover, combining information from different surveys or from a census and a survey may address the problem (Chatterjee, Lahiri and Li, 2008, Kim and Rao, 2012, among others). This is the less expensive tool for increasing sample sizes in order to improve the precision of the estimates (Costa, Satorra and Ventura, 2006; Lee et al., 2007).

The methods appropriate for this type of analysis are termed small area estimation methods. The word "small" relates to the sample size of the domain, rather than to the geographical size of the domain, or the size of the domain's population. Ghosh and Rao (1994); Pfeffermann (2002); Rao (2003a) and Pfeffermann (2013) have reviewed these methods in detail.

Furthermore, Fay and Herriot (1979) introduced a basic area level model to obtain model based estimates for per-capita income for small places in the United States. This model is useful in cases where auxiliary data are available at the area level or when it is not possible to link the information of the sample units with census data or unit level data, which might not be available due to confidentiality issues (Diao, Smith, Datta, Maiti, and Opsomer, 2013).

The Fay-Herriot (FH) model is the best known small area estimation model. It combines an estimate for the i^{th} area mean using the survey data (direct estimate) with an estimate from the census data based on a regression equation derived from the survey data (synthetic estimate). The final estimate is a weighted combination of the direct and synthetic estimates. Details of this model are given in sections 1.5 and 1.6.

Using the combined point estimate from a small area model, one can obtain confidence interval (CI), which is the area of interest for this thesis. In addition to this, Datta, Ghosh, Smith and Lahiri (2002); Basu, Ghosh and Mukerjee (2003) among others have derived CIs using the Taylor series expansion method. Hall and Maiti (2006); Chatterjee et al. (2008) and Kubokawa and Nagashima (2012) also addressed it using the parametric bootstrap method (Resampling approach). Resampling is a method that uses only the sample data and resamples this data without replacement, to give subsets of the original sample (jackknifing), or sampling with replacement from the available sample (bootstrapping) to create different realizations of the experimental results.

Traditionally, we construct the CIs using the point estimate $\pm$ q $\times$ standard error, where q is the Z (standard normal) or t cut-off point. However, this method does not produce improved coverage probability (CP) (the percent of CIs containing the true value compared to the desired confidence level) of the intervals (Hall and Maiti, 2006; Li, 2006; Dass, Maiti, Ren and Sinha, 2012). This is because such CIs do not take into account the additional variability resulting from estimation of unknown parameters in the model (Chatterjee et al., 2008).

This study aims to construct improved CIs by including the additional variability due to the estimation of the unknown parameters under both levels of the FH model using the Taylor series expansion method. According to Diao et al. (2013), CIs based on the Taylor series expansion method are easy to implement in practice. This is because the model parameters need to be estimated only once using standard software, and can then be used in the construction of plug-in type CIs.

Furthermore, the performances of our methods are numerically studied in comparison with the naive CIs (CIs that ignore the uncertainty in parameter estimates). This is done using extensive simulations based on the simulation design of Datta et al. (2005) with some modifications.

Moreover, we apply the methods developed to the Ethiopian 2007 census (Federal Democratic Republic of Ethiopia Population and Census Commission, 2008) and 2010/11 HCE data, to estimate poverty and percentage of food expenditure at the zone level. The details of the estimation of poverty and percentage of food expenditure measures are given in Chapter 6.

1.2 General Overview of Confidence Interval Estimation

Apart from the methods for obtaining the EBLUP and weighted estimates, a further issue is that of obtaining CIs for the population small area means based on EBLUP and weighted estimates. The CI provides a range of values, and states as a percentage of our degree of confidence that the true value of a parameter is within that range. The CI for θ_i (1.5.1) which is based only on the sampling model (1.5.2) is not efficient, since it has large AL. In addition to this, an interval based only on linking model (1.5.3) neglects the most important area specific data that is modeled using sampling model. Thus, efficient CIs for the small area parameter based on both levels of the FH model (1.5.4) provide CIs with smaller lengths when compared to the standard method (Chatterjee et al., 2008). These CIs can be achieved by acquiring information from similar areas (Nandram, 1999; Li, 2006; Chatterjee et al., 2008; Li and Lahiri, 2010; Kubokawa and Nagashima, 2012). Furthermore, a common procedure is to employ statistical models, which make it possible to borrow information for the estimation by utilizing data from similar or neighbouring areas or from earlier similar surveys to improve confidence as we do to improve the small area mean estimates. For example, Basu et al. (2003) constructed a CI which is accurate to terms $O(m^{-1})$ using the Taylor series expansion. Li (2006) and Li and Lahiri (2010) derived a CI which is correct up to $O(m^{-3/2})$ using the parametric bootstrap method, where m represents the number of sampled areas.

Chatterjee et al. (2008) and Datta and Ghosh (2012) state that Cox (1975) made the first attempt to construct EB CIs in the context of small area estimation. He also motivated the idea of obtaining an EB CI in the small area estimation terminology. In addition to this, he developed approximate CIs that are accurate to the order of $O(m^{-1})$ for a small area mean θ_i for the balanced (equal number of observations in the different small areas) FH model without any covariates. Furthermore, Morris (1983b) obtained a CI that is correct to the order of $O(m^{-1})$ for an EB estimator that includes the additional uncertainty due to the estimation of the unknown parameters for the FH model under the assumption of equal sampling variances. Additionally, Prasad and Rao (1990) discussed details of small area models incorporating covariates by assuming that the error variances are equal. Hulting and Harville (1991) further compared frequentist and Bayesian methods for developing prediction intervals for a small area mean under the assumption of equal sampling variances. Nandram (1999) also developed CIs for small area means for a type of unit-level small area model for a large number of areas without covariates under equal sampling variances. Furthermore, Jiang and Zhang (2002) developed prediction intervals for a distribution free non-Gaussian linear mixed model under large samples for each small area. Moreover, Datta et al. (2002) developed EB CIs that are correct up to the order of $O(m^{-3/2})$ for equal sampling variances across small areas.

Considering this, Hall and Maiti (2006) proposed interval estimation based on a double bootstrap calibrated sample which is centred at the regression estimate rather than at the EBLUP estimate. Their prediction intervals have coverage accuracy $O(m^{-3})$. Chatterjee et al. (2008) further proposed a CI which is based on a single bootstrap sample for area level means using a normal approximation. Their bootstrap histogram differs from the true EBLUP distribution by only $O(d^3 n^{-3/2})$ where d is the number of parameters and n is the number of observations within the areas. The coverage error is defined in terms of n, sample sizes within the areas, rather than m, the number of sampled areas, as under the Hall and Maiti (2006) approach. This method necessitates repeated generation of a pivotal quantity from several bootstrap samples. Kubokawa and Nagashima (2012) have also obtained CIs based on the parametric bootstrap methods.

Morris (1983b), Prasad and Rao (1990), Datta et al. (2002) amongst others constructed their CIs for a special case of equal sampling variances. However, in small area estimation, we are interested in small areas and hence the sampling error variance estimates ψ_i in the FH model can be unstable (Boonstra and Buelens, 2011). Furthermore, the recent study by Datta and Ghosh (2012) argued that equal sampling variances seldom arise for small area problems (Rao, 2012). Some studies have also been conducted considering the non-constant variance component assumptions. Additionally, Chatterjee et al. (2008) considered two different ψ_i's patterns and A = 1 for the first pattern and A = 2 for the second pattern; where A is the variance of the random effect defined in section 1.5. Kubokawa and Nagashima (2012) also considered two different ψ_i's patterns, but A = 1 for both patterns.

The resampling methods and the Taylor series expansion methods are common approaches for estimating MSEs and developing CIs when exact formulas are not available. However, there are criticisms on both the resampling and the Taylor series expansion methods in the literature.

As we have stated earlier, MSE and CI estimations have been derived based on the Taylor series expansion method by Datta et al. (2002); Basu et al. (2003); Kubokawa (2011) and others. However, it is difficult to approximate MSE and construct CIs based on the EBLUP with the desired level of accuracy. It is also difficult to determine the variance, covariance, second-order bias and partial derivatives of certain matrices in relation to parameters which are not known (Kubokawa and Nagashima, 2012). As a solution, Butar and Lahiri (2003) developed the resampling method that does not need computation of the bias, variance, covariance and partial derivatives.

Calibration and computational difficulties of bootstrap methods have been discussed by Hall, Lee and Young (2000) and Nankervis (2005). In addition to this, the coverage accuracy of CIs based on bootstrap method has been improved through calibration (Chatterjee et al., 2008). Chatterjee et al. (2008, page 7) state that "however, it is not always clear what property of an interval, that is, length, coverage, end points or some other characteristic, ought to be calibrated". "In addition to the computational difficulties especially in the case of calibration, these methods have other issues, for example, choosing between equal tail or shortest interval quantile points, and so on" (Diao et al., 2013, page 498). There are also questions on the use of pivotal statistics and calibration (Chatterjee et al., 2008). Considering this, it is found that calibration is both time and computational effort consuming and the results often lack straightforward interpretability (Chatterjee et al., 2008). It is also not clear on how to construct CIs for the difference of two small area means using the resampling approach (Diao et al., 2013).

It should also be noted that CIs for θ_i are sometimes known by the term prediction intervals when the quantities θ_i's are random. The parameters θ_i's are random variables whose distributions depend on a p-variate auxiliary variable $\mathbf{x}_i$ known for the m small areas, $\theta_i = \mathbf{x}_i'\beta + v_i$ in a Bayesian setting. Thus, prediction intervals are confidence intervals under this setting. Prediction intervals can be used to measure the prediction uncertainty in financial departments. They are more informative and helpful in many aspects of small area estimation. Such intervals combine both point prediction and hypothesis testing (Chatterjee et al., 2008). For instance, prediction intervals may help to establish if different zones in Ethiopia have similar resources and needs, or if certain subpopulation groups are highly affected by poverty or food insecurity. The issue of prediction intervals has been addressed by Hulting and Harville (1991); Jiang and Zhang (2002); Hall and Maiti (2006); Chatterjee et al. (2008) and others.

As we have discussed earlier, CIs have been addressed for θ_i based on EBLUP and EB estimates. There has been no literature on CIs based on weighted estimates of the form (1.5.4). A drawback of CIs based on EBLUP estimates is that it cannot give an interval when $\mathrm{MSE}(\hat{\theta}_i^{\mathrm{EB}})$ takes a negative value (Kubokawa and Nagashima, 2012). To find a solution, we propose CIs based on the weighted estimator with fixed weights. Datta et al. (2011) derived MSE for the weighted estimators with fixed weights which cannot reveal negative value for the FH model. We further use their MSE estimators to derive improved CIs. "Note that the problem of constructing a good EBs prediction interval is much more difficult than the problem of obtaining an accurate MSE" (Nandram, 1999, page 328).

In summary, construction of CIs has received little attention when compared to point estimation in the issue of small area estimation methodology. Construction of CIs has also been limited and restricted to individual small area means. Furthermore, interval estimation for two or more population means is still an area to which very little attention has been paid. Most recently, Diao et al. (2013) proposed CIs for a small area mean as well as for the difference of two small area means for unequal variances under the FH model which is correct up to the order $O(m^{-2})$ using the Taylor series expansion. "We note that, since the small area estimators are not independent, this extension from a single prediction CI to one for differences (or more generally, linear combinations) of predictions is not immediate" (Diao et al., 2013, page 4). In this thesis we proposed CIs for a small area mean and for the difference of two small area means based on EBLUP and weighted estimates with fixed weights for a coverage accuracy of order $O(m^{-3/2})$ under unequal sampling variances for the FH model.

1.3 Statement of the Problem

Mean squared errors (MSEs) have been used to assess the accuracy of estimates. One of the drawbacks of MSEs is that they are not area specific since they do not involve the direct estimator in its expression (Rao, 2003b). As a solution, we propose area specific MSE estimators under the area level model.

Construction of CIs based on the empirical best linear unbiased predictor (EBLUP) have been developed for linear mixed models in small area estimation. However, when we use different methods of estimating variance components, it is possible to get negative values for these components. For example, if the estimated variance of the random component (under the linear mixed model) is negative, Prasad and Rao (1990) set it equal to zero. Those zero estimates will cause problem in the computing procedure and also lead the EBLUP estimator to run into difficulties (Chatterjee and Lahiri, 2002; Li, 2006). In addition, the contribution of the MSE estimate, assuming all parameters are known, becomes zero. This affects the accuracy of both the MSE and CI estimates. Considering this, Pryseley, Tchonlafi, Verbeke and Molenberghs (2011) argue that negative estimates of variance components in the generalized linear model can occur in practice and can be estimated using statistical packages. These negative variance component estimates result due to the method of estimating them (Brown and Prescott, 2006). Such an estimate will usually be an underestimate of a variance component whose true value is small or zero.

In addition, the estimated MSE of EBLUP sometimes takes negative values. A drawback of CIs based on is that it cannot give an interval when the estimates of MSE are negative (and are set to zero) (Kubokawa and Nagashima, 2012). Thus, we proposed alternative CIs based on the MSE of weighted estimator with fixed weights. The MSE estimates under this approach are more stable and always positive (Datta, Kubokawa, Molina and Rao, 2011).

However, all the CIs based on the weighted estimator with fixed weights are free from such drawbacks. In summary, in this thesis, we focus on the EBLUP and weighted estimators of a small area mean, associated MSE estimates and asymptotically improved CIs under different methods of estimating the variance components.

1.4 Objectives of the Study

The main aim of this thesis is to derive area specific MSEs and CIs for a small area mean. The specific objectives of the study are

- to provide different area specific MSE estimators under the FH model and compare these through simulation
- to introduce improved CIs for a single small area mean as well as for the difference of two small area means
- to illustrate the application of our methods using real data sets.

1.5 The Fay-Herriot Model

Let m denote the number of small areas (zones) and θ_i represent the ith small area mean. Suppose y_i $(i = 1, ..., m)$ represents the direct survey estimator of θ_i (obtained from the vector of observations $(y_{i1}, ..., y_{in_i})'$ for area i); n_i denotes the sample size within a small area; $\mathbf{x_i}$=$(x_{1i}, ..., x_{ki})'$ is a $k \times 1$ vector of known covariates for area i related to the target parameters and $\beta = (\beta_1, ..., \beta_k)'$ is a $k \times 1$ vector of regression coefficients. The aim is to estimate the population small area means

$$\theta_i = \mathbf{x_i'}\beta + v_i,\ 1, ..., m \tag{1.5.1}$$

where the random effects v_i have $E(v_i) = 0$ and $Var(v_i) = A(\geq 0)$. Under this set-up, the FH model, assuming the normality, can be written as

Level 1: Sampling Model:

$$y_i|\theta_i \overset{iid}{\sim} N[\theta_i, \psi_i],\ i = 1, ..., m. \tag{1.5.2}$$

ψ_i represents the sampling variance of y_i to be estimated from the sample survey.

Level 2: Linking Model:

$$\theta_i \overset{ind}{\sim} N[\mathbf{x_i'}\beta, A],\ i = 1, ..., m. \tag{1.5.3}$$

Level 1 and Level 2 models are based on the observed data and the borrowed strength component, respectively (Chatterjee et al., 2008). According to Jiang and Lahiri (2006), Level 1 is used to account for the sampling distribution of the direct estimates y_i from the survey, which are weighted averages for small area i and Level 2 is used to link the small area means θ_i to a vector of an area dependent auxiliary variables $\mathbf{x}_i$ obtained from the census and other administrative records. The unknown parameters β and A of Level 2 are estimated from the available data. The above two level model can be rewritten as the following mixed (containing both fixed and random effects) model:

$$y_i = \mathbf{x_i'}\beta + v_i + e_i, i = 1, ..., m \tag{1.5.4}$$

where $\mathbf{x_i'}\beta$ is a fixed effect. A fixed effect is one where the levels of the effect in the sample are the only ones of interest. For example, if we are using sex as a variable for predicting poverty measures, we will treat this as a fixed effect variable as we are assuming this to be the data of interest. Note that the area specific random effects v_i are used to relate θ_i to $\mathbf{x_i'}$. Nissinen (2009, page 14) states that, “the random area effect, v_i, arises as the residual of a linear model specified for the target variable”. The presence of the random area effect, v_i, differentiates the FH model from the linear regression model. The variations in the area estimates which are not explained by the auxiliary variables or the sampling errors are accounted for by the random area effect, v_i, of the FH model. e_i is the sampling error for the ith small area. Both v_i and e_i are mutually independently distributed random variables.

The FH model can also be expressed in matrix notation as

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{v} + \mathbf{e} \tag{1.5.5}$$

where $\mathbf{X} = (x_1, ..., x_m)'$, $\mathbf{y} = (y_1, ..., y_m)'$, $\mathbf{v} = (v_1, ..., v_m)$, $\mathbf{e} = (e_1, ..., e_m)$, the variance-covariance matrix of $\mathbf{y}$ is $\Sigma(A) = \mathrm{diag}(\psi_1, ..., \psi_m) + A\mathbf{I}_m$. We assume that rank $(\mathbf{X}) = k$.

Direct estimators (estimators for an area that use values of the variable of interest, y, only from the sample units in the domain) are unreliable since the sample size available from the particular small geographical area of interest is small (Rao, 2003a). This is why the FH model is considered to combine data from the related administrative records to derive reliable small geographical area level estimators (Datta and Lahiri, 2000, Li, 2006). The FH model borrows information from related local areas to obtain estimates which are more reliable (Prasad and Rao, 1990, Rao, 2003a, Rao, 2003b).

1.6 Estimation of the Small Area Mean in the FH Model

Point estimators have been found to be very useful in small area estimation and related problems. They are widely discussed in the literature. Both Bayesian and non-Bayesian methods were considered to address the point estimation and the associated problem of measuring uncertainty of the point estimator. Furthermore, a substantial amount of literature has evolved on a wide application of small area estimation models (see for example, Fay and Herriot, 1979; Morris, 1983a; Morris, 1983b; Prasad and Rao, 1990; Carlin and Louis, 1996; Datta and Lahiri, 2000 ; Chatterjee and Lahiri, 2002; Jiang and Zhang, 2002; Pfefferman, 2002; Rao, 2003a; Rao, 2003b; Hall and Maiti, 2006; Li and Lahiri, 2010; Boonstra and Buelens, 2011; Jiang and Tang, 2011; Datta and Ghosh, 2012; Pfefferman, 2013 and others).

1.6.1 Best Linear Unbiased Predictor (BLUP) Estimator

Small area means or totals can be expressed as linear combinations of fixed and random effects. The best known method for the prediction of mixed effects is the BLUP. The BLUP estimator was first originated by Henderson (1950) and used by many authors in different applications. It is a weighted combination of the direct estimator, y_i and the regression synthetic estimator, $\mathbf{x}_i'\hat{\beta}$ (Pfeffermann, 2002; Rao, 2003a). One of the most commonly used procedure of deriving the BLUP of θ_i is the following. First, we derive the best predictor of the small area means of interest. Then, we replace the vector β of the fixed effects by its ML or weighted least squares (WLS) or ordinary least squares (OLS) estimators, assuming that the variance components are known. At this stage we can obtain the BLUP of $\theta_i = \mathbf{x}_i'\beta + v_i$. The interest in small area estimation is to estimate or predict the small area means θ_i. The BLUP for θ_i

for the FH model is given by

$$\tilde{\theta}_i^B = (1-\gamma_i)y_i + \gamma_i \mathbf{x}'_i\tilde{\beta} \tag{1.6.1}$$

where $\gamma_i = \frac{\psi_i}{A+\psi_i}$ and $\tilde{\beta} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y}$ with $\mathbf{X} = (\mathbf{x}'_i)_{1\le i\le m}$, $\mathbf{y} = (y'_i)_{1\le i\le m}$, $\Sigma = A\mathbf{I}_m + \Psi$ and $\Psi = \mathrm{diag}(\psi_i, 1 \le i \le m)$. $\mathbf{I}_m$ is the identity matrix. Note that the proofs of some of the properties of BLUP are given in section 3.1 of Chapter 3.

1.6.2 Empirical Best Linear Unbiased Predictor (EBLUP) Estimator

Das, Jiang and Rao (2004, page 818) state that "the term 'empirical predictor' refers to a two-stage predictor of a linear combination of fixed and random effects. In the first stage, a predictor is obtained but it involves unknown parameters; thus, in the second stage, the unknown parameters are replaced by their estimates". However, the EBLUP estimator is a linear combination of the direct and synthetic estimators. The BLUP is not usable until we estimate the model parameters (β and A). Furthermore, a number of estimation procedures for the model variance of the FH model are available in the literature. For example, the method of moments used by Prasad and Rao (PR) (1990), Maximum Likelihood (ML) and Restricted Maximum Likelihood (REML) used by Datta and Lahiri (2000), the Fay and Herriot (FH) (1979) estimation method and so on. When the unknown parameters are replaced by their estimators, then we will have the EBLUP of θ_i which is given by

$$\hat{\theta}_i^{EB} = (1-\hat{\gamma}_i)y_i + \hat{\gamma}_i \mathbf{x}'_i\hat{\beta}, \quad i = 1, ..., m \tag{1.6.2}$$

where $\hat{\beta} = (\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{y}$ and $\hat{\gamma}_i = \frac{\psi_i}{\psi_i+\hat{A}}$.

For an area k without sample data (i.e., $n_k = 0$ for some k), we use the regression synthetic estimator of θ_k given by

$$\hat{\theta}_k^{Syn} = \mathbf{x}'_k\hat{\beta}, \quad i = 1, ..., m. \tag{1.6.3}$$

1.6.3 Weighted Estimator with Fixed Weights

Besides the EBLUP estimator, here we consider another form of estimator which is called a weighted estimator with fixed weights. The EBLUP estimator reduces to a synthetic estimator $\mathbf{x}'_i\hat{\beta}$ regardless of the sampling variance, ψ_i, of the direct estimator y_i when $\hat{A}$ is negative. This makes the contribution of the MSE estimate, assuming all parameters are known, to become zero. The following alternative weighted estimator with fixed weights will be used to overcome this problem. The weighted estimator with fixed weights $1-w_i$ and w_i (Rao, 2003a, Datta et al., 2011) is given by

$$\hat{\theta}_i^{w} = (1-w_i)y_i + w_i \mathbf{x}'_i\hat{\beta}, \quad i = 1, ..., m, \tag{1.6.4}$$

where w_i has a value $0 \leq w_i \leq 1$. It can be chosen as $w_i = 1/_2$ or can be determined from past knowledge. As discussed by Datta et al. (2011), $\hat{\theta}_i^w$ and $\hat{\theta}_i^{EB}$ are special cases of the general estimator of the form $\hat{\theta}_i = y_i + h_i(y)$ which is proposed by Rivest and Belmonte (2000). Datta et al. (2011) derived MSEs and nearly unbiased estimators of MSE for the weighted estimator with fixed weights. In this thesis, we use their MSE estimators of the weighted estimator with fixed weights.

1.6.4 Empirical Bayes Estimator

The EBLUP method is applicable for linear mixed models that cover the basic area level and u-nit level models. On the other hand, the Empirical Bayes (EB) method is more generally applicable, covering generalized linear mixed models that are used to handle categorical and count data. For the Fay-Herriot model, the EB approach may be summarized as follows (Rao, 2003a):

Step 1. Obtain the posterior density, $\phi(y_i \mid \hat{\theta}_i, \beta, A, \psi_i)$, of the parameters of interest, given the data θ_i.

Step 2. Estimate the model parameters (β, A, ψ_i) from the marginal distribution of the data, $\phi(y_i \mid \beta, A, \psi_i)$.

Step 3. Use the posterior density, $\phi(y_i \mid \hat{\theta}_i, \hat{\beta}, \hat{A}, \hat{\psi}_i)$, for making inferences about θ_i, where $(\hat{\beta}, \hat{A}, \hat{\psi}_i)$ are estimators of unknown parameters.

It should also noted that EB and EBLUP estimators defined in sections 1.6.2 and 1.6.4 are identical under the FH model with the normal errors and quadratic loss. Harville (1991) considered a one-way random effects model and showed that, in this case, the EBLUP is identical to an EB estimator. EB can be Empirical Best for frequentist consideration, thus we are using EB in place of EBLUP throughout this thesis.

1.7 Thesis Chapter Headings

This thesis has the following chapters and specific descriptions of each chapter are given as follows.

Chapter 2 gives a literature review of the existing approaches to small area estimation. The main purpose of this chapter is to review and discuss some of the important developments of interval estimation in small area estimation methods. We briefly review variance component estimation, MSE approximation and estimation, and interval estimation in the small area estimation literature. There are point and interval estimation based on resampling techniques as well as the Taylor series approximation for small area problems.

In Chapter 3, we derive the MSE approximation based on the EBLUP and weighted estimators with fixed weights. We present the MSE estimation for different methods of variance component estimation. In this chapter we also propose area specific MSE estimators for the FH model.

Chapter 4 discusses the construction of CIs based on the EBLUP and the weighted estimator with fixed weights. We propose different CIs for a small area mean as well as for the difference of two small area means under the FH model.

Chapter 5 contains a simulation study to examine the finite sample performance and the accuracy of each MSE estimator of the EBLUP estimator. It also contains an extensive simulations study to evaluate the performance of the proposed CIs in comparison with the naive methods. It presents the discussion of the details of results for MSE and CIs.

Chapter 6 provides applications of the methodologies to the Ethiopian 2007 census and 2010/11 HCE data sets. To examine the performance of our method, and to compare it with others, we use an example of real data. We obtain poverty and percentage of food expenditure measures at the zone levels of Ethiopia with their model diagnostics, and discuss these.

Chapter 7 completes the thesis by summarizing the main findings of the thesis, giving some concluding remarks as well suggestions as to further research areas.

Chapter 2

Literature Review

This chapter reviews: variance component estimators, measures of uncertainty and interval estimation in small estimation.

2.1 Review of Literature Concerning Estimation of Variance Components

Variance components are not the main areas of interest in small area estimation. However, we need to estimate them accurately in order to get reliable small area means, which are the main interest (Rao, 2003a; Jiang and Lahiri, 2006). Variance components estimation is an important problem in the FH model. In practice, the variance components are usually unknown. Accurate estimation of variance components are necessary in order to obtain efficient point and interval estimates for the small area means θ_i. A number of variable component methods exist in the literature.

2.1.1 Estimation of Sampling Variance

To use the FH model, an estimate of the within-area sampling variance, ψ_i, must be available. In this section, we review some of the commonly used methods for estimating ψ_i in the literature.

Fay and Herriot (1979) introduced the Generalized Variance Function (GVF) method in order to estimate the sampling error variance, ψ_i. GVF is a method that explains the relationship between a statistic and its corresponding variance, and traditionally it has been used for variance estimates. Although methods based on GVFs have been used to obtain smoothed estimates of the ψ_i's, the assumption of known sampling variances, ψ_i, is restrictive (Rao, 2003b; Ha, Lahiri and Parsons, 2011).

Boonstra and Buelens (2011) proposed the pooled estimates of the sampling error variances. They determine the sampling variance for areas together, or even all the areas together. Their design variance, assuming an unrestricted random sample, is given by:

$$\psi_i = \frac{1}{n_i}(1 - \frac{n_i}{N_i})s_i^2$$

where the sample variance is $s_i^2 = \frac{1}{n_i - 1}\sum_{i \in s_i}(y_i - \bar{y})^2$, $\bar{y}$ is the sample average i. The pooled sample variance is, $s^2_{pooled} = \frac{1}{n-m}\sum_{i=1}^{m}(n_i - 1)s_i^2$, and we use this instead of the area specific variances s_i^2.

2.1.2 Estimation of Model Variance

Different model variance estimators are proposed in the literature. These are the Prasad-Rao (PR) moment estimator, FH moment estimator, ML and REML estimators. let us discuss each estimation methods as follows.

The Fay and Herriot (1979) moment estimator of A_{FH} is based on the WLS residual sum of squares. We can get A_{FH} estimator by simplifying the following equation iteratively:

$$\sum_{i=1}^{m} \frac{(y_i - \mathbf{x}'_i \tilde{\beta})^2}{\psi_i + A_{FH}} = m - p \tag{2.1.1}$$

where m is the number of small areas, p is the dimension of the vector of auxiliary variables x_i and

$$\tilde{\beta} = \left\{ \sum_{i=1} \frac{\mathbf{x_i}\mathbf{x_i'}}{\psi_i + A_{FH}} \right\}^{-1} \left\{ \sum_{i=1} \frac{\mathbf{x_i} y_i}{\psi_i + A_{FH}} \right\}. \tag{2.1.2}$$

The equation is iteratively solved subject to the condition $A_{FH} \geq 0$ (Rao, 2003a).

Prasad and Rao (1990) proposed a simple method of moments to estimate the model variance. The estimator is given by

$$A_{PR} = \frac{1}{m-p}[\mathbf{r}'\mathbf{r} - \sum \psi_i(1 - h_{ii})] \tag{2.1.3}$$

where the matrix $\mathbf{X}$ is full column rank, $\mathbf{r} = \mathbf{y} - \mathbf{X}\hat{\beta}$ is a vector of residuals, $h_{ii} = \mathbf{x}_i(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i$ diagonal entries of the hat matrix given by $\mathbf{H} = \mathbf{X}(\mathbf{X}\mathbf{X}')^{-1}\mathbf{X}'$ and

$$\hat{\beta} = (\mathbf{X}\mathbf{X}')^{-1}\mathbf{X}'\mathbf{y} \tag{2.1.4}$$

is the OLS estimator of β. On the other hand, $\hat{\beta} = \tilde{\beta}(\hat{A})$ is a WLS estimator of β and is given by

$$\tilde{\beta}(A) = \left[\sum_{i=1}^{m} \frac{\mathbf{x_i}\mathbf{x_i'}}{A + \psi_i} \right]^{-1} \left(\sum_{i=1}^{m} \frac{\mathbf{x}_i y_i}{A + \psi_i} \right). \tag{2.1.5}$$

If A_{PR} is negative, Prasad and Rao (1990) set it equal to zero. However, the probability of having a negative estimate of A_{PR} goes to zero very fast as $m \to \infty$.

The asymptotic variance of the ML estimator of A is given by $\hat{A}_{ML} = \max(0, A^*_{ML})$, where A^*_{ML} is gained by solving the following non-linear equation in A iteratively:

$$\sum_{i=1}^{m}(A+\psi_i)^{-1} = \sum_{i=1}^{m}(A+\psi_i)^{-2}\left\{y_i - \mathbf{x}_i'\tilde{\beta}(A)\right\}^2. \tag{2.1.6}$$

For REML estimation, we also solve the following non-linear equation in A to obtain the REML estimator, A^*_{REML}:

$$\sum_{i=1}^{m}\frac{\left\{y_i - \mathbf{x}_i'\tilde{\beta}(A)\right\}^2}{(A+\psi_i)^2} = \sum_{i=1}^{m}(A+\psi_i)^{-1} - \sum_{i=1}^{m}\frac{\mathbf{x}_i'\left\{\sum_{k=1}^{m}(A+\psi_k)^{-1}\mathbf{x}_k\mathbf{x}_k'\right\}^{-1}\mathbf{x}_i}{(A+\psi_i)^2}. \tag{2.1.7}$$

The REML estimator is given by $\hat{A}_{REML} = \max(0, A^*_{REML})$. For a given A, the ML estimator of β is obtained as

$$\tilde{\beta}(A) = \left\{\sum_{i=1}^{m}(A+\psi_i)^{-1}\mathbf{x}_i\mathbf{x}_i'\right\}^{-1}\sum_{i=1}^{m}(A+\psi_i)^{-1}\mathbf{x}_i\mathbf{y}_i. \tag{2.1.8}$$

The ML and REML methods are developed under the normality assumption. However, the normality assumption is likely to be violated in real life. For the normal error model, ML and REML methods are widely used to estimate the unknown variance components. But, the MLE works poorly when the number of small areas is limited (Li, 2006). Saei and Chambers (2003) designed a REML approach to reduce the bias of the ML estimate of the variance components.

2.1.3 Asymptotic Variance Estimators

Datta et al. (2005) noted that an EBLUP based on the ML/REML estimators of the model variance will have the smallest approximate MSE since the ML estimators have the smallest asymptotic variance.

Asymptotic variances of the ML and/or REML estimators are given by

$$V(\hat{A}_{ML}) = V(\hat{A}_{REML}) = 2\left[\sum_{i=1}^{m}\frac{1}{(A+\psi_i)^2}\right]^{-1} \tag{2.1.9}$$

where $\hat{A}_{ML}$ or $\hat{A}_{RE}$ are used to estimate A (Datta and Lahiri, 2000).

Prasad and Rao (1990) derived asymptotic variances of the PR moment estimator which is given by

$$V(\hat{A}_{PR}) = \frac{2}{m^2}\sum_{i=1}^{m}(A+\psi_i)^2. \tag{2.1.10}$$

Datta, Rao and Smith (2002) derived the asymptotic variance of the FH moment estimator which is given by

$$V(\hat{A}_{FH}) = 2m\left[\sum_{i=1}^{m}\frac{1}{A+\psi_i}\right]^{-2}. \qquad (2.1.11)$$

The next section reviews the literature concerning the MSE of a small area mean.

2.2 Review of Literature Concerning MSE

2.2.1 Approximation of the MSE of the EBLUP Estimator

The most common practical problem in small area estimation is measuring the variability associated with the EBLUP. MSE is used as a measure of variability under the EBLUP estimator in small area estimation. Methods of approximating and estimating $\mathrm{MSE}(\hat{\theta}_i^{EB})$ that include the uncertainty when estimating $\hat{\beta}$ and $\hat{A}$ have been studied extensively in the literature (e.g., Basu et al., 2003; Rao, 2003a; Hall and Maiti, 2006; Chatterjee et al., 2008; Li and Lahiri, 2010; Datta and Ghosh, 2012 and others). The MSE of $\hat{\theta}_i^{EB}$ defined as $\mathrm{MSE}(\hat{\theta}_i^{EB}) = E\left\{\hat{\theta}_i^{EB} - \theta_i\right\}^2$. It can be decomposed as:

$$\mathrm{MSE}(\hat{\theta}_i^{EB}) = E(\tilde{\theta}_i - \theta_i)^2 + E(\hat{\theta}_i^{EB} - \tilde{\theta}_i)^2 + 2E[(\tilde{\theta}_i - \theta_i)(\hat{\theta}_i^{EB} - \tilde{\theta}_i)]. \qquad (2.2.1)$$

The cross-product term in the above equation is zero for any even and translation invariant estimator $\hat{A}$ under the normality of random effects v_i and e_i (Kackar and Harville, 1981). Kackar and Harville (1984) also derived the approximations to the MSE of the EBLUP under the normality assumptions. Thus, the MSE of the EBLUP can be decomposed as:

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

Furthermore, Henderson (1975) showed that, $\mathrm{MSE}(\tilde{\theta}_i) = g_{1i}(A) + g_{2i}(A)$. Therefore, $\mathrm{MSE}(\hat{\theta}_i^{EB})$ can be expressed into three components as:

$$\mathrm{MSE}(\hat{\theta}_i^{EB}) = g_{1i}(A) + g_{2i}(A) + E\left\{\hat{\theta}_i^{EB} - \tilde{\theta}_i\right\}^2,$$

where

$g_{1i}(A) = A\psi_i(A+\psi_i)^{-1}$, is the contribution to MSE or error assuming all parameters are known.

$g_{2i}(A) = \frac{\psi_i^2}{(A+\psi_i)^2}\mathbf{x}_i'\left\{\sum_{u=1}^{m}(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$, ψ_u is the sampling variances for all areas u and ψ_i is the area specific sampling variance as defined in the above section. This component is the uncertainty (or excess MSE) due to estimation of β.

$g_{3i}(A) = E\left\{\hat{\theta}_i^{EB} - \tilde{\theta}_i\right\}^2$, is an additional uncertainty (or excess MSE) due to estimation of the variance component A. Generally, there is no closed-form expression available for this specific component of

MSE of $\hat{\theta}_i^{EB}$, which creates difficulties in obtaining an approximation of the MSE of EBLUP. There are several important analytical articles concerning methods which handle this problem for different specific models and variance components estimation methods (e.g., Kackar and Harville, 1984; Prasad and Rao, 1990; Datta and Lahiri, 2000; Datta et al., 2005; Kubokawa, 2010). Alternatively, resampling methods have also been considered to estimate the MSE of the EBLUP (e.g., Butar and Lahiri, 2003; Hall and Maiti, 2006; Li, 2006; Chatterjee et al., 2008; Li and Lahiri, 2010; Kubokawa and Nagashima, 2012).

Prasad and Rao (1990), extending the work of Kackar and Harvile (1984), approximate the true MSE of the EBLUP under the normality of the two error terms and for the case where A is estimated by the analysis of variance (ANOVA) method as,

$$\mathrm{MSE}(\hat{\theta}_i^{EB}) = \mathrm{E}\left\{\hat{\theta}_i^{EB} - \theta_i\right\}^2 = g_{1i}(A_{PR}) + g_{2i}(A_{PR}) + g_{3i}(A_{PR}) + o(m^{-1}). \tag{2.2.2}$$

The first two terms in the above equation remain the same for all methods for estimating A considered here, but the third term, approximated by $g_{3i}(A)$, depends on $V(\hat{A})$. Prasad and Rao (1990) obtained $g_{3i}(A_{PR})$ for the moment estimator $\hat{A}_{PR}$ as,

$$g_{3i}(A_{PR}) = 2\psi_i^2(\psi_i + A_{PR})^{-3} m^{-2} \sum_{u=1}^{m} (\psi_u + A_{PR})^2.$$

Following Prasad and Rao (1990); Datta and Lahiri (2000), obtained a second-order approximation to $g_{3i}(A)$ by ignoring all $o(m^{-1})$ terms for large m. Datta and Lahiri (2000) derived identical $g_{3i}(A)$ for the ML and REML estimators of A:

$$g_{3i}(A_{ML}) = 2\psi_i^2(\psi_i + A_{ML})^{-3} \left\{\sum_{u=1}^{m} (\psi_u + A_{ML})^{-2}\right\}^{-1}.$$

Datta et al. (2005) derived $g_{3i}(A)$ for the FH estimator of A:

$$g_{3i}(A_{FH}) = 2m\psi_i^2(\psi_i + A_{FH})^{-3} \left\{\sum_{u=1}^{m} (\psi_u + A_{FH})^{-1}\right\}^{-2}.$$

2.2.2 Approximation of the MSE of the Weighted Estimator with Fixed Weights

Datta et al. (2011) derived the MSE of $\hat{\theta}_i^{w}$ under the normality as shown below

$$\mathrm{MSE}(\hat{\theta}_i^{w}) = g_{1i}(A) + g_{2i}(A) + g_{3wi}(A), \tag{2.2.3}$$

where

$$g_{3wi}(A) = (\gamma_i(A) - w_i)^2 \left[(A + \psi_i) - \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i\right] + o(m^{-1}). \tag{2.2.4}$$

Their MSE of $\hat{\theta}_i^{w}$ can also be written as

$$\mathrm{MSE}(\hat{\theta}_i^{w}) = h_{1i}(A) + h_{2i}(A) + o(m^{-1}), \tag{2.2.5}$$

where $h_{1i}(A) = \psi_i(1 - w_i)^2 + Aw_i^2$ and $h_{2i}(A) = \left[g_{2i}(A) - (\gamma_i(A) - w_i)^2\,\mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i\right]$.

2.2.3 Estimation of the MSE of the EBLUP Estimator

This section reviews the second-order unbiased estimators of the MSE of the EBLUP $\hat{\theta}_i^{EB}$ for different methods of estimating the variance components. If $E[\widehat{MSE}(\hat{\theta}_i^{EB}) - MSE(\hat{\theta}_i^{EB})] = o(m^{-1})$, then $\widehat{MSE}(\hat{\theta}_i^{EB})$ is a second-order unbiased estimator of $MSE(\hat{\theta}_i^{EB})$. Ghosh and Rao (1994) and Rao (2003a) presented the theoretical results about estimating the MSE of EBLUP estimators.

If β and A are known, the variance of BLUP is given by:

$$V(\hat{\theta}_i^{EB}) = g_{1i}(A). \tag{2.2.6}$$

Under the EBLUP approaches, β and A are replaced by sample estimates and a naive variance estimator is obtained by replacing A by $\hat{A}$ in g_{1i} given by (Datta and Ghosh, 2012),

$$\widehat{MSE}_{N1}(\hat{\theta}_i^{EB}) = g_{1i}(\hat{A}) \tag{2.2.7}$$

$$= \frac{\hat{A}\psi_i}{\hat{A}+\psi_i}.$$

This estimator ignores the variability of $\hat{\beta}$ and $\hat{A}$ and hence under-estimates the true variance.

If β is unknown but A is known, another naive MSE estimator is obtained by estimating the MSE of the BLUP and is given by (Datta and Lahiri, 2000; Rao, 2003b):

$$\widehat{MSE}_{N2}(\hat{\theta}_i^{EB}) = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) \tag{2.2.8}$$

$$= \frac{\hat{A}\psi_i}{\hat{A}+\psi_i} + \frac{\psi_i^2}{(\hat{A}+\psi_i)^2}\mathbf{x}_i'\left\{\sum_{u=1}^m \frac{\mathbf{x}_u\mathbf{x}_u'}{\hat{A}+\psi_u}\right\}^{-1}\mathbf{x}_i.$$

This MSE estimator is likely to underestimate the true MSE since it does not incorporate the additional uncertainty due to the estimation of the model variance.

Prasad and Rao (1990) derived second-order corrected MSE estimator of the EBLUP under the normal models with asymptotic bias $b_{\hat{A}_{PR}} = 0$ for the ANOVA estimators $\hat{A}_{PR}$ of the variance components. It can be expressed by,

$$\widehat{MSE}_{PR}(\hat{\theta}_i^{EB}) = g_{1i}(\hat{A}_{PR}) + g_{2i}(\hat{A}_{PR}) + 2g_{3i}(\hat{A}_{PR}) \tag{2.2.9}$$

$$= \frac{\hat{A}_{PR}\psi_i}{\hat{A}_{PR}+\psi_i} + \frac{\psi_i^2}{(\hat{A}_{PR}+\psi_i)^2}\mathbf{x}_i'\left\{\sum_{u=1}^m \frac{\mathbf{x}_u\mathbf{x}_u'}{\hat{A}_{PR}+\psi_u}\right\}^{-1}\mathbf{x}_i$$

$$+ \frac{4}{m^2}\frac{\psi_i^2}{(\psi_i+\hat{A}_{PR})^3}\sum_{u=1}^m(\psi_u + \hat{A}_{PR})^2$$

is a second-order unbiased estimator of $MSE(\hat{\theta}_i^{EB})$.

Datta and Lahiri (2000) derived the second-order MSE approximations for the ML and REML estimators of the unknown parameters. Let the asymptotic bias of $\hat{A}$ up to the order of $o(m^{-1})$ be $b_{\hat{A}}$. Let also denote the partial derivatives of $g_{1i}(A)$ with respect to A by $\nabla g_{1i}(A)$. The MSE estimator of $\hat{\theta}_i^{EB}$ for ML estimator $\hat{A}_{ML}$ is given by

$$\widehat{MSE}_{ML}(\hat{\theta}_i^{EB}) = g_{1i}(\hat{A}_{ML}) + g_{2i}(\hat{A}_{ML}) + 2g_{3i}(\hat{A}_{ML}) - b_{\hat{A}_{ML}}(\hat{A}_{ML})\nabla g_{1i}(\hat{A}_{ML}) \tag{2.2.10}$$

$$= \frac{\hat{A}_{ML}\psi_i}{\hat{A}_{ML}+\psi_i} + \frac{\psi_i^2}{(\hat{A}_{ML}+\psi_i)^2}\mathbf{x}_i' \left\{\sum_{u=1}^m (\hat{A}_{ML}+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$$

$$+4\psi_i^2(\psi_i+\hat{A}_{ML})^{-3}\left\{\sum_{u=1}^m(\psi_u+\hat{A}_{ML})^{-2}\right\}^{-1}$$

$$-\frac{-\mathrm{tr}\left\{[\sum_{u=1}^m(\hat{A}+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u']^{-1}[\sum_{u=1}^m(\hat{A}_{ML}+\psi_u)^{-2}\mathbf{x}_u\mathbf{x}_u']\right\}}{\sum_{u=1}^m(\hat{A}_{ML}+\psi_u)^{-2}}\left(\frac{\psi_i}{\hat{A}_{ML}+\psi_i}\right)^2$$

where $\nabla g_{1i}(\hat{A}_{ML}) = (\frac{\psi_i}{\hat{A}_{ML}+\psi_i})^2 > 0$. For the ML estimator $\hat{A}_{ML}$, the estimated bias of $\hat{A}_{ML}$ is given by

$$b_{\hat{A}_{ML}}(\hat{A}_{ML}) = \frac{-\mathrm{tr}\left\{[\sum_{u=1}^m(\hat{A}+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u']^{-1}[\sum_{u=1}^m(\hat{A}_{ML}+\psi_u)^{-2}\mathbf{x}_u\mathbf{x}_u']\right\}}{\sum_{u=1}^m(\hat{A}_{ML}+\psi_u)^{-2}}.$$

The bias of the REML estimator $\hat{A}_{RE}$, $b_{\hat{A}}(\hat{A}_{RE}) \doteq 0$, if terms of order $o(m^{-1})$ is ignored. Therefore, the MSE estimator of $\hat{\theta}_i^{EB}$ is given by

$$\widehat{MSE}_{RE}(\hat{\theta}_i^{EB}) = g_{1i}(\hat{A}_{RE}) + g_{2i}(\hat{A}_{RE}) + 2g_{3i}(\hat{A}_{RE}), \tag{2.2.11}$$

$$= \frac{\hat{A}_{RE}\psi_i}{\hat{A}_{RE}+\psi_i} + \frac{\psi_i^2}{(\hat{A}_{RE}+\psi_i)^2}\mathbf{x}_i'\left\{\sum_{u=1}^m(\hat{A}_{RE}+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$$

$$+\frac{4\psi_i^2}{(\psi_i+\hat{A}_{RE})^3}\left\{\sum_{u=1}^m(\psi_u+\hat{A}_{RE})^{-2}\right\}^{-1}.$$

Datta et al. (2005) derived the MSE estimator of $\hat{\theta}_i^{EB}$, for the FH estimator $\hat{A}_{FH}$. Thus, using the FH estimator $\hat{A}_{FH}$, the MSE estimator of $\hat{\theta}_i^{EB}$ is given by

$$\widehat{MSE}_{FH}(\hat{\theta}_i^{EB}) = g_{1i}(\hat{A}_{FH}) + g_{2i}(\hat{A}_{FH}) + 2g_{3i}(\hat{A}_{FH}) - b_{\hat{A}_{FH}}(\hat{A}_{FH})\left\{\gamma_i(\hat{A}_{FH})\right\}^2 \tag{2.2.12}$$

$$= \frac{\hat{A}_{FH}\psi_i}{\hat{A}_{FH}+\psi_i} + \frac{\psi_i^2}{(\hat{A}_{FH}+\psi_i)^2}\mathbf{x}_i'\left\{\sum_{u=1}^m(\hat{A}_{FH}+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$$

$$+\frac{4m\psi_i^2}{(\psi_i+\hat{A}_{FH})^3}\left\{\sum_{u=1}^m(\psi_u+\hat{A}_{FH})^{-1}\right\}^{-2}$$

$$-\frac{2\left\{m\sum_{u=1}^m(\hat{A}_{FH}+\psi_u)^{-2}-\left\{\sum_{u=1}^m(\hat{A}_{FH}+\psi_u)^{-1}\right\}^2\right\}}{\sum_{u=1}^m(\hat{A}_{FH}+\psi_u)^{-2}}\left\{\frac{\psi_i}{\hat{A}_{FH}+\psi_i}\right\}^2$$

where $\gamma_i(\hat{A}_{FH}) = \frac{\psi_i}{\hat{A}_{FH}+\psi_i}$. The bias to terms of order $O(m^{-1})$ for the FH model is given by

$$b_{\hat{A}_{FH}}(\hat{A}_{FH}) = \frac{2\left\{m\sum_{u=1}^m(\hat{A}_{FH}+\psi_u)^{-2}-\left\{\sum_{u=1}^m(\hat{A}_{FH}+\psi_u)^{-1}\right\}^2\right\}}{\sum_{u=1}^m(\hat{A}_{FH}+\psi_u)^{-2}}.$$

The sampling variance, ψ_i is unknown and thus has to be estimated along the regression coefficient and the variance of the random effect (Rivest and Vandal, 2003). They adjusted the Prasad and Rao (1990)

MSE estimator for the estimation of ψ_i. It is assumed that s_i^2, the estimator of ψ_i, is a statistic with a $N(\psi_i, \phi_i)$ distribution. The adjustment is done by adding the term $\frac{2\phi_i \hat{A}^2}{s_i^2+\hat{A}}$, thus giving the generalized PR estimator:

$$\widehat{\text{MSE}}_{\text{PRg}}(\hat{\theta}_i^{\text{EB}}) = \frac{\hat{A}_{\text{PR}} s_i^2}{\hat{A}_{\text{PR}} + s_i^2} + \frac{s_i^4 \mathbf{x}_i' \hat{F}^{-1} \mathbf{x}_i}{(\hat{A}_{\text{PR}} + s_i^2)^2} + 2\frac{\widehat{\text{Var}}(\hat{A}_{\text{PR}}) + \hat{A}_{\text{PR}}^2 \hat{\phi}_i}{(\hat{A}_{\text{PR}} + s_i^2)^3} \tag{2.2.13}$$

where $\hat{F} = \frac{\sum \mathbf{x}_i \mathbf{x}_i'}{\hat{A}+s_i^2}$ and $\widehat{V}(\hat{A}_{\text{PR}}) = \frac{2}{m^2}\sum_{i=1}(A+\psi_i)^2$, $\phi_i = \frac{2\psi_i^2}{n_i-1}$.

2.2.4 Estimation of the MSE of the Weighted Estimator with Fixed Weights

A nearly unbiased estimator of $\text{MSE}(\hat{\theta}_i^{\text{w}})$ using REML estimator $\hat{A}_{\text{REML}}$ is given by (Datta et al., 2011)

$$\begin{aligned}\widehat{\text{MSE}}_{\text{RE}}(\hat{\theta}_i^{\text{w}}) &= h_{1i}(\hat{A}_{\text{RE}}) + h_{2i}(\hat{A}_{\text{RE}}) \\ &= \psi_i(1-w_i)^2 + A_{\text{RE}} w_i^2 \\ &\quad + \frac{\psi_i^2}{(\hat{A}_{\text{RE}}+\psi_i)^2}\mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m (\hat{A}_{\text{RE}}+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i \\ &\quad -(\hat{\gamma}_i - w_i)^2\,\mathbf{x}_i'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{x}_i.\end{aligned} \tag{2.2.14}$$

Similarly for the FH moment estimator $\hat{A}_{\text{FH}}$ a nearly unbiased estimator of $\text{MSE}(\hat{\theta}_i^{\text{w}})$ is given by

$$\begin{aligned}\widehat{\text{MSE}}_{\text{FH}}(\hat{\theta}_i^{\text{w}}) &= h_{1i}(\hat{A}_{\text{FH}}) + h_{2i}(\hat{A}_{\text{FH}}) - b_{\hat{A}_{\text{FH}}}(\hat{A}_{\text{FH}}) w_i^2 \\ &= \psi_i(1-w_i)^2 + A_{\text{FH}} w_i^2 - (\hat{\gamma}_i - w_i)^2\,\mathbf{x}_i'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{x}_i \\ &\quad + \frac{\psi_i^2}{(\hat{A}_{\text{FH}}+\psi_i)^2}\mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m (\hat{A}_{\text{FH}}+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i \\ &\quad - \frac{2\left\{m\sum_{u=1}^m(\hat{A}_{\text{FH}}+\psi_u)^{-2} - \left\{\sum_{u=1}^m(\hat{A}_{\text{FH}}+\psi_u)^{-1}\right\}^2\right\}}{\sum_{u=1}^m(\hat{A}_{\text{FH}}+\psi_u)^{-2}} w_i^2.\end{aligned} \tag{2.2.15}$$

2.2.5 MSE Estimation of the Non-sampled Areas

The MSE for the non-sampled areas are given by

$$\begin{aligned}\text{MSE}(\hat{\theta}_k^{\text{Syn}}) &= E(\mathbf{x}_k'\hat{\beta} - \theta_k)^2 \\ &\approx A\mathbf{x}_k'\left\{\textstyle\sum_{i=1}^m (A+\psi_i)^{-1}\mathbf{x}_i\mathbf{x}_i'\right\}^{-1}\mathbf{x}_k + A.\end{aligned}$$

A nearly unbiased MSE estimator for the non-sampled areas are given by

$$\widehat{\text{MSE}}(\hat{\theta}_k^{\text{Syn}}) = \hat{A}\mathbf{x}_k'\left\{\sum_{i=1}^m (\hat{A}+\psi_i)^{-1}\mathbf{x}_i\mathbf{x}_i'\right\}^{-1}\mathbf{x}_k + \hat{A} \tag{2.2.16}$$

where i stands for the sampled areas whereas k stands for the non-sampled areas.

The next section reviews CIs in small area estimation. It presents a review of CIs through the Taylor series expansion method and resampling methods.

2.3 Review of Literature Concerning Confidence Interval Estimation of a Small Area Mean

In this section we give a detailed review of CI estimation for small area mean in small area estimation.

The Naive (conventional) CI based on the sample mean y_i is given by:

$$CI_i : y_i \pm z_{\alpha/2}\sqrt{\psi_i},$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percent point of $N(0,1)$. The CP of this Naive CI is identical to the confidence coefficient $1-\alpha$. However, the width of the CI of CI_i is large for small sample size, since y_i has large standard error (Kubokawa, 2010).

Cox (1975) suggests generating the following EB interval:

$$CI_i : (1-\hat{\gamma}_i)Y_i + \hat{\gamma}_i \mathbf{x}_i'\hat{\beta} \pm z_{\alpha/2}\psi^{1/2}(1-\hat{\gamma}_i)^{1/2}, \tag{2.3.1}$$

where $\hat{\gamma}_i$ and $\hat{\beta}$ are estimators of $\gamma_i = \frac{\psi_i}{\psi_i + A}$ and β, respectively. Under standard regularity conditions,

$$P(\theta_i \in CI_i) = 1-\alpha + O(m^{-1}).$$

This CI attains its desired CP asymptotically with margin of error is of the order $O(m^{-1})$.

Morris (1983a) was the first to consider CIs for small area means for the FH model with covariates in the small area estimation context. Morris (1983a) developed the CP for θ_i with a target CP of at least $1-\alpha$ given by

$$\hat{\theta}_i^{EB} \pm z s_i$$

where z is the upper $\alpha/_2$ point of a standard normal variate; $s_i^2 = \psi(1-(m-r)/m)\hat{\gamma} + v(y_i - z_i'b)^2$; $v = \frac{2}{m-r-2}\hat{\gamma}^2$. The Morris (1983a) interval can also be rewritten as:

$$P(\hat{\theta}_i^{EB} - z s_i \leq \theta_i \leq \hat{\theta}_i^{EB} + z s_i) = 1-\alpha - 2m^{-1} z\phi(z) h_{im}(z;\rho) + o(m^{-1}),$$

when $r = 1$ and $\mathbf{z}_m$ equals the $m \times 1$ vector of 1's, $\rho = \gamma/_{(1-\gamma)}$, $\hat{\rho} = \hat{\gamma}/_{(1-\hat{\gamma})}$ and $h_{im}(z;\rho) = \frac{1}{4}(z^2+1)\rho^2$.

Based on the simulation study conducted by Morris (1983a), the average CP was 0.97 compared to a nominal 0.95 for 100 simulation runs. The AL of the EB CIs was 35% smaller than the CIs based on the direct estimates. He also noted that the CP tends to $1-\alpha$ as the number of areas tends to ∞ or the sampling variance tends to zero.

2.3.1 Confidence Intervals Based on the Taylor Series Expansion Methods

As we have mentioned in Chapter 1, CIs have been addressed by using the Taylor series expansion method (e.g., Prasad and Rao, 1990; Datta et al., 2002; Basu et al., 2003; among others). In this section, we review the CIs obtained using the Taylor series expansion techniques.

Basu et al. (2003) proposed two EB prediction intervals under a normal regression model. The CPs and ALs of their CIs are discussed via appropriate higher order asymptotics using the Taylor series expansion. Basu et al. (2003) discussed that the EB CI developed by Morris (1983a) does not increase the CP in the sense that the coverage of order remains $O(m^{-1})$. They also studied the CI developed by Carlin and Louis (1996), and discussed that their CIs has a CP of order $O(m^{-1})$. Their modification and analytic implementation achieve a margin of error $O(m^{-1})$.

They made a simple modification of Morris (1983a) interval that attains a CP $1-\alpha$ with margin of error $O(m)^{-1}$. This is given by:

$$I = \left[\hat{\theta}_i^{EB} - z\left\{1 + m^{-1}h_{im}(z;\hat{\rho})\right\} s_i,\ \hat{\theta}_i^{EB} + z\left\{1 + m^{-1}h_{im}(z;\hat{\rho})\right\} s_i\right]$$

where $h_{im}(z;\rho) = \frac{1}{2}(c_{im} - r)\rho + \frac{1}{4}(z^2+1)\rho^2$.

They obtained a clear expression for another prediction interval for θ_i with the same asymptotic CP based on the idea of Carlin and Louis (1996). Their approach revealed a CI for θ_i of the form

$$\left[\theta_i^* - z\left\{1 + m^{-1}h_{ik}^*(z;\rho^*)\right\}\{\psi(1-B^*)\}^{1/2},\ \theta_i^* + z\left\{1 + m^{-1}h_{im}^*(z;\rho^*)\right\}\{\psi(1-B^*)\}^{1/2}\right],$$

where $\rho^* = \frac{B^*}{1-B^*}$ if $B^* < 1$ and $B^* = \frac{\psi}{\max(\psi, \frac{m-r}{m}s^2)}$ is the ML estimator of B. Both of these two intervals attain the target CP with margin of error $o(m^{-1})$.

Datta et al. (2002) proposed CIs for the balanced FH model in order to attain asymptotic coverage correct up to $O(m^{-1})$. They studied conditional CPs of such intervals for unknown A by conditioning on a suitable ancillary statistic. They also derived the EB CIs based on unconditional CPs. They conducted a simulation study to assess the performance of the approximate CIs. They used a simple simulation design with $m = 30$ small areas with no covariates. They found little qualitative difference between the conditional and unconditional CPs.

Diao et al. (2013) developed closed form CIs for the small area and for the difference of two small area means with unequal sampling variances which are second order correct up to $O(m^{-2})$ using the Taylor series approximation. This is the procedure adopted in our CIs derivation procedure. They conducted a simulation study to investigate the finite sample performance of their CIs. They adopted the simulation set-up of Datta et al. (2005). For all variance patterns, their methods results in CP larger than the naive method.

2.3.2 Confidence Intervals Based on Resampling Approaches

In this section, we review CIs for a small area mean based on the parametric bootstrap method. Recently researchers applied this method to the construction of CIs.

Chatterjee and Lahiri (2002) used the parametric bootstrap method to construct CIs for a small area mean in the FH model which is accurate for CP up to order of $O(m^{-3/2})$. They examined the performance of their proposed parametric bootstrap method for small $m = 10$, arbitrarily through the simulation. They found that the naive method has severe under-coverage problem.

Chatterjee et al. (2008) proposed the parametric bootstrap prediction interval for the FH and general linear mixed models. They conducted a simulation study to compare the performance of their methods with the naive methods. They considered the FH model with $m = 15$ and $\mathbf{x}_i'\beta = 0$ and five groups of small areas with three areas in each group under two different sampling variance patterns. The performance of their parametric bootstrap prediction interval performs better than the Prasad and Rao (1990) CIs.

Li (2006) provided the prediction interval under a mixed regression model. He employed a similar method to Chatterjee et al. (2008) for the general linear mixed model. They argued that the AL of the synthetic parametric bootstrap prediction interval is always larger than the one based on the conditional distribution of θ_i given Y which is derived in Chatterjee et al. (2008). Li (2006) proposed a CI based on the parametric bootstrap prediction interval which is correct up to $O(m^{-3/2})$. They conducted a simulation study to compare the performance of their prediction intervals under the FH model.

Li and Lahiri (2010) proposed an adjusted ML estimator of the variance of the random effect that maximizes an adjusted likelihood function. An adjusted likelihood function is defined as a product of the variance of the random effect, A and a standard likelihood function. They extended the findings of Chatterjee et al. (2008) to the case where β is estimated by the WLS method and the variance of the random effect A by the adjusted ML method. They constructed the parametric bootstrap prediction intervals for the small area means under the FH model. Their method required the use of a strictly positive consistent estimate of the variance of the random effect A. They compared CPs and ALs of seven different CIs of θ_i.

Kubokawa and Nagashima (2012) developed CIs based on the parametric bootstrap method. For the simulation purposes, they have applied the FH model with $m = 15, A = 1$ and two ψ_i patterns: (a) 0.7, 0.6, 0.5, 0.4, 0.3; (b) 4.0, 0.6, 0.5, 0.4, 0.1. They considered the case when $\mathbf{x}_i'\beta = 0$. In their particular set-up the number of simulation runs is 10,000 and the size of the bootstrap sample is 1000. The simulation results showed that their proposed CI performs well for homogeneous pattern, but for the heterogeneous pattern, their CI CPs are much smaller than the 95% nominal value.

Chapter 3

BLUP and Assessing Uncertainty of Point Estimators

This chapter presents the measures of uncertainty of both EBLUP and weighted estimators with fixed weights. We also present some of the theorems that have been proved by different experts. We present the proofs of these for the sake of clarity and simplicity.

3.1 BLUP Estimator

Ghosh and Rao (1994) gave a comprehensive review of the model based procedures for estimating small area parameters. The BLUP estimator was developed by Henderson (1953) for linear mixed models. This estimator does not depend on the normality assumptions of the random effects. However, it depends on the variance of the random effects (Rao, 2003a). The following theorem states and proves the two common properties of the BLUP estimator. The proof of Theorem 3.1.1 is provided for the sake of completeness. For example, Datta and Ghosh (2012) proved the following theorems.

Theorem 3.1.1. The BLUP of θ_i and some of its properties are given by the following theorem stated as follows:

(a) The BLUP of $\theta_i = \mathbf{x_i'}\beta + v_i$ is given by

$$\tilde{\theta}_i^B = (1-\gamma_i)y_i + \gamma_i \mathbf{x_i'}\tilde{\beta}$$

where $\gamma_i = \frac{\psi_i}{A+\psi_i}$ and $\tilde{\beta} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y}$ with $\mathbf{X} = (\mathbf{x}_i')_{1\le i\le m}$, $\mathbf{y} = (y_i')_{1\le i\le m}$, $\Sigma = A\mathbf{I}_m + \Psi$ and

$\Psi = \text{diag}(\psi_i, 1 \le i \le m)$. $\mathbf{I}_m$ is the identity matrix.

(b) The BLUP of θ which is $\tilde{\theta}^B$ is unbiased in the sense that $E[\tilde{\theta}^B - \theta] = 0$ under the FH model.

Proof

a) This part of the theorem is proved by Kurnia and Notodiputro (2006) under the Bayesian consideration.

Consider the following expressions:

$f(y_i|\theta_i) = \frac{1}{\sqrt{2\pi\psi_i}} \exp\left(-\frac{1}{2\psi_i}(y_i - \theta_i)^2\right)$ and

$\pi(\theta_i) = \frac{1}{\sqrt{2\pi A}} \exp\left(-\frac{1}{2A}(\theta_i - \mathbf{x}_i'\beta)^2\right)$

$f(\mathbf{y}, \theta|\beta, A) = \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\psi_i}} \exp\left(-\frac{1}{2\psi_i}(y_i - \theta_i)^2\right) \frac{1}{\sqrt{2\pi A}} \exp\left(-\frac{1}{2A}(\theta_i - \mathbf{x}_i'\beta)^2\right)$

for $\mathbf{y} = (y_1, ..., y_m)'$, $\theta = (\theta_1, ..., \theta_m)'$.

Looking at the two exponential functions without the $(-1/_2)$ factor from $f(\mathbf{y}, \theta|\beta, A)$,

$$\frac{1}{\psi_i}(y_i - \theta_i)^2 + \frac{1}{A}(\theta_i - \mathbf{x}_i'\beta)^2$$

$$= \frac{1}{\psi_i}(y_i^2 - 2y_i\theta_i + \theta_i^2) + \frac{1}{A}(\theta_i^2 - 2\mathbf{x}_i'\beta\theta_i + (\mathbf{x}_i'\beta)^2)$$

$$= \left(\frac{1}{\psi_i} + \frac{1}{A}\right)\theta_i^2 - 2\left(\frac{y_i}{\psi_i} + \frac{\mathbf{x}_i'\beta}{A}\right)\theta_i + c_i$$

$$= \left(\frac{1}{\psi_i} + \frac{1}{A}\right)\left\{\theta_i^2 - 2\theta_i\left(\frac{y_i}{\psi_i} + \frac{\mathbf{x}_i'\beta}{A}\right)/_{\left(\frac{1}{\psi_i}+\frac{1}{A}\right)} + \left(\frac{y_i}{\psi_i} + \frac{\mathbf{x}_i'\beta}{A}\right)/_{\left(\frac{1}{\psi_i}+\frac{1}{A}\right)} \times \left(\frac{y_i}{\psi_i} + \frac{\mathbf{x}_i'\beta}{A}\right)/_{\left(\frac{1}{\psi_i}+\frac{1}{A}\right)}\right\} + c_i'$$

$$= \left(\frac{1}{\psi_i} + \frac{1}{A}\right)\left\{\theta_i - \left(\frac{y_i}{\psi_i} + \frac{\mathbf{x}_i'\beta}{A}\right)/_{\left(\frac{1}{\psi_i}+\frac{1}{A}\right)}\right\}^2 + c_i'$$

where c_i and c_i' are constants and independent from θ_i. So,

$$(\theta_i|y_i, \beta, A) \sim N\left(\left(\frac{y_i}{\psi_i} + \frac{\mathbf{x}_i'\beta}{A}\right)/_{\left(\frac{1}{\psi_i}+\frac{1}{A}\right)}, \left(\frac{1}{\psi_i} + \frac{1}{A}\right)^{-1}\right)$$

$$= N\left(\mathbf{x}_i'\beta + \frac{A}{A+\psi_i}(y_i - \mathbf{x}_i'\beta), \frac{A\psi_i}{A+\psi_i}\right).$$

Thus, for the FH model, the BLUP of $\theta_i = \mathbf{x}_i'\beta + v_i$ if A and β known is given by

$\tilde{\theta}_i^B = \tilde{\theta}_i(y_i|\beta, A) = (1-\gamma_i)y_i + \gamma_i\mathbf{x}_i'\tilde{\beta}$, where $\gamma_i = \frac{\psi_i}{A+\psi_i}$, which is the same as EB or EBLUP.

Proof

b) For the given

$\theta = \mathbf{X}\beta + \mathbf{v}$, $\mathbf{y} = \mathbf{X}\beta + \mathbf{v} + \mathbf{e}$, $\tilde{\beta}_{GLS} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y}$ and $\tilde{\theta}^B = (1-\gamma)\mathbf{y} + \gamma\mathbf{X}\tilde{\beta}_{GLS}$

$$E[\tilde{\theta}^B - \theta] = E[(1-\gamma)\mathbf{y} + \gamma\mathbf{X}\tilde{\beta}_{GLS} - \mathbf{X}\beta - \mathbf{v}]$$

$$= E[(1-\gamma)(\mathbf{X}\beta + \mathbf{v} + \mathbf{e}) + \gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y} - \mathbf{X}\beta - \mathbf{v}]$$

$$= E[(1-\gamma)(\mathbf{X}\beta + \mathbf{v} + \mathbf{e}) + \gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}(\mathbf{X}\beta + \mathbf{v} + \mathbf{e}) - \mathbf{X}\beta - \mathbf{v}]$$

$=\mathrm{E}[\mathbf{X}\beta+\mathbf{v}+\mathbf{e}-\gamma\mathbf{X}\beta-\gamma\mathbf{v}-\gamma\mathbf{e}+\gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{X}\beta+\gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{v}+\gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{e}-$
$\mathbf{X}\beta-\mathbf{v}]$

$=\mathrm{E}[\mathbf{v}+\mathbf{e}-\gamma\mathbf{X}\beta-\gamma\mathbf{v}-\gamma\mathbf{e}+\gamma\mathbf{X}\beta+\gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{v}+\gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{e}-\mathbf{v}]$

$=\mathrm{E}[\mathbf{e}-\gamma\mathbf{v}-\gamma\mathbf{e}+\gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{v}+\gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{e}]$

$=\mathrm{E}(\mathbf{e})-\gamma\mathrm{E}(\mathbf{v})-\gamma\mathrm{E}(\mathbf{e})+\gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathrm{E}(\mathbf{v})+\gamma\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathrm{E}(\mathbf{e})$

$=0.$

Thus, $\tilde{\theta}^{\mathrm{B}}$ is an unbiased estimator of θ.

In addition, the BLUP is best in the sense that its MSE, defined as the mean squared difference between the BLUP and its mean, $\mathrm{MSE}(\tilde{\theta}_i^{\mathrm{EB}}) = \mathrm{E}[\tilde{\theta}_i^{\mathrm{EB}}-\theta_i]^2$ is the minimum among all other linear unbiased predictors (Searle, Casella and McCulloch, 1992).

3.2 MSE Estimator

We reviewed methods of approximating and estimating $\mathrm{MSE}(\hat{\theta}_i^{\mathrm{EB}})$ that account for the variability of $\hat{\beta}$ and $\hat{\mathrm{A}}$ in Chapter 2. The following regularity conditions are used for approximating MSE estimators (Theorem 3.2.1 and Theorem 3.2.3) and deriving confidence intervals (Theorem 4.1.1, Theorem 4.1.2, Theorem 4.2.2 and Theorem 4.2.3).

Assumption 1: The quantities $\mathbf{x}_i$, A, ψ_i are bounded.

Assumption 2: v_i and e_i are mutually independent with $v_i \overset{\mathrm{ind}}{\sim} \mathrm{N}[0,\mathrm{A}]$ and $e_i \overset{\mathrm{ind}}{\sim} \mathrm{N}[0,\psi_i]$, $i=1,...,m$.

Assumption 3: $\hat{\mathrm{A}}=\hat{\mathrm{A}}(\mathbf{y})$ is an estimator of A which satisfies

(i) they are even functions of $\mathbf{y}$, so that $\hat{\mathrm{A}}(-\mathbf{y})=\hat{\mathrm{A}}(\mathbf{y})$.

(ii) they are translation invariant functions, so that $\hat{\mathrm{A}}(\mathbf{y}+\mathbf{X}d)=\hat{\mathrm{A}}(\mathbf{y})$

for any $d \in R^p$ and for all $\mathbf{y}$.

Assumption 4: $\max_{1\le i\le m}\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i \to 0$ as $m\to\infty$, or $\max_{1\le i\le m}\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i = O(m^{-1})$ (see Prasad and Rao, 1990) where $\mathbf{x}_i$ is the i^{th} row of $\mathbf{X}$. All conditions are very often assumed in the small area estimation literature. In this section, we derive the the second order MSE approximation using the Stein identity. Stein (1981) showed that

$$\mathrm{E}\left[h_i(\mathbf{y})(y_i-\mathbf{x}_i'\beta)\right]=(\psi_i+\mathrm{A})\mathrm{E}\left[\frac{\partial h_i(\mathbf{y})}{\partial y_i}\right] \qquad (3.2.1)$$

where $h_i(\mathbf{y})$ is an absolutely continuous function. This identity is called the Stein identity (Kubokawa, 2010, Datta et al., 2011).

3.2.1 Approximation of $\text{MSE}(\hat{\theta}_i^{EB})$

Theorem 3.2.1. An asymptotic expression of the MSE which is accurate to the order $O(m^{-1})$ under the assumptions 2, 3 and 4 is given by

$\text{MSE}(\hat{\theta}_i^{EB}) = g_{1i}(A) + g_{2i}(A) + g_{3i}(A) + O(m^{-1})$

where, $g_{1i}(A) = \frac{A\psi_i}{A+\psi_i}$, $g_{2i}(A) = \gamma_i^2 \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i$ and $g_{3i}(A) = \frac{\psi_i^2}{(A+\psi_i)^3} V(\hat{A}) + o(m^{-1})$.

Proof. Note that this theorem has been proved by Prasad and Rao (1990), Datta and Lahiri (2000) and others for different methods of estimation. However in this thesis, we include the proof using the approximation based on an orthogonal (non correlated) decomposition of the MSE $E(\hat{\theta}_i^{EB} - \theta_i)^2$ and the Stein identity given by equation (3.2.1) for the sake of completeness. Specifically, we use the decomposition

$\hat{\theta}_i^{EB} - \theta_i = \left\{(1-\gamma_i)y_i + \gamma_i \mathbf{x}_i'\beta - \theta_i\right\} + \left\{\hat{\theta}_i^{B} - (1-\gamma_i)y_i - \gamma_i \mathbf{x}_i'\beta\right\} + \left\{\hat{\theta}_i^{EB} - \hat{\theta}_i^{B}\right\}$

(see Datta and Ghosh, 2012).

If we let the components be defined as follows

$a = \left\{(1-\gamma_i)y_i + \gamma_i \mathbf{x}_i'\beta - \theta_i\right\}$, $b = \left\{\hat{\theta}_i^{B} - (1-\gamma_i)y_i - \gamma_i \mathbf{x}_i'\beta\right\}$

and $c = \left\{\hat{\theta}_i^{EB} - \hat{\theta}_i^{B}\right\}$.

Then,

$(a+b+c)^2 = a^2 + b^2 + c^2 + 2ab + 2ac + 2bc$. Since the first component is always orthogonal to the second and the third components, $2E(ab) = 2E(ac) = 0$. The orthogonality of the last two components holds only for certain specific estimators of A under assumption 3. Thus, under this assumption, $2E(bc) = 0$ (Datta and Ghosh, 2012).

We now evaluate the three components of

$$E\left\{\hat{\theta}_i^{EB} - \theta_i\right\}^2 = E\left\{(1-\gamma_i)y_i + \gamma_i \mathbf{x}_i'\beta - \theta_i\right\}^2 + E\left\{\hat{\theta}_i^{B} - (1-\gamma_i)y_i - \gamma_i \mathbf{x}_i'\beta\right\}^2 + E\left\{\hat{\theta}_i^{EB} - \hat{\theta}_i^{B}\right\}^2.$$

Simple calculation of the first component shows that (see Jiang and Tang, 2011)

$E\left\{(1-\gamma_i)y_i + \gamma_i \mathbf{x}_i'\beta - \theta_i\right\}^2 = (1-\gamma_i)\psi_i = \frac{A\psi_i}{A+\psi_i}$.

Since $\gamma_i = \frac{\psi_i}{\psi_i + A}$,

$(1-\gamma_i)\psi_i = (1 - \frac{\psi_i}{\psi_i+A})\psi_i = \frac{A\psi_i}{A+\psi_i} = g_{1i}(A)$; is the contribution to MSE or error assuming all parameters are known.

Derivation of the second component is given by

$E\left\{\hat{\theta}_i^{B} - (1-\gamma_i)y_i - \gamma_i \mathbf{x}_i'\beta\right\}^2$.

Substituting $\hat{\theta}_i^{B}$ reveals that

$$E\left\{\hat{\theta}_i^{B} - (1-\gamma_i)y_i - \gamma_i \mathbf{x}_i'\beta\right\}^2 = E\left\{(1-\gamma_i)y_i + B_i \mathbf{x}_i'\hat{\beta}_{GLS} - (1-\gamma_i)y_i - \gamma_i \mathbf{x}_i'\beta\right\}^2$$

$$= E\left\{\gamma_i \mathbf{x}_i'\hat{\beta}_{GLS} - \gamma_i \mathbf{x}_i'\beta\right\}^2$$

$$=\mathrm{E}\left\{\gamma_i \mathbf{x}_i'(\hat{\beta}_{GLS}-\beta)\right\}^2$$

$$=\gamma_i^2 \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i = g_{2i}(A),$$ is the uncertainty (or excess MSE) due to estimation of β.

Let the third component be denoted by $g_{3i}(A) = \mathrm{E}\left\{\hat{\theta}_i^{EB} - \hat{\theta}_i^{B}\right\}^2$. This is the excess of MSE due to estimation of the variance component A.

Using the Taylor series expansion, following Kubokawa (2010), we have

$$\hat{\theta}_i^{EB} = \hat{\theta}_i^{B} + \frac{\partial \hat{\theta}_i^{B}}{\partial A}(\hat{A}-A)$$

$$\hat{\theta}_i^{EB} - \hat{\theta}_i^{B} = \frac{\partial \hat{\theta}_i^{B}}{\partial A}(\hat{A}-A)$$

$$\left(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B}\right)^2 = \left(\frac{\partial \hat{\theta}_i^{B}}{\partial A}\right)^2(\hat{A}-A)^2$$

$$\mathrm{E}\left(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B}\right)^2 = \mathrm{E}\left[\left(\frac{\partial \hat{\theta}_i^{B}}{\partial A}\right)^2(\hat{A}-A)^2\right]$$

$\frac{\partial \hat{\theta}_i^{B}}{\partial A} = \frac{\psi_i}{(A+\psi_i)^2}(y_i - \mathbf{x}_i'\tilde{\beta}(A)) + \gamma_i \mathbf{x}_i'\frac{\partial \tilde{\beta}(A)}{\partial A}$, since $\frac{\partial \tilde{\beta}(A)}{\partial A} = O_P(m^{-1/2})$ (see Datta et al., 2011).

$$\frac{\partial \hat{\theta}_i^{B}}{\partial A} = \frac{\psi_i}{(A+\psi_i)^2}(y_i - \mathbf{x}_i'\beta) + o(m^{-1}).$$

Substituting reveals that

$$\mathrm{E}\left(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B}\right)^2 = \mathrm{E}\left[\left(\frac{\partial \hat{\theta}_i^{B}}{\partial A}\right)^2(\hat{A}-A)^2\right]$$

$$= \frac{\psi_i^2}{(A+\psi_i)^4}\mathrm{E}\left[(y_i - \mathbf{x}_i'\beta)^2(\hat{A}-A)^2\right] + o(m^{-1}).$$

Using the Stein identity given by equation (3.2.1)

$$\mathrm{E}\left(\hat{\theta}_i^{EB} - \hat{\theta}_i^{B}\right)^2 = \frac{\psi_i^2}{(A+\psi_i)^4}(A+\psi_i)\mathrm{E}\left[\frac{\partial}{\partial y_i}(y_i - \mathbf{x}_i'\beta)(\hat{A}-A)^2\right] + o(m^{-1})$$

$= \frac{\psi_i^2}{(A+\psi_i)^3}\mathrm{E}\left[2(y_i - \mathbf{x}_i'\beta)(\hat{A}-A)\frac{\partial \hat{A}}{\partial y_i} + (\hat{A}-A)^2\right]$, but $\frac{\partial \hat{A}}{\partial y_i} = O_P(m^{-1})$

$$= \frac{\psi_i^2}{(A+\psi_i)^3}\mathrm{E}(\hat{A}-A)^2 + o(m^{-1}).$$

Thus,

$$g_{3i}(A) = \frac{\psi_i^2}{(A+\psi_i)^3}V(\hat{A}),$$

where $\mathrm{E}(\hat{A}-A)^2 = V(\hat{A})$.

This is the $g_{3i}(A)$ which was proved by Prasad and Rao (1990), Datta and Lahiri (2000) and others.

□

Note that according to Rao (2001), $g_{3i}(A)$ can also be expressed as

$$\tilde{g}_{3i}(A) = \frac{\psi_i^2}{(\psi_i + \hat{A})^4}(y_i - \mathbf{x}_i'\tilde{\beta})^2 V(\hat{A}). \tag{3.2.2}$$

Corollary 3.2.2. We propose another $\tilde{g}_{3JY1i}(A)$ which includes area dependent auxiliary variables as follows

$$\text{JY1}: \qquad \tilde{g}_{3JY1i}(A) = g_{3i}(A) - g_{5i}(A) \tag{3.2.3}$$

where $\quad g_{5i}(A) = \frac{g_{2i}(A)}{(A+\psi_i)^2} V(\hat{A})$.

Note that this estimator does not include y_i, but it includes the area specific auxiliary variables unlike the usual $g_{3i}(A)$.

Proof. In (3.2.2), $E\left\{y_i - \mathbf{x}_i'\tilde{\beta}(A)\right\}^2$ can be simplified as follows.

$$E\left\{y_i - \mathbf{x}_i'\tilde{\beta}(A)\right\}^2 = E\left[(y_i - \mathbf{x}_i'\beta) - \mathbf{x}_i'\left\{\tilde{\beta}(A) - \beta\right\}\right]^2$$

$$= (A + \psi_i) - \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}(A)\mathbf{X})^{-1}\mathbf{x}_i \quad \text{(see Datta et al., 2011)}$$

$$= (A + \psi_i) - \mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m (A + \psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}\mathbf{x}_i.$$

Thus, $\tilde{g}_{3JY1i}(A)$ can be written as:

$$\tilde{g}_{3JY1i}(A) = \frac{\psi_i^2}{(A+\psi_i)^3}V(\hat{A}) - \frac{\psi_i^2}{(A+\psi_i)^4}\mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}(A)\mathbf{X})^{-1}\mathbf{x}_i V(\hat{A})$$

$$= g_{3i}(A) - \frac{g_{2i}(A)}{(A+\psi_i)^2}V(\hat{A})$$

$$= g_{3i}(A) - g_{5i}(A), \text{ where } g_{5i}(A) = \frac{g_{2i}(A)}{(A+\psi_i)^2}V(\hat{A}). \qquad \square$$

Theorem 3.2.3. Under the assumptions 1, 2, 3 and 4,
$E|\tilde{g}_{3JY1i}(\hat{A}) - \tilde{g}_{3JY1i}(A)| = o(m^{-1})$.

Proof. $\tilde{g}_{3JY1i}(\hat{A}) - \tilde{g}_{3JY1i}(A) = (g_{3i}(\hat{A}) - g_{3i}(A)) - (g_{5i}(\hat{A}) - g_{5i}(A))$

$$= I_1 - I_2.$$

According to Prasad and Rao (1990),
$E|I_1| = |g_{3i}(\hat{A}) - g_{3i}(A)| = o(m^{-1})$.
Let us show that $E|I_2| = o(m^{-1})$.

$$g_{5i}(A) = \frac{g_{2i}(A)}{(A+\psi_i)^2}V(\hat{A})$$

$$= \frac{\psi_i^2}{(A+\psi_i)^4}\mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i V(\hat{A}).$$

For example, for the PR moment estimator, $V(\hat{A}) = \frac{2}{m^2}\sum_{i=1}^m (A + \psi_i)^2$. Substituting reveals that

$g_{5i}(A) = \frac{2\psi_i^2 \sum_{i=1}^m (A+\psi_i)^2}{m^2(A+\psi_i)^4} \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i$.

$I_2 = g_{5i}(\hat{A}) - g_{5i}(A)$

$$= \frac{2\psi_i^2 \sum_{i=1}^m (\hat{A}+\psi_i)^2}{m^2(\hat{A}+\psi_i)^4} \mathbf{x}_i'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{x}_i - \frac{2\psi_i^2 \sum_{i=1}^m (A+\psi_i)^2}{m^2(A+\psi_i)^4} \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i.$$

Rewriting I_2

$$I_2 = \frac{2\psi_i^2 \sum_{i=1}^m (\hat{A}+\psi_i)^2}{m^2(\hat{A}+\psi_i)^4} \mathbf{x}_i'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{x}_i - \frac{2\psi_i^2 \sum_{i=1}^m (\hat{A}+\psi_i)^2}{m^2(\hat{A}+\psi_i)^4} \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i$$

$$+ \frac{2\psi_i^2 \sum_{i=1}^m (\hat{A}+\psi_i)^2}{m^2(\hat{A}+\psi_i)^4} \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i - \frac{2\psi_i^2 \sum_{i=1}^m (A+\psi_i)^2}{m^2(A+\psi_i)^4} \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i$$

$$= (2/m)\psi_i^2(\hat{A}+\psi_i)^{-4}(\hat{A}^2 + 2\hat{A}\bar{\psi} + \sum \psi_i^2/m) \left[\mathbf{x}_i'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{x}_i - \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i\right]$$

$$+ \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i \left[(2/m^2)\psi_i^2(\hat{A}+\psi_i)^{-4}\sum_{i=1}^m(\hat{A}+\psi_i)^2\right.$$

$$\left. - (2/m^2)\psi_i^2(A+\psi_i)^{-4}\sum_{i=1}^m(A+\psi_i)^2\right].$$

Applying the Taylor series expansion around A, we get

$$I_3 = \mathbf{x}_i'(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{x}_i - \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i$$

$$= (\hat{A} - A)\left[\partial \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i/\partial A\right]_{A=A^*}$$

$$= -(\hat{A} - A)\mathbf{x}_i'(\mathbf{X}'\Sigma^{*-1}\mathbf{X})^{-1}(\mathbf{X}'\Sigma^{*-2}\mathbf{X})^{-1}(\mathbf{X}'\Sigma^{*-1}\mathbf{X})^{-1}\mathbf{x}_i.$$

$E|I_3| = o(m^{-1})$ (see Prasad and Rao, 1990).

It can also been shown that

$$\mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i \leq (1/m)(A+\psi_U)(m \max_i \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i) = O(m^{-1}).$$

$$I_4 = (2/m^2)\psi_i^2\left[(\hat{A}+\psi_i)^{-4}\sum_{i=1}^m(\hat{A}+\psi_i)^2 - (A+\psi_i)^{-4}\sum_{i=1}^m(A+\psi_i)^2\right]$$

$$= (2/m)\psi_i^2(\hat{A}+\psi_i)^{-4}\left[(\hat{A}-A)^2 + 2(A+\bar{\psi})(\hat{A}-A)\right]$$

$$+ (2/m)\psi_i^2(A^2 + 2A\bar{\psi} + \sum \psi_i^2/m)\left[(\hat{A}+\psi_i)^{-4} - (A+\psi_i)^{-4}\right]$$

$= I_5 + I_6$, where $\bar{\psi} = \sum_{i=1}^m \psi_i/m$.

Using a similar argument to Prasad and Rao (1990), it can be shown that

$E|I_5| = o(m^{-1})$

and

$E|I_6| = o(m^{-1})$.

Thus, $E|\tilde{g}_{3JY1i}(\hat{A}) - \tilde{g}_{3JY1i}(A)| = o(m^{-1})$. □

3.2.2 Estimation of $\mathrm{MSE}(\hat{\theta}_i^{EB})$

The MSE estimator based on the PR moment estimator is given by:

$$\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB}) = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}(\hat{A})$$

where $g_{1i}(\hat{A}) = \hat{A}\psi_i(\hat{A}+\psi_i)^{-1}$, $g_{2i}(\hat{A}) = \frac{\psi_i^2}{(\hat{A}+\psi_i)^2}\mathbf{x}_i'\left\{\sum_{u=1}^{m}(\hat{A}+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$
$g_{3i}(\hat{A}) = 2\psi_i^2(\psi_i+\hat{A})^{-3}m^{-2}\sum_{u=1}^{m}(\psi_u+\hat{A})^2$.

Datta et al. (2005) obtained the MSE estimator of the EBLUP which is given by:

$$\widehat{\mathrm{MSE}}(\hat{\theta}_{FHi}^{EB}) = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3FHi}(\hat{A}) - b_{\hat{A}}(\hat{A})\left\{B_i(\hat{A})\right\}^2$$

where $g_{3FHi}(\hat{A}) = 2m\psi_i^2(\psi_i+\hat{A})^{-3}\left\{\sum_{u=1}^{m}(\psi_u+\hat{A})^{-1}\right\}^{-2}$, for the FH estimator $\hat{A}_{FH}$, we need the following expression for its bias to terms of order $O(m^{-1})$:
$b_{FH}(A) = \frac{2\left[m\sum_{i=1}^{m}(A+\psi_i)^{-2}-\left\{\sum_{i=1}^{m}(A+\psi_i)^{-1}\right\}^2\right]}{\left\{\sum_{i=1}^{m}(A+\psi_i)^{-1}\right\}^3}$.

3.2.3 Area Specific MSE Estimation for EBLUP

One of the criticisms of the above MSE estimators is that they are not area-specific in the sense that they do not explicitly depend on the direct estimator although the area-specific auxiliary data is involved in the component of the uncertainty due to estimation of the regression coefficient (Rao, 2003b). Thus, in this thesis we proposed alternative area specific MSE estimators by extending the Rao (2001) area specific estimators. Thus an area specific estimator of $g_{3i}(A)$ is simply obtained by writing $g_{3i}(A)$ as

$$g_{3i}^*(A, y_i) = \psi_i^2(\psi_i+A)^{-4}(y_i - \mathbf{x}_i'\tilde{\beta})^2 V(A), \tag{3.2.4}$$

where the middle expectation term can be estimated by $(y_i - \mathbf{x}_i'\hat{\beta})^2$, as noted by Rao (2001). The resulting estimator $g_{3i}^*(\hat{A})$ is area-specific since it involves the direct estimator. It is less stable than the Prasad and Rao $g_{3i}(\hat{A})$ even if its instability will not affect the corresponding MSE estimator. As a result, the coefficient of variation (CV) of $\widehat{\mathrm{MSE}}^*(\hat{\theta}_i^{EB})$ should be comparable to the CV of $\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB})$, at least for moderate to large m (Rao, 2003b).

a) An area specific estimator of $g_{3i}(\hat{A})$ according to Rao (2001) is given by (see: Rao, 2003a, Rao, 2003b):

$$\text{Rao2001}: \qquad g_{3Raoi}(\hat{A}, y_i) = \frac{\psi_i^2}{(\psi_i+A)^4}(y_i - \mathbf{x}_i'\hat{\beta})^2 V(\hat{A}). \tag{3.2.5}$$

b) We propose another area specific estimator of $g_{3i}(\hat{A})$ which is an extension of Rao (2001)'s area specific $g_{3Raoi}(\hat{A}, y_i)$ given by

$$\text{JY}: \qquad g_{3JYi}(\hat{A}, y_i) = \frac{\psi_i^2}{(\psi_i+A)^4}\frac{(y_i - \mathbf{x}_i'\hat{\beta})^2}{\hat{A}+\psi_i - \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i}V(\hat{A}), \tag{3.2.6}$$

where $\mathrm{Var}(y_i - \mathbf{x}_i'\tilde{\beta}) = A + \psi_i - \mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i$ (see Corollary 3.2.2). This proposal standardizes the residuals in order to make them scale free (Ananya, 2007).

3.2.3.1 Area Specific MSE Estimation Based on the ML or REML Estimators

According to Rao (2003a), the two alternative area specific estimators based on the ML Estimators (Datta and Lahiri, 2000) are given by:

$$\widehat{\mathrm{MSE}}_1^*(\hat{\theta}_i^{EB}) = g_{1i}(\hat{A}) - b_{\hat{A}}(\hat{A})\nabla g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}^*(\hat{A}, y_i),$$

and

$$\widehat{\mathrm{MSE}}_2^*(\hat{\theta}_i^{EB}) = g_{1i}(\hat{A}) - b_{\hat{A}}(\hat{A})\nabla g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g_{3i}^*(\hat{A}, y_i).$$

Similarly, the two alternative area specific estimators based on the REML Estimators (Datta and Lahiri, 2000) are given by:

$$\widehat{\mathrm{MSE}}_1^*(\hat{\theta}_i^{EB}) \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}^*(\hat{A}, y_i),$$

and

$$\widehat{\mathrm{MSE}}_2^*(\hat{\theta}_i^{EB}) \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g_{3i}^*(\hat{A}, y_i),$$

where $\hat{A}$ is chosen as $\hat{A}_{ML}$ or $\hat{A}_{RE}$.

3.2.3.2 Area Specific MSE Estimation Based on Method of Moments

The two alternative area specific estimators based on the FH Estimator (Datta et al., 2005) are given by:

$$\widehat{\mathrm{MSE}}_1^*(\hat{\theta}_i^{EB}) \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}^*(\hat{A}, y_i) - b_{\hat{A}}(\hat{A})\left\{B_i(\hat{A})\right\}^2$$

and

$$\widehat{\mathrm{MSE}}_2^*(\hat{\theta}_i^{EB}) \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g_{3i}^*(\hat{A}, y_i) - b_{\hat{A}}(\hat{A})\left\{B_i(\hat{A})\right\}^2.$$

For the FH estimator $\hat{A}_{FH}$, we need the following expression for its bias to terms of order $O(m^{-1})$:
$b_{FH}(A) = \frac{2\left[m\sum_{i=1}^m (A+\psi_i)^{-2} - \left\{\sum_{i=1}^m (A+\psi_i)^{-1}\right\}^2\right]}{\left\{\sum_{i=1}^m (A+\psi_i)^{-1}\right\}^3}$.
Note that $g_{3i}^*(\hat{A}, y_i)$ can be an area specific estimator of $g_{3i}(\hat{A})$. See Rao (2003a) for the detailed discussion about extension of MSE estimation to area specific MSE estimation.

3.2.4 MSE of the Weighted Estimator with Fixed Weights

3.2.4.1 MSE Approximation for the Weighted Estimator with Fixed Weights

Theorem 3.2.4. Under the assumptions 2, 3 and 4, the MSE of the weighted estimator with fixed weights, $\mathrm{MSE}(\hat{\theta}_i{}^{w})$ can be approximated as follows (see Datta et al., 2011)

$$\mathrm{MSE}(\hat{\theta}_i{}^{w}) = g_{1i}(A) + g_{2i}(A) + g_{3wi}(A) + o(m^{-1}).$$

Proof. This theorem is proved in a similar manner with Theorem 3.2.1 applying the decomposition for the sake of completeness.

$$\hat{\theta}_i{}^{w} - \theta_i = \left\{(1-B_i)y_i + B_i\mathbf{x}_i'\beta - \theta_i\right\} + \left\{\hat{\theta}_i{}^{B} - (1-B_i)y_i - B_i\mathbf{x}_i'\beta\right\} + \left\{\hat{\theta}_i{}^{w} - \hat{\theta}_i{}^{B}\right\}.$$

In Theorem 3.2.1, we have derived the first two components of $\mathrm{MSE}(\hat{\theta}_i{}^{w})$. Thus,

$$\mathrm{MSE}(\hat{\theta}_i{}^{w}) = g_{1i}(A) + g_{2i}(A) + E\left\{\hat{\theta}_i{}^{w} - \hat{\theta}_i{}^{B}\right\}^2.$$

Now let us evaluate the third component as follows.

$$g_{3i}(A) = E\left\{\hat{\theta}_i{}^{w} - \hat{\theta}_i{}^{B}\right\}^2 = E\left\{-w_i(y_i - \mathbf{x}_i'\hat{\beta}) + \gamma_i(y_i - \mathbf{x}_i'\tilde{\beta}(A))\right\}^2.$$

Using the Taylor series expansion,

$$\hat{\beta} = \tilde{\beta}(A) + \frac{\partial\tilde{\beta}(A)}{\partial A}(\hat{A} - A).$$

Substituting reveals that

$$\begin{aligned}
g_{3wi}(A) &= E\left\{-w_i\left(y_i - \mathbf{x}_i'\left(\tilde{\beta}(A) + \frac{\partial\tilde{\beta}(A)}{\partial A}(\hat{A}-A)\right)\right) + \gamma_i(y_i - \mathbf{x}_i'\tilde{\beta}(A))\right\}^2 \\
&= E\left\{(\gamma_i - w_i)(y_i - \mathbf{x}_i'\tilde{\beta}(A)) - \mathbf{x}_i'w_i\frac{\partial\tilde{\beta}(A)}{\partial A}(\hat{A}-A)\right\}^2 \\
&= (\gamma_i - w_i)^2 E\left(y_i - \mathbf{x}_i'\tilde{\beta}(A)\right)^2 - 2w_i(\gamma_i - w_i)E\left((y_i - \mathbf{x}_i'\tilde{\beta}(A))\mathbf{x}_i'\frac{\partial\tilde{\beta}(A)}{\partial A}(\hat{A}-A)\right) \\
&= (\gamma_i - w_i)^2\left\{(A + \psi_i) - \mathbf{x}_i'\left(\mathbf{X}'\Sigma^{-1}(A)\mathbf{X}\right)^{-1}\mathbf{x}_i\right\} + o(m^{-1}) \quad \text{(see Datta et al., 2011).}
\end{aligned}$$

Combination of the above results reveals that

$$\mathrm{MSE}(\hat{\theta}_i{}^{w}) = g_{1i}(A) + g_{2i}(A) + g_{3wi}(A) + o(m^{-1}).$$

□

3.2.4.2 MSE Estimation for the Weighted Estimator with Fixed Weights

According to Datta et al. (2011), for example, the nearly unbiased estimator of the $\mathrm{MSE}(\hat{\theta}_i{}^{w})$ under the FH moment estimator $\hat{A}$ is given by

$$\widehat{\mathrm{MSE}}(\hat{\theta}_i{}^{w}) = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g_{3wi}(\hat{A}) - \hat{b}_{FH}g_{1i}(\hat{A})w_i^2, \tag{3.2.7}$$

where $\hat{b}_{FH}(\hat{A}) = \frac{2\left[m\sum_{i=1}^{m}(\hat{A}+\psi_i)^{-2} - \left\{\sum_{i=1}^{m}(\hat{A}+\psi_i)^{-1}\right\}^2\right]}{\left\{\sum_{i=1}^{m}(\hat{A}+\psi_i)^{-1}\right\}^3}$.

A nearly unbiased estimator of the MSE($\hat{\theta}_i^{w}$) under the PR moment and REML estimators of $\hat{A}$ are given by

$$\widehat{\mathrm{MSE}}(\hat{\theta}_i^{w}) = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g_{3wi}(\hat{A}). \tag{3.2.8}$$

It should be clearly noted that: `MSE_ML` and `MSE_REML` estimators are proposed by Datta and Lahiri (2000) for the ML and REML estimators and `MSE_FH` is proposed by Datta et al. (2005) for the FH moment estimator under the FH model. The rest of the MSE estimators (`MSE_Rao`, `MSE_JY` and `MSE_JY1`) are the modified version of the MSE estimators for the FH model.

Chapter 4

Confidence Intervals with Improved Coverage Probability

4.1 Confidence Intervals for a Small Area Mean

In this section we propose CIs for $\theta_i = \mathbf{x}_i\beta + v_i$ based on the EBLUP and weighted estimators with fixed weights under the FH model. This is an extension of Diao et al. (2013) CIs for θ_i based on EBLUP and weighted estimator with fixed weights. As we have clearly stated in Chapter 1, Diao et al. (2013) proposed CIs based on EBLUP estimators which is correct up to the order $O(m^{-2})$. However, their CIs have limitation when either $\hat{A}$ or $\widehat{MSE}(\hat{\theta}_i^{EB})$ or both are negative. Their CIs can also be criticized as they are not area specific CIs. To fill this gap, we propose different CIs based on EBLUP and weighted estimator with fixed weights.

4.1.1 Confidence Intervals for a Small Area Mean Based on the Weighted Estimator with Fixed Weights

The following theorem states the improved CIs based on weighted estimator with fixed weights.

Theorem 4.1.1. For any real $z > 0$ and under assumptions 1, 2 and 4
$P[\theta_i \in I^{CW}(\hat{A})] = 2\Phi(z) - 1 - z\phi(z)\eta_w + O(m^{-3/2})$

where
$\eta_w = (z^2+1)\frac{\psi_i^2}{4A^2(A+\psi_i)^2}V(\hat{A}) + \frac{\psi_i^2+(\gamma_i-w_i)^2(A+\psi_i)^2}{A\psi_i(A+\psi_i)}\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i.$

Proof. This Theorem is proved similarly to Diao et al. (2013) and Datta et al. (2002). The conditional distribution of θ_i given y_i is given by $\theta_i \mid y_i \sim N(\hat{\theta}_i^w, g_{1i}(A))$, where $\hat{\theta}_i^w = (1-w_i)y_i + w_i\mathbf{x}_i'\hat{\beta}$, $\tilde{\theta}_i^B = (1-\gamma_i)y_i + \gamma_i\mathbf{x}_i'\tilde{\beta}$, $g_{1i}(A) = \frac{A\psi_i}{A+\psi_i}$.

Consider first that

$$
\begin{aligned}
&P\left[\theta_i \le \hat{\theta}_i^w + z\times\sqrt{\widehat{MSE}(\hat{\theta}_i^w)}\right]\\
&= P\left[\frac{\theta_i-\hat{\theta}_i^B}{\sqrt{E(\theta_i-\hat{\theta}_i^B)^2}} \le \frac{\hat{\theta}_i^w-\hat{\theta}_i^B+z\times\sqrt{\widehat{MSE}(\hat{\theta}_i^w)}}{\sqrt{E(\theta_i-\hat{\theta}_i^B)^2}}\right]\\
&= P\left[\frac{\theta_i-\hat{\theta}_i^B}{\sqrt{g_{1i}(A)}} \le \frac{\hat{\theta}_i^w-\hat{\theta}_i^B+z\times\sqrt{\widehat{MSE}(\hat{\theta}_i^w)}}{\sqrt{g_{1i}(A)}}\right]\\
&= P\left[\frac{\theta_i-\hat{\theta}_i^B}{\sqrt{g_{1i}(A)}} \le \frac{z\times\sqrt{g_{1i}(A)}}{\sqrt{g_{1i}(A)}} + \frac{z\times\sqrt{\widehat{MSE}(\hat{\theta}_i^w)}-z\times\sqrt{g_{1i}(A)}+\hat{\theta}_i^w-\hat{\theta}_i^B}{\sqrt{g_{1i}(A)}}\right]\\
&= P\left[\frac{\theta_i-\hat{\theta}_i^B}{\sqrt{g_{1i}(A)}} \le \frac{z\times\sqrt{g_{1i}(A)}}{\sqrt{g_{1i}(A)}} + \frac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w)}-\sqrt{g_{1i}(A)}\right)+\hat{\theta}_i^w-\hat{\theta}_i^B}{\sqrt{g_{1i}(A)}}\right]\\
&= P\left[\frac{\theta_i-\hat{\theta}_i^B}{\sqrt{g_{1i}(A)}} \le z + \frac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w)}-\sqrt{g_{1i}(A)}\right)+\hat{\theta}_i^w-\hat{\theta}_i^B}{\sqrt{g_{1i}(A)}}\right]\\
&= P\left[\frac{\theta_i-\hat{\theta}_i^B}{\sqrt{g_{1i}(A)}} \le z+F(z)\right]\\
&= E[\Phi(z+F(z))], \text{ where } F(z) = \frac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w)}-\sqrt{g_{1i}(A)}\right)+\hat{\theta}_i^w-\tilde{\theta}_i^B}{\sqrt{g_{1i}(A)}}.
\end{aligned}
$$

Thus, the CP of $I^{CW}(\hat{A})$ is written as

$$
\begin{aligned}
P[\theta_i \in I^{CW}(\hat{A})] &= P[-z+F(-z) < \frac{\theta_i-\hat{\theta}_i^w}{\sqrt{g_{1i}(A)}} < z+F(z)]\\
&= E[\Phi(z+F(z))-\Phi(-z+F(-z))].
\end{aligned}
$$

Using the Taylor series expansion with an integral remainder term, $\Phi(z+F(z))$ is evaluated as

$$
\begin{aligned}
\Phi(z+F(z)) &= \Phi(z)+F(z)\phi(z)+\tfrac{1}{2}F^2(z)\phi'(z)+\tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2\phi''(x)dx\\
&= \Phi(z)+[F(z)-\tfrac{z}{2}F^2(z)]\phi(z)+\tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2(x^2-1)\phi(x)dx.
\end{aligned}
$$

Taking expectations on both sides reveals that

$$
\begin{aligned}
E(\Phi(z+F(z))) = \Phi(z)+E\Big[&F(z)\phi(z)-\tfrac{z}{2}F^2(z)\phi(z)\\
&+\tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2(x^2-1)\phi(x)dx\Big].
\end{aligned}
$$

But

$$
\begin{aligned}
\hat{\theta}_i^w-\tilde{\theta}_i^B &= (1-w_i)y_i+w_i\mathbf{x}_i'\hat{\beta}-\left((1-\gamma_i)y_i+\gamma_i\mathbf{x}_i'\tilde{\beta}\right)\\
&= y_i-w_i(y_i-\mathbf{x}_i'\hat{\beta})-y_i+\gamma_i(y_i-\mathbf{x}_i'\tilde{\beta}(A))
\end{aligned}
$$

$$= \gamma_i(y_i - \mathbf{x}_i'\tilde{\beta}(A)) - w_i(y_i - \mathbf{x}_i'\hat{\beta}).$$

Using the Taylor series expansion we have

$$\hat{\beta} = \tilde{\beta}(A) + \tilde{\beta}^{(1)}(A)(\hat{A} - A)$$

$$\hat{\theta}_i^w - \tilde{\theta}_i^B = \gamma_i(y_i - \mathbf{x}_i'\tilde{\beta}(A)) - w_i\left(y_i - \mathbf{x}_i'(\tilde{\beta}(A) + \tilde{\beta}^{(1)}(A)(\hat{A} - A))\right)$$

$$= \gamma_i(y_i - \mathbf{x}_i'\tilde{\beta}(A)) - w_i(y_i - \mathbf{x}_i'\tilde{\beta}(A)) - w_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A} - A))$$

$$= (\gamma_i - w_i)(y_i - \mathbf{x}_i'\tilde{\beta}(A)) - w_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A} - A)).$$

Substituting reveals that

$$F(z) = \frac{z \times \left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w)} - \sqrt{g_{1i}(A)}\right) + \hat{\theta}_i^w - \tilde{\theta}_i^B}{\sqrt{g_{1i}(A)}}$$

$$= \frac{z \times \left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w)} - \sqrt{g_{1i}(A)}\right)}{\sqrt{g_{1i}(A)}} + \frac{(\gamma_i - w_i)(y_i - \mathbf{x}_i'\tilde{\beta}(A))}{\sqrt{g_{1i}(A)}} - \frac{w_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A} - A))}{\sqrt{g_{1i}(A)}}$$

$$= S_1 + R_2 + R_3.$$

Using the Taylor series expansion it can be shown that

$$ES_1 = z\tfrac{1}{2}\left(g_{1i}(A)\right)^{-1} E\left(\widehat{MSE}(\hat{\theta}_i^w) - g_{1i}(A)\right)$$

$$-z\tfrac{1}{8}\left(g_{1i}(A)\right)^{-2} E\left(\widehat{MSE}(\hat{\theta}_i^w) - g_{1i}(A)\right)^2$$

$$+z\tfrac{3}{16}\left(g_{1i}(A)\right)^{-1/2} E\int_{g_{1i}(A)}^{\widehat{MSE}(\hat{\theta}_i^w)} x^{-5/2}\left(\widehat{MSE}(\hat{\theta}_i^w) - g_{1i}(A)\right)^2 dx.$$

Let us compute the expectations of each term in turn

$E\left(\widehat{MSE}(\hat{\theta}_i^w) - g_{1i}(A)\right) = E(\widehat{MSE}(\hat{\theta}_i^w)) - E(g_{1i}(A))$.

$E(\widehat{MSE}(\hat{\theta}_i^w)) = MSE(\hat{\theta}_i^w)$, since $E(\widehat{MSE}(\hat{\theta}_i^w))$ is nearly a second order unbiased estimator of $MSE(\hat{\theta}_i^w)$.

Using the Taylor series expansion, the second term can be expressed as

$\widehat{MSE}(\hat{\theta}_i^w) = MSE(\hat{\theta}_i^w(A)) + MSE(\hat{\theta}_i^w(A^*))^{(1)}(\hat{A} - A) + O(m^{-3/2})$.

This can be rewritten as

$\widehat{MSE}(\hat{\theta}_i^w) - g_{1i}(A) = g_{1i}^{(1)}(A^*)(\hat{A} - A)$,

$E\left(\widehat{MSE}(\hat{\theta}_i^w) - g_{1i}(A)\right)^2 = (g_{1i}^{(1)}(A^*))^2 E(\hat{A} - A)^2$,

where $g_{1i}(A) = \frac{A\psi_i}{A+\psi_i}$, $g_{1i}^{(1)}(A) = \frac{\psi_i}{A+\psi_i} - \frac{A\psi_i}{(A+\psi_i)^2} = \frac{\psi_i^2}{(A+\psi_i)^2}$, $(g_{1i}^{(1)}(A))^2 = \frac{\psi_i^4}{(A+\psi_i)^4}$.

Substituting reveals that

$E\left(\widehat{MSE}(\hat{\theta}_i^w) - g_{1i}(A)\right)^2 = \frac{\psi_i^4}{(A+\psi_i)^4} E(\hat{A} - A)^2 + O(m^{-3/2})$,

where, $E(\hat{A} - A)^2 = V(\hat{A}) + O_p(m^{-1})$, $V(\hat{A})$ can be the asymptotic variance of the FH, PR, ML and REML estimators of A.

Let us compute the integral term using the same approach by Datta et al. (2002); Chatterjee and Lahiri (2002) and Diao et al. (2013) as follows.

$$E\int_{g_{1i}(A)}^{\widehat{MSE}(\hat{\theta}_i^w)} x^{-5/2}\left(\widehat{MSE}(\hat{\theta}_i^w)-g_{1i}(A)\right)^2 dx = EI_1 + EI_2.$$

$$EI_1 = \left[E\int_{g_{1i}(A)}^{\widehat{MSE}(\hat{\theta}_i^w)} x^{-5/2}\left(\widehat{MSE}(\hat{\theta}_i^w)-g_{1i}(A)\right)^2 dx\right] I_{\{\widehat{MSE}(\hat{\theta}_i^w)\geq g_{1i}(A)\}}$$

and

$$EI_2 = \left[E\int_{g_{1i}(A)}^{\widehat{MSE}(\hat{\theta}_i^w)} x^{-5/2}\left(\widehat{MSE}(\hat{\theta}_i^w)-g_{1i}(A)\right)^2 dx\right] I_{\{\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)\}}.$$

But

$$\begin{aligned}
EI_1 &= \left[E\int_{g_{1i}(A)}^{\widehat{MSE}(\hat{\theta}_i^w)} x^{-5/2}\left(\widehat{MSE}(\hat{\theta}_i^w)-g_{1i}(A)\right)^2 dx\right] I_{\{\widehat{MSE}(\hat{\theta}_i^w)\geq g_{1i}(A)\}}\\
&\leq \left[E\int_{g_{1i}(A)}^{\widehat{MSE}(\hat{\theta}_i^w)} (g_{1i}(A))^{-5/2}\left(\widehat{MSE}(\hat{\theta}_i^w)-g_{1i}(A)\right)^2 dx\right] I_{\{\widehat{MSE}(\hat{\theta}_i^w)\geq g_{1i}(A)\}}\\
&\leq \left[\tfrac{1}{3}(g_{1i}(A))^{-5/2} E\left(\widehat{MSE}(\hat{\theta}_i^w)-g_{1i}(A)\right)^3\right]\\
&= O(m^{\frac{-3}{2}}),
\end{aligned}$$

using the Taylor series expansion $E\left(\widehat{MSE}(\hat{\theta}_i^w)-g_{1i}(A)\right)^3 = O(m^{\frac{-3}{2}})$.

In order to evaluate EI_2, we choose $\varepsilon_m = m^{-\alpha}$ $(0<\alpha<1/2)$ (see Datta et al., 2002). According to Diao et al. (2013) we can rewrite $I_{\{\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)\}}$ as follows: $I_{\{\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)\}} = I_{\{\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)-\varepsilon_m\}} + I_{\{g_{1i}(A)-\varepsilon_m\leq\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)\}}$.

Then,

$$\begin{aligned}
&\left[E\int_0^{g_{1i}(A)-\widehat{MSE}(\hat{\theta}_i^w)} x^2\left(\widehat{MSE}(\hat{\theta}_i^w)+x\right)^{-5/2} dx\right] I_{\{\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)-\varepsilon_m\}}\\
&\leq \left[E\int_0^{g_{1i}(A)-\widehat{MSE}(\hat{\theta}_i^w)} x^{4/2-5/2} dx\right] I_{\{\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)-\varepsilon_m\}}\\
&= E\left[2\left\{g_{1i}(A)-\widehat{MSE}(\hat{\theta}_i^w)\right\}^{1/2} I_{\{\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)-\varepsilon_m\}}\right]\\
&\leq 2\sqrt{E|g_{1i}(A)-\widehat{MSE}(\hat{\theta}_i^w)|}\sqrt{P\left(g_{1i}(A)-\widehat{MSE}(\hat{\theta}_i^w)\geq\varepsilon_m\right)}\\
&\leq O(m^{-1/2})E\sqrt{\frac{\left(g_{1i}(A)-\widehat{MSE}(\hat{\theta}_i^w)\right)^3}{\varepsilon_m^3}}\\
&= O(m^{\frac{-3}{2}}),
\end{aligned}$$

using the Taylor series expansion $E\left(\widehat{MSE}(\hat{\theta}_i^w)-g_{1i}(A)\right)^3 = O(m^{\frac{-3}{2}})$.

$$\begin{aligned}
&E\left[\int_{\widehat{MSE}(\hat{\theta}_i^w)}^{g_{1i}(A)} x^{-5/2}\left(x-\widehat{MSE}(\hat{\theta}_i^w)\right)^2 dx I_{\{g_{1i}(A)-\varepsilon_m\leq\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)\}}\right]\\
&\leq (g_{1i}(A)-\varepsilon_m)^{-5/2}E\left[\int_{\widehat{MSE}(\hat{\theta}_i^w)}^{g_{1i}(A)}\left(x-\widehat{MSE}(\hat{\theta}_i^w)\right)^2 dx I_{\{g_{1i}(A)-\varepsilon_m\leq\widehat{MSE}(\hat{\theta}_i^w)\leq g_{1i}(A)\}}\right]\\
&\leq \tfrac{1}{3}(g_{1i}(A)-\varepsilon_m)^{-5/2}E\left(g_{1i}(A)-\widehat{MSE}(\hat{\theta}_i^w)\right)^3
\end{aligned}$$

$= O(m^{\frac{-3}{2}}),$

using the Taylor series expansion $E\left(\widehat{\mathrm{MSE}}(\hat{\theta}_i^{w}) - g_{1i}(A)\right)^3 = O(m^{\frac{-3}{2}})$.
Thus, $EI_1 + EI_2 = O(m^{\frac{-3}{2}})$.
Combination of all the above expressions gives

$$ES_1 = \frac{z}{2}\left(g_{1i}(A)\right)^{-1}(g_{2i}(A) + g_{3wi}(A)) - \frac{z}{8}\left(g_{1i}(A)\right)^{-2}\frac{\psi_i^4}{(A+\psi_i)^4}V(\hat{A}) + O(m^{-3/2})$$

$$= \frac{z}{2}\left(\frac{A\psi_i}{A+\psi_i}\right)^{-1}\left[\frac{\psi_i^2}{(A+\psi_i)^2}\mathbf{x}_i'(\textstyle\sum_{i=1}^{m}\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i})^{-1}\mathbf{x}_i\right.$$
$$\left. + (\gamma_i - w_i)^2\left((A+\psi_i) - \mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i\right)\right]$$
$$- \frac{z}{8}\left(\frac{A\psi_i}{A+\psi_i}\right)^{-2}\frac{\psi_i^4}{(A+\psi_i)^4}V(\hat{A}) + O(m^{-3/2}).$$

R_2 can be written as

$$R_2 = g_{1i}^{-1/2}(A)(\gamma_i - w_i)(y_i - \mathbf{x}_i'\beta) + o(m^{-1}).$$

Using the Stein identity given by equation (3.2.1),

$$ER_2 = g_{1i}^{-1/2}(A)(A+\psi_i)\frac{\partial(\gamma_i - w_i)}{\partial y_i} = 0.$$

$ER_3 = 0$ since $\hat{A} - A = O_p(m^{-1/2})$ (see Datta et al., 2011).

$F(z)^2$ can be written as:

$$F(z)^2 = (S_1 + R_2 + R_3)^2$$
$$= S_1^2 + R_2^2 + R_3^2 + S_1R_2 + S_1R_3 + R_2R_3.$$

Evaluating the expected value of each term turn by turn reveals that

$$ES_1^2 = \frac{z^2}{4}E\left(g_{1i}^{-1}(A)(g_{1i}^{1}(A)(\hat{A} - A))\right)^2 + O(m^{-3/2})$$
$$= \frac{z^2}{4}g_{1i}^{-2}(A)(g_{1i}^{1}(A))^2E(\hat{A} - A)^2 + O(m^{-3/2})$$
$$= \frac{z^2\psi_i^2}{4A^2(A+\psi_i)^2}V(\hat{A}) + O(m^{-3/2}).$$

$$R_2^2 = \left(\frac{(\gamma_i - w_i)(y_i - \mathbf{x}_i'\tilde{\beta}(A))}{\sqrt{g_{1i}(A)}}\right)^2$$
$$= g_{1i}^{-1}(A)(\gamma_i - w_i)^2(y_i - \mathbf{x}_i'\beta)^2 + o(m^{-1}).$$

Using the Stein identity (equation (3.2.1)) we have,

$$ER_2^2 = g_{1i}^{-1}(A)(\gamma_i - w_i)^2(A+\psi_i)\frac{\partial}{\partial y_i}(y_i - \mathbf{x}_i'\beta) + o(m^{-1})$$
$$= g_{1i}^{-1}(A)(\gamma_i - w_i)^2(A+\psi_i) + o(m^{-1}).$$

R_3^2 can be simplified as:

$$R_3^2 = \left(-\frac{w_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A} - A))}{\sqrt{g_{1i}(A)}}\right)^2$$

$$= g_{1i}^{-1}(A)(w_i \mathbf{x}_i' \tilde{\beta}^{(1)}(A))^2(\hat{A}-A))^2.$$

Taking expectations on both sides

$$ER_3^2 = g_{1i}^{-1}(A)(w_i \mathbf{x}_i' \tilde{\beta}^{(1)}(A))^2 E(\hat{A}-A))^2$$

$$= g_{1i}^{-1}(A)(w_i \mathbf{x}_i' \tilde{\beta}^{(1)}(A))^2 V(\hat{A}),$$

since $\tilde{\beta}^{(1)}(A) = O_P(m^{-1/2})$ (Datta et al., 2011).

$$S_1R_2 = \frac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w)}-\sqrt{g_{1i}(A)}\right)}{\sqrt{g_{1i}(A)}} \frac{(\gamma_i-w_i)(y_i-\mathbf{x}_i'\tilde{\beta}(A))}{\sqrt{g_{1i}(A)}}$$

$$= g_{1i}^{-1} z(\gamma_i - w_i)\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w)} - \sqrt{g_{1i}(A)}\right)(y_i - \mathbf{x}_i'\tilde{\beta}(A)).$$

Using the Taylor series expansion it follows that

$$S_1R_2 = g_{1i}^{-1} z(\gamma_i - w_i)\left(g_{1i}^{(1)}(A)(\hat{A}-A)\right)(y_i - \mathbf{x}_i'\tilde{\beta}(A))$$

$$= g_{1i}^{-1} z(\gamma_i - w_i)\left(g_{1i}^{(1)}(A)(\hat{A}-A)\right)(y_i - \mathbf{x}_i'\beta) + o(m^{-1}).$$

Applying the Stein identity which is given by equation (3.2.1)

$ES_1R_2 = g_{1i}^{-1} z(\gamma_i - w_i) g_{1i}^{(1)}(A)(A+\psi_i)\frac{\partial \hat{A}}{\partial y_i} + o(m^{-1})$, since $\frac{\partial \hat{A}}{\partial y_i} = O_p(m^{-1})$.

Similarly R_2R_3 can be simplified as:

$$R_2R_3 = \frac{(\gamma_i-w_i)(y_i-\mathbf{x}_i'\tilde{\beta}(A))}{\sqrt{g_{1i}(A)}} \frac{w_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A))}{\sqrt{g_{1i}(A)}}$$

$$= g_{1i}^{-1}(A)(\gamma_i - w_i)(y_i - \mathbf{x}_i'\beta) w_i \mathbf{x}_i' \tilde{\beta}^{(1)}(A)(\hat{A}-A)) + o(m^{-1}).$$

Using the Stein identity (see equation (3.2.1))

$$ER_2R_3 = g_{1i}^{-1}(A) w_i(\gamma_i - w_i)(A+\psi_i) E\frac{\partial}{\partial y_i}\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A))\right\}.$$

Note that

$$\frac{\partial}{\partial y_i}\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)\right\} = \frac{\partial}{\partial y_i}\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)\right\}(\hat{A}-A) + \mathbf{x}_i'\tilde{\beta}^{(1)}(A)\frac{\partial \hat{A}}{\partial y_i}$$

where

$$\frac{\partial}{\partial y_i}\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)\right\} = \frac{\partial}{\partial A}\mathbf{x}_i'\left\{\frac{\partial}{\partial y_i}\tilde{\beta}(A)\right\}$$

$$= \frac{\partial}{\partial A}\left\{\frac{\tilde{h}_{ii}}{A+\psi_i}\right\} + O(m^{-1}),$$

where $\tilde{h}_{ii} = \mathbf{x}_i'\left\{\sum_{j=1}^m (A+\psi_j)^{-1}\mathbf{x}_j\mathbf{x}_j'\right\}^{-1}$ is $O(m^{-1})$, $\frac{\partial \hat{A}}{\partial y_i} = O_P(m^{-1})$ and $\frac{\partial}{\partial y_i}(\hat{\beta}(\hat{A})) = O_P(m^{-1})$ (see Kubokawa, 2010, Datta et al., 2011).

Hence it follows that

$$\frac{\partial}{\partial y_i}\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A))\right\} = O(m^{-1}).$$

Similarly S_1R_3 can be simplified as:

$$S_1R_3 = -\frac{z\times\left(\sqrt{\widehat{\text{MSE}}(\hat{\theta}_i^w)}-\sqrt{g_{1i}(A)}\right)}{\sqrt{g_{1i}(A)}}\frac{w_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A))}{\sqrt{g_{1i}(A)}}$$

$$= -g_{1i}^{-1}(A)zw_i\left(\sqrt{\widehat{\text{MSE}}(\hat{\theta}_i^w)}-\sqrt{g_{1i}(A)}\right)\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A).$$

Using the Taylor series expansion we have,

$$S_1R_3 = -g_{1i}^{-1}(A)zw_ig_{1i}^{(1)}(A)(\hat{A}-A)\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)$$

$$= -g_{1i}^{-1}(A)zw_ig_{1i}^{(1)}(A)\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)^2.$$

Taking expectations on both sides

$$ES_1R_3 = -g_{1i}^{-1}(A)zw_ig_{1i}^{(1)}(A)\mathbf{x}_i'\tilde{\beta}^{(1)}(A)E(\hat{A}-A)^2$$

$$= -g_{1i}^{-1}(A)zw_ig_{1i}^{(1)}(A)\mathbf{x}_i'\tilde{\beta}^{(1)}(A)V(\hat{A}),$$

since $\tilde{\beta}^{(1)}(A) = O_P(m^{-1/2})$ (Datta et al., 2011).

Combination of all the above expressions gives

$$EF(z) = E(S_1)+E(R_2)+E(R_3)$$

$$= \frac{z}{2}\left(\frac{A\psi_i}{A+\psi_i}\right)^{-1}\left[\frac{\psi_i^2}{(A+\psi_i)^2}\mathbf{x}_i'(\textstyle\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i})^{-1}\mathbf{x}_i\right.$$

$$\left.+\ (\gamma_i-w_i)^2\left((A+\psi_i)-\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i\right)\right]$$

$$-\ \frac{z}{8}\left(\frac{A\psi_i}{A+\psi_i}\right)^{-2}\frac{\psi_i^4}{(A+\psi_i)^4}V(\hat{A})+O(m^{-3/2}).$$

$$E(F(z)-F(-z)) = z\frac{\psi_i}{A(A+\psi_i)}\mathbf{x}_i'(\textstyle\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i})^{-1}\mathbf{x}_i$$

$$+\ z\left(\frac{A\psi_i}{A+\psi_i}\right)^{-1}(\gamma_i-w_i)^2(A+\psi_i))$$

$$-\ z\left(\frac{A\psi_i}{A+\psi_i}\right)^{-1}(\gamma_i-w_i)^2\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i$$

$$-\ \frac{z}{4}\frac{\psi_i^2}{A^2(A+\psi_i)^2}V(\hat{A})+O(m^{-3/2}).$$

$$EF(z)^2 = ES_1^2+ER_2^2+ER_3^2+ES_1R_2+ES_1R_3+ER_2R_3$$

$$= \frac{z^2\psi_i^2}{4A^2(A+\psi_i)^2}V(\hat{A})+g_{1i}^{-1}(A)(\gamma_i-w_i)^2(A+\psi_i)+O(m^{-3/2})$$

$$E(F(z)^2+F(-z)^2) = \frac{z^2\psi_i^2}{2A^2(A+\psi_i)^2}V(\hat{A})+2\left(\frac{A\psi_i}{A+\psi_i}\right)^{-1}(\gamma_i-w_i)^2(A+\psi_i)$$

$$+\ O(m^{-3/2}).$$

$$E\left[F(z)-F(-z)-\frac{z}{2}(F(z)^2+F(-z)^2\right]$$

$$= -z\phi(z)\left\{(z^2+1)\frac{\psi_i^2}{4A^2(A+\psi_i)^2}V(\hat{A})-\frac{\psi_i^2+(\gamma_i-w_i)^2(A+\psi_i)^2}{A\psi_i(A+\psi_i)}\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i\right\}.$$

The remainder term $\int_z^{z+F(z)}(z+F(z)-x)^2(x^2-1)\phi(x)dx$ can be simplified using the approach by Datta et al. (2002), Chatterjee and lahiri (2002) and Diao et al. (2013) as follows. Since within the limits of the integral $0 < |z+F(z)-x| < |F(z)|$ and $|(x^2-1)\phi(x)| < 2\phi(\sqrt{3})$, it follows that

$$
\begin{aligned}
& E|\int_z^{z+F(z)}(x^2-1)(z+F(z)-x)^2\phi(x)dx| \\
& \leq E|F(z)|^2|\int_z^{z+F(z)}2\phi(\sqrt{3})dx| \\
& = 2E|F(z)|^3\phi(\sqrt{3}) \\
& = O(m^{-3/2}).
\end{aligned}
$$

Combination and simplifications of the above results give the CP of $I^{CW}(\hat{A})$ as

$$
\begin{aligned}
P[\theta_i \in I^{CW}(\hat{A})] &= E[\Phi(z+F(z))-\Phi(-z+F(-z))] \\
&= 2\Phi(z)-1-z\phi(z)\eta_w+O(m^{-3/2})
\end{aligned}
$$

where

$$\eta_w = (z^2+1)\frac{\psi_i^2}{4A^2(A+\psi_i)^2}V(\hat{A})+\frac{\psi_i^2+(\gamma_i-w_i)^2(A+\psi_i)^2}{A\psi_i(A+\psi_i)}\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i. \qquad \square$$

4.1.2 Confidence Intervals Based on the EBLUP

Theorem 4.1.2. For any real $z > 0$,

$$P\left[\theta_i \in \hat{\theta}_i^{EB} \pm z\times\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}\right] = 2\Phi(z)-1-\frac{z\phi(z)}{4}\eta+O(m^{-3/2})$$

where $\eta = (z^2+1)\frac{\psi_i^2}{A^2(A+\psi_i)^2}V(\hat{A})$ will give an EBLUP CI that matches the nominal CP to the order of $O(m^{-3/2})$. These CIs based on the EBLUP can also be written as:

$$\hat{\theta}_i^{EB} \pm t_\alpha^*\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}, \tag{4.1.1}$$

where $t_\alpha^* = z_{\alpha/2}(1+h(\hat{A}))$, $h(\hat{A}) = (z_{\alpha/2}^2+1)\frac{\psi_i^2}{8\hat{A}^2(\hat{A}+\psi_i)^2}V(\hat{A})$.

This is similar to the second order improved CI by Kubokawa (2011). Note that the proof of this theorem is given in the Appendix A. With this in mind, the recent CIs based on the EBLUP are stated as follows:

$$\text{Our method}: \quad \hat{\theta}_i^{EB} \pm t_\alpha^*\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}, \tag{4.1.2}$$

where $t_\alpha^* = z_{\alpha/2}+(z_{\alpha/2}^3+z_{\alpha/2})\frac{\psi_i^2}{8A^2(A+\psi_i)^2}V(\hat{A})$.

$$\text{Diao et al. (2013)}: \quad \hat{\theta}_i^{EB} \pm t_\alpha\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}, \tag{4.1.3}$$

where $t_\alpha = z+\frac{z^3+z}{8h_i^4}\frac{g_{3i}\psi_i^2}{\hat{A}+\psi_i}$, and $h_i^2 = g_{1i}+g_{2i}$ for the PR and REML estimators.

Both (4.1.2) and (4.1.3) are constructed by using the non area specific MSEs. In addition, Theorem 4.1.2 is based on the usual non area specific $g_{3i}(A)$. However, the following corollaries are obtained by using the proposed area specific MSEs. Once Theorem 4.1.2 is proved, Corollary 4.1.3, Corollary 4.1.4 and Corollary 4.1.5 can be easily obtained through replacing the usual $g_{3i}(A)$ by area specific $g_{3i}(A, y_i)$'s given by (3.2.5), (3.2.6) and (3.2.3) respectively. The corresponding area specific CIs are also given in (5.2.8), (5.2.9) and (5.2.10) respectively. Consider the following three corollaries which follows Theorem 4.1.2 by the Taylor series expansion.

Corollary 4.1.3. The CP of $I^{CEBRao}(\hat{A})$ based on the Rao (2001) is written as

$$P[\theta_i \in I^{CEBRao}(\hat{A})] = 2\Phi(z) - 1 - \frac{z\phi(z)}{4}\eta_{Rao} + O(m^{-3/2})$$

where $\eta_{Rao} = (z^2+1)\frac{\psi_i^2}{\hat{A}^2(\hat{A}+\psi_i)^3}(y_i - \mathbf{x}_i'\hat{\beta})^2 V(\hat{A})$.

Proof: See Theorem 4.1.2.

Corollary 4.1.4. The CP of $I^{CEBJY}(\hat{A})$ based on JY is written as

$$P[\theta_i \in I^{CEBJY}(\hat{A})] = 2\Phi(z) - 1 - \frac{z\phi(z)}{4}\eta_{JY} + O(m^{-3/2})$$

where $\eta_{JY} = (z^2+1)\frac{\psi_i^2}{\hat{A}^2(\hat{A}+\psi_i)^3}\frac{(y_i-\mathbf{x}_i'\tilde{\beta})^2}{\hat{A}+\psi_i-\mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i}V(\hat{A})$.

Proof: See Theorem 4.1.2.

Corollary 4.1.5. The CP of $I^{CEBJY1}(\hat{A})$ based on JY1 is written as

$$P[\theta_i \in I^{CEBJY1}(\hat{A})] = 2\Phi(z) - 1 - \frac{z\phi(z)}{4}\eta_{JY1} + O(m^{-3/2})$$

where $\eta_{JY1} = (z^2+1)\frac{\psi_i^2}{\hat{A}^2(\hat{A}+\psi_i)^3}\left(\hat{A}+\psi_i-\mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i\right)V(\hat{A})$.

Proof: See Theorem 4.1.2.

4.2 Confidence Intervals for the Difference between Two Small Area Means

In this section we are interested in the difference of two small area means rather than in the specific values of the small area means themselves. The difference between estimated means (i.e., direct or EBLUP) gives information about the difference between population means. This is also an extension of Diao et al. (2013) CIs for $\theta_i - \theta_j$ where $i \neq j$ based on EBLUP and weighted estimator with fixed weights. Diao et al. (2013) constructed CIs for $\theta_i - \theta_j$ based on $\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}$ with coverage accuracy $O(m^{-2})$

for unequal sampling error variances. In this thesis, we develop closed form CIs for $\theta_i - \theta_j$ based on $\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}$ and $\hat{\theta}_i^{w} - \hat{\theta}_j^{w}$ under unequal sampling error variances and CP is correct to $O(m^{-3/2})$. We derive the theorem similarly to Diao et al. (2013).

Theorem 4.2.1. $\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB} \pm \hat{\theta}_j^{EB}) = \widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB}) + \widehat{\mathrm{MSE}}(\hat{\theta}_j^{EB})$

$$\pm \frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)} \mathbf{x}_i' \left(\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right)^{-1} \mathbf{x}_j.$$

The proof for this theorem is given in the Appendix B.

4.2.1 Confidence Intervals for $\theta_i - \theta_j$ based on the Weighted Estimator with Fixed Weights

Theorem 4.2.2. For any $z > 0$ and under the assumptions 1, 2 and 4, the CP of the improved CIs for $\theta_i - \theta_j$ where $i \neq j$ based on $\hat{\theta}_i^{w} - \hat{\theta}_j^{w}$ is given by

$$P\left[\theta_i - \theta_j \in \hat{\theta}_i^{w} - \hat{\theta}_j^{w} \pm z \times \sqrt{\widehat{\mathrm{MSE}}(\hat{\theta}_i^{w} - \hat{\theta}_j^{w})}\right]$$
$$= 2\Phi(z) - 1 - z\phi(z)\eta_{WD} + O(m^{-3/2})$$

where

$$\eta_{WD} = \frac{z^2+1}{4}\left(\frac{A\psi_i}{A+\psi_i} + \frac{A\psi_j}{A+\psi_j}\right)^{-2}\left(\frac{\psi_i^2}{(A+\psi_i)^2} + \frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A})$$
$$+ \left(\frac{A\psi_i}{A+\psi_i} + \frac{A\psi_j}{A+\psi_j}\right)^{-1}\left[(\gamma_i - w_i)^2 \mathbf{x}_i' \left(\sum \frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1} \mathbf{x}_i\right.$$
$$- \frac{\psi_j^2}{(A+\psi_j)^2}\mathbf{x}_i' \left(\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i}\right)^{-1} \mathbf{x}_i + (\gamma_j - w_j)^2 \mathbf{x}_i' \left(\sum \frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1} \mathbf{x}_i$$
$$\left. + \frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i' \left(\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right)^{-1} \mathbf{x}_j\right].$$

$\Phi(z)$ and $\phi(z)$ are the CDF and pdf of the normal distribution functions respectively and z is an upper percentile of the standard normal distribution function.

Proof. This Theorem is also proved similarly to the previous Theorem. The CP of $I^{CWD}(\hat{A})$ is written as

$$P\left[\theta_i - \theta_j \in I^{CWD}(\hat{A})\right] = P\left[-z + F(-z) < \frac{\theta_i - \theta_j - (\hat{\theta}_i^{w} - \hat{\theta}_j^{w})}{\sqrt{g_{1i}(A) + g_{1j}(A)}} < z + F(z)\right]$$
$$= E[\Phi(z + F(z)) - \Phi(-z + F(-z))]$$

where $F(z) = \dfrac{z \times \left(\sqrt{\widehat{\mathrm{MSE}}(\hat{\theta}_i^{w} - \hat{\theta}_j^{w})} - \sqrt{g_{1i}(A) + g_{1j}(A)}\right) + \hat{\theta}_i^{w} - \tilde{\theta}_i^{B} - (\hat{\theta}_j^{w} - \tilde{\theta}_j^{B})}{\sqrt{g_{1i}(A) + g_{1j}(A)}}$,

$g_{1i}(A) = \frac{A\psi_i}{A+\psi_i}$, $g_{1j}(A) = \frac{A\psi_j}{A+\psi_j}$.

Using the Taylor series expansion with an integral remainder term, $\Phi(z+G(z))$ is evaluated as

$$\Phi(z+F(z)) = \Phi(z)+F(z)\phi(z)+\tfrac{1}{2}F^2(z)\phi'(z)+\tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2\phi''(x)dx$$

$$= \Phi(z)+[F(z)-\tfrac{z}{2}F^2(z)]\phi(z)-\tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2(1-x^2)\phi(x)dx.$$

Now observe that

$$E(\Phi(z+F(z))) = \Phi(z)+E\left[F(z)\phi(z)-\tfrac{z}{2}F^2(z)\phi(z)\right]$$
$$-E\left[\tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2(1-x^2)\phi(x)dx\right].$$

Similarly to a one population case, we have

$$\hat{\theta}_i^w-\tilde{\theta}_i^B-\left\{\hat{\theta}_j^w-\tilde{\theta}_j^B\right\} = (1-w_i)y_i+w_i\mathbf{x}_i'\hat{\beta}-(1-\gamma_i)y_i-\gamma_i\mathbf{x}_i'\tilde{\beta}$$
$$-\left\{(1-w_j)y_j+w_j\mathbf{x}_j'\hat{\beta}-(1-\gamma_j)y_j-\gamma_j\mathbf{x}_j'\tilde{\beta}\right\}$$
$$= (\gamma_i-w_i)(y_i-\mathbf{x}_i'\tilde{\beta}(A))-w_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)$$
$$-(\gamma_j-w_j)(y_j-\mathbf{x}_j'\tilde{\beta}(A))+w_j\mathbf{x}_j'\tilde{\beta}^{(1)}(A)(\hat{A}-A)$$
$$= \left((\gamma_i-w_i)(y_i-\mathbf{x}_i'\tilde{\beta}(A))-(\gamma_j-w_j)(y_j-\mathbf{x}_j'\tilde{\beta}(A))\right)$$
$$+\left(w_j\mathbf{x}_j'-w_i\mathbf{x}_i'\right)\tilde{\beta}^{(1)}(A)(\hat{A}-A).$$

Thus F(z) can be decomposed into three components as follows:

$$F(z) = \frac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)}-\sqrt{g_{1i}(A)+g_{1j}(A)}\right)}{\sqrt{g_{1i}(A)+g_{1j}(A)}}+\frac{(\gamma_i-w_i)(y_i-\mathbf{x}_i'\tilde{\beta}(A))-(\gamma_j-w_j)(y_j-\mathbf{x}_j'\tilde{\beta}(A))}{\sqrt{g_{1i}(A)+g_{1j}(A)}}$$
$$+\frac{\left(w_j\mathbf{x}_j'-w_i\mathbf{x}_i'\right)\tilde{\beta}^{(1)}(A)(\hat{A}-A)}{\sqrt{g_{1i}(A)+g_{1j}(A)}}$$
$$= Q_1+Q_2^*+Q_3^* \quad \text{say.}$$

$$Q_1 = \frac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)}-\sqrt{g_{1i}(A)+g_{1j}(A)}\right)}{\sqrt{g_{1i}(A)+g_{1j}(A)}}.$$

Applying the Taylor series expansion

$$Q_1 = (g_{1i}(A)+g_{1j}(A))^{-1}\tfrac{z}{2}\left(\widehat{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)$$
$$-(g_{1i}(A)+g_{1j}(A))^{-2}\tfrac{z}{8}\left(\widehat{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)^2$$
$$+(g_{1i}(A)+g_{1j}(A))^{-3}\tfrac{z}{16}\left(\widehat{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)^3 + R.$$

Taking expectations on both sides

$$EQ_1 = (g_{1i}(A)+g_{1j}(A))^{-1}\tfrac{z}{2}E\left(\widehat{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)$$
$$-(g_{1i}(A)+g_{1j}(A))^{-2}\tfrac{z}{8}E\left(\widehat{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)^2$$
$$+(g_{1i}(A)+g_{1j}(A))^{-3}\tfrac{z}{16}E\left(\widehat{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)^3.$$

Let us evaluate each term turn by turn

$$E\left(\widehat{\mathrm{MSE}}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)$$

$$= \mathrm{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)-g_{1i}(A)-g_{1j}(A)$$

$$= \mathrm{MSE}(\hat{\theta}_i^w)+\mathrm{MSE}(\hat{\theta}_j^w)-\frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j - g_{1i}(A)-g_{1j}(A)$$

$$= g_{1i}(A)+g_{2i}(A)+g_{3wi}(A)+g_{1j}(A)+g_{2j}(A)+g_{3wj}(A)$$

$$- g_{1i}(A)-g_{1j}(A)-\frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j$$

$$= g_{2i}(A)+g_{3wi}(A)+g_{2j}(A)+g_{3wj}(A)-\frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j.$$

The second term can be evaluated as follows

$$E\left(\widehat{\mathrm{MSE}}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)^2.$$

Using the Taylor series expansion

$$\widehat{\mathrm{MSE}}(\hat{\theta}_i^w-\hat{\theta}_j^w) = g_{1i}(A)+g_{1j}(A)+(g_{1i}(A)+g_{1j}(A))^{(1)}(\hat{A}-A)$$

$$\widehat{\mathrm{MSE}}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A)) = (g_{1i}(A)+g_{1j}(A))^{(1)}(\hat{A}-A).$$

Substituting reveals that

$$E\left(\widehat{\mathrm{MSE}}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)^2$$

$$= E\left((g_{1i}^{(1)}(A)+g_{1j}^{(1)}(A))(\hat{A}-A)\right)^2$$

$$= \left(g_{1i}^{(1)}(A)+g_{1j}^{(1)}(A)\right)^2 E(\hat{A}-A)^2$$

$g_{1i}^{(1)}(A) = \frac{\psi_i^2}{(A+\psi_i)^2}$, $g_{1j}^{(1)}(A) = \frac{\psi_j^2}{(A+\psi_j)^2}$ and $E(\hat{A}-A)^2 = V(\hat{A})$

$$E\left(\widehat{\mathrm{MSE}}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)^2$$

$$= \left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 E(\hat{A}-A)^2$$

$$= \left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A}).$$

Similarly, the third term can be simplified as

$$(g_{1i}(A)+g_{1j}(A))^{-3}\frac{z}{16}E\left(\widehat{\mathrm{MSE}}(\hat{\theta}_i^w-\hat{\theta}_j^w)-(g_{1i}(A)+g_{1j}(A))\right)^3 = O(m^{-3/2}).$$

Combining all the above expressions reveals that

$$EQ_1 = (g_{1i}(A)+g_{1j}(A))^{-1}\frac{z}{2}\left(g_{2i}(A)+g_{3wi}(A)+g_{2j}(A)+g_{3wj}(A)\right.$$

$$\left. -\frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j\right)$$

$$-(g_{1i}(A)+g_{1j}(A))^{-2}\frac{z}{8}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A})+O(m^{-3/2}).$$

Let us now evaluate Q_1^2 as follows

$$Q_1^2 = \left(\frac{z \times \left(\sqrt{\widehat{\mathrm{MSE}}(\hat{\theta}_i^w - \hat{\theta}_j^w)} - \sqrt{g_{1i}(A) + g_{1j}(A)} \right)}{\sqrt{g_{1i}(A) + g_{1j}(A)}} \right)^2.$$

Using the Taylor series expansion and taking expectations on both sides, we have

$$E(Q_1^2) = (g_{1i}(A) + g_{1j}(A))^{-2} \tfrac{z^2}{4} E\left(\widehat{\mathrm{MSE}}(\hat{\theta}_i^w - \hat{\theta}_j^w) - (g_{1i}(A) + g_{1j}(A)) \right)^2$$

$$= \left(\tfrac{A\psi_i}{(A+\psi_i)} + \tfrac{A\psi_j}{(A+\psi_j)}\right)^{-2} \tfrac{z^2}{4} \left(\tfrac{\psi_i^2}{(A+\psi_i)^2} + \tfrac{\psi_j^2}{(A+\psi_j)^2} \right)^2 V(\hat{A}) + O(m^{-3/2}).$$

$$Q_2^* = \frac{(\gamma_i - w_i)(y_i - \mathbf{x}_i'\tilde{\beta}(A)) - (\gamma_j - w_j)(y_j - \mathbf{x}_j'\tilde{\beta}(A))}{\sqrt{g_{1i}(A) + g_{1j}(A)}}$$

$$= \frac{(\gamma_i - w_i)(y_i - \mathbf{x}_i'\beta) - (\gamma_j - w_j)(y_j - \mathbf{x}_j'\beta)}{\sqrt{g_{1i}(A) + g_{1j}(A)}} + o(m^{-1}).$$

Using the Stein identity, $EQ_2^* = 0$

$$Q_2^{*2} = \left(\frac{(\gamma_i - w_i)(y_i - \mathbf{x}_i'\beta) - (\gamma_j - w_j)(y_j - \mathbf{x}_j'\beta)}{\sqrt{g_{1i}(A) + g_{1j}(A)}} \right)^2$$

$$= \left(g_{1i}(A) + g_{1j}(A) \right)^{-1} \left((\gamma_i - w_i)(y_i - \mathbf{x}_i'\beta) - (\gamma_j - w_j)(y_j - \mathbf{x}_j'\beta) \right)^2$$

$$= \left(g_{1i}(A) + g_{1j}(A) \right)^{-1} (\gamma_i - w_i)^2 (y_i - \mathbf{x}_i'\beta)^2$$

$$- 2 \left(g_{1i}(A) + g_{1j}(A) \right)^{-1} (\gamma_i - w_i)(\gamma_j - w_j)(y_i - \mathbf{x}_i'\beta)(y_j - \mathbf{x}_j'\beta)$$

$$+ \left(g_{1i}(A) + g_{1j}(A) \right)^{-1} (\gamma_j - w_j)^2 (y_j - \mathbf{x}_j'\beta)^2$$

$$= Q_{21}^{*2} + Q_{22}^{*2} + Q_{23}^{*2}.$$

The three terms can be simplified using the Stein identity as follows

$$EQ_{21}^{*2} = \left(g_{1i}(A) + g_{1j}(A) \right)^{-1} (\gamma_i - w_i)^2 (A + \psi_i)$$

$$EQ_{23}^{*2} = \left(g_{1i}(A) + g_{1j}(A) \right)^{-1} (\gamma_j - w_j)^2 (A + \psi_j)$$

and

$$EQ_{22}^{*2} = 0$$

$$EQ_2^{*2} = EQ_{21}^{*2} + EQ_{22}^{*2} + EQ_{23}^{*2}$$

$$= \left(g_{1i}(A) + g_{1j}(A) \right)^{-1} \left((\gamma_i - w_i)^2 (A + \psi_i) + (\gamma_j - w_j)^2 (A + \psi_j) \right)$$

$$Q_3^* = \frac{\left(w_j \mathbf{x}_j' - w_i \mathbf{x}_i' \right) \tilde{\beta}^{(1)}(A)(\hat{A} - A)}{\sqrt{g_{1i}(A) + g_{1j}(A)}}.$$

Similarly we can easily show that the expected value of Q_3^*, $EQ_3^* = 0$.

$$Q_3^{*2} = \left(\frac{\left(w_j \mathbf{x}_j' - w_i \mathbf{x}_i' \right) \tilde{\beta}^{(1)}(A)(\hat{A} - A)}{\sqrt{g_{1i}(A) + g_{1j}(A)}} \right)^2$$

$$= \left(g_{1i}(A) + g_{1j}(A) \right)^{-1} \left(\left(w_j \mathbf{x}_j' - w_i \mathbf{x}_i' \right) \tilde{\beta}^{(1)}(A)(\hat{A} - A) \right)^2$$

$$= \left(g_{1i}(A)+g_{1j}(A)\right)^{-1}\left(\left(w_j\mathbf{x}_j'-w_i\mathbf{x}_i'\right)\tilde{\beta}^{(1)}(A)\right)^2(\hat{A}-A)^2.$$

$$EQ_3^{*2} = \left(g_{1i}(A)+g_{1j}(A)\right)^{-1}\left(\left(w_j\mathbf{x}_j'-w_i\mathbf{x}_i'\right)\tilde{\beta}^{(1)}(A)\right)^2 E(\hat{A}-A)^2$$

$$= \left(g_{1i}(A)+g_{1j}(A)\right)^{-1}\left(\left(w_j\mathbf{x}_j'-w_i\mathbf{x}_i'\right)\tilde{\beta}^{(1)}(A)\right)^2 V(\hat{A}),$$

since $\tilde{\beta}^{(1)}(A) = O_P(m^{-1/2})$ (Datta et al., 2011).

$$Q_1Q_2^* = \frac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^w-\hat{\theta}_j^w)}-\sqrt{g_{1i}(A)+g_{1j}(A)}\right)}{\sqrt{g_{1i}(A)+g_{1j}(A)}}\frac{(\gamma_i-w_i)(y_i-\mathbf{x}_i'\tilde{\beta}(A))-(\gamma_j-w_j)(y_j-\mathbf{x}_j'\tilde{\beta}(A))}{\sqrt{g_{1i}(A)+g_{1j}(A)}}.$$

Using the Taylor series expansion

$$Q_1Q_2^* = \left(g_{1i}(A)+g_{1j}(A)\right)^{-1}\left((\gamma_i-w_i)(y_i-\mathbf{x}_i'\beta)-(\gamma_j-w_j)(y_j-\mathbf{x}_j'\beta)\right)$$

$$\times z\times\left(g_{1i}^{(1)}(A)+g_{1j}^{(1)}(A)\right)(\hat{A}-A)+o(m^{-1})$$

$$= \left(g_{1i}(A)+g_{1j}(A)\right)^{-1}\Big[(\gamma_i-w_i)(y_i-\mathbf{x}_i'\beta)z\times\left(g_{1i}^{(1)}(A)+g_{1j}^{(1)}(A)\right)(\hat{A}-A)$$

$$-(\gamma_j-w_j)(y_j-\mathbf{x}_j'\beta)z\times\left(g_{1i}^{(1)}(A)+g_{1j}^{(1)}(A)\right)(\hat{A}-A)\Big].$$

Using the Stein identity, we have

$$EQ_1Q_2^* = E\left(g_{1i}(A)+g_{1j}(A)\right)^{-1}z\Big[(\gamma_i-w_i)(A+\psi_i)\times\left(g_{1i}^{(1)}(A)+g_{1j}^{(1)}(A)\right)\frac{\partial\hat{A}}{\partial y_i}$$

$$-(\gamma_j-w_j)(A+\psi_j)\times\left(g_{1i}^{(1)}(A)+g_{1j}^{(1)}(A)\right)\frac{\partial\hat{A}}{\partial y_j}\Big],$$

since $\frac{\partial\hat{A}}{\partial y_j} = O_P(m^{-1})$. Similarly the contributions of $Q_1Q_3^*$ and $Q_2Q_3^*$ are negligible. Note that the remainder and the integral terms can be simplified using a similar argument to Theorem 4.1.1. Thus, combination of all the above expressions reveals that

$$E(F(z)-F(-z)) = 2\left(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j}\right)^{-1}\frac{z}{2}\frac{\psi_i^2}{(A+\psi_i)^2}\mathbf{x}_i'(\textstyle\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i})^{-1}\mathbf{x}_i$$

$$+2\frac{z}{2}\left(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j}\right)^{-1}\Big[(\gamma_i-w_i)^2((A+\psi_i)$$

$$-(\gamma_i-w_i)^2\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i$$

$$+\frac{\psi_j^2}{(A+\psi_j)^2}\mathbf{x}_i'(\textstyle\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i})^{-1}\mathbf{x}_i+(\gamma_j-w_j)^2((A+\psi_j)$$

$$-(\gamma_j-w_j)^2\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i$$

$$-\frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\textstyle\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j$$

$$-\frac{1}{4}(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j})^{-1}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A})\Big]$$

$$+O(m^{-3/2}).$$

$$-\frac{z}{2}E(F(z)^2+F(-z)^2) = -z\left(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j}\right)^{-2}\frac{z^2}{4}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A})$$

$$-\, z\left(\frac{A\psi_i}{(A+\psi_i)}+\frac{A\psi_j}{(A+\psi_j)}\right)^{-1}\left[(\gamma_i-w_i)^2(A+\psi_i)+(\gamma_j-w_j)^2(A+\psi_j)\right]$$
$$+O(m^{-3/2}).$$

Combination of all the above results gives as

$$P\left[\theta_i-\theta_j \in \hat{\theta}_i^{w}-\hat{\theta}_j^{w}\pm z\times\sqrt{\widehat{MSE}(\hat{\theta}_i^{w}-\hat{\theta}_j^{w})}\right]=2\Phi(t)-1-z\phi(z)\eta_{WD}+O(m^{-3/2})$$

where

$$\eta_{WD}=\frac{z^2+1}{4}\left(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j}\right)^{-2}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A})$$
$$+\left(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j}\right)^{-1}\Bigg[(\gamma_i-w_i)^2\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i$$
$$-\frac{\psi_j^2}{(A+\psi_j)^2}\mathbf{x}_i'(\textstyle\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i})^{-1}\mathbf{x}_i+(\gamma_j-w_j)^2\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i$$
$$+\frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\textstyle\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j\Bigg].$$

□

4.2.2 Confidence Intervals for $\theta_i-\theta_j$ based on EBLUP

Similarly, we also obtain improved CIs for the difference of two small area population means, $\theta_i-\theta_j$ where $i\neq j$ based on EBLUP estimator. The CI for the difference of two small area means will be given by the following theorem.

Theorem 4.2.3. For any $z>0$

$$P\left[\theta_i-\theta_j \in I^{CEBD}(\hat{A})\right]=P\left[\theta_i-\theta_j \in \hat{\theta}_i^{EB}-\hat{\theta}_j^{EB}\pm z\times\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB}-\hat{\theta}_j^{EB})}\right]$$
$$=2\Phi(z)-1-\frac{z\phi(z)}{4}\eta_D+O(m^{-3/2})$$

where

$$\eta_D=(z^2+1)(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j})^{-2}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A}).$$

The proof for this theorem is also given in the Appendix C.

Consider the following three corollaries which follow Theorem 4.2.3 by the Taylor series expansion. These are the area specific CIs for $\theta_i-\theta_j$ based on $\hat{\theta}_i^{EB}-\hat{\theta}_j^{EB}$, where $i\neq j$.

Corollary 4.2.4. The CP of $I^{CEBDRao}(\hat{A})$ based on the Rao (2001) is written as

$$P[\theta_i-\theta_j \in I^{CEBDRao}(\hat{A})]=2\Phi(z)-1-\frac{z\phi(z)}{4}\eta_{RaoD}+O(m^{-3/2})$$

where

$$\eta_{RaoD}=\frac{(z_{\alpha/2}^2+1)}{8}\left(g_{1i}(A)+g_{1j}(A)\right)^{-2}\frac{\left((A+\psi_i)g_{3Raoi}(A)+(A+\psi_j)g_{3Raoj}(A)\right)^2}{V(A)}$$

$g_{3Raoi}(\hat{A}) = \frac{\psi_i^2}{(\hat{A}+\psi_i)^4}(y_i - \mathbf{x}_i'\hat{\beta})^2 V(\hat{A})$; $g_{3Raoj}(\hat{A}) = \frac{\psi_j^2}{(\hat{A}+\psi_j)^4}(y_j - \mathbf{x}_j'\hat{\beta})^2 V(\hat{A})$.

Proof: See Theorem 4.2.3.

Corollary 4.2.5. The CP of $I^{CEBDJY}(\hat{A})$ based on JY is written as

$$P[\theta_i - \theta_j \in I^{CEBDJY}(\hat{A})] = 2\Phi(z) - 1 - \frac{z\phi(z)}{4}\eta_{JYD} + O(m^{-3/2})$$

where

$$\eta_{JYD} = \frac{(z_{\alpha/2}^2+1)}{8}\left(g_{1i}(A) + g_{1j}(A)\right)^{-2}\frac{\left((A+\psi_i)g_{3JYi}(A)+(A+\psi_j)g_{3JYj}(A)\right)^2}{V(A)}$$

$$g_{3JYi}(\hat{A}) = \frac{\psi_i^2}{(\hat{A}+\psi_i)^4}\frac{(y_i-\mathbf{x}_i'\tilde{\beta})^2}{\hat{A}+\psi_i-\mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i}V(\hat{A});$$

$$g_{3JYj}(\hat{A}) = \frac{\psi_j^2}{(\hat{A}+\psi_j)^4}\frac{(y_j-\mathbf{x}_j'\tilde{\beta})^2}{\hat{A}+\psi_j-\mathbf{x}_j'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_j}V(\hat{A}).$$

Proof: See Theorem 4.2.3.

Corollary 4.2.6. The CP of $I^{CEBDJY1}(\hat{A})$ based on JY1 is written as

$$P[\theta_i - \theta_j \in I^{CEBDJY1}(\hat{A})] = 2\Phi(z) - 1 - \frac{z\phi(z)}{4}\eta_{JY1D} + O(m^{-3/2})$$

where

$$\eta_{JY1D} = \frac{(z_{\alpha/2}^2+1)}{8}\left(g_{1i}(A) + g_{1j}(A)\right)^{-2}\frac{\left((A+\psi_i)g_{3JY1i}(A)+(A+\psi_j)g_{3JY1j}(A)\right)^2}{V(A)}$$

$$\tilde{g}_{3JY1i}(A) = g_{3i}(A) - g_{5i}(A), \textit{ where } \quad g_{5i}(A) = \frac{g_{2i}(A)}{(A+\psi_i)^2}V(\hat{A})$$

$$\tilde{g}_{3JY1j}(A) = g_{3j}(A) - g_{5j}(A), \textit{ where } \quad g_{5j}(A) = \frac{g_{2j}(A)}{(A+\psi_j)^2}V(\hat{A}).$$

Proof: See Theorem 4.2.3.

4.3 Confidence Intervals Based on EBLUP vs Weighted Estimator with Fixed Weights

4.3.1 Confidence Intervals for a Small Area Parameter Based on EBLUP vs Weighted Estimator with Fixed Weights

The symmetric CI for θ_i based on $\hat{\theta}_i^{EB}$ is written as:

$$\text{Method III}: \hat{\theta}_i^{EB} \pm t_\alpha^* \sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}, \tag{4.3.1}$$

where $t_\alpha^* = z_{\alpha/2}(1 + h(\hat{A}))$, $h(\hat{A}) = (z_{\alpha/2}^2 + 1)\frac{\psi_i^2}{8\hat{A}^2(\hat{A}+\psi_i)^2}V(\hat{A})$.

The symmetric CI for θ_i based on $\hat{\theta}_i^w$ is given by

$$\text{Method IIIW}: \hat{\theta}_i^w \pm t_{\alpha w}\sqrt{\widehat{\text{MSE}}(\hat{\theta}_i^w)}, \tag{4.3.2}$$

where $t_{\alpha w} = t_\alpha^* + t_w$,
$t_w = z_{\alpha/2}\left\{\frac{\psi_i^2+(\gamma_i-w_i)^2(A+\psi_i)^2}{2A\psi_i(A+\psi_i)}\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i\right\}$.

4.3.2 Confidence Intervals for the Difference between Two Small Area Means Based on EBLUP vs Weighted Estimator with Fixed Weights

The improved CIs for $\theta_i-\theta_j$ based on $\hat{\theta}_i^{EB}-\hat{\theta}_j^{EB}$ where $i\neq j$ is written as:

$$\text{Method IIID}: \hat{\theta}_i^{EB}-\hat{\theta}_j^{EB} \pm t_{\alpha_D}\sqrt{\widehat{\text{MSE}}(\hat{\theta}_i^{EB})-\widehat{\text{MSE}}(\hat{\theta}_j^{EB})}, \tag{4.3.3}$$

where $t_{\alpha_D} = z_{\alpha/2}\left[\frac{(z_{\alpha/2}^2+1)}{8}\left(g_{1i}(A)+g_{1j}(A)\right)^{-2}\frac{\left((A+\psi_i)g_{3i}(A)+(A+\psi_j)g_{3j}(A)\right)^2}{V(A)}+1\right]$.

The improved CI for $\theta_i-\theta_j$ based on $\hat{\theta}_i^w-\hat{\theta}_j^w$ where $i\neq j$ is written as:

$$\text{Method IIIWD}: \hat{\theta}_i^w-\hat{\theta}_j^w \pm t_{\alpha_{WD}}\sqrt{\widehat{\text{MSE}}(\hat{\theta}_i^w)-\widehat{\text{MSE}}(\hat{\theta}_j^w)}, \tag{4.3.4}$$

where $t_{\alpha_{WD}} = t_{\alpha_D} + t_{wD}$,

$$t_{wD} = +\, z_{\alpha/2}\left(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j}\right)^{-1}\left[(\gamma_i-w_i)^2\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i\right.$$
$$-\frac{\psi_j^2}{(A+\psi_j)^2}\mathbf{x}_i'\left(\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i}\right)^{-1}\mathbf{x}_i+(\gamma_j-w_j)^2\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i$$
$$\left.+\ \frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j\right].$$

Chapter 5

Finite Sample Performance of MSEs and CIs

This chapter discusses the main results on area specific MSEs, improved CIs, extension to CIs for the difference between two small area means and comparative discussion of the proposed CIs in relation to some existing CIs in the literature.

5.1 Comparison of the MSE Estimators

MSE is the most commonly used method of assessing uncertainty of the EBLUP estimator. It is an important component of small area estimation research (Diao et al., 2013). In this section we perform extensive simulations to compare the proposed methods with some of the MSEs in the literature.

5.1.1 Simulation Study I

The RB and the RR expressions of MSE estimator are given by

$$\mathrm{RB_i} = 100 \times \frac{\mathrm{E}\left[\widehat{\mathrm{MSE}}(\hat{\theta}_\mathrm{i}^{\mathrm{EB}}) - \mathrm{MSE}(\hat{\theta}_\mathrm{i}^{\mathrm{EB}})\right]}{\mathrm{MSE}(\hat{\theta}_\mathrm{i}^{\mathrm{EB}})}, \tag{5.1.1}$$

$$\mathrm{RR_i} = 100 \times \frac{\mathrm{E}\left\{\widehat{\mathrm{MSE}}(\hat{\theta}_\mathrm{i}^{\mathrm{EB}}) - \mathrm{MSE}(\hat{\theta}_\mathrm{i}^{\mathrm{EB}})\right\}^2}{\left\{\mathrm{MSE}(\hat{\theta}_\mathrm{i}^{\mathrm{EB}})\right\}^2} \tag{5.1.2}$$

where $E(\widehat{MSE}(\hat{\theta}_i^{EB}))$ can be estimated empirically as the average value of $\widehat{MSE}(\hat{\theta}_i^{EB})$ over replicates. Note that equations (5.1.1) and (5.1.2) are equations given by Kubokawa and Nagashima (2012) multiplied by 100.

Likewise, $MSE(\hat{\theta}_i^{EB})$ can be defined as the average value of $(\hat{\theta}_i^{EB}-\theta_i)^2$ over replicates. $\widehat{MSE}(\hat{\theta}_i^{EB}) = \frac{1}{R}\sum_{i=1}^{R}\widehat{MSE}(\hat{\theta}_i^{EB(r)})$, the true MSE of $\hat{\theta}_i^{EB}$ is approximated as $MSE(\hat{\theta}_i^{EB}) = \frac{1}{R}\sum_{i=1}^{R}(\theta_i^{EB(r)}-\theta_i^{(r)})^2$, $R = 10,000$, number of replications. We calculate the true values of $MSE(\hat{\theta}_i^{EB})$ in advance, which can be obtained based on 10,000 simulated data sets. A simulation study is conducted using SAS software package to investigate the MSE estimators. We adopt the SAS simulation codes by Li (2007) with modification to check the performance of MSE estimators. Li (2007) studied the adjusted density method (ADM) following Morris' comments under the Fay-Herriot model. They proposed EBLUP estimator of θ_i under the use of ADM estimators with their associated MSE estimators using the Taylor series expansion. They also proposed a new parametric bootstrap prediction interval method using their proposed ADM estimator. Next, we describe the simulation set up of the study and subsequently present the results obtained.

5.1.2 Design of Simulation Study

We consider the simulation set-up by Datta et al. (2005) with minor modifications. For the FH model with the number of areas, m = 10, 15, 30, 45 and 60, $\mathbf{x}_i = (1, \mathbf{x}_i)'$, where $\mathbf{x}_i; i = 1, ..., m$ were generated i.i.d. from $N(0,1)$, $\beta = (1,1)'$ and $A = 1$. The sampling errors, e_i, are generated from $N(0, \psi_i)$ for ψ_i specified by the following three different variance patterns: pattern (I) 0.7, 0.6, 0.5, 0.4, 0.3, pattern (II) 2, 0.6, 0.5, 0.4, 0.2 and pattern (III) 4, 0.6, 0.5, 0.4, 0.1. According to Rao (2003a) pattern(I) is nearly balanced ($\frac{\max(\psi_i)}{\min(\psi_i)} = 2.3$), pattern(II) has intermediate variability ($\frac{\max(\psi_i)}{\min(\psi_i)} = 10$) and pattern(III) has the highest variability ($\frac{\max(\psi_i)}{\min(\psi_i)} = 40$). The random effects, v_i are generated from two different distributions, namely the normal $N(0,1)$ and the Laplace $(0,1)$ distributions. We considered two different cases; (i) without covariates: $x_i\beta = 0$, we generate 10,000 data sets from $y_i = v_i + e_i$, $(i = 1, \cdots, m)$ and (ii) with covariates: $\beta = (1,1)'$, similarly, we generate 10,000 data sets from $y_i = x_i\beta + v_i + e_i$, $(i = 1, \cdots, m)$ to account for the estimation of the unknown regression parameters which comes from the real data applications. "The small areas are divided into five equal sized groups, and the ψ_i's remain the same within each group" (Diao et al., 2013, page 7).

"We simulated the MSEs using the 10000 data sets for each combination of method of estimating A and distribution of v_i, and averaged over the small areas with the same ψ_i values within each pattern" (Datta et al., 2005, page 188). The MSE estimators were also calculated and their RBs and RRs over Monte Carlo samples were computed.

In this section the small sample properties of MSE estimators for the normal and the Laplace random effects distribution are investigated by means of a simulation study. The main goal of the simulation study is to assess the performances of different MSE estimators for different variance patterns using empirical measures of RB and RR. We present the four different MSE estimators based on the REML, ML and FH methods as follows.

The four different MSE estimators based on the REML estimators are given by:

$$\texttt{MSE\_Rao} \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g^*_{3Raoi}(\hat{A}, y_i) \tag{5.1.3}$$

$$\texttt{MSE\_JY} \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g^*_{3JYi}(\hat{A}, y_i) \tag{5.1.4}$$

$$\texttt{MSE\_JY1} \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g^*_{3JY1i}(\hat{A}) \tag{5.1.5}$$

$$\texttt{MSE\_REML} \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}(\hat{A}) \tag{5.1.6}$$

where $\hat{A}$ is the estimate of the variance of the random effect A based on the REML estimator, $g^*_{3Raoi}(\hat{A}, y_i) = \frac{\psi_i^2}{(\psi_i+\hat{A})^4}(y_i - \mathbf{x}_i'\tilde{\beta})^2 V(\hat{A})$, $g^*_{3JYi}(\hat{A}, y_i) = \frac{\psi_i^2}{(\psi_i+\hat{A})^4}\frac{(y_i-\mathbf{x}_i'\tilde{\beta})^2}{\hat{A}+\psi_i-\mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i} V(\hat{A})$, $g^*_{3JY1i}(\hat{A}) = g_{3i}(\hat{A}) - \frac{g_{2i}(\hat{A})}{(A+\psi_i)^2} V(\hat{A})$, $g_{3i}(\hat{A}) = \frac{\psi_i^2}{(A+\psi_i)^3} V(\hat{A})$. Note that `MSE_Rao`, `MSE_JY`, `MSE_JY1` and `MSE_REML` are the MSE estimators based on the Rao (2001), JY, JY1 and REML Estimators (Datta and Lahiri, 2000) $g_{3i}(A)'$s respectively. `MSE_Rao` and `MSE_JY` are area specific MSE estimators.

The four different MSE estimators based on the ML Estimators (Datta and Lahiri, 2000) are given by:

$$\texttt{MSE\_Rao} \approx g_{1i}(\hat{A}) - b_{\hat{A}}(\hat{A})\nabla g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g^*_{3Raoi}(\hat{A}, y_i) \tag{5.1.7}$$

$$\texttt{MSE\_JY} \approx g_{1i}(\hat{A}) - b_{\hat{A}}(\hat{A})\nabla g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g^*_{3JYi}(\hat{A}, y_i) \tag{5.1.8}$$

$$\texttt{MSE\_JY1} \approx g_{1i}(\hat{A}) - b_{\hat{A}}(\hat{A})\nabla g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g^*_{3JY1i}(\hat{A}) \tag{5.1.9}$$

$$\texttt{MSE\_ML} \approx g_{1i}(\hat{A}) - b_{\hat{A}}(\hat{A})\nabla g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}(\hat{A}) \tag{5.1.10}$$

where $\nabla g_{1i}(\hat{A}) = (\frac{\psi_i}{\hat{A}_{ML}+\psi_i})^2 > 0$. For the ML estimator $\hat{A}_{ML}$, the estimated bias of $\hat{A}_{ML}$ is given by $b_{\hat{A}}(\hat{A}) = \frac{-tr\{[\sum_{u=1}^m(\hat{A}+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u']^{-1}[\sum_{u=1}^m(\hat{A}+\psi_u)^{-2}\mathbf{x}_u\mathbf{x}_u']\}}{\sum_{u=1}^m(\hat{A}+\psi_u)^{-2}}$.

The four different MSE estimators based on the FH Estimator (Datta et al., 2005) are given by:

$$\texttt{MSE\_Rao} \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g^*_{3Raoi}(\hat{A}, y_i) - b_{\hat{A}_{FH}}(\hat{A}_{FH})\left\{B_i(\hat{A}_{FH})\right\}^2 \tag{5.1.11}$$

$$\texttt{MSE\_JY} \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g^*_{3JYi}(\hat{A}, y_i) - b_{\hat{A}_{FH}}(\hat{A}_{FH})\left\{B_i(\hat{A}_{FH})\right\}^2 \tag{5.1.12}$$

$$\texttt{MSE\_JY1} \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g^*_{3JY1i}(\hat{A}) - b_{\hat{A}_{FH}}(\hat{A}_{FH})\left\{B_i(\hat{A}_{FH})\right\}^2 \tag{5.1.13}$$

$$\texttt{MSE\_FH} \approx g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}(\hat{A}) - b_{\hat{A}_{FH}}(\hat{A}_{FH})\left\{B_i(\hat{A}_{FH})\right\}^2 \tag{5.1.14}$$

For the FH estimator $\hat{A}_{FH}$, we need the following expression for its bias to terms of order $O(m^{-1})$: $b_{\hat{A}_{FH}}(\hat{A}_{FH}) = \frac{2\left[m\sum_{i=1}^m(\hat{A}+\psi_i)^{-2} - \left\{\sum_{i=1}^m(\hat{A}+\psi_i)^{-1}\right\}^2\right]}{\left\{\sum_{i=1}^m(\hat{A}+\psi_i)^{-1}\right\}^3}$.

5.1.3 Simulation Results without Covariates

Following Datta et al. (2005), we used a model with no auxiliary variables for the Tables 5.1 - 5.3. However, Tables 5.4 and 5.5 shows results with covariates. These results are shown separately for each of the sampling variances ψ_i, distinguishing among the cases $\hat{A} = \hat{A}_{ML}, \hat{A} = \hat{A}_{REML}$ and $\hat{A} = \hat{A}_{FH}$. The FH, ML and REML corresponds the FH, ML and REML estimators of the MSE of the EBLUP, respectively.

Table 5.1 reports the simulated values of MSE values, averaged over areas within G1, G2, G3, G4 and G5 for the normal distribution without covariates. The values are reported for m = 15, A = 1, and three ψ_i patterns given above. This shows that all the methods produced similar results with respect to MSE estimations for pattern (I).

Table 5.2 shows the RB values of the MSE estimators that were obtained through the simulation study for m = 15 without covariates. Table 5.3 presents the summary of the RR of MSE estimators when m = 15 without covariates. RR is a number that measures the risk of using MSE estimators under certain circumstances. The lower risk MSE estimators perform better than their higher risk counterparts. The RR of MSE_Rao, MSE_JY and MSE_JY1 are nearly zero for the ML, REML and FH methods for all patterns. As shown in Tables 5.2 and 5.3, in terms of RB and RR the MSE estimators based on the ML, REML and FH behave similarly for pattern (I), with no particular one emerging as clearly better than the other three. When m = 15, estimators based on the FH method have very small bias for all variance patterns. The bias of MSE_JY1 estimator is smaller than MSE_Rao and MSE_JY estimators for the ML and REML methods. The bias and risk of all the estimators are less than 7% when $\hat{A} = \hat{A}_{FH}$ for both patterns (II) and (III) are negligible for group G5 and the normal distribution. In both tables, we considered the Laplace distribution in order to assess the robustness of the MSE estimators to possible deviations from the normality assumptions. The RR of MSE estimators based on the FH method are less than 7% in all cases.

Table 5.2 and 5.3 show the RB and RR of MSE estimates based on the FH, ML and REML estimators of A under the three different sampling variances when m = 15 respectively.

Table 5.1: Simulated values of $\widehat{\mathrm{MSE}}(\hat{\theta}_{\mathrm{i}}^{\mathrm{EB}})$ multiplied by 100 for m = 15, A = 1, ψ_{i} patterns (I), (II) and (III) and for the normal random effects distribution

Groups	Pattern I			Pattern II			Pattern III			Pattern I			Pattern II			Pattern III		
	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH
	Usual MSEs									`MSE_JY`								
G1	43.2	42.6	45.7	65.3	70.4	74.7	77.8	85.7	89.4	46.4	45.9	45.8	74.5	75.2	75	88.1	90.1	90
G2	39.4	38.8	41.5	37	38.7	41.7	36.5	38.5	41	42.3	41.8	41.7	42.5	42	41.9	41.8	41.8	41.2
G3	35	34.4	36.9	33	34.3	37	32.5	34.1	36.3	37.7	37.1	37.1	37.9	37.3	37.3	37.2	37	36.5
G4	30	29.5	31.6	28.4	29.4	31.7	27.9	29.2	31	32.4	31.8	31.8	32.6	32	32	31.9	31.7	31.2
G5	24.1	23.8	25.5	16.9	17.2	18.6	9.2	9.3	10	26.4	25.7	25.8	19.7	18.8	19.2	12.3	10.2	10.1
	`MSE_Rao`									`MSE_JY1`								
G1	46.1	45.7	45.6	74.2	74.8	74.7	87.8	89.8	89.7	45.7	45.5	45.5	73.5	74.4	89.3	87.2	89.2	89.3
G2	42.1	41.6	41.5	42.2	41.8	41.7	41.6	41.6	41	41.6	41.3	41.4	41.4	41.3	40.8	40.6	40.9	40.8
G3	37.5	36.9	36.9	37.6	37.1	37.1	37	36.8	36.3	37	36.7	36.7	36.8	36.6	36.1	36	36.2	36.1
G4	32.2	31.6	31.7	32.4	31.8	31.8	31.7	31.5	31.1	31.8	31.4	31.4	31.6	31.3	30.8	30.8	30.9	30.8
G5	26.2	25.6	25.7	19.6	18.7	19.1	12.2	10.2	10.1	25.8	25.3	25.4	18.7	18.2	9.9	10	9.7	9.9

Note that:

- `MSE_Rao`, `MSE_JY`, `MSE_JY1`, `MSE_REML`, `MSE_ML` and `MSE_FH` referring the area specific MSE estimator values given by equations (5.1.3 - 5.1.14).
- Methods of estimating A: ML, maximum likelihood; REML, restricted maximum likelihood; FH, Fay and Herriot (1979).
- Groups: G1, areas 1-3; G2, areas 4-6; G3, areas 7-9; G4, areas 10-12, G5, areas 13-15.

Table 5.2: Simulated values of the RB of $\widehat{\mathrm{MSE}}(\hat{\theta}_{\mathrm{i}}^{\mathrm{EB}})$ for A = 1, m = 15, ψ_{i} pattern (I), (II), (III) and for the normal and the Laplace random effects distributions

Groups	Normal Distribution									Laplace Distribution								
	Pattern I			Pattern II			Pattern III			Pattern I			Pattern II			Pattern III		
	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH
	Usual MSEs																	
G1	-12.4	-7.7	-0.8	-15.4	-8.5	-2.9	-15.6	-6.9	-4.8	-16.7	-9.1	1.9	-20	-9.5	-3.8	-21.9	-7.5	-9.4
G2	-10.9	-6.5	0.4	-12.8	-7.6	-0.4	-14.8	-8.4	-3.3	-15.8	-8.4	3.1	-20	-10.8	-0.6	-27.6	-14.1	-12.3
G3	-11.4	-7.2	-0.6	-13.2	-8.3	-1.3	-15.3	-9	-4.6	-14.5	-7.6	4.3	-19.2	-10.2	0.3	-27.4	-13.7	-12.9
G4	-8.7	-4.8	1.6	-10.6	-5.8	0.9	-13.1	-6.7	-3	-13.5	-6.9	5.8	-18.4	-9.6	1.4	-27.8	-13.8	-13.8
G5	-7.2	-3.8	2.4	-6.5	-2.7	3.5	-10.5	-1.5	-2.5	-9.4	-3.2	10.9	-9.5	-0.7	16.7	-23.8	-1	-1.5
	MSE_Rao																	
G1	-1.1	-1.1	-0.9	-4	-2.7	-3	-4.7	-2.4	-4.5	1.6	1.5	1.7	-5.3	-2.6	-3.5	-9.6	-3	-8.9
G2	0.5	0.3	0.3	-0.5	-0.1	-0.3	-3	-1	-3.3	3.3	2.8	2.8	-0.5	0.9	-0.8	-11.1	-3.7	-12.3
G3	-0.1	-0.5	-0.5	-1	-0.8	-1.1	-3.7	-1.8	-4.5	5.9	4.7	4.2	1.8	2.6	0.3	-10.1	-2.6	-12.7
G4	2.7	2.1	1.9	1.9	1.9	1.3	-1.3	0.7	-2.7	8.8	6.7	6.1	4.6	4.6	1.9	-9.2	-1.8	-13.5
G5	4.5	3.3	3.2	8.3	5.8	6.1	18.2	7.9	-1.5	16.9	12.7	12.1	30.4	21	23.8	63.2	28.7	-1
	MSE_JY																	
G1	-0.6	-0.6	-0.4	-3.6	-2.3	-2.5	-4.4	-2.1	-4.1	2.4	2.2	2.4	-4.8	-2.1	-3	-9.3	-2.7	-8.5
G2	1	0.8	0.8	0.1	0.4	0.2	-2.4	-0.5	-2.8	4.1	3.6	3.6	0.4	1.8	0.1	-10.4	-3	-11.7
G3	0.4	0	0	-0.4	-0.3	-0.6	-3.1	-1.2	-4	6.8	5.5	5.1	2.8	3.5	1.2	-9.2	-1.8	-12.2
G4	3.2	2.6	2.4	2.5	2.4	1.8	-0.7	1.2	-2.2	10	7.7	7.1	5.8	5.6	2.9	-8.3	-1	-12.9
G5	5.2	3.8	3.7	9.1	6.4	6.9	20.1	8.5	-1.2	18.3	13.8	13.3	32.8	22.6	25.9	69	30.8	-0.6
	MSE_JY1																	
G1	-2.1	-1.6	-1.1	-4.8	-3.3	-4.9	-5.4	-3.1	-4.9	-0.1	0.5	1.3	-6.4	-3.6	-4	-10.4	-3.8	-9.5
G2	-0.7	-0.3	-0.1	-2.4	-1.3	-3.8	-5.3	-2.6	-3.8	1.5	1.8	2.4	-3.8	-1.4	-1.2	-14.9	-6.6	-12.8
G3	-1.4	-1.1	-1	-3.1	-2.1	-5	-6.1	-3.4	-5	3.7	3.5	3.5	-2.3	-0.1	-0.5	-14.5	-5.9	-13.4
G4	1.2	1.3	1.1	-0.5	0.4	-3.5	-4.1	-1.2	-3.5	5.9	5.2	4.8	-0.5	1.3	0.3	-14.7	-5.8	-14.3
G5	2.6	2.3	1.9	3.2	3.2	-2.9	-3	2.8	-2.9	12.7	11	9.4	14.8	14.7	13.9	-3.2	11.8	-2.1

Note that:

- MSE_Rao, MSE_JY, MSE_JY1, MSE_REML, MSE_ML and MSE_FH referring the MSE estimator values given by equations (5.1.3 - 5.1.14).
- Methods of estimating A: ML, maximum likelihood; REML, restricted maximum likelihood; FH, Fay and Herriot (1979).
- Groups: G1, areas 1-3; G2, areas 4-6; G3, areas 7-9; G4, areas 10-12, G5, areas 13-15.

Table 5.3: Simulated values of the RR of $\widehat{\text{MSE}}(\hat{\theta}_i^{EB})$ for A = 1, m = 15, ψ_i pattern (I), (II), (III) and for the normal and the Laplace random effects distributions

Groups	Normal Distribution									Laplace Distribution								
	Pattern I			Pattern II			Pattern III			Pattern I			Pattern II			Pattern III		
	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH
	Usual MSEs																	
G1	1.5	0.6	0	2.4	0.7	0.1	2.4	0.5	0.2	2.8	0.8	0.1	4	0.9	0.2	4.8	0.6	0.9
G2	1.2	0.4	0	1.6	0.6	0	2.2	0.7	0.1	2.5	0.7	0.1	4	1.2	0	7.6	2	1.5
G3	1.3	0.5	0	1.8	0.7	0	2.4	0.8	0.2	2.1	0.6	0.2	3.7	1	0	7.5	1.9	1.7
G4	0.8	0.2	0	1.1	0.3	0	1.7	0.5	0.1	1.8	0.5	0.3	3.4	0.9	0	7.7	1.9	1.9
G5	0.5	0.2	0.1	0.4	0.1	0.1	1.1	0	0.1	0.9	0.1	1.2	0.9	0	2.8	5.7	0	0
	MSE_Rao																	
G1	0	0	0	0.2	0.1	0.1	0.2	0.1	0.2	0	0	0	0.3	0.1	0.1	0.9	0.1	0.8
G2	0	0	0	0	0	0	0.1	0	0.1	0.1	0.1	0.1	0	0	0	1.2	0.1	1.5
G3	0	0	0	0	0	0	0.2	0.1	0.2	0.4	0.2	0.2	0	0.1	0	1	0.1	1.6
G4	0.1	0.1	0	0	0	0	0	0	0.1	0.8	0.5	0.4	0.2	0.2	0	0.9	0	1.8
G5	0.2	0.1	0.1	0.7	0.3	0.4	3.3	0.6	0	2.9	1.6	1.5	9.3	4.4	5.7	39.9	8.2	0
	MSE_JY																	
G1	0	0	0	0.1	0.1	0.1	0.2	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.1	0.9	0.1	0.7
G2	0	0	0	0	0	0	0.1	0	0.1	0.2	0.1	0.1	0	0	0	1.1	0.1	1.4
G3	0	0	0	0	0	0	0.1	0	0.2	0.5	0.3	0.3	0.1	0.1	0	0.9	0	1.5
G4	0.1	0.1	0.1	0.1	0.1	0	0	0	0.1	1	0.6	0.5	0.3	0.3	0.1	0.7	0	1.7
G5	0.3	0.2	0.2	0.9	0.4	0.5	4.1	0.7	0	3.4	1.9	1.8	10.8	5.1	6.7	47.6	9.5	0
	MSE_JY1																	
G1	0.1	0	0	0.2	0.1	0.3	0.3	0.1	0.3	0	0	0	0.4	0.1	0.2	1.1	0.1	0.9
G2	0	0	0	0.1	0	0.1	0.3	0.1	0.1	0	0	0.1	0.2	0	0	2.2	0.4	1.6
G3	0	0	0	0.1	0.1	0.3	0.4	0.1	0.3	0.1	0.1	0.1	0.1	0	0	2.1	0.4	1.8
G4	0	0	0	0	0	0.1	0.2	0	0.1	0.4	0.3	0.2	0	0	0	2.2	0.3	2.1
G5	0.1	0.1	0	0.1	0.1	0.1	0.1	0.1	0.1	1.6	1.2	0.9	2.2	2.2	1.9	0.1	1.4	0

Note that:

- MSE_Rao, MSE_JY, MSE_JY1, MSE_REML, MSE_ML and MSE_FH referring the MSE estimator values given by equations (5.1.3 - 5.1.14).
- Methods of estimating A: ML, maximum likelihood; REML, restricted maximum likelihood; FH, Fay and Herriot (1979).
- Groups: G1, areas 1-3; G2, areas 4-6; G3, areas 7-9; G4, areas 10-12, G5, areas 13-15.

5.1.4 Simulation Results with Covariates

For the FH model incorporating covariates, the RB and RR of MSE estimators are presented in Tables 5.4 and 5.5. As shown in Table 5.5, for pattern (II), the average RR of MSE_ML, MSE_Rao, MSE_JY and MSE_JY1 are 6.32%, 1.22%, 1.66% and 1.38% respectively.

Table 5.4: Simulated values of the RB of $\widehat{\text{MSE}}(\hat{\theta}_i^{EB})$ for A = 1, m = 15, ψ_i pattern (I), (II), (III) and for the normal random effects distribution with covariates

Groups	Pattern I			Pattern II			Pattern III			Pattern I			Pattern II			Pattern III		
	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH
	Usual MSEs									MSE_JY								
G1	-18.3	-8	-1.4	-26.7	-9.4	-4.3	-34.9	-7.9	-6.1	-1.5	-0.5	-0.3	-10.9	-3.2	-3.4	-21.5	-3.1	-5.1
G2	-16.2	-6.3	0.3	-24	-7.5	-0.7	-39.6	-8.4	-4	1.2	1.5	1.5	-5.7	1.1	0.7	-24.8	-0.1	-2.8
G3	-16.4	-7.1	-0.6	-24.8	-8.2	-1.6	-41.1	-9.1	-5.1	1.2	0.9	0.9	-5.8	0.6	0.1	-25.9	-0.7	-3.7
G4	-13.5	-4.4	2	-22.9	-5.5	1	-41.4	-6.6	-3.1	5.3	4	3.8	-1.8	3.9	3.2	-24.8	2.2	-1.5
G5	-11.7	-3.3	3	-27.1	-2.1	4.9	-54.9	-0.9	1.4	9.7	5.8	5.8	25.5	10.5	11.6	125.3	16.6	6.4
	MSE_Rao									MSE_JY1								
G1	-2.8	-1.6	-1.4	-11.8	-4.1	-4.4	-22	-3.8	-5.8	-4.7	-2.5	-2.1	-12.9	-4.8	-4.7	-22.8	-4.5	-6.3
G2	-0.2	0.4	0.4	-7.2	-0.2	-0.5	-26	-1.3	-3.9	-2.4	-0.7	-0.4	-10.6	-1.9	-1.6	-29.3	-3.3	-4.8
G3	-0.3	-0.3	-0.3	-7.5	-0.7	-1.2	-27.2	-2	-4.8	-3	-1.5	-1.4	-11.6	-2.6	-2.5	-31.2	-4.1	-5.9
G4	3.6	2.8	2.5	-3.9	2.5	1.7	-26.4	0.9	-2.7	0.2	1.3	1	-9.5	0.2	-0.1	-31.7	-1.6	-4.1
G5	7.4	4.4	4.3	18.5	8.4	9.3	98.6	13.3	4.8	2	2.3	1.9	-13.3	3.4	3.2	-44.1	3.1	0.3

Note that:

- MSE_Rao, MSE_JY, MSE_JY1, MSE_REML, MSE_ML and MSE_FH referring the MSE estimator values given by equations (5.1.3 - 5.1.14).
- Methods of estimating A: ML, maximum likelihood; REML, restricted maximum likelihood; FH, Fay and Herriot (1979).
- Groups: G1, areas 1-3; G2, areas 4-6; G3, areas 7-9; G4, areas 10-12, G5, areas 13-15.

Table 5.5: Simulated values of the RR of $\widehat{\text{MSE}}(\hat{\theta}_i^{EB})$ for A = 1, m = 15, ψ_i pattern (I), (II), (III) and for the normal random effects distribution with covariates

Groups	Pattern I			Pattern II			Pattern III			Pattern I			Pattern II			Pattern III		
	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH	ML	REML	FH
	Usual MSEs									MSE_JY								
G1	3.4	0.7	-1.4	7.1	0.9	0.2	12.2	0.6	0.4	0	0	0	1.2	0.1	0.1	4.6	0.1	0.3
G2	2.6	0.4	0.3	5.8	0.6	0	15.7	0.7	0.2	0	0	0	0.3	0	0	6.1	0	0.1
G3	2.7	0.5	-0.6	6.1	0.7	0	16.9	0.8	0.3	0	0	0	0.3	0	0	6.7	0	0.1
G4	1.8	0.2	2	5.3	0.3	0	17.2	0.4	0.1	0.3	0.2	0.2	0	0.2	0.1	6.2	0.1	0
G5	1.4	0.1	3	7.3	0.1	0.3	30.1	0	0	1	0.4	0.4	6.5	1.1	1.4	157	2.8	0.4
	MSE_Rao									MSE_JY1								
G1	0.1	0	0	1.4	0.2	0.2	4.9	0.2	0.4	0.2	0.1	0.1	1.7	0.2	0.2	5.2	0.2	0.4
G2	0	0	0	0.5	0	0	6.8	0	0.2	0.1	0	0	1.1	0	0	8.6	0.1	0.2
G3	0	0	0	0.6	0	0	7.4	0.1	0.2	0.1	0	0	1.4	0.1	0.1	9.7	0.2	0.4
G4	0.1	0.1	0.1	0.2	0.1	0	7	0	0.1	0	0	0	0.9	0	0	10.1	0	0.2
G5	0.6	0.2	0.2	3.4	0.7	0.9	97.2	1.8	0.3	0.1	0.1	0.1	1.8	0.1	0.1	19.5	0.1	0

Note that:

- MSE_Rao, MSE_JY, MSE_JY1, MSE_REML, MSE_ML and MSE_FH referring the MSE estimator values given by equations (5.1.3 - 5.1.14).
- Methods of estimating A: ML, maximum likelihood; REML, restricted maximum likelihood; FH, Fay and Herriot (1979).
- Groups: G1, areas 1-3; G2, areas 4-6; G3, areas 7-9; G4, areas 10-12, G5, areas 13-15.

Figures 5.1-5.3 plots the average RB over the number of small areas, m. We have simulated the RB for the 5 groups within each pattern and number of small areas. Then we took the average of the absolute values of the RB over groups for each pattern and number of small areas. We have presented the findings for pattern (I), (II) and (III) in Figures 5.1, 5.2 and 5.3 respectively, as shown below. The green, blue, red and black colours indicate that the MSE estimators based on $g_{3i}(\hat{A})$'s (Rao2001, JY, JY1 and ML or REML or FH) respectively.

The key findings in Figure 5.1 are:

- For the homogeneous ψ_i pattern, the average RB of all the MSE estimators is less than 10%.
- The average RB of MSE_Rao, MSE_JY and MSE_JY1 are less than the MSE of the ML and REML estimators.
- All the MSE estimators based on the FH method are identical. They are nearly unbiased.

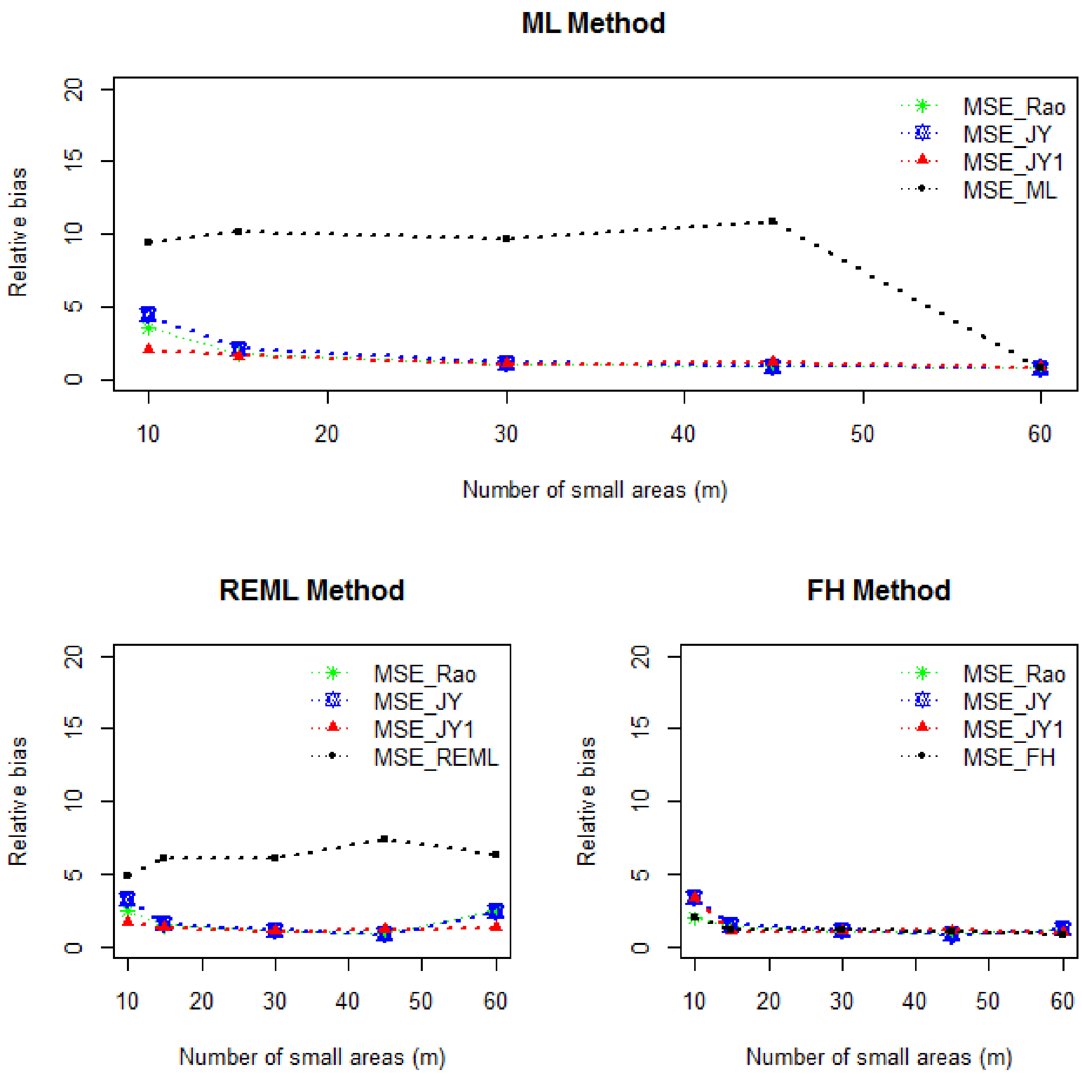

Figure 5.1: Simulated values of the average of the absolute value of the RB of $\widehat{\mathrm{MSE}}(\hat{\theta}_i^{\mathrm{EB}})$ for A = 1 and ψ_i pattern (I) and for the normal random effects distribution. The points are connected with lines for visibility purposes.

The key findings in Figure 5.2 are:

- For the ψ_i pattern (II), the average RB of all the MSE estimators are less than 15% .
- The average RB of MSE estimators based on the FH method is less than 5% over all domains. MSE estimators based FH method are nearly unbiased.
- For the ML and REML estimators MSE_JY1 remains nearly unbiased over all areas.

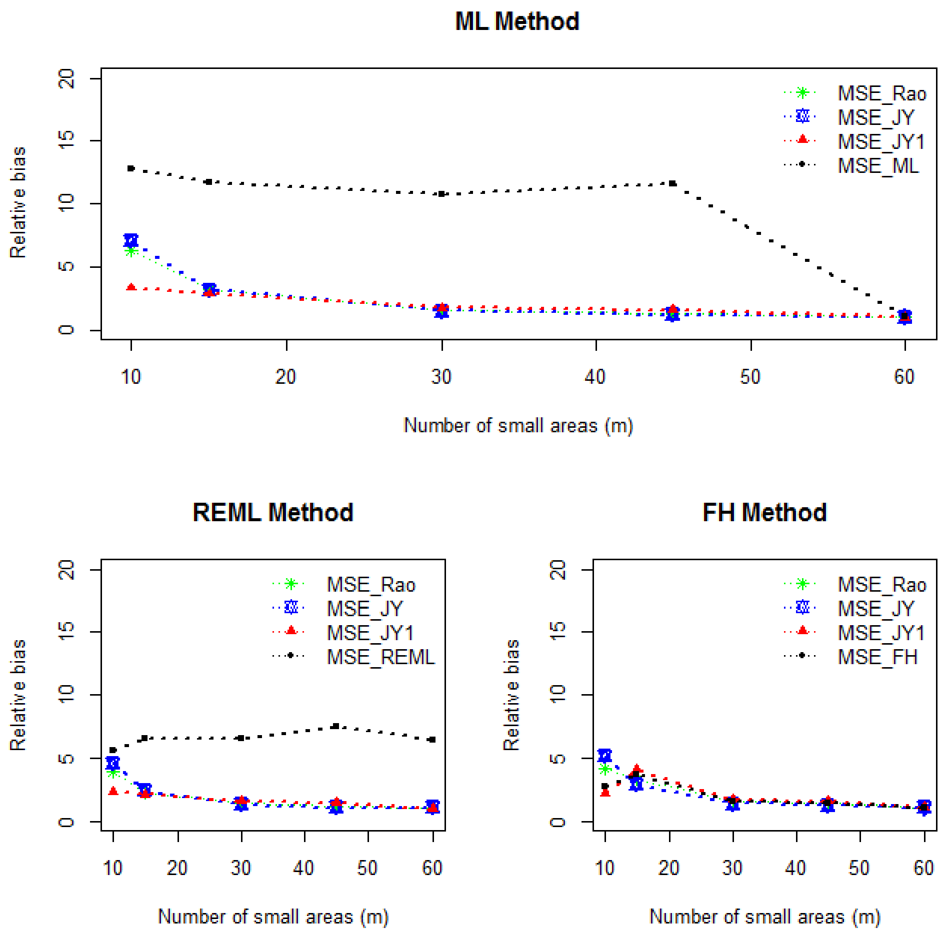

Figure 5.2: Simulated values of the average of the absolute value of the RB of $\widehat{\text{MSE}}(\hat{\theta}_i^{\text{EB}})$ for A = 1 and ψ_i pattern (II) and for the normal random effects distribution. The points are connected with lines for visibility purposes.

The key findings in Figure 5.3 are:

- The average RB of MSE estimators decreases as m increases.
- When m < 30, the average RB of MSE estimators based on MSE_JY1 is small when compared to its competitors for the ML and REML methods.
- MSE estimators are approximately unbiased for the FH method in all cases.

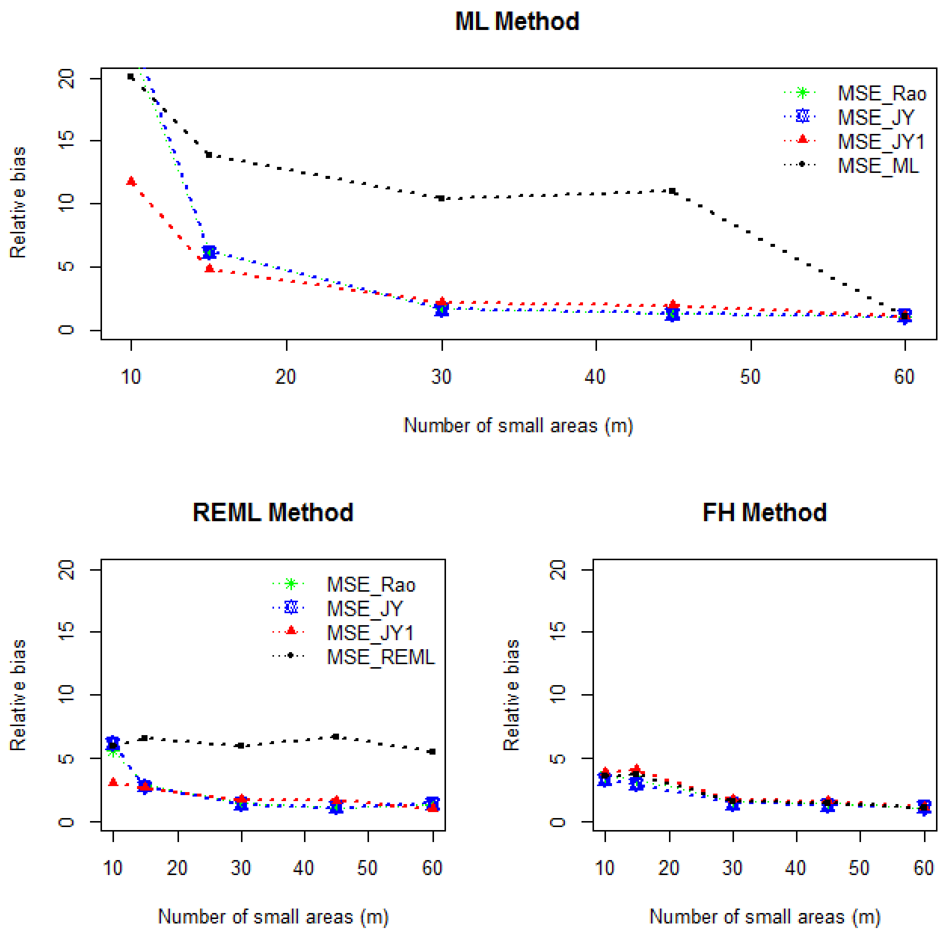

Figure 5.3: Simulated values of the average of the absolute value of the RB of $\widehat{\mathrm{MSE}}(\hat{\theta}_i^{\mathrm{EB}})$ for A = 1 and ψ_i pattern (III) and for the normal random effects distribution. The points are connected with lines for visibility purposes.

5.1.5 Discussion of Simulation Results

In this section we discuss results from the simulation studies of the area specific MSE estimators. We investigate the performances of the proposed area specific MSE estimators using simulation studies. First we consider the simulation study without covariates as shown in Tables 5.1 - 5.3. Table 5.1 show that all the methods produced similar results with respect to MSE estimations for pattern (I). MSE estimates based on the ML, REML and FH methods are nearly equal for all patterns. For pattern (III) estimates based on the ML, REML and FH are nearly equal. The MSE results based on the FH method are not sensitive to the sampling variance patterns. Our findings were consistent with Datta et al. (2005).

As shown in Tables 5.2 and 5.3, in terms of RB and RR the MSE estimators based on the ML, REML and FH behave similarly for pattern (I), with no particular one emerging as clearly better than the other three. When m = 15, estimators based on the FH method have very small bias for all variance patterns. For the same pattern the bias of MSE estimators are generally positive and thus accumulative at higher levels of aggregation. The bias of MSE_JY1 estimator is smaller than MSE_Rao and MSE_JY estimators for the ML and REML methods. All the estimators have small biases and risks for the case where $\hat{A} = \hat{A}_{FH}$ and the biases are generally smaller than the biases obtained when $\hat{A} = \hat{A}_{ML}$ for all variance patterns, particularly for the smaller variance ψ_i. The bias and risk of all the estimators are less than 7% when $\hat{A} = \hat{A}_{FH}$ for both patterns (II) and (III) are negligible for group G5. This result is in agreement with the previous results of Datta et al. (2005) and Datta and Lahiri (2000). Thus, the FH method performed better than the ML and REML methods for patterns (II) and (III). As shown by Table 5.3, the RR of MSE_Rao, MSE_JY and MSE_JY1 are nearly zero for the ML, REML and FH methods for all patterns. In terms of RR, for pattern (I), all the MSE estimators are robust. In addition, in terms of RR all the MSE estimators are robust for all patterns.

We considered the Laplace distribution in order to assess the robustness of the MSE estimators to possible deviations from the normality assumptions. As shown in Table 5.2 and 5.3, the RB and RR are generally higher than the RB and RR when the random effects are generated from the normal distributions. The MSE estimator based on the ML method has large negative bias. For example, the RB of the MSE estimators are -23.8%, 63.2%, 69% and -3.2% for MSE_ML, MSE_Rao, MSE_JY and MSE_JY1, respectively when $\hat{A} = \hat{A}_{ML}$ for pattern (III) and $\psi_i = 0.1$ (small sampling error variance). All the methods are sensitive to the deviation from normality of the random effects distribution considered in the simulation for all patterns with $\psi_i = 0.1$ except the FH method for pattern (III). The RB of MSE_JY1 is smaller when $\hat{A} = \hat{A}_{FH}$ for pattern (III) and $\psi_i = 0.1$. The RR of MSE estimators based on the FH method are less than 7% in all cases.

For the FH model incorporating covariates, the RB and RR of MSE estimators are presented in

Tables 5.4 and 5.5. For the simulation exercise with covariates MSE estimators based on ML methods are biased for all variance patterns. The RB of MSE estimators based on ML method are -54.9%, 98.6%, 125.3% and -44.1% for MSE_ML, MSE_Rao, MSE_JY and MSE_JY1 respectively for G5 (the smallest variance in pattern (III)). MSE_JY1 performs well for the ML, REML and FH methods. As shown in Table 5.5, the RR of MSE_ML, MSE_Rao, MSE_JY and MSE_JY1 are 30.1%, 97.2%, 157% and 19.5%, respectively for G5. For pattern (II), the average RR of MSE_ML, MSE_Rao, MSE_JY and MSE_JY1 are 6.32%, 1.22%, 1.66% and 1.38% respectively. However, the RR of MSE estimators based on the REML and FH methods are negligible (less than 5%) for all patterns. MSE estimators based on the REML and FH methods are robust. Thus, it is recommended to use MSE estimators based on these two methods.

5.2 Comparison of Confidence Intervals: Simulation Study II

In this study, we are adopting the same simulation setting used in Study I. We assess the performance of the two-sided CIs based on the CPs and ALs. In order to evaluate the CPs and ALs of the proposed method, we perform a simulation study using the PR, FH, ML and REML estimators for all variance patterns. The CP and AL of CIs can be expressed as:

$$\mathrm{CP_i} = \sum_{i=1}^{m} \frac{|\theta_i \in \mathrm{CI_i}|}{m}, \quad \mathrm{AL_i} = \sum_{i=1}^{m} \frac{|\text{length of } \mathrm{CI_i}|}{m} \tag{5.2.1}$$

where , i=1,...,m, $\mathrm{CI_i}$ is the CI of a certain method.

5.2.1 Comparison of Confidence Intervals Based on the Weighted Estimator with Fixed Weights

The purpose of this section is to compare the performance of the proposed methods using models under different sampling variances.

Cox (1975) proposed the following EB interval:

$$\text{Method IW(MetIW)}: \hat{\theta}_i^w \pm z_{\alpha/2}\psi_i^{1/2}(1-\hat{\gamma}_i)^{1/2}. \tag{5.2.2}$$

Prasad and Rao (1990) proposed the following EB interval:

$$\text{Method IIW(MetIIW)}: \hat{\theta}_i^w \pm z_{\alpha/2}\sqrt{\widehat{\mathrm{MSE}}(\hat{\theta}_i^w)}. \tag{5.2.3}$$

The improved CI of the ith small area mean $\theta_i = x_i'\beta + v_i$ based on the weighted estimator with fixed weights is given by

$$\text{Method IIIW(MetIIIW)}: \hat{\theta}_i^w \pm t_{\alpha w}\sqrt{\widehat{\mathrm{MSE}}(\hat{\theta}_i^w)}, \tag{5.2.4}$$

where $t_{\alpha w} = z_{\alpha/2}(z^3_{\alpha/2} + z_{\alpha/2})\frac{\psi_i^2}{8A^2(A+\psi_i)^2}V(\hat{A})+$

$$+ z_{\alpha/2}\left\{\frac{\psi_i^2+(\gamma_i-w_i)^2(A+\psi_i)^2}{2A\psi_i(A+\psi_i)}\mathbf{x}_i'\left(\sum\frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i\right\}.$$

For $t_{\alpha w}$, the improved CI given in (4.1.7) satisfies that
$P\left[\theta_i \in \{\hat{\theta}_i^w \pm t_{\alpha w}\sqrt{\widehat{\text{MSE}}(\hat{\theta}_i^w)}\}\right] = 1-\alpha+O(m^{-3/2})$.

5.2.2 Discussion of Simulation Results

In this section we consider three different CIs given in (5.2.2), (5.2.3) and (5.2.4), which are referred as Method IW, Method IIW and Method IIIW, respectively. In each table and figure, Method IW, Method IIW and Method IIIW are the Cox (1975), Prasad and Rao (1990) and our proposed method respectively. We have carried out more extensive simulations which reflect the general pattern of performance of Method IW, Method IIW and Method IIIW reported here. Based on 10,000 simulations we estimated the CPs and ALs of the three methods for all selected small areas and variance patterns. The results are reported in Tables 5.6 and 5.7. These results are dependent on the sampling variance patterns (Chatterjee et al., 2008).

The main goal of our study is to develop closed form improved CIs for a small area mean based on the EBLUP and weighted estimator with fixed weights using the Taylor series expansion methods under the FH model. We explored the CPs and ALs of the proposed CIs compared to Method IW and Method IIW.

Method IW is easy to construct, but its CP is open to question. This is the first attempt by Cox (1975) to construct EB CIs in small area estimation problems. It has a coverage accuracy of order $O(m^{-1})$ (Diao et al., 2013; Chatterjee et al., 2008). For variance patterns (II) and (III), this method truly underestimates the CPs. For example, as shown in Table 5.6, Method IW has under-coverage problem (it could be as low as 69.2% compared to the nominal value of 95% for the normal random effects).

Prasad and Rao (1990) proposed Method IIW by incorporating covariates when the error variances are equal. However, in this study Method IIW is used for unequal sampling variances. Since this method has a coverage accuracy of order $O(m^{-1})$, similar to Method IW, it cannot be an appropriate choice in small area estimation problems (Diao et al., 2013). Thus, we proposed method (Method IIIW) that is accurate to the order of $O(m^{-3/2})$ under unequal sampling variances. Our methods provide an improved coverage in the sense that the margin of error of the CP is of order $O(m^{-3/2})$. Method IIIW performed uniformly well throughout. When $m = 20$, the CP of Method IIIW never differed from its nominal value by more than 1.3%, 9.2%, 2.2% and 1.3% for the ML, PR, FH and REML estimators respectively for the normal distribution. These are values slightly below the 95% nominal value.

Considering the Laplace random effects distribution helps us to investigate the robustness of our proposed methods when the data violate the normality assumptions. As we can see the CPs of Method IIIW are improved when compared to Method IW and Method IIW. The CPs of Method IW becomes worse when dealing with the Laplace distribution.

We also examine the performance of our proposed method over a range of number of small areas, m, for all variance patterns as shown by Figures 5.4, 5.5 and 5.6. The CPs of Method IW, Method IIW and Method IIIW are obtained through the above simulation exercise for the number of areas such as m=10, 15, 20, 30, 40, 50, 60, 80 and 100. When the number of small areas, m, increases, the CPs of Method IIIW are close to the nominal level for all patterns. When $m \geq 60$, there is no clear difference between Method IIW and Method IIIW in terms of CPs and ALs under the normal distribution, even though Method IIIW continues to be closer to the nominal level, hence it shows its consistency (Diao et al., 2013). The percentage difference between CPs and ALs are negligible. We thus conclude that Method IIIW is superior to the rest of the methods in terms of the CP.

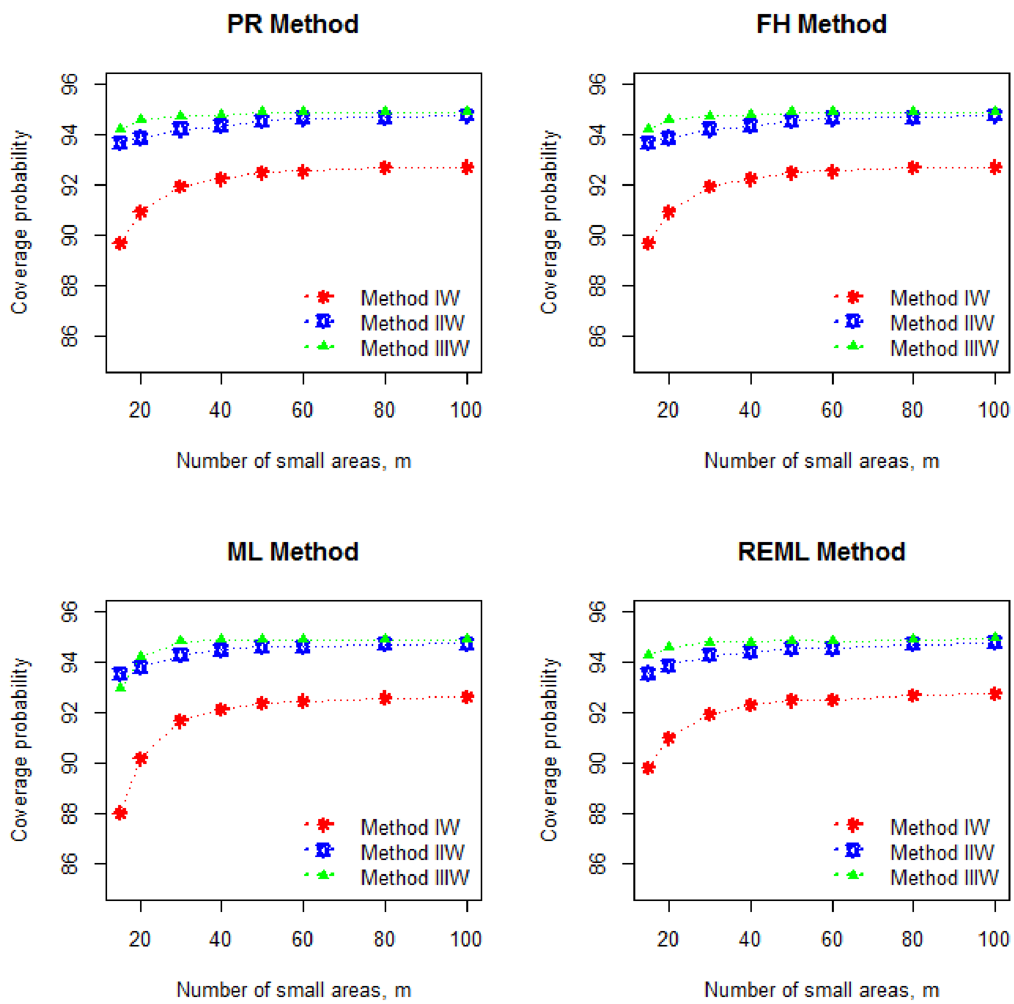

Figure 5.4: Simulated values of the CPs of Method IW, Method IIW and Method IIIW for nominal 95% CIs for ψ_i pattern (I). The points are connected with lines for visibility purposes.

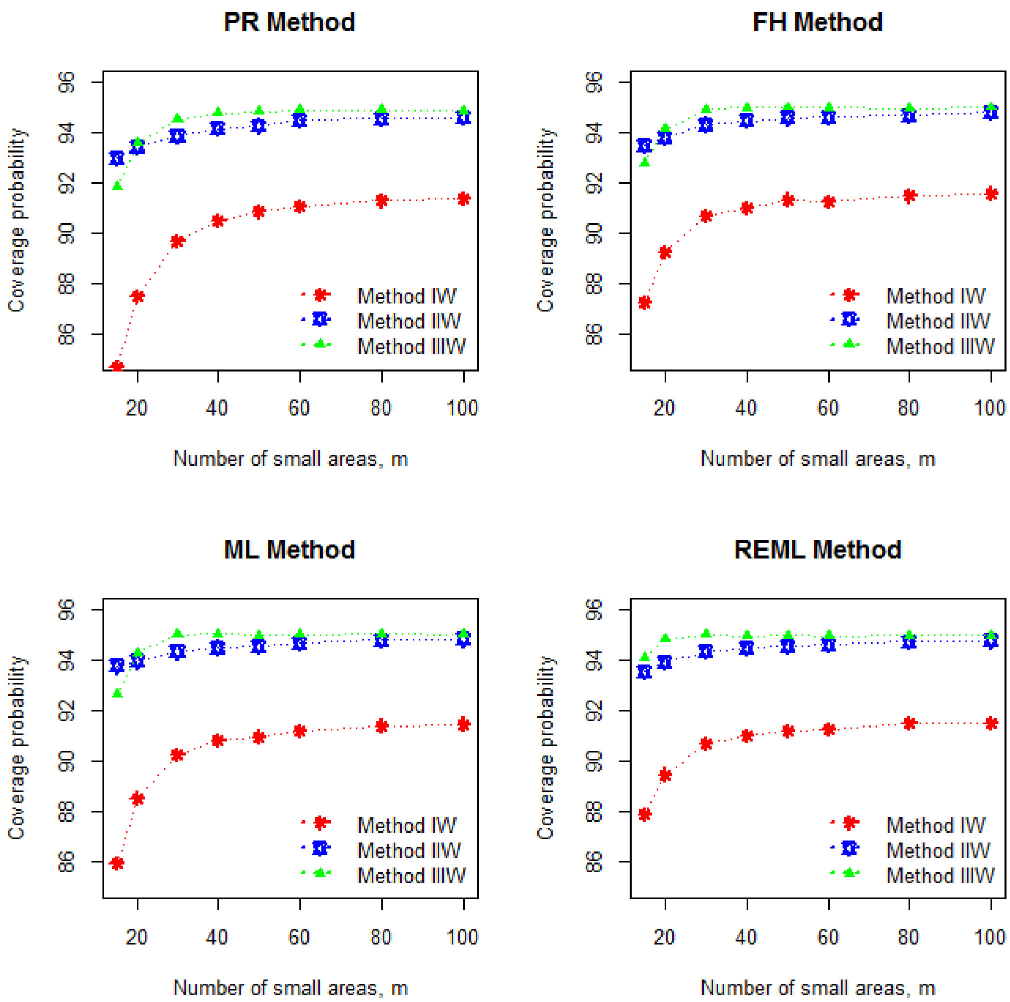

Figure 5.5: Simulated values of the CPs of Method IW, Method IIW and Method IIIW for nominal 95% CIs for ψ_i pattern (II). The points are connected with lines for visibility purposes.

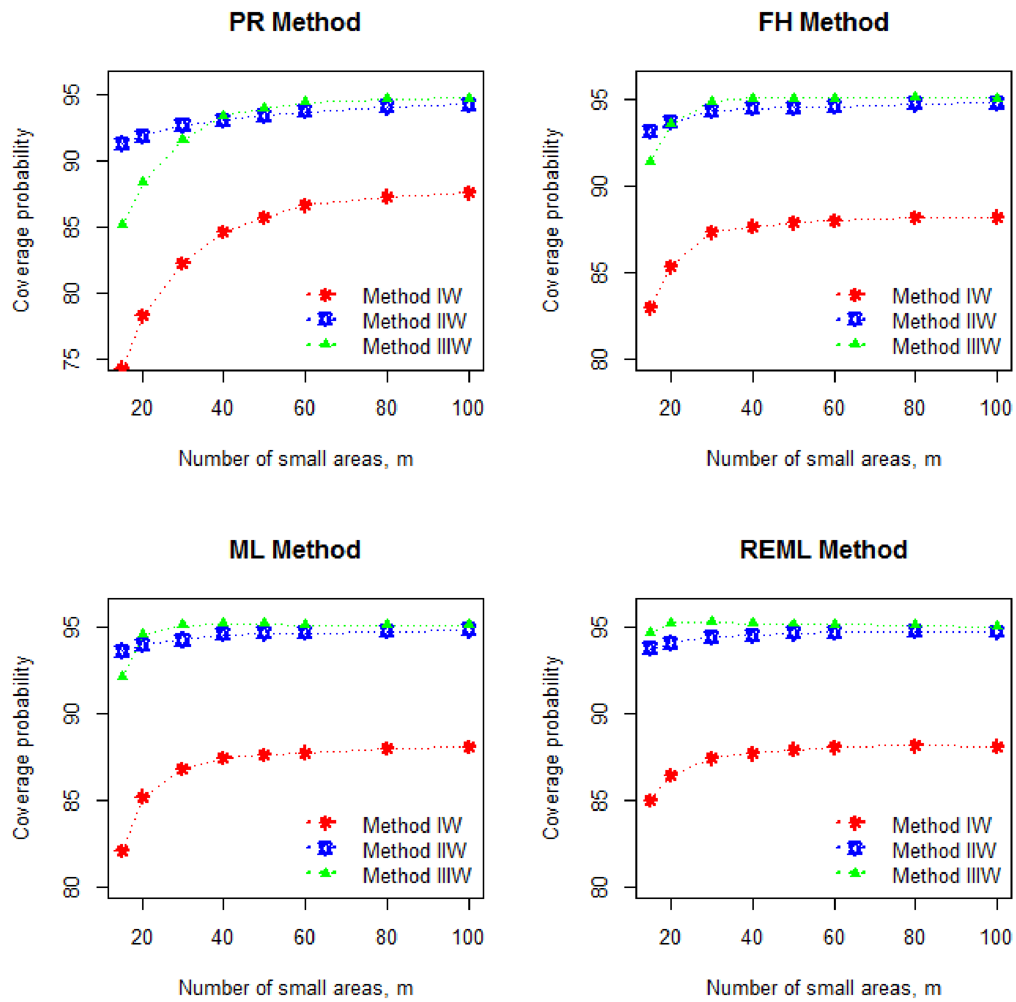

Figure 5.6: Simulated values of the CPs of Method IW, Method IIW and Method IIIW for nominal 95% CIs for ψ_i pattern (III). The points are connected with lines for visibility purposes.

Table 5.6: CPs and ALs of a small area mean for nominal 95% CIs for ψ_i pattern (I), (II) and (III). Results obtained for the normal random effects distributions and m = 20.

Groups	CP											
	PR			FH			ML			REML		
	Pattern I											
	MetIW	MetIIW	MetIIIW	MetIW	MetIIW	MetIIIW	MetIW	MetIIW	MetIIIW	MetIW	MetIIW	MetIIIW
G1	92.9	94.3	95.0	92.9	94.2	94.9	92.5	94.3	95.0	93.0	94.2	95.0
G2	92.5	94.0	94.7	92.7	94.3	94.8	92.2	94.2	94.8	92.7	94.2	94.9
G3	92.2	94.3	94.8	92.2	94.2	94.7	91.9	94.2	94.8	92.3	94.2	94.8
G4	91.2	94.0	94.6	91.1	94.0	94.5	91.0	94.1	94.7	91.3	94.0	94.7
G5	89.0	93.9	94.4	89.3	94.1	94.5	88.9	94.0	94.5	89.2	93.9	94.5
	Pattern II											
G1	89.0	94.6	96.0	90.9	94.9	96.1	89.5	94.7	95.8	90.7	94.6	96.1
G2	91.3	94.0	94.5	92.4	94.1	94.6	91.6	93.9	94.6	92.7	94.3	94.9
G3	90.8	93.7	94.1	91.8	94.0	94.4	91.3	94.1	94.6	92.2	94.2	94.8
G4	89.8	93.5	93.9	91.0	94.0	94.3	90.6	94.2	94.6	91.2	94.0	94.6
G5	83.9	93.1	93.4	85.0	93.8	94.3	84.8	94.1	94.5	85.0	93.9	94.5
	Pattern III											
G1	78.3	95.1	93.8	84.8	95.0	96.2	83.6	94.9	96.4	85.1	94.9	96.7
G2	85.9	92.6	90.7	92.0	94.0	94.0	91.8	94.0	94.6	92.5	94.0	94.8
G3	85.2	92.1	90.3	91.5	93.8	93.8	91.3	93.9	94.4	92.0	94.0	94.6
G4	84.8	91.9	89.9	90.3	93.6	93.5	90.8	94.2	94.6	91.2	94.3	94.8
G5	69.2	89.4	87.9	73.2	93.4	94.1	73.7	94.2	95.2	74.4	94.4	95.7
	AL											
	Pattern I											
G1	2.44	2.57	2.42	2.44	2.57	3.15	2.41	2.56	3.12	2.44	2.57	3.00
G2	2.33	2.48	2.32	2.33	2.48	3.06	2.30	2.48	3.04	2.33	2.48	2.92
G3	2.20	2.40	2.21	2.20	2.40	2.98	2.18	2.39	2.96	2.20	2.39	2.83
G4	2.04	2.30	2.10	2.04	2.30	2.91	2.02	2.30	2.88	2.04	2.30	2.75
G5	1.84	2.21	1.98	1.84	2.21	2.84	1.83	2.20	2.81	1.84	2.20	2.67
	Pattern II											
G1	3.03	3.44	4.04	3.10	3.45	3.66	3.01	3.44	3.55	3.09	3.44	3.74
G2	2.28	2.48	2.88	2.33	2.49	2.52	2.28	2.48	2.40	2.33	2.48	2.61
G3	2.16	2.39	2.79	2.20	2.40	2.42	2.16	2.39	2.30	2.20	2.39	2.52
G4	2.00	2.29	2.70	2.04	2.31	2.32	2.01	2.30	2.19	2.04	2.30	2.42
G5	1.54	2.09	2.55	1.57	2.10	2.12	1.55	2.09	1.97	1.57	2.10	2.24
	Pattern III											
G1	3.15	4.43	5.96	3.39	4.44	4.95	3.29	4.43	5.16	3.40	4.44	4.98
G2	2.14	2.45	3.08	2.32	2.48	2.60	2.29	2.48	2.85	2.33	2.49	2.67
G3	2.02	2.36	2.99	2.19	2.39	2.50	2.17	2.39	2.76	2.20	2.40	2.57
G4	1.88	2.26	2.90	2.03	2.30	2.40	2.01	2.30	2.68	2.04	2.31	2.48
G5	1.09	1.92	2.80	1.16	1.97	2.17	1.16	1.98	2.53	1.17	1.98	2.27

Table 5.7: CPs and ALs of a small area mean for nominal 95% CIs for ψ_i pattern (I), (II) and (III). Results obtained for the Laplace (double exponential) random effects distributions and m = 20.

Groups	CP											
	PR			FH			ML			REML		
	Pattern I											
	MetIW	MetIIW	MetIIIW	MetIW	MetIIW	MetIIIW	MetIW	MetIIW	MetIIIW	MetIW	MetIIW	MetIIIW
G1	87.0	92.1	94.2	86.7	90.5	94.2	84.8	89.8	94.3	87.7	92.8	94.4
G2	87.4	92.0	94.1	87.1	90.4	94.2	85.4	89.8	94.3	87.8	92.6	94.0
G3	88.0	92.0	94.1	87.5	90.1	94.1	86.0	89.7	94.3	88.1	92.3	94.0
G4	87.8	91.5	93.6	87.3	90.0	93.9	85.6	89.2	93.6	88.6	92.3	93.8
G5	87.8	91.4	93.5	87.1	89.6	93.5	85.7	89.2	93.6	88.3	91.9	93.6
	Pattern II											
G1	73.8	87.2	95.0	77.4	87.7	94.9	76.8	90.2	94.9	79.8	93.8	94.9
G2	80.5	86.0	93.6	84.0	86.7	93.8	84.7	89.4	94.1	87.8	92.6	94.1
G3	80.9	85.8	93.6	84.4	86.6	93.8	85.0	89.1	93.9	88.0	92.3	93.7
G4	81.0	85.5	93.1	84.3	86.3	93.4	85.1	88.8	93.5	88.4	92.2	93.7
G5	79.9	84.7	91.9	82.9	85.8	92.6	84.1	88.5	93.3	87.1	91.6	93.1
	Pattern III											
G1	56.7	76.8	95.5	65.5	82.1	95.0	66.7	89.9	95.0	70.5	94.6	94.8
G2	70.0	75.4	93.1	79.1	81.0	93.7	84.1	88.7	94.0	88.4	93.4	94.2
G3	70.2	75.3	92.6	79.1	80.6	93.2	84.5	88.6	93.7	88.9	93.3	94.0
G4	70.3	74.8	92.1	79.1	80.5	93.0	84.7	88.3	93.4	88.8	92.8	93.6
G5	65.9	73.3	88.4	73.2	79.8	90.4	79.4	87.8	92.7	83.9	92.1	92.8
	AL											
	Pattern I											
G1	1.92	2.17	2.34	1.93	2.11	2.18	1.85	2.17	2.30	1.93	2.05	2.18
G2	1.86	2.08	2.24	1.86	2.00	2.08	1.79	2.08	2.19	1.87	1.93	2.08
G3	1.78	1.98	2.13	1.78	1.87	1.98	1.72	1.97	2.08	1.79	1.80	1.98
G4	1.68	1.87	2.01	1.68	1.74	1.87	1.62	1.86	1.96	1.69	1.66	1.87
G5	1.55	1.75	1.89	1.55	1.60	1.75	1.50	1.75	1.84	1.56	1.51	1.75
	Pattern II											
G1	2.14	3.14	3.61	2.22	3.14	3.24	2.16	3.14	3.32	2.27	3.15	3.35
G2	1.73	2.07	2.39	1.81	1.97	2.07	1.78	2.08	2.11	1.86	2.08	2.16
G3	1.65	1.97	2.28	1.73	1.84	1.96	1.71	1.98	1.99	1.78	1.98	2.04
G4	1.56	1.86	2.16	1.63	1.70	1.85	1.61	1.87	1.87	1.68	1.87	1.92
G5	1.25	1.60	1.93	1.31	1.38	1.60	1.31	1.59	1.61	1.36	1.61	1.66
	Pattern III											
G1	2.14	4.22	5.04	2.27	4.19	4.80	2.26	4.19	4.50	2.41	4.20	4.43
G2	1.56	2.09	2.13	1.72	2.05	2.45	1.77	2.07	2.14	1.87	2.02	2.07
G3	1.48	1.98	1.98	1.65	1.94	2.34	1.69	1.96	2.03	1.79	1.90	1.97
G4	1.39	1.82	1.87	1.55	1.83	2.24	1.60	1.85	1.91	1.69	1.77	1.86
G5	0.85	1.24	1.44	0.95	1.40	1.95	1.01	1.44	1.50	1.06	1.31	1.45

5.2.3 Comparison of Confidence Intervals Based on EBLUP

Consider the following six different CIs based on EBLUP:

Cox (1975) suggests generating the following EB interval:

$$\text{Method I (Meth I)}: \hat{\theta}_i^{EB} \pm z_{\alpha/2}\psi_i^{1/2}(1-\hat{\gamma}_i)^{1/2}. \tag{5.2.5}$$

Prasad and Rao (1990) proposed a CI which is written as:

$$\text{Method II (Meth II)}: \hat{\theta}_i^{EB} \pm z_{\alpha/2}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}. \tag{5.2.6}$$

The improved CI based on the EBLUP is written as:

$$\text{Method III (Meth III)}: \hat{\theta}_i^{EB} \pm t_\alpha^* \sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}, \tag{5.2.7}$$

where $t_\alpha^* = z_{\alpha/2}(1+h(\hat{A}))$, $h(\hat{A}) = (z_{\alpha/2}^2+1)\frac{\psi_i^2}{8\hat{A}^2(\hat{A}+\psi_i)^2}V(\hat{A})$.

The corresponding area specific versions of (4.1.10) are written as follows:

$$\text{CIRao}: \hat{\theta}_i^{EB} \pm t_{Rao}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})} \tag{5.2.8}$$

where $t_{Rao} = z_{\alpha/2} + (z_{\alpha/2}^3 + z_{\alpha/2})\frac{g_{3Raoi(A+\psi_i)}}{8A^2}$, $g_{3Raoi}(\hat{A}) = \frac{\psi_i^2}{(A+\psi_i)^4}(y_i - \mathbf{x}_i'\hat{\beta})^2 V(\hat{A})$.

$$\text{CIJY}: \hat{\theta}_i^{EB} \pm t_{JY}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})} \tag{5.2.9}$$

where $t_{JY} = z_{\alpha/2} + (z_{\alpha/2}^3 + z_{\alpha/2})\frac{g_{3JYi(A+\psi_i)}}{8A^2}$, $g_{3JYi}(\hat{A}) = \frac{\psi_i^2}{(\psi_i+\hat{A})^4}\frac{(y_i-\mathbf{x}_i'\tilde{\beta})^2}{\hat{A}+\psi_i-\mathbf{x}_i'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_i}V(\hat{A})$.

Another improved CI based on the EBLUP is written as:

$$\text{CIJY1}: \hat{\theta}_i^{EB} \pm t_{JY1}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})} \tag{5.2.10}$$

where $t_{JY1} = z_{\alpha/2} + (z_{\alpha/2}^3 + z_{\alpha/2})\frac{g_{3JY1i(A+\psi_i)}}{8A^2}$, $g_{3JY1i}(\hat{A}) = g_{3i}(A) - \frac{g_{2i}(A)}{(A+\psi_i)^2}V(\hat{A})$.

Note that for t_α^*, t_{Rao}, t_{JY} and t_{JY1} the improved CIs given in (5.2.7), (5.2.8), (5.2.9) and (5.2.10) satisfies

$$P\left[\theta_i \in \{\hat{\theta}_i^{EB} \pm t_\alpha^*\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}\}\right] = 1-\alpha+O(m^{-3/2}). \tag{5.2.11}$$

$$P\left[\theta_i \in \{\hat{\theta}_i^{EB} \pm t_{Rao}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}\}\right] = 1-\alpha+O(m^{-3/2}). \tag{5.2.12}$$

$$P\left[\theta_i \in \{\hat{\theta}_i^{EB} \pm t_{JY}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}\}\right] = 1-\alpha+O(m^{-3/2}). \tag{5.2.13}$$

$$P\left[\theta_i \in \{\hat{\theta}_i^{EB} \pm t_{JY1}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}\}\right] = 1-\alpha+O(m^{-3/2}). \tag{5.2.14}$$

5.2.4 Discussion of Simulation Results

In this section we compare six different CIs given in (5.2.5), (5.2.6), (5.2.7), (5.2.8), (5.2.9) and (5.2.10), which are referred as Method I, Method II, Method III, CIRao, CIJY and CIJY1, respectively. Method I and Method II represents the Cox (1975) and Prasad and Rao (1990) CIs for θ_i. Both of them are naive methods and have a coverage accuracy $O(m^{-1})$. However, Method III, CIRao, CIJY and CIJY1 are improved CIs with coverage accuracy $O(m^{-3/2})$. We improved the coverage in the sense that the margin of error of the CP is of order $O(m^{-3/2})$. The CPs and ALs of the proposed methods are comparable with Diao et al. (2013)'s CIs. However, CIRao and CIJY are constructed based on the proposed area specific MSEs.

The simulation results in Tables 5.8 and 5.9 show the estimated CPs and ALs of CIs for θ_i under the normal and the Laplace random effects distributions for number of small areas, m = 20. For example, when $\hat{A} = \hat{A}_{PR}$ and pattern (III), Method I has a CP of 80.2%, 81.8%, 81.8%, 82.0% and 84.1% for G1, G2, G3, G4 and G5 respectively for normal random effects distribution. The CPs of the proposed methods are larger than Method I and Method II for all patterns. Method III, CIRao, CIJY and CIJY1 meet the nominal coverage rate more frequently than Method I and Method II. Thus, Method III, CIRao, CIJY and CIJY1 perform better than Method I and Method II in terms of coverage accuracy. However, the AL of these Methods is a bit larger than Method I and Method II for large sampling variances (G1 and G2). The proposed methods extend the widths of Method I and Method II so as to satisfy the nominal confidence level (Kubokawa, 2011).

We have also considered the Laplace distribution to assess the robustness of the methods to possible deviations from the normality assumptions. In terms of CPs the proposed methods perform better than Method I and Method II. Method I has got a serious under-coverage problem (as low as 66.6%, 66.4%, 66.6%, 66.6% and 67.7% for G1, G2, G3, G4 and G5 respectively, values which are far below the 95% nominal value) for pattern (III) and when $A = \hat{A}_{PR}$ under the Laplace distribution.

Similar to the weighted estimator with fixed weights case, Figures 5.7 and 5.8 report CPs and ALs of CIs over a range of number of small areas, m, for pattern (II). The CPs and ALs of CIs of all the methods are obtained through the above simulation exercise for the number of areas such as m=15, 20, 30, 40, 50, 60, 80 and 100. As the number of small areas, m, increases the CPs increase and the ALs decrease for all the methods and patterns. When $m \geq 60$, there is very little difference introduced by using any of the methods in terms of AL and CP for all patterns. All the methods except Method I perform better any time in terms of CPs and ALs. There is also no clear difference among the proposed methods in terms of CPs and ALs under the normal distribution. The percentage difference between CPs and ALs are negligible.

Table 5.8: CPs of a small area mean for nominal 95% CIs for ψ_i pattern (II). Results obtained for the normal and the Laplace (double exponential) random effects distributions and $m = 20$.

Groups	CP											
	Normal Distribution											
	PR						REML					
	Meth I	Meth II	Meth III	CIRao	CIJY	CIJY1	Meth I	Meth II	Meth III	CIRao	CIJY	CIJY1
	Pattern I											
G1	91.4	93.1	93.7	94.6	94.6	94.7	91.7	94.1	94.6	94.9	94.7	94.5
G2	91.9	93.6	94.0	94.7	94.7	94.8	92.0	94.3	94.8	94.9	94.9	94.7
G3	92.0	93.6	94.0	94.7	94.9	95.0	92.3	94.5	94.9	94.9	94.7	94.7
G4	92.5	94.0	94.3	95.0	95.0	94.9	92.5	94.7	94.9	95.1	95.1	95.0
G5	92.9	94.3	94.5	95.1	95.2	95.0	92.9	95.0	95.1	95.1	95.1	95.0
	Pattern II											
G1	87.6	91.5	93.0	92.8	92.5	94.3	90.6	93.1	93.8	94.0	93.7	93.8
G2	89.1	94.4	93.9	94.9	94.8	95.6	91.8	94.2	94.5	94.6	94.5	94.7
G3	89.3	94.7	94.3	94.8	95.1	95.8	91.9	94.4	94.8	94.7	94.7	94.7
G4	89.3	95.0	94.2	95.3	95.2	96.0	92.1	94.6	94.8	94.8	94.6	95.0
G5	90.3	95.8	95.1	95.8	95.9	96.1	93.1	95.0	95.2	95.1	95.3	95.0
	Pattern III											
G1	80.2	87.7	92.2	89.1	89.5	94.6	89.9	92.0	93.5	93.5	93.7	93.7
G2	81.8	95.1	96.7	96.2	96.2	98.1	91.9	94.0	94.4	94.4	94.5	94.4
G3	81.8	95.2	96.7	96.3	96.7	98.1	92.2	94.3	94.6	94.7	94.8	94.4
G4	82.0	95.9	96.9	96.9	96.9	97.9	92.4	94.4	94.6	94.7	94.9	94.7
G5	84.1	96.4	96.5	96.6	96.7	97.0	93.9	95.1	95.2	95.2	95.2	95.1
	Laplace Distribution											
	Pattern I											
G1	85.3	91.9	92.3	93.8	93.8	94.0	85.6	93.0	93.3	93.5	93.6	93.3
G2	85.3	92.2	92.6	94.3	94.3	94.3	85.9	93.4	93.7	93.9	94.0	93.6
G3	85.4	92.4	92.8	94.4	94.4	94.4	86.0	93.8	94.1	94.1	94.1	94.0
G4	85.3	92.8	93.1	95.0	95.1	94.8	86.2	94.3	94.6	94.8	94.8	94.5
G5	86.1	93.8	94.0	96.0	96.1	95.6	86.9	95.4	95.6	95.7	95.8	95.5
	Pattern II											
G1	77.3	89.7	90.5	91.0	91.1	93.1	84.4	89.9	90.2	90.5	90.5	90.2
G2	77.7	95.1	94.6	95.9	95.9	96.9	85.3	93.0	92.9	93.3	93.4	92.8
G3	77.7	95.5	94.9	96.2	96.2	97.1	85.5	93.4	93.4	93.6	93.7	93.3
G4	77.8	96.3	95.3	96.8	96.9	97.4	85.5	94.1	93.8	94.3	94.4	93.8
G5	78.7	97.7	96.8	97.9	97.9	97.9	86.9	96.1	95.9	96.2	96.3	95.8
	Pattern III											
G1	66.6	87.5	91.4	88.2	88.2	94.8	84.9	88.4	88.9	89.0	89.0	88.9
G2	66.4	97.0	98.2	97.9	98.0	99.2	86.0	91.3	91.6	92.2	92.2	91.5
G3	66.6	97.5	98.4	98.2	98.3	99.1	86.2	91.8	92.1	92.6	92.6	92.0
G4	66.6	97.6	98.3	98.4	98.5	99.1	86.3	92.1	92.3	93.0	93.1	92.2
G5	67.7	98.0	98.1	98.2	98.2	98.4	88.9	95.8	95.8	96.1	96.1	95.8

Table 5.9: ALs of a small area mean for nominal 95% CIs for ψ_i pattern (II). Results obtained for the normal and the Laplace (double exponential) random effects distributions and $m = 20$.

Groups	AL											
	Normal Distribution											
	PR						REML					
	Meth I	Meth II	Meth III	CIRao	CIJY	CIJY1	Meth I	Meth II	Meth III	CIRao	CIJY	CIJY1
	Pattern I											
G1	2.42	2.54	2.60	2.71	2.72	2.71	2.42	2.60	2.66	2.71	2.73	2.70
G2	2.31	2.42	2.47	2.58	2.59	2.58	2.31	2.48	2.53	2.58	2.59	2.57
G3	2.18	2.29	2.32	2.42	2.44	2.42	2.18	2.34	2.38	2.42	2.43	2.41
G4	2.02	2.12	2.14	2.24	2.25	2.24	2.02	2.17	2.19	2.23	2.24	2.22
G5	1.83	1.90	1.92	2.02	2.02	2.00	1.83	1.95	1.96	2.00	2.01	1.99
	Pattern II											
G1	2.98	3.28	3.49	3.54	3.54	3.63	3.06	3.30	3.51	3.53	3.54	3.50
G2	2.24	2.52	2.52	2.58	2.59	2.65	2.30	2.49	2.53	2.54	2.54	2.53
G3	2.12	2.39	2.37	2.44	2.44	2.50	2.17	2.35	2.38	2.39	2.39	2.37
G4	1.97	2.23	2.20	2.26	2.27	2.32	2.02	2.17	2.19	2.20	2.20	2.19
G5	1.52	1.75	1.71	1.77	1.77	1.80	1.56	1.65	1.66	1.66	1.66	1.65
	Pattern III											
G1	3.06	3.45	4.39	4.46	4.53	5.09	3.36	3.58	3.93	4.17	4.20	4.08
G2	2.07	2.67	2.89	3.00	3.01	3.69	2.30	2.47	2.52	2.58	2.59	2.55
G3	1.95	2.57	2.76	2.91	2.92	3.59	2.18	2.33	2.37	2.42	2.43	2.39
G4	1.81	2.45	2.62	2.81	2.81	3.47	2.02	2.16	2.18	2.23	2.24	2.20
G5	1.05	1.87	2.01	2.54	2.51	3.15	1.16	1.21	1.21	1.24	1.25	1.22
	Laplace Distribution											
	Pattern I											
G1	1.89	2.12	2.16	2.26	2.27	2.27	1.90	2.21	2.25	3.10	3.12	3.68
G2	1.82	2.05	2.09	2.19	2.20	2.21	1.83	2.14	2.18	2.82	2.83	3.70
G3	1.75	1.98	2.01	2.11	2.11	2.12	1.76	2.07	2.09	2.82	2.82	3.72
G4	1.65	1.88	1.90	2.01	2.02	2.02	1.66	1.96	1.99	2.83	2.84	3.76
G5	1.52	1.74	1.76	1.87	1.88	1.87	1.52	1.83	1.84	3.26	3.28	4.39
	Pattern II											
G1	2.11	2.55	2.68	2.76	2.76	2.86	2.23	2.51	2.60	2.61	2.62	2.60
G2	1.69	2.29	2.27	2.38	2.38	2.54	1.82	2.13	2.15	2.16	2.17	2.15
G3	1.62	2.24	2.21	2.31	2.32	2.48	1.74	2.05	2.07	2.08	2.08	2.06
G4	1.52	2.18	2.12	2.24	2.25	2.40	1.64	1.95	1.96	1.97	1.98	1.95
G5	1.22	1.99	1.89	2.04	2.05	2.19	1.33	1.61	1.61	1.62	1.62	1.60
	Pattern III											
G1	2.10	2.69	3.36	2.25	2.25	2.24	2.37	2.58	2.72	2.73	2.73	2.72
G2	1.51	2.71	3.04	2.18	2.18	2.18	1.82	2.07	2.10	2.12	2.12	2.09
G3	1.44	2.72	3.03	2.09	2.10	2.09	1.75	1.98	2.01	2.03	2.03	2.00
G4	1.34	2.73	3.04	1.98	1.99	1.98	1.65	1.88	1.90	1.92	1.93	1.89
G5	0.81	3.05	3.47	1.84	1.84	1.83	1.04	1.20	1.20	1.22	1.23	1.19

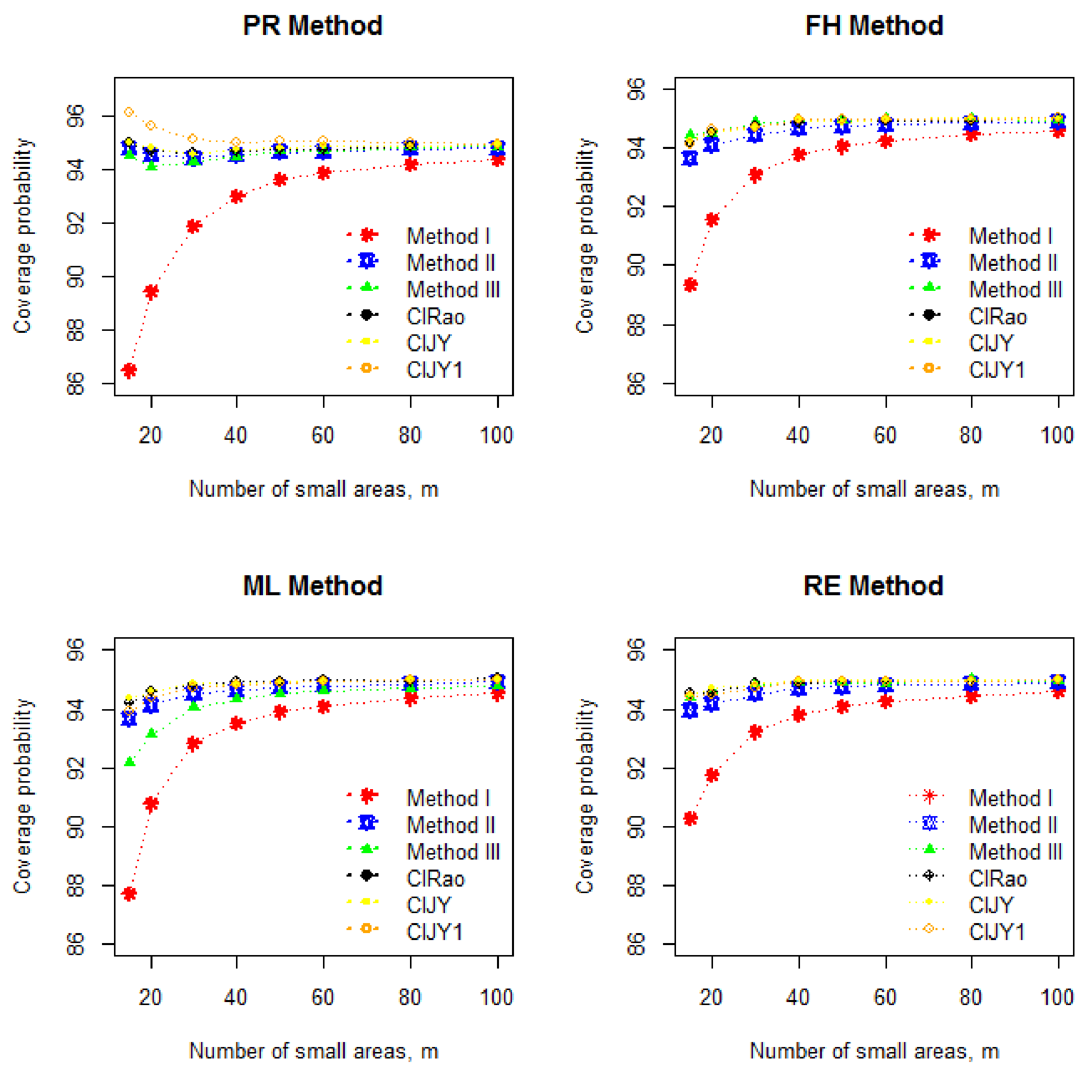

Figure 5.7: Simulated values of the CPs of Method I, Method II, Method III, CIRao, CIJY and CIJY1 for nominal 95% CIs for ψ_i pattern (II). The points are connected with lines for visibility purposes.

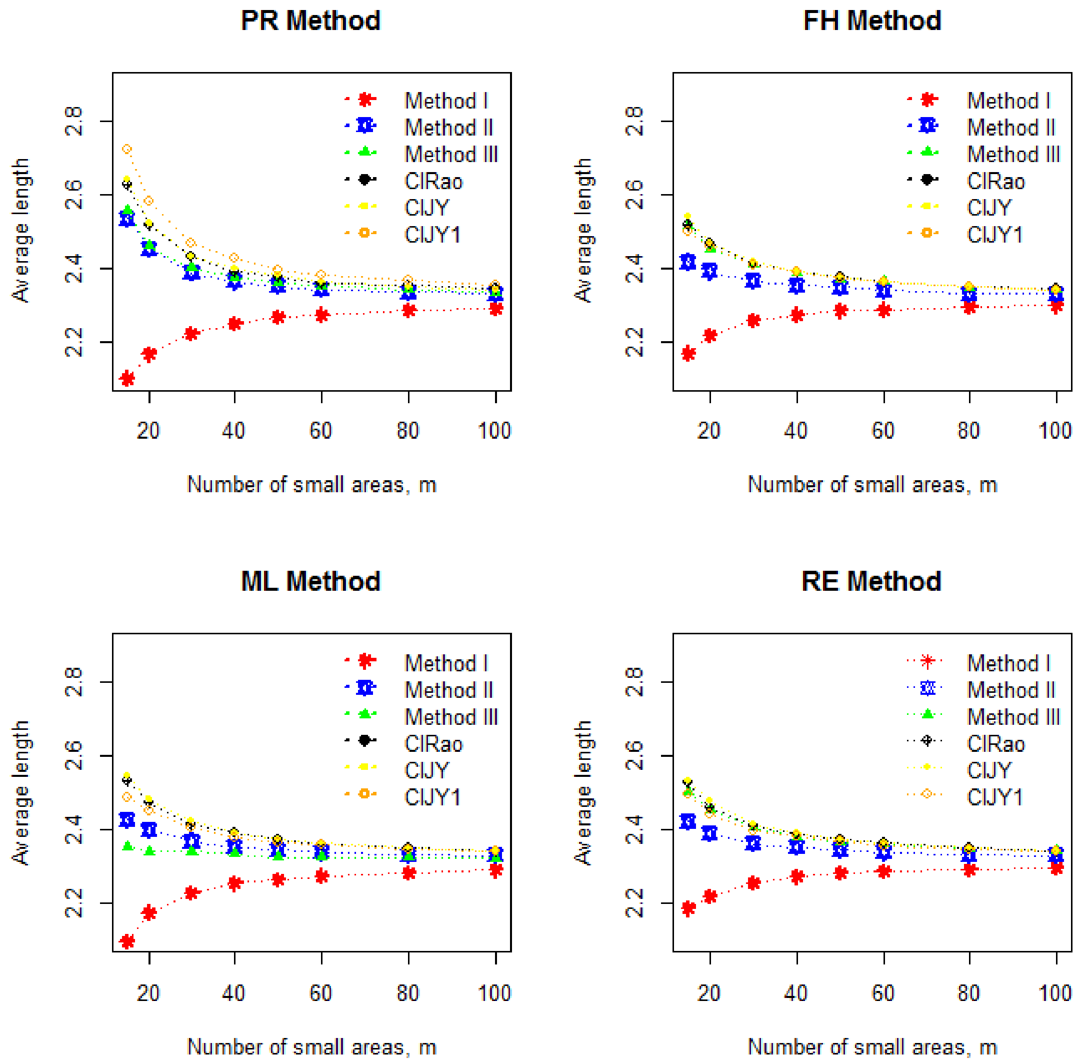

Figure 5.8: Simulated values of the ALs of Method I, Method II, Method III, CIRao, CIJY and CIJY1 for nominal 95% CIs for ψ_i pattern (II). The points are connected with lines for visibility purposes.

5.3 Comparison of Confidence Intervals for the Difference of Two Small Area Means: Simulation Study III

A simulation study is conducted to investigate the performance of the proposed methods with the Cox (1975) and Prasad and Rao (1990) type CIs. The performance measures are the CP and AL for the 95% CIs. The results are based on the 10000 simulations under the three different sampling variance patterns provided by Datta et al. (2005). We also adopt the simulation set up by Datta et al. (2005) with some modification. We wrote an extension of SAS codes written by Li (2007) for CIs for a small area mean to the difference of two small area means.

5.3.1 Comparison of Confidence Intervals Based on the Weighted Estimator with Fixed weights

The modified Cox (1975) EB interval for $\theta_i - \theta_j$ based on $\hat{\theta}_i^w - \hat{\theta}_j^w$ where $i \neq j$ is given by:

$$\text{Method IWD (MIWD)} : \hat{\theta}_i^w - \hat{\theta}_j^w \pm z_{\alpha/2}\left(\psi_i^{1/2}(1-\hat{\gamma}_i)^{1/2} + \psi_j^{1/2}(1-\hat{\gamma}_j)^{1/2}\right). \tag{5.3.1}$$

The modified Prasad and Rao (1990) EB interval for $\theta_i - \theta_j$ based on $\hat{\theta}_i^w - \hat{\theta}_j^w$ where $i \neq j$ is written as:

$$\text{Method IIWD (MIIWD)} : \hat{\theta}_i^w - \hat{\theta}_j^w \pm z_{\alpha/2}\sqrt{\widehat{\text{MSE}}(\hat{\theta}_i^w) - \widehat{\text{MSE}}(\hat{\theta}_j^w)}. \tag{5.3.2}$$

The improved CI for $\theta_i - \theta_j$ based on $\hat{\theta}_i^w - \hat{\theta}_j^w$ where $i \neq j$ is written as:

$$\text{Method IIIWD (MIIIWD)} : \hat{\theta}_i^w - \hat{\theta}_j^w \pm t_{\alpha_{WD}}\sqrt{\widehat{\text{MSE}}(\hat{\theta}_i^w) - \widehat{\text{MSE}}(\hat{\theta}_j^w)}, \tag{5.3.3}$$

where $t_{\alpha_{WD}} = z_{\alpha/2}(1 + h(\hat{A}))$.

$$\begin{aligned}
h(\hat{A}) = &\frac{z^2_{\alpha/2}+1}{8}\left(\frac{A\psi_i}{A+\psi_i} + \frac{A\psi_j}{A+\psi_j}\right)^{-2}\left(\frac{\psi_i^2}{(A+\psi_i)^2} + \frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A}) \\
&+ \left(\frac{A\psi_i}{A+\psi_i} + \frac{A\psi_j}{A+\psi_j}\right)^{-1}\Bigg[(\gamma_i - w_i)^2 \mathbf{x}_i'\left(\sum \frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i \\
&- \frac{\psi_j^2}{(A+\psi_j)^2}\mathbf{x}_i'\left(\textstyle\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i}\right)^{-1}\mathbf{x}_i + (\gamma_j - w_j)^2\mathbf{x}_i'\left(\sum \frac{\mathbf{x}_i'\mathbf{x}_i}{A+\psi_i}\right)^{-1}\mathbf{x}_i \\
&+ \frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\textstyle\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j\Bigg].
\end{aligned}$$

Note that for $t_{\alpha_{WD}}$ the improved CI given in (4.2.3) satisfies that
$P\left[\theta_i - \theta_j \in \{\hat{\theta}_i^w - \hat{\theta}_j^w \pm t_{\alpha_{WD}}\sqrt{\widehat{\text{MSE}}(\hat{\theta}_i^w) - \widehat{\text{MSE}}(\hat{\theta}_j^w)}\}\right] = 1 - \alpha + O(m^{-3/2})$.

5.3.2 Discussion of Simulation Results

Similar to the one population case, in this section we compare the three different CIs given in (5.3.1), (5.3.2) and (5.3.3), which are referred as Method IWD, Method IIWD and Method IIIWD, respectively. The estimated CPs and ALs of these CIs from the simulation study are summarized in Table 5.10 and Figures 5.9-5.11. We obtain the estimated results based on the 10,000 simulation runs.

We proposed the improved CIs for the difference between two population means based on the weighted estimator with fixed weights. As seen in Table 5.10, Method IIIWD has larger CP than Method IWD and Method IIWD, indicating that the coverage is quite accurate and about right. However, the AL of Method IIIWD is a bit larger than Method IWD and Method IIWD. This is because wider CIs have higher CPs while narrower CIs have lower CPs. The PR method seems to produce a little bit lower CP than the ML, REML and FH methods for pattern III. Method IWD with nominal 95% coverage may have 67.5% coverage for pattern III if the PR method is used. As we can see the CPs of Method IIIWD are quite close to the nominal confidence coefficient when both m = 30 and m = 60.

Figures 5.9-5.11 plot the estimated CPs of different CIs over a range of number of small areas, m. As seen in these Figures, when the number of small areas, m, increases, the CPs of the Method IIWD and Method IIIWD are close to the nominal level. However, when the number of small areas, m, is not large enough, the CPs of these CIs may be a bit far away from the 95% nominal value.

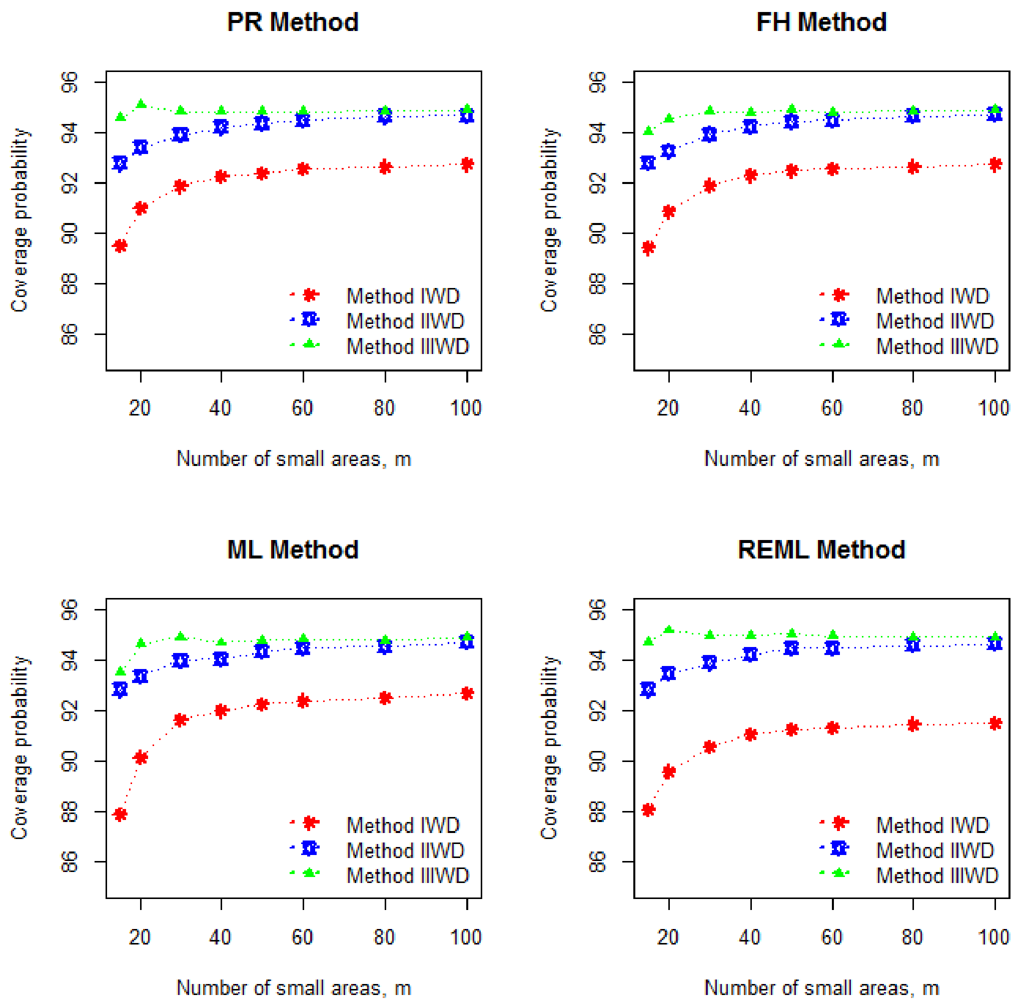

Figure 5.9: Simulated values of the CPs of Method IWD, Method IIWD and Method IIIWD for nominal 95% CIs for ψ_i pattern I. The points are connected with lines for visibility purposes.

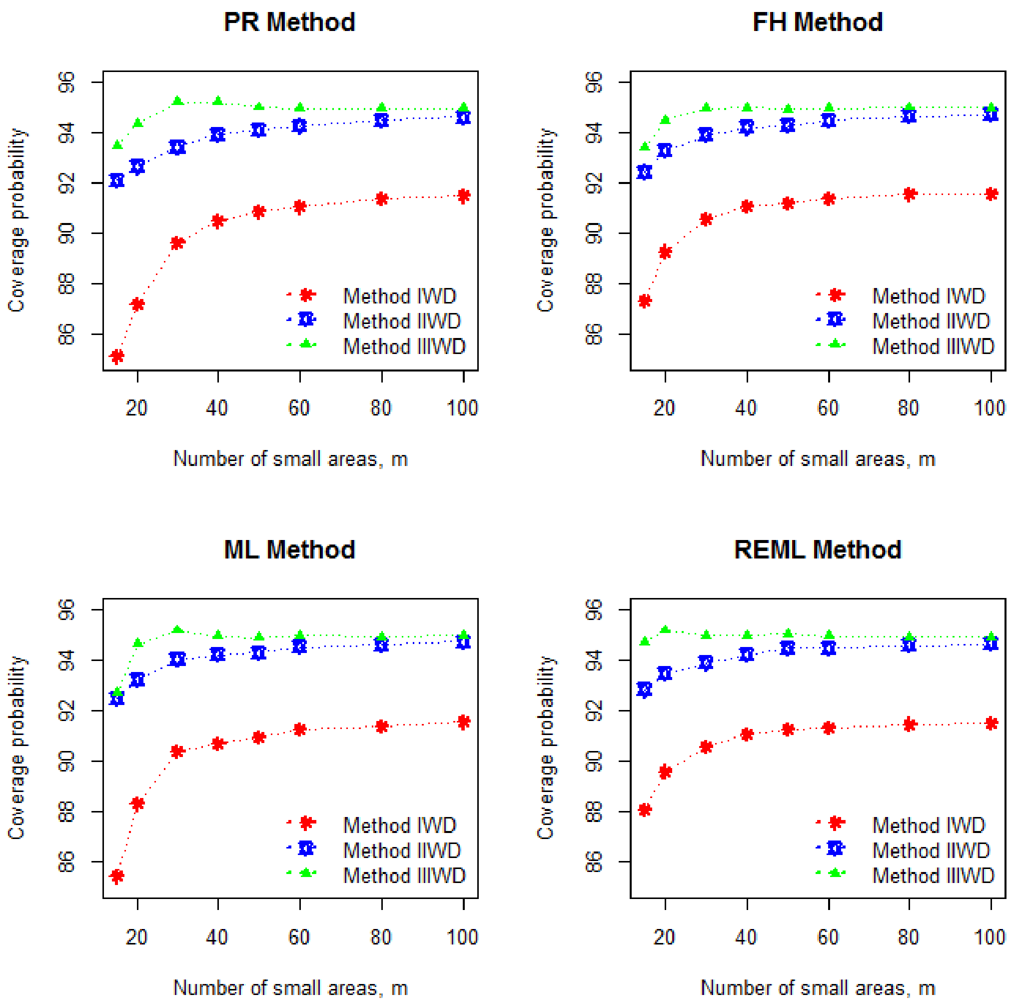

Figure 5.10: Simulated values of the CPs of Method IWD, Method IIWD and Method IIIWD for nominal 95% CIs for ψ_i pattern II. The points are connected with lines for visibility purposes.

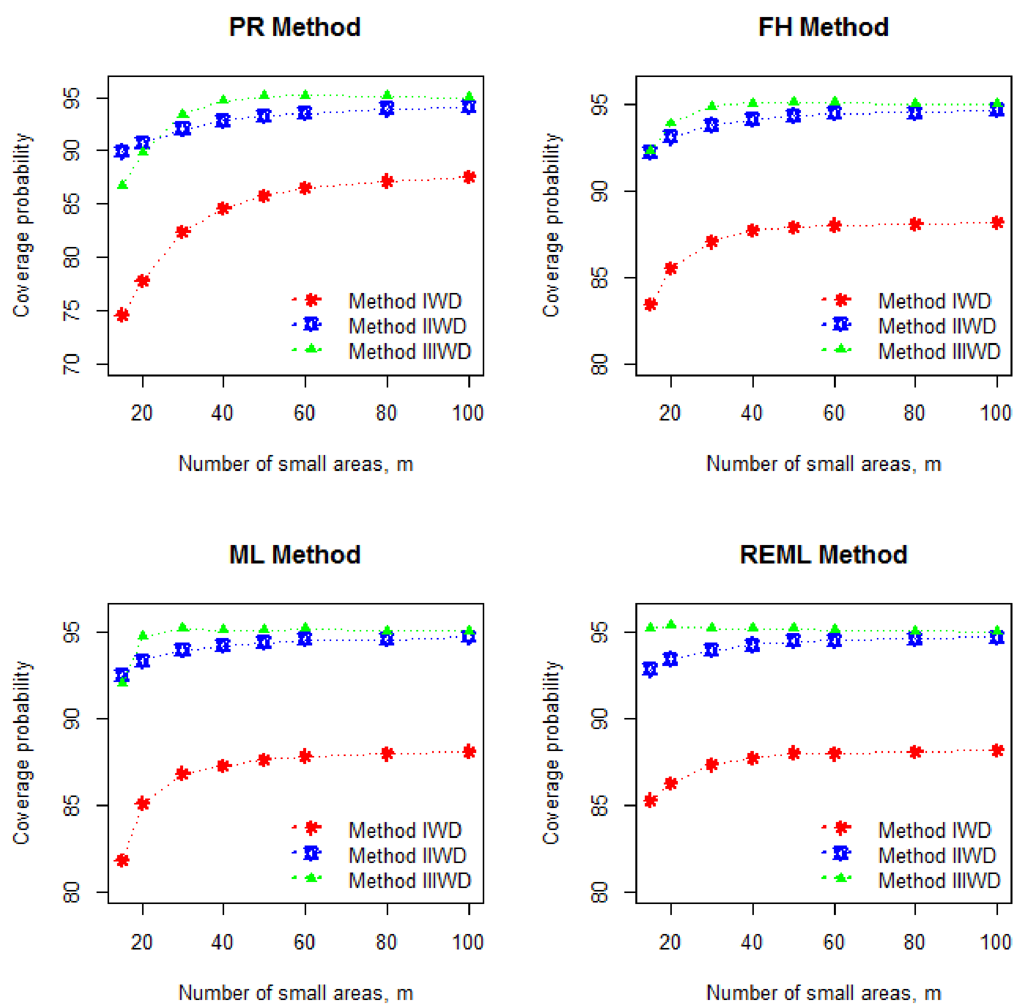

Figure 5.11: Simulated values of the CPs of Method IWD, Method IIWD and Method IIIWD for nominal 95% CIs for ψ_i pattern III. The points are connected with lines for visibility purposes.

Table 5.10: CPs and ALs of the difference between two small area means for nominal 95% CIs for ψ_i pattern (I), (II) and (III). Results obtained for the normal random effects distributions and m = 20.

Groups	CP											
	PR			FH			ML			REML		
	Pattern I											
	MIWD	MIIWD	MIIIWD	MIWD	MIIWD	MIIIWD	MIWD	MIIWD	MIIIWD	MIWD	MIIWD	MIIIWD
G1	92.1	93.4	95.3	91.9	93.3	94.8	91.2	93.4	94.9	92.1	93.5	95.2
G2	91.8	93.3	95.1	91.8	93.3	94.6	91.1	93.6	95.0	92.3	93.7	95.2
G3	91.4	93.4	95.0	91.4	93.3	94.6	90.6	93.4	94.6	91.4	93.3	94.9
G4	90.7	93.5	95.1	90.6	93.3	94.4	89.8	93.2	94.4	90.7	93.3	94.8
G5	89.0	93.3	94.8	88.5	93.0	94.3	88.1	93.3	94.4	88.9	93.4	94.7
	Pattern II											
G1	87.0	94.1	95.7	89.6	94.1	95.9	87.9	94.0	96.0	89.6	94.2	96.5
G2	89.5	92.7	94.5	91.4	93.3	94.4	90.5	93.2	94.5	92.0	93.6	95.1
G3	88.8	92.4	94.1	91.0	93.2	94.3	90.0	93.2	94.5	91.4	93.3	94.9
G4	88.2	92.4	94.0	90.2	93.1	94.0	89.1	92.9	94.2	90.4	93.3	94.7
G5	82.4	91.5	93.4	84.1	92.6	93.7	84.0	92.9	94.0	84.6	93.0	94.8
	Pattern III											
G1	75.1	94.5	91.0	83.6	94.4	95.7	81.9	94.4	96.5	84.1	94.4	97.1
G2	82.6	91.2	89.9	90.8	93.1	93.5	90.7	93.4	94.4	91.8	93.4	95.0
G3	82.3	90.7	89.8	90.4	93.0	93.3	90.3	93.1	94.1	91.6	93.4	94.9
G4	81.4	89.9	89.6	89.6	93.0	93.3	89.4	93.1	94.0	90.6	93.3	94.7
G5	67.5	87.0	88.5	73.3	91.9	93.6	73.4	92.9	94.8	74.4	92.9	95.5
	AL											
	Pattern I											
G1	3.42	3.58	3.54	3.42	3.57	4.36	3.35	3.57	3.95	3.42	3.58	3.71
G2	3.27	3.46	3.39	3.26	3.46	4.27	3.20	3.46	3.83	3.27	3.46	3.58
G3	3.09	3.34	3.24	3.08	3.34	4.18	3.03	3.34	3.72	3.09	3.34	3.45
G4	2.86	3.22	3.09	2.86	3.21	4.10	2.82	3.21	3.60	2.87	3.22	3.32
G5	2.58	3.08	2.94	2.58	3.08	4.05	2.55	3.08	3.51	2.59	3.08	3.20
	Pattern II											
G1	4.19	4.76	5.38	4.33	4.78	5.25	4.17	4.77	5.26	4.33	4.78	5.36
G2	3.16	3.43	3.59	3.25	3.45	3.63	3.18	3.44	3.30	3.26	3.46	3.15
G3	2.98	3.31	3.44	3.07	3.34	3.50	3.01	3.33	3.13	3.08	3.34	2.95
G4	2.77	3.18	3.30	2.85	3.21	3.37	2.80	3.20	2.95	2.86	3.22	2.73
G5	2.14	2.89	3.06	2.19	2.93	3.17	2.17	2.92	2.60	2.20	2.93	2.28
	Pattern III											
G1	4.33	6.17	7.84	4.76	6.18	7.09	4.56	6.17	7.07	4.76	6.18	7.01
G2	2.92	3.39	3.83	3.24	3.45	3.39	3.18	3.45	3.95	3.26	3.46	3.86
G3	2.76	3.27	3.68	3.06	3.33	3.24	3.01	3.33	3.84	3.08	3.34	3.75
G4	2.56	3.13	3.54	2.84	3.21	3.08	2.80	3.20	3.74	2.86	3.21	3.64
G5	1.49	2.64	3.45	1.62	2.75	2.78	1.62	2.75	3.82	1.64	2.77	3.69

5.3.3 Comparison of Confidence Intervals Based on EBLUP

The performances of the four proposed methods are compared to the performance of the two naive methods under different sampling variances. The main purpose of this section is to find out which method performs best under the normal and the non normal random effects distributions.

The modified Cox (1975) EB interval for $\theta_i - \theta_j$ based on $\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}$ where $i \neq j$ is written as:

$$\text{Method ID (MetID)}: \hat{\theta}_i^{EB} - \hat{\theta}_j^{EB} \pm z_{\alpha/2}\left(\psi_i^{1/2}(1-\hat{\gamma}_i)^{1/2} + \psi_j^{1/2}(1-\hat{\gamma}_j)^{1/2}\right). \quad (5.3.4)$$

The modified Prasad and Rao (1990) EB interval for $\theta_i - \theta_j$ based on $\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}$ where $i \neq j$ is written as:

$$\text{Method IID (MetIID)}: \hat{\theta}_i^{EB} - \hat{\theta}_j^{EB} \pm z_{\alpha/2}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB}) - \widehat{MSE}(\hat{\theta}_j^{EB})}. \quad (5.3.5)$$

The improved CIs for $\theta_i - \theta_j$ based on $\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}$ where $i \neq j$ is written as:

$$\text{Method IIID (MetIIID)}: \hat{\theta}_i^{EB} - \hat{\theta}_j^{EB} \pm t_{\alpha_D}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB}) - \widehat{MSE}(\hat{\theta}_j^{EB})}, \quad (5.3.6)$$

where $t_{\alpha_D} = z_{\alpha/2}\left[\frac{(z_{\alpha/2}^2+1)}{8}\left(g_{1i}(A)+g_{1j}(A)\right)^{-2}\frac{\left((A+\psi_i)g_{3i}(A)+(A+\psi_j)g_{3j}(A)\right)^2}{V(A)}+1\right]$.

The corresponding area specific versions of (4.2.6) are written as follows:

$$\text{CIRaoD}: \hat{\theta}_i^{EB} - \hat{\theta}_j^{EB} \pm t_{RaoD}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB}) - \widehat{MSE}(\hat{\theta}_j^{EB})} \quad (5.3.7)$$

$$\text{where } t_{RaoD} = z_{\alpha/2}\left[\frac{(z_{\alpha/2}^2+1)}{8}\left(g_{1i}(A)+g_{1j}(A)\right)^{-2}\frac{\left((A+\psi_i)g_{3Raoi}(A)+(A+\psi_j)g_{3Raoj}(A)\right)^2}{V(A)}+1\right].$$

$$\text{CIJYD}: \hat{\theta}_i^{EB} - \hat{\theta}_j^{EB} \pm t_{JYD}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB}) - \widehat{MSE}(\hat{\theta}_j^{EB})} \quad (5.3.8)$$

where $t_{JYD} = z_{\alpha/2}\left[\frac{(z_{\alpha/2}^2+1)}{8}\left(g_{1i}(A)+g_{1j}(A)\right)^{-2}\frac{\left((A+\psi_i)g_{3JYi}(A)+(A+\psi_j)g_{3JYj}(A)\right)^2}{V(A)}+1\right]$.

Another improved CIs for $\theta_i - \theta_j$ based on $\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}$ where $i \neq j$ is written as:

$$\text{CIJY1D}: \hat{\theta}_i^{EB} - \hat{\theta}_j^{EB} \pm t_{JY1D}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB}) - \widehat{MSE}(\hat{\theta}_j^{EB})} \quad (5.3.9)$$

where $t_{JY1D} = z_{\alpha/2}\left[\frac{(z_{\alpha/2}^2+1)}{8}\left(g_{1i}(A)+g_{1j}(A)\right)^{-2}\frac{\left((A+\psi_i)g_{3JY1i}(A)+(A+\psi_j)g_{3JY1j}(A)\right)^2}{V(A)}+1\right]$.

Note that for t_{α_D}, t_{RaoD}, t_{JYD} and t_{JY1D} the improved CIs given in (5.3.6), (5.3.7), (5.3.8) and (5.3.9) satisfies that

$$P\left[\theta_i - \theta_j \in \{\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB} \pm t_{\alpha_D}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB}) - \widehat{MSE}(\hat{\theta}_j^{EB})}\}\right] = 1 - \alpha + O(m^{-3/2}). \quad (5.3.10)$$

$$P\left[\theta_i-\theta_j \in \{\hat{\theta}_i^{EB}-\hat{\theta}_j^{EB} \pm t_{RaoD}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})-\widehat{MSE}(\hat{\theta}_j^{EB})}\}\right] = 1-\alpha+O(m^{-3/2}). \qquad (5.3.11)$$

$$P\left[\theta_i-\theta_j \in \{\hat{\theta}_i^{EB}-\hat{\theta}_j^{EB} \pm t_{JYD}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})-\widehat{MSE}(\hat{\theta}_j^{EB})}\}\right] = 1-\alpha+O(m^{-3/2}). \qquad (5.3.12)$$

$$P\left[\theta_i-\theta_j \in \{\hat{\theta}_i^{EB}-\hat{\theta}_j^{EB} \pm t_{JY1D}\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})-\widehat{MSE}(\hat{\theta}_j^{EB})}\}\right] = 1-\alpha+O(m^{-3/2}). \qquad (5.3.13)$$

5.3.4 Discussion of Simulation Results

Similar to the previous case, in this section we compare six different CIs given in (5.3.4), (5.3.5), (5.3.6), (5.3.7), (5.3.8) and (5.3.9), which are referred as Method ID, Method IID, Method IIID, CIRaoD, CIJYD and CIJY1D, respectively. Tables 5.11 and 5.12 show the estimated CPs and ALs of all the CIs under the FH model for the normal and the Laplace (double exponential) random effects distribution for pattern (II). We have investigated the numerical performances of these methods under the FH model without covariates based on 10,000 simulation runs. We see that the CPs of Method IIID, CIRaoD, CIJYD and CIJY1D behaved satisfactorily. They never differed from their 95% nominal value by more than 1.9%, 0.8%, 0.8% and 0.2% for the FH, PR, ML and REML estimators respectively for pattern (II) and the normal random effects distribution. Like in the previous section, in the normal and the Laplace cases, the ALs of the proposed methods are a bit larger than Method ID and Method IID especially for large sampling variances (G1 and G2) and small sample sizes. This is to be expected since Method IIID, CIRaoD, CIJYD and CIJY1D increases the width of the intervals to meet the nominal coverage.

As far as the Laplace distribution is concerned, we see that for m = 20 there is an under-coverage problem for all the methods, but the under-coverage problem becomes more serious for Method ID. For this distribution, the CPs of the proposed methods never differed from their 95% nominal value by more than 3.9%, 4.2%, 3.2% and 3.5% (slightly below the 95% nominal value) for the FH, PR, ML and REML estimators respectively for pattern (II). These findings are sensible for this heavy tail distribution.

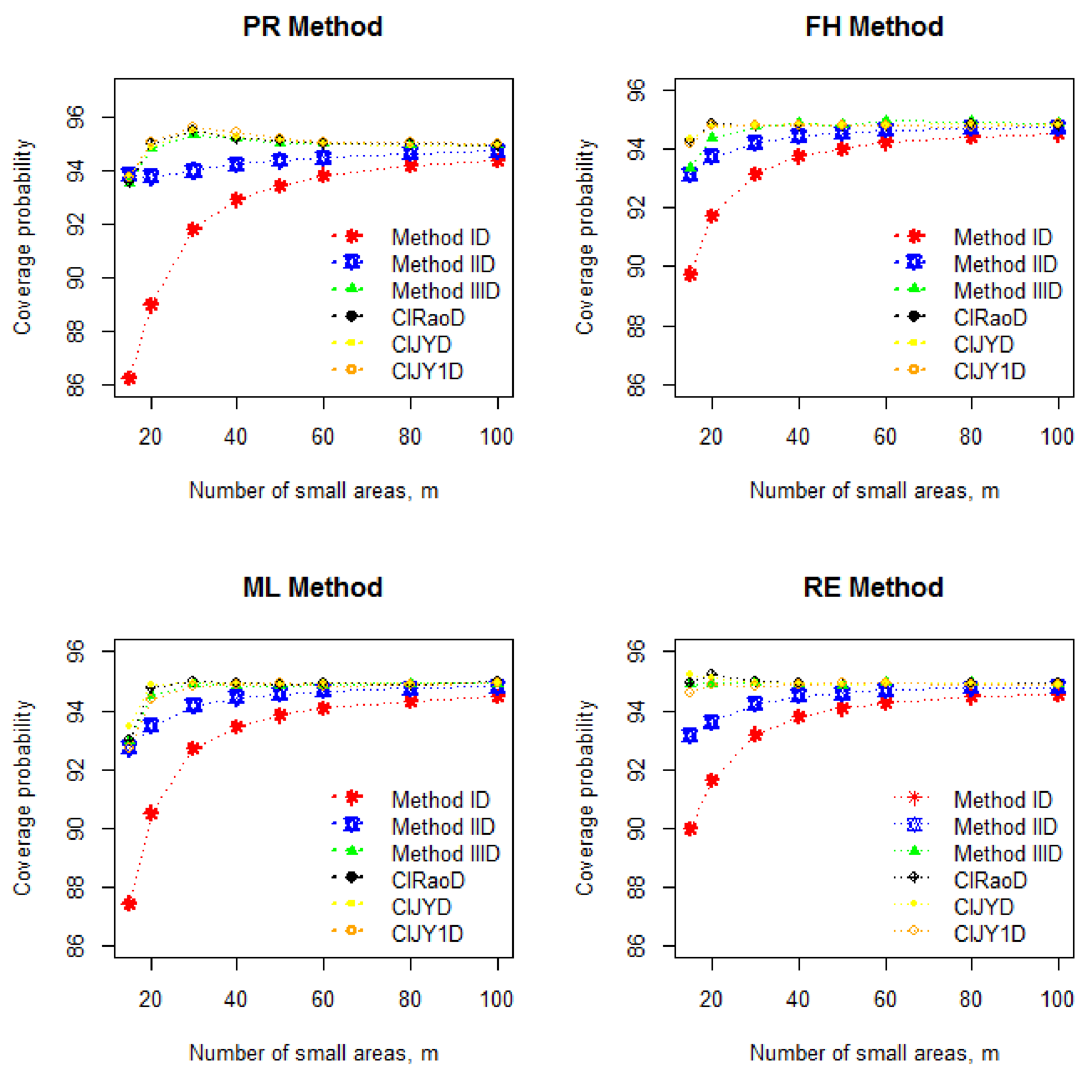

Figure 5.12: Simulated values of the CPs of Method ID, Method IID, Method IIID, CIRaoD, CIJYD and CIJY1D for nominal 95% CIs for ψ_i pattern II. The points are connected with lines for visibility purposes.

Table 5.11: CPs of the difference between two small area means for nominal 95% CIs for ψ_i pattern (II). Results obtained for the normal and the Laplace (double exponential) random effects distributions and m = 20.

Groups	CP											
	Normal Distribution											
	REML						FH					
	MetID	MetIID	MetIIID	CIRaoD	CIJYD	CIJY1D	MetID	MetIID	MetIIID	CIRaoD	CIJYD	CIJY1D
G1	90.5	92.1	94.9	95.2	95.1	94.8	92.5	93.9	94.4	95.9	95.9	96.0
G2	91.5	93.7	94.9	95.2	95.2	94.9	91.8	93.8	94.4	94.9	95.1	95.0
G3	91.9	94.2	94.9	95.4	95.2	95.0	92.2	94.3	94.4	95.2	94.9	95.0
G4	92.1	94.3	95.1	95.2	95.0	94.8	92.2	94.5	94.4	95.1	94.8	94.8
G5	92.7	95.0	94.8	95.2	95.1	94.8	90.7	93.1	94.2	93.1	93.1	93.1
	ML						PR					
G1	89.1	91.7	94.6	94.6	94.8	94.5	87.7	90.7	95.6	95.6	95.6	95.8
G2	90.2	93.6	94.5	94.8	94.8	94.4	88.7	93.5	95.0	95.1	94.8	95.2
G3	90.8	94.0	94.6	94.9	94.9	94.2	89.0	94.0	94.7	95.0	94.9	95.0
G4	91.0	94.2	94.5	94.8	95.1	94.3	89.3	94.5	94.6	94.7	94.8	94.8
G5	91.9	95.1	94.2	94.7	94.9	94.4	90.3	95.7	94.2	94.5	94.3	94.5
	Laplace Distribution											
	REML						FH					
G1	82.8	86.1	94.2	91.6	92.6	92.2	88.5	91.0	89.6	91.3	91.2	90.9
G2	83.9	90.1	94.4	91.8	92.8	92.2	87.9	91.9	90.7	93.6	93.8	93.3
G3	83.9	90.7	94.5	91.8	92.8	92.1	88.1	92.6	91.1	94.4	94.5	94.0
G4	84.1	91.5	94.4	91.9	92.7	92.3	88.3	93.3	92.1	95.1	95.1	94.8
G5	85.4	94.0	94.4	91.5	92.5	91.8	87.3	93.5	93.8	95.8	95.9	95.6
	ML						PR					
G1	85.4	89.3	91.7	92.0	91.9	92.3	81.9	88.0	91.1	91.1	90.8	90.8
G2	86.3	91.9	91.9	92.3	92.5	92.4	82.4	90.7	92.3	92.9	92.8	94.1
G3	86.3	92.3	92.0	92.3	92.3	92.5	82.6	90.7	93.1	93.8	93.7	94.9
G4	86.8	93.1	92.2	93.1	93.2	92.5	82.7	90.4	93.8	94.6	94.6	95.5
G5	87.7	95.2	93.9	95.2	95.0	93.7	83.3	90.1	95.5	96.5	96.6	96.5

Table 5.12: ALs of the difference between two small area means for nominal 95% CIs for ψ_i pattern (II). Results obtained for the normal and the Laplace (double exponential) random effects distributions and m = 20.

Groups	AL											
	REML						FH					
	MetID	MetIID	MetIIID	CIRaoD	CIJYD	CIJY1D	MetID	MetIID	MetIIID	CIRaoD	CIJYD	CIJY1D
Normal Distribution												
G1	4.32	4.54	6.94	5.40	5.79	4.52	4.32	4.54	4.53	4.32	4.72	4.69
G2	3.25	3.46	4.04	3.72	3.95	3.45	3.25	3.46	3.48	3.59	3.67	3.73
G3	3.07	3.26	3.71	3.46	3.62	3.25	3.07	3.27	3.28	3.43	3.65	3.48
G4	2.85	3.02	3.34	3.26	3.26	3.02	2.85	3.03	3.04	3.14	3.47	3.19
G5	2.20	2.31	2.41	2.36	2.37	2.31	2.19	2.32	2.32	2.56	2.62	2.60
	ML						PR					
G1	4.19	4.53	4.36	4.93	3.91	5.25	4.22	4.53	5.30	5.11	5.54	5.25
G2	3.19	3.46	3.43	3.57	3.95	3.72	3.17	3.50	3.73	3.64	3.82	3.75
G3	3.02	3.27	3.25	3.34	3.69	3.47	2.99	3.32	3.49	3.40	3.55	3.50
G4	2.81	3.03	3.03	3.29	3.25	3.18	2.78	3.09	3.19	3.39	3.26	3.21
G5	2.17	2.33	2.32	2.38	2.37	2.36	2.15	2.42	2.37	3.11	2.37	2.38
Laplace Distribution												
	REML						FH					
G1	3.27	3.42	4.06	4.89	4.47	4.10	3.04	3.20	4.30	3.93	4.29	3.98
G2	2.68	2.89	3.24	3.59	4.12	3.31	2.68	2.92	3.35	3.28	3.36	3.52
G3	2.57	2.78	3.08	3.37	2.85	3.15	2.57	2.81	3.17	2.98	3.18	3.34
G4	2.43	2.63	2.87	3.10	3.14	2.94	2.42	2.67	2.94	2.88	2.93	3.16
G5	1.96	2.15	2.24	2.33	2.57	2.28	2.09	2.37	2.27	2.83	2.61	2.82
	ML						PR					
G1	3.38	3.11	4.47	3.83	4.91	3.84	3.40	3.09	4.45	4.63	4.93	4.54
G2	2.90	2.57	3.42	2.93	3.87	3.15	3.02	2.51	3.46	3.55	3.63	3.59
G3	2.80	2.47	3.23	2.79	3.06	3.01	2.93	2.40	3.27	3.52	3.61	3.39
G4	2.66	2.33	2.99	2.30	2.35	2.82	2.83	2.27	3.04	3.25	3.32	3.15
G5	2.22	1.89	2.28	2.26	2.30	2.22	2.49	1.83	2.32	2.37	3.05	2.40

5.4 Comparison of the Proposed Methods with the Existing Methods

5.4.1 Comparison of Confidence Intervals for a Small Area Mean

Here we consider two existing methods namely Diao et al. (2013)'s method (based on the Taylor serious expansion) and Chatterjee et al. (2008)'s method (based on bootstrapping). We adopt the simulation set up of Chatterjee et al. (2008) given by Pattern (IV) (4.0, 0.6, 0.5, 0.4, 0.2). Table 5.13 and Table 5.14 show the CPs and ALs for nominal 95% CIs for Pattern (IV) under the normal and the chi-square distributions. Note that Met III, CIRao, CIJY, CIJY1 and Met IIIW represent our proposed methods whereas PB-ET represents the parametric bootstrap corrected CI proposed by Chatterjee et al. (2008). Diao-EB represents the corrected CI proposed by Diao et al. (2013).

Table 5.13: CPs and ALs of a small area mean for nominal 95% CIs for ψ_i pattern (IV). Results obtained for the normal random effects distributions and $m = 15$. Note that the first value in the cell is the proposed method and the value in the parenthesis is the naive method.

ψ_i		EBLUP								Weighted		Diao-EB		Chatterjee
		Met III		CIRao		CIJY		CIJY1		Met IIIW				
		FH	REML	FH	REML	FH	REML	FH	REML	FH	REML	FH	REML	PB-ET
							Normal							
4.0	CP	93.7	93.4	93.8	93.7	93.6	93.6	93.9	94.1	96.1	97.1	96.0	95.3	95.7
		(91.2)	(91.7)	(90.9)	(92.3)	(91.1)	(91.9)	(91.4)	(92.6)	(95.0)	(94.2)	(91.9)	(90.7)	
	AL	4.00	3.91	4.02	4.01	3.95	3.99	4.01	4.04	4.80	5.16	4.26	4.10	4.73
		(3.57)	(3.54)	(3.59)	(3.64)	(3.53)	(3.62)	(3.58)	(3.66)	(4.45)	(4.45)	(3.79)	(3.62)	
0.6	CP	94.4	93.7	94.2	95.1	93.8	95.2	94.5	95.2	94.3	94.6	96.4	96.1	95.6
		(93.9)	(93.2)	(93.7)	(94.8)	(93.3)	(94.8)	(94.0)	(94.8)	(93.6)	(93.8)	(94.5)	(93.7)	
	AL	2.54	2.47	2.54	2.61	2.51	2.58	2.54	2.60	2.48	2.71	2.73	2.69	2.95
		(2.49)	(2.42)	(2.49)	(2.56)	(2.45)	(2.53)	(2.49)	(2.55)	(2.33)	(2.47)	(2.57)	(2.52)	
0.5	CP	94.4	93.9	94.5	95.4	94.3	95.0	94.4	95.4	94.0	94.5	96.3	96.1	95.5
		(93.9)	(93.6)	(94.0)	(95.0)	(93.9)	(94.7)	(94.0)	(95.1)	(93.3)	(93.8)	(94.8)	(94.2)	
	AL	2.39	2.32	2.39	2.45	2.36	2.42	2.39	2.44	2.39	2.61	2.56	2.53	2.70
		(2.34)	(2.28)	(2.35)	(2.41)	(2.32)	(2.39)	(2.34)	(2.41)	(2.21)	(2.38)	(2.42)	(2.37)	
0.4	CP	94.6	94.1	94.4	95.5	94.3	95.0	94.5	95.4	93.8	94.4	96.4	96.2	95.6
		(94.3)	(93.8)	(94.1)	(95.3)	(94.0)	(94.8)	(94.2)	(95.1)	(93.6)	(93.8)	(95.1)	(94.6)	
	AL	2.20	2.14	2.20	2.26	2.18	2.23	2.20	2.25	2.29	2.51	2.36	2.34	2.49
		(2.17)	(2.11)	(2.18)	(2.23)	(2.15)	(2.21)	(2.17)	(2.22)	(2.09)	(2.29)	(2.24)	(2.20)	
0.2	CP	95.2	94.7	95.0	95.5	95.1	95.5	95.1	95.6	93.8	94.6	95.9	95.8	95.9
		(95.1)	(94.6)	(94.9)	(95.5)	(95.0)	(95.4)	(95.0)	(95.6)	(93.3)	(93.9)	(95.4)	(95.1)	
	AL	1.67	1.62	1.68	1.70	1.66	1.69	1.67	1.69	2.07	2.33	1.81	1.79	1.83
		(1.66)	(1.61)	(1.67)	(1.70	(1.66)	(1.68)	(1.66)	(1.69)	(1.84)	(2.07)	(1.70)	(1.68)	

Table 5.14: CPs and ALs of a small area mean for nominal 95% CIs for ψ_i pattern (IV). Results obtained for the chi-square random effects distributions and m = 15. Note that the first value in the cell is the proposed method and the value in the parenthesis is the naive method.

ψ_i		EBLUP								Weighted		Diao-EB		Chatterjee
		Met III		CIRao		CIJY		CIJY1		Met IIIW				
		FH	REML	FH	REML	FH	REML	FH	REML	FH	REML	FH	REML	PB-ET
Centered chi-square														
4	CP	93.2	93.5	93.4	93.6	93.3	93.5	93.0	93.7	95.6	97.2	91.6	92.3	91.7
		(91.6)	(92.1)	(91.8)	(92.4)	(91.6)	(92.0)	(91.5)	(92.3)	(94.7)	(95.1)	(85.3)	(87.7)	
	AL	5.17	4.24	5.21	4.27	5.14	4.21	5.20	4.24	4.82	4.89	3.60	3.51	4.35
		(4.24)	(3.98)	(4.26)	(4.00)	(4.21)	(4.00)	(4.26)	(3.99)	(4.61)	(4.45)	(2.99)	(2.87)	
0.6	CP	94.5	94.6	95.1	95.5	94.6	95.1	94.4	94.7	93.9	94.8	95.4	94.5	92.3
		(94.0)	(94.1)	(94.8)	(95.1)	(94.1)	(94.7)	(93.9)	(94.3)	(93.1)	(94.0)	(91.6)	(90.9)	
	AL	2.65	2.59	2.68	2.62	2.64	2.58	2.65	2.59	3.03	2.48	2.94	2.75	2.71
		(2.59)	(2.43)	(2.62)	(2.43)	(2.58)	(2.44)	(2.59)	(2.43)	(2.22)	(2.47)	(2.18)	(2.10)	
0.5	CP	94.7	94.8	95.4	95.5	94.9	95.1	94.6	94.8	93.7	94.5	95.8	95.1	92.4
		(94.4)	(94.4)	(95.1)	(95.2)	(94.6)	(94.7)	(94.3)	(94.4)	(92.9)	(93.8)	(92.1)	(91.5)	
	AL	2.47	2.42	2.50	2.45	2.46	2.42	2.47	2.42	2.95	2.39	2.93	2.72	2.45
		(2.43)	(2.28)	(2.45)	(2.28)	(2.42)	(2.28)	(2.43)	(2.28)	(2.11)	(2.36)	(2.10)	(2.02)	
0.4	CP	94.8	94.8	95.6	96.0	95.0	95.3	94.7	95.1	93.7	94.7	96.3	96.0	93.3
		(94.5)	(94.5)	(95.3)	(95.8)	(94.7)	(95.0)	(94.4)	(94.9)	(92.9)	(94.1)	(93.1)	(92.4)	
	AL	2.26	2.22	2.28	2.25	2.26	2.23	2.26	2.22	2.87	2.29	2.94	2.69	2.34
		(2.23)	(2.09)	(2.25)	(2.10)	(2.23)	(2.10)	(2.23)	(2.10)	(1.99)	(2.25)	(2.01)	(1.94)	
0.2	CP	95.2	95.3	96.1	96.0	95.6	95.6	95.2	95.4	93.6	94.4	97.5	97.6	94.4
		(95.1)	(95.3)	(96.0)	(95.9)	(95.5)	(95.5)	(95.1)	(95.3)	(93.0)	(93.7)	(96.3)	(96.6)	
	AL	1.69	1.67	1.72	1.70	1.71	1.68	1.69	1.67	2.69	2.08	3.36	2.91	1.73
		(1.68)	(1.58)	(1.71)	(1.58)	(1.70)	(1.58)	(1.68)	(1.58)	(1.73)	(2.03)	(1.78)	(1.71)	

5.4.2 Comparison of Confidence Intervals for the Difference of Two Small Area Means

Table 5.15 shows the CPs and ALs for nominal 95% CIs for the difference of two small area means for Pattern (I) and the normal distribution. This table presents the comparison between the proposed methods and Diao et al. (2013) method.

Table 5.15: CPs and ALs of the difference between two small area means for nominal 95% CIs for ψ_i pattern (I). Results obtained for the normal random effects distributions and m = 15. Note that the first value in the cell is the proposed method and the value in the parenthesis is the naive method.

Pattern (I)		EBLUP								Weighted		Diao-EB	
		Met IIID		CIRaoD		CIJYD		CIJY1D		Met IIIWD			
		FH	REML	FH	REML	FH	REML	FH	REML	FH	REML	FH	REML
0.7	CP	94.4	95.0	94.6	95.2	94.0	95.1	94.2	95.0	94.4	95.0	95.2	95.2
		(92.8)	(92.7)	(93.1)	(93.3)	(92.6)	(92.9)	(92.7)	(92.7)	(93.0)	(92.9)	(90.0)	(89.8)
	AL	3.89	4.02	3.89	3.95	3.8	3.99	4.09	4.13	3.54	3.91	4.27	4.22
		(3.64)	(3.64)	(3.65)	(3.66)	(3.61)	(3.62)	(3.63)	(3.63)	(3.56)	(3.55)	(3.48)	(3.48)
0.6	CP	94.5	95.2	94.8	95.3	94	94.9	94.3	95.0	94.1	94.8	94.8	94.7
		(93.2)	(93.4)	(93.7)	(93.5)	(93.0)	(93.2)	(93.0)	(92.9)	(93.0)	(92.8)	(91.0)	(90.8)
	AL	3.69	3.8	3.58	3.71	3.72	3.82	3.85	3.88	3.38	3.78	3.99	3.95
		(3.48)	(3.48)	(3.50)	(3.50)	(3.46)	(3.47)	(3.47)	(3.47)	(3.44)	(3.44)	(3.39)	(3.36)
0.5	CP	94.3	95.1	94.6	95.4	94.1	95	94.2	94.8	94.1	94.7	95.2	95.1
		(93.3)	(93.4)	(93.8)	(94.1)	(93.5)	(93.6)	(93.3)	(93.4)	(92.9)	(92.7)	(92.0)	(91.9)
	AL	3.45	3.54	3.32	3.45	3.38	3.55	3.58	3.6	3.21	3.66	3.85	3.80
		(3.29)	(3.29)	(3.30)	(3.31)	(3.28)	(3.28)	(3.28)	(3.28)	(3.32)	(3.32)	(3.22)	(3.22)
0.4	CP	94.4	94.7	94.4	95.2	94.1	95.0	94.3	94.9	94.3	94.4	95.6	95.6
		(93.9)	(93.5)	(94.2)	(94.3)	(93.8)	(94.1)	(93.8)	(93.8)	(92.4)	(92.6)	(93.8)	(94.0)
	AL	3.17	3.23	3.09	3.23	3.16	3.2	3.26	3.28	3.04	3.54	3.48	3.45
		(3.06)	(3.06)	(3.06)	(3.08)	(3.06)	(3.06)	(3.05)	(3.05)	(3.19)	(3.19)	(3.07)	(3.06)
0.3	CP	94.4	94.7	94.8	95.1	94.5	94.8	94.5	94.7	94.2	94.3	95.3	95.3
		(94.4)	(94.1)	(94.4)	(94.7)	(93.9)	(94.6)	(94.4)	(94.0)	(92.2)	(92.5)	(94.0)	(94.3)
	AL	2.82	2.85	2.79	2.82	2.76	2.82	2.87	2.89	2.87	3.44	3.26	3.21
		(2.76)	(2.76)	(2.78)	(2.79)	(2.77)	(2.77)	(2.75)	(2.75)	(3.05)	(3.05)	(2.78)	(2.77)

Our proposed method based on the weighted estimator with fixed weights has slightly lower CPs especially for lower sampling variances. However, the CIs proposed by Diao et al. (2013) and Chatterjee et al. (2008) are slightly more conservative than the proposed method. In terms of AL, the CIs proposed by Diao et al. (2013) and Chatterjee et al. (2008) performed slightly better than our proposed method. Our method performs well when the sampling variance is high when compared to the model variance.

Furthermore, the CPs of the proposed methods based on EBLUP are slightly smaller than the nominal level of 0.95. However the CPs of Diao et al. (2013) and Chatterjee et al. (2008) are slightly greater than the nominal level when the sampling variance, ψ_{i} is 4. In general, the CPs and ALs of the proposed methods are comparable to Diao et al. (2013) and Chatterjee et al. (2008). Moreover, the extent of underestimation to the CPs of the proposed methods is small and ignorable.

Chapter 6

Application to Real Data

In this chapter, we illustrate the proposed methods in an application to a real data set relating to the estimation of poverty and percentage of food expenditure measures in all zones of Ethiopia. To examine the performance of our method, and to compare it with others, we use an example of real data. One of the aims in this thesis is to derive indirect estimators which borrow strength from related zones, based on a model linking all zones called the FH model. This thesis is the first to report model-based (EBLUP) estimates of poverty and percentage of food expenditure measures at the zone levels in Ethiopia. We have adapted the part of the codes from package 'sae' by Molina and Marhuenda (2013) with modification for the real data analysis.

6.1 Data Sources

6.1.1 Household Consumption Expenditure Survey

The HCE survey was conducted by the government statistical agency, Central Statistical Authority (CSA) in 2010/11. The survey covered the population in sedentary areas of all regions that included the rural and selected urban areas.

Sampling Design

In order to obtain representative sample at the national level, four different categories are considered in the HCE survey such as rural, major urban centers and other urban centers (medium and small size town's category). Let us briefly state each category as follows.

Category I - Rural: - this category consists of 86 zones and special woredas within the 8 regions of Ethiopia. There are 10 reporting levels in this category. The sampling methods that were implemented under this category was two stage cluster sampling by considering enumeration areas as the primary sampling units and 12 households as the secondary sampling units (CSA, 2012).

Category II - Major urban centers:- this category consists of 15 urban centers. There are 24 reporting levels (14 urban centers and 10 Sub cities of Addis Ababa administration) in this category. Similarly, the sampling methods that were implemented under this category was two stage cluster sampling by considering enumeration areas as the primary sampling units and 16 households as the secondary sampling units (CSA, 2012).

Category III and IV - Other urban centers:- this category includes other urban centers for eight regions excluding Harari, Addis Ababa and Dire Dawa since they don't have urban centers other than those grouped under category II. In this case, the sampling methods that were implemented under this category was a stratified three stage cluster sampling by considering urban centers as the primary sampling units, enumeration areas as the second stage sampling units and 16 households are selected from each of the selected enumeration areas as the third stage sampling units (CSA, 2012).

Sample Size

At country level in 2010/11 HCE, a total of 28032 households (10,368 in rural and 17,664 in urban areas) were planned to be covered (CSA, 2012). However, 202 households were not covered due to various reasons. Thus, 27830 households were successfully covered by the survey. In order to fill the possible gap in the survey data we used the auxiliary information from the 2007 Ethiopian household and census data.

6.1.2 The 2007 Population and Housing Census of Ethiopia

The census provides basic data for small areas and small population groups at all levels. The 2007 census of Ethiopia is the most recent census. It provides comprehensive information on housing and on household socio-demographic conditions, along with the characteristics of individuals household members such as sex, age, marital status, educational attainment, employment status, and household size, thus allowing for the finest geographical dis-aggregation (Federal Democratic Republic of Ethiopia Population and Census Commission, 2008).

6.2 Small Area Estimation of Percentage of Food Expenditure in Ethiopia, 2011

In this section, we illustrate the proposed methods to estimate the percentage of food expenditure by using the 2010/11 HCE and the 2007 census data sets of Ethiopia. Expenditure refers to total expenditure, including accommodation, food, purchases, travel, leisure activities and miscellaneous expenditure. The proposed CIs are used to estimate the percentage of food expenditure relative to total expenditure in 94 small areas (zones) in Ethiopia. In 2012, only 6.6% of consumer expenditure goes to the cost of food consumed at home relative to total consumption expenditures in America. The amount of income spent on food fell from 17.5% to 9.7% between 1960 and 2007 in America (Morrison, 2014). However, countries such as Spain, Norway, Belgium and France spend twice as much on food compared to America. Whereas countries like Mexico, China and Turkey spend three times as much on food compared to America. In contrast, people in poor countries are forced to spend the largest share of income on food (The Economist, 2013). The percentage of food expenditure spent on food in 2012 for some African countries such as Ethiopia, Kenya, Niger, Namibia and South Africa are 39.84%, 47.44%, 44.99%, 24.36% and 26.04%, respectively ("Household consumption expenditure", 2014). Considering this, food prices started increasing in 2007 and exerting pressure on household budgets while low income households spend a greater percentage of income on food (Carr, Lee, Scaife and Hayes, 2014). If you spend very little on food, that leaves more money for things like health care, education, energy, home and saving.

The percentage of food expenditure can be calculated as

$$\mathrm{y_i} = \left(\frac{\text{food expenditure}}{\text{total expenditure}}\right) * 100\%. \tag{6.2.1}$$

6.3 Small Area Estimation of Poverty in Ethiopia, 2011

Poverty estimation at zonal level will be accomplished by using the 2010/11 HCE survey and 2007 census data sources. A set of variables common to the survey and census are required so that small area estimates can be produced at appropriate geographic levels.

6.3.1 Method of Measuring Poverty in Small Areas

The generic dictionary meaning of the term "poor" is a lack of means for comfortable subsistence and "poverty" is the state of being poor (Ray, Mazhari, Passah and Pandey, 2000). The Ministry of Finance and Economic Development (MoFED, 2012) in Ethiopia studied the prevalence of poverty at the national and regional levels in Ethiopia using the 2010/11 HCE survey. According to the MoFED (2012) report, the prevalence of poverty at the national level is 29.6% in 2010/11.

Consider a finite population D of size N partitioned into m areas D_i, $i = 1,...,m$. The response variable is the estimate of poverty measures based on 2010/11 HCE survey data. An Ethiopian household is considered in poverty when its consumption expenditure is below a poverty line per adult person per year (Birr) 3,781 (MoFED, 2012). The Birr is the unit of currency in Ethiopia, and a poverty line can be explained as the level of income sufficient to purchase the minimum standard of nutrition/food (Citro and Kalton, 2000, Ray et al., 2000).

The most widely used poverty measures are the percentage of the poor (head count ratio), the aggregate poverty gap (poverty gap index), and the distribution of income among the poor (poverty severity index). These measures can be defined in terms of the well-known Foster, Greer, and Thorbecke (1984) P_α class of poverty measures. These methods can be modified for small area estimation as follows. Let $(E_{i1},...,E_{in_i})'$ be a vector of real per-adult household expenditure for area i, $Z > 0$ is poverty line and n_i is the total number of households in area i. For $\alpha \geq 0$, let $P_{\alpha i}$ be defined by

$$P_{\alpha i} = \frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{Z-E_{ij}}{Z}\right)^{\alpha} I(E_{ij} < Z); \quad i = 1,...,m,\ j = 1,...,n_i, \tag{6.3.1}$$

where $I(E_{ij} < Z) = 1$ if $E_{ij} < Z$ (person under poverty) and $I(E_{ij} < Z) = 0$ otherwise (person not under poverty), $P_{\alpha i}$ is the poverty measures in area i.

If $\alpha = 0$, the poverty index is called the head count ratio or poverty prevalence (P_{0i}). Hence P_{0i} corresponds to the proportion of individuals under poverty in small area i.

If $\alpha = 1$, the poverty index is called the poverty gap index (P_{1i}) and measures the area mean of the relative distance to the poverty line (the poverty gap) of each individual. Poverty gap ratio can also be interpreted as an indicator of potentials for eliminating poverty by targeting transfers to the poor.

If $\alpha = 2$, the poverty index is called the poverty severity index and it measures the squared proportional shortfalls from the poverty line. This measure squares, and large values of P_{2i} point out to areas with severe level of poverty.

Let $S \subset D$ be a sample drawn from the population and let $S_i = S \bigcap D_i$ be the sample from area i. A direct estimator of $P_{\alpha i}$ for a sampled domain is the unweighted sample mean

$$\hat{P}_{\alpha i} = \frac{1}{n_i}\sum_{j \in S_i}\left(\frac{Z-E_{ij}}{Z}\right)^{\alpha} I(E_{ij} < Z); \quad i = 1,...,m,\ j = 1,...,n_i, \alpha = 0,1,2. \tag{6.3.2}$$

Let w_{ij} be the sampling weight of individual j from sampled area i. Then an approximately design-unbiased estimator of $P_{\alpha i}$ based on the 2010/11 HCE survey is the weighted sample mean given by

$$\hat{P}_{\alpha i} = \frac{1}{\hat{N}_i} \sum_{j \in S_i} w_{ij} \left(\frac{Z - E_{ij}}{Z} \right)^{\alpha} I(E_{ij} < Z); \quad i = 1, ..., m, \; j = 1, ..., N_i, \alpha = 0, 1, 2 \tag{6.3.3}$$

where $\hat{N}_i = \sum_{j \in S_i} w_{ij}$ is a design-unbiased estimator of the population size N_i of sampled area i. If the sampling weights w_{ij} do not depend on the unit j, as it occurs for example under simple random sampling within areas where $w_{ij} = \frac{n_i}{N_i}$, $j = 1, ..., N_i$, then (6.3.3) reduces to the unweighted sample mean (6.3.2) (Molina and Rao, 2010).

One of the aims of this thesis is to derive estimators of poverty measures which “borrow strength” from related zones, based on a model linking all zones using the FH model. Thus, the FH model can be used with

$$\theta_i = P_{\alpha i}, \; y_i = \hat{P}_{\alpha i}, \quad \psi_i = Var(\hat{P}_{\alpha i}) = \hat{V}(y_i), \tag{6.3.4}$$

where the design-based variances of these estimators can be approximated by (see Esteban, Morales, Perez and Santamaria, 2012)

$$\hat{V}(y_i) = \frac{1}{\hat{N}_i^2} \sum_{j \in S_i} w_{ij}(w_{ij} - 1)(E_{ij} - y_i)^2. \tag{6.3.5}$$

6.4 Specification of Variables

In the Fay-Herriot model (Fay and Herriot, 1979), the response variables (since we consider three response variables in this thesis) are the direct estimators of percentage of food expenditure, total poverty proportion and food poverty proportion from the 2010/11 HCE survey. The area level covariates are the area proportions of individuals in each category of the auxiliary variables (since in our case all auxiliary variables are categorical). Furthermore, the process of choosing the auxiliary variables is crucial since the small area estimation methods are mainly based on the development of a regression model. Good auxiliary information related to the variables of interest plays a vital role in determining suitable linking models (Rao, 2003b). These variables should meet the following criteria:

- available in both the household survey and the census
- household survey and census are comparable
- sufficiently correlated with household expenditure.

This includes explanatory variables namely, sex, age, marital status, educational attainment, employment and household size. In the case of strongly correlated explanatory variables, conditions being equal, we select the ones with a higher level of correlation with the poverty or percentage of food expenditure measures to estimate.

Moreover, zones within regions are unplanned domains and the sample sizes for zones in Ethiopia range from 0 households in the (Degahabor, Fik, Korahe, Gode, Warder, Afdera, Zone 2 and Zone 4) zones to 883 households in Semen Gonder zone with an average sample size of 303 households. We consider 86 sampled zones and 8 non sampled zones. The quartiles of the distribution of the 86 sampled zones sample sizes are $q_0 = 48.0$, $q_1 = 196.0$, $q_2 = 309.0$, $q_3 = 384.0$ and $q_4 = 883.0$, so the sample sizes are too small to employ direct estimators to obtain both percentage of food expenditure and poverty estimates in all the zones (Esteban et al., 2012). The population size is 37.2 million for males and 36.5 million for females, with a total population size of over 73.7 million.

6.5 Procedures for Obtaining EBLUP Estimates in the Fay-Herriot Model

In many practical applications, it is very hard to retrieve information at the individual (respondent) level due to confidentiality concerns (Chatterjee and Lahiri, 2002). Diao et al. (2013, page 2) state that "another important advantage of the Fay-Herriot model is that it only requires summary data, and not unit-level data that might be unavailable to the analyst, because of confidentiality concerns". Since the FH model deals with area level summary data and not unit-level data, the BLUP estimator is applicable for general sampling designs (Ghosh and Rao, 1994; Rao, 2003a). In summary, the FH model is developed on the small area estimates themselves (Chatterjee and Lahiri, 2002). Having these considerations in mind, the procedures for obtaining EBLUP estimates under the FH model are outlined as follows.

- In the FH model, the direct estimate for θ_i is based on the 2010/11 HCE survey data $\{\mathbf{y_i} = (y_{i1}, ..., y_{in_i})', n_i : n_i > 0, i = 1, 2, \cdots, m\}$, where y_i is the annual household consumption expenditure for households in the ith zone that can be obtained by dividing the total HCE at the household level by the household size. For example, the poverty proportion and the percentage of consumption expenditure measures for the m zones are calculated by using (6.2.1) and (6.3.3) from the given data sets. Thus, in our case, the target variable would be $\bar{y}_{1\times 86} = (\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_{86})$.
- $\bar{X}_i$ is the known population mean of the explanatory variable for the ith zone from 2007 census data of Ethiopia. We get an average for the household e.g. number of females divided by the number of people, number of males divided by the number of people, the number of people less than or equal to 30 divided by the number of people and so on. The population means of the $p = 12$

explanatory variables for the $m = 86$ zones from 2007 census data of Ethiopia are used as auxiliary variables.

- If the variance components were known, the regression coefficients could be estimated by using the weighted least squares estimator: $\tilde{\beta} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y}$, $\Sigma = A\mathbf{I}_m + \Psi$. $\mathbf{I}_m$ is the identity matrix, $\mathbf{X}$ is a matrix of area dependent auxiliary variables of size mxp.

- In our case the variance components are unknown. The variance of the random effect (A) can be estimated using the FH moment estimator, PR moment estimator, ML and REML methods using both the HCE and census data sets. We present the application results based on the FH moment estimator since the FH estimator performs better than other estimators through a simulation study. It does not also require the normality assumptions. For the detailed discussion of the estimation of variance of the random effects see section (2.1.2). Thus, the estimated variance of the random effects ($\hat{A}$) for total poverty, food poverty and percentage of food expenditure are (0.0218, 0.0238 and 28.790) respectively.

- The sampling variance $\Psi = \text{diag}(\psi_i, 1 \leq i \leq m)$, can be estimated by using the GVF using the HCE survey data. GVF defined in sections 2.1.1 and 6.6 are used in order to smooth out the unreliable and noisy design based estimated variances. We should also note that both the direct and the sampling variance estimates are obtained by incorporating the sampling weights in the analysis using the unit level data from the HCE (Rao, 2003a; Hawalla and Lahiri, 2010; Esteban et al., 2012).

- The estimated regression coefficients for total poverty, food poverty and percentage of food expenditure are

$$\begin{aligned}
\hat{\beta}_{\text{total_poverty}} &= (-0.3269, -0.1352, 0.6614, 0.3286, -0.8367, 0.2379) \\
\hat{\beta}_{\text{food_poverty}} &= (-0.3274, 0.07539, 0.5708, 0.5511, -1.0985, 0.4051) \text{ and} \\
\hat{\beta}_{\text{percentage_food_expe}} &= (0.8813, 26.5901, 22.0419, 1.7313, -8.2381, 7.7227)
\end{aligned}$$

- The estimated shrinkage factor can easily be obtained using $\hat{\gamma}_i = \frac{\psi_i}{\psi_i + \hat{A}}$.

- Replacing β and γ_i by $\hat{\beta}$ and $\hat{\gamma}_i$, we get the EBLUP estimator of θ_i : $\hat{\theta}_i^{EB} = (1 - \hat{\gamma}_i)y_i + \hat{\gamma}_i \mathbf{x}_i'\hat{\beta}$, $i = 1, ..., m$, see section (1.6.2) for detailed discussions about EBLUP.

- In summary, there are 86 zones which are common to both the 2010/11 HCE survey and 2007 census data sets. For the common zones ($m = 86$), we obtain the direct estimate from the HCE survey data and regression estimate by applying the regression equation to the summary data. Thus, we have $m = 86$ sampled zones with estimates, where the estimates are based on data from two different data sets. However, for the $m = 8$ non-sampled zones, we have regression estimates by using equation 1.6.3.

Note that regression coefficients are estimated at the country level models instead of regional or zonal level models. This is because the sample size at the country level is higher when compared to region and zone level. Fitting separate regional or zonal level regression models may reduce the bias for each region or zone even though there are small sample sizes when compared to the country level model. However, Fig 6.4 shows that the EBLUP versus direct estimates have linear relationship.

6.6 Generalized Variance Function

We use the GVF to smooth out the uncertainty of the design based variance estimate. The GVF will be used in the small area estimation of poverty and percentage of food expenditure measures. According to Hawalla and Lahiri (2010) GVF was probably introduced first by Fay and Herriot (1979) for their model in order to smooth out the sampling error variances in a complex survey setting. In our work we consider the bias corrected GVF by fitting the following simple linear regression model

$$\log(\hat{V}(y_i)) = b_0 + b_1 y_i + \varepsilon_i, \tag{6.6.1}$$

where $\log(\hat{V}(y_i))$ is the dependent variable, y_i as independent variable, $\hat{V}(y_i)$ is the design-based variance given (6.3.5) for poverty measures, b_0 and b_1 are the least squares estimates and $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$, $i = 1, ..., m$. Then the GVF motivated from the above model is given by

$$\widehat{GVF} = \exp(\frac{\hat{\sigma}^2}{2})\exp(\hat{b}_0 + \hat{b}_1 y_i). \tag{6.6.2}$$

Note that a similar procedure can be considered to obtain the GVF of the percentage of food expenditure. The factors $\left(\exp(\frac{\hat{\sigma}^2_{food}}{2}), \exp(\frac{\hat{\sigma}^2_{total}}{2}), \exp(\frac{\hat{\sigma}^2_{percentage}}{2})\right) = (1.139798, 1.155714, 1.290708)$ are the bias-correction terms in the log-linear analysis for the food poverty, total poverty and percentage of food expenditure estimates, respectively. Underestimation of the true variances will occur when we ignore the correction term in GVF method (Rivest and Belmonte, 2000, Esteban et al., 2012). Figure 6.1 (right three) further illustrates the scatter plot of $\log(\hat{V}(y_i))$ versus y_i in a dispersion graph. Figure 6.1 (left three) also illustrates the $\widehat{GVF}$ and $\hat{V}(y_i)$ in dots and stars respectively versus food poverty, total poverty and percentage of food expenditure measures. From the figures we can easily observe that how the GVF smooth out the unreliable and noisy design based estimated variances (Esteban et al., 2012, Hawala and Lahiri, 2010).

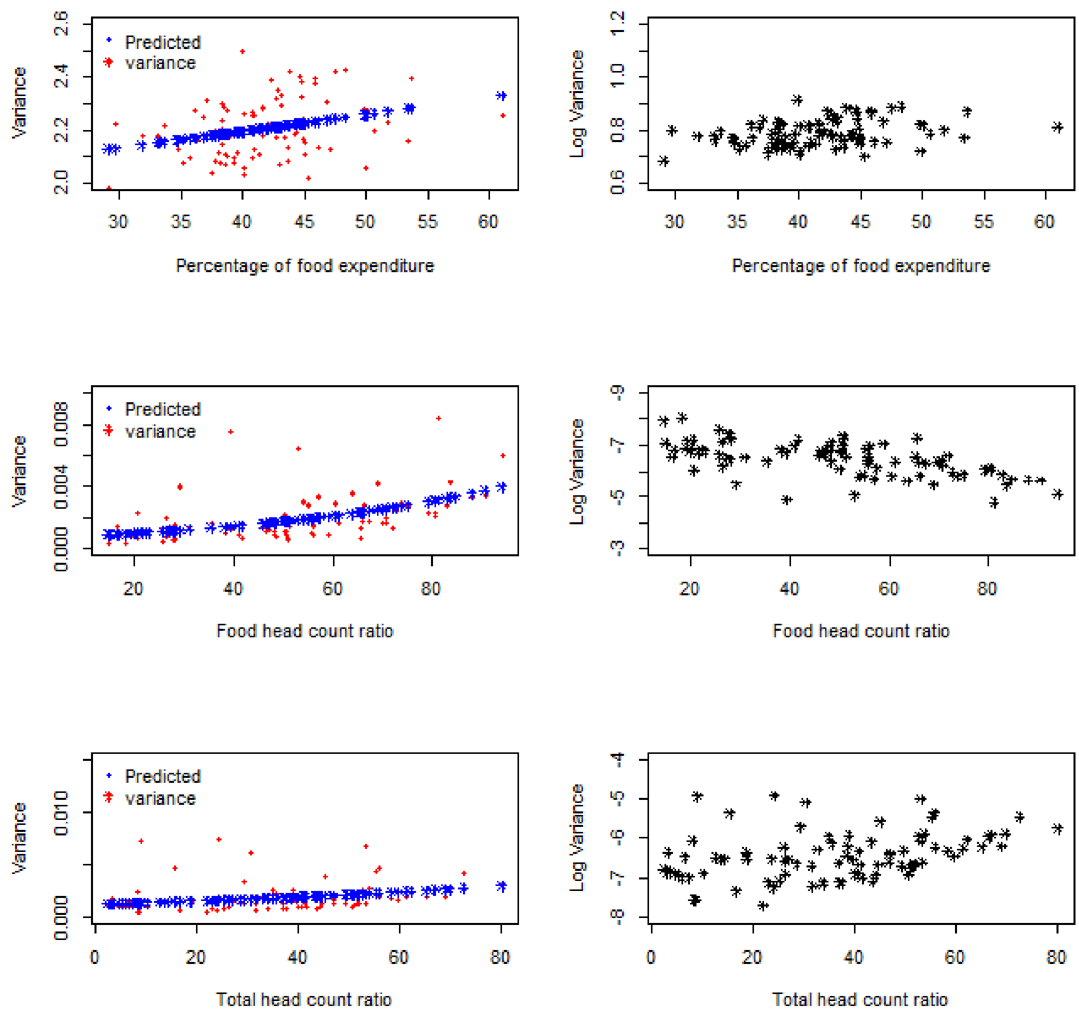

Figure 6.1: Dispersion plots for GVF in 2010/11 using the HCE survey

6.7 Percentage of Food Expenditure Estimates

All the application results are based on the EBLUP estimators which are given by equation (1.5.2). This is because the estimated variance of the random effects, $\hat{A}$ for total poverty (0.0218), food poverty (0.0238) and percentage of food expenditure (28.790) based on the FH method are positive. Furthermore, we have obtained CIs for the population small area means based on the EBLUP estimates for the percentage of food expenditure. Additionally, CIs provide a range of values of the true value of the population percentage of food expenditure measure.

A complete list of the direct, EBLUP, standard errors, CV and CI estimates at the zone level are given by Table D.2 in the Appendix D. Zone3 50.062% (95% CI=47.203%–52.921%), Borena 51.262% (95% CI=48.402%–54.122%), Itang 52.878% (95% CI=50.004%–55.752%), Nuer 53.00% (95% CI=50.129%–55.872%) and Basketo 59.30% (95% CI=56.399%–62.202%) spent the highest share of income on food (above 50%). "This result matches with other economical indicators, as they are quite deprived provinces where the percentage of food expenditure is important" (Militino, Goicoa and Ugarte, 2012, page 2941). In contrast, zones such as Zone5 29.679% (95% CI=26.892%–32.465%), Konso 30.323% (95% CI=27.554%–33.092%), Mekelle 32.155% (95% CI=29.37%–34.94%), Nifas Silk 33.368% (95% CI=30.582%–36.153%), Bahirdar 33.145% (95% CI=30.352%–35.937%), Akaki 34.181% (95% CI=31.377%–36.985%) and Adama 34.697% (95% CI=31.905%–37.489%) spent the smallest share of income on food.

Figure 6.2 plots the EBLUP estimates of the percentage of food expenditure together with their corresponding 95% CIs and their widths for all sampled zones of Ethiopia. As we observe from the plots, direct estimators have wide CIs. This is because the small sample size will lead the direct estimators to have low precision and this low precision will be reflected in wide CIs (Tanton, 2007). Nandram (1999, page 326) states that "if a prediction interval for the finite population mean of a small area is based only on its own data, it is likely to be too wide".

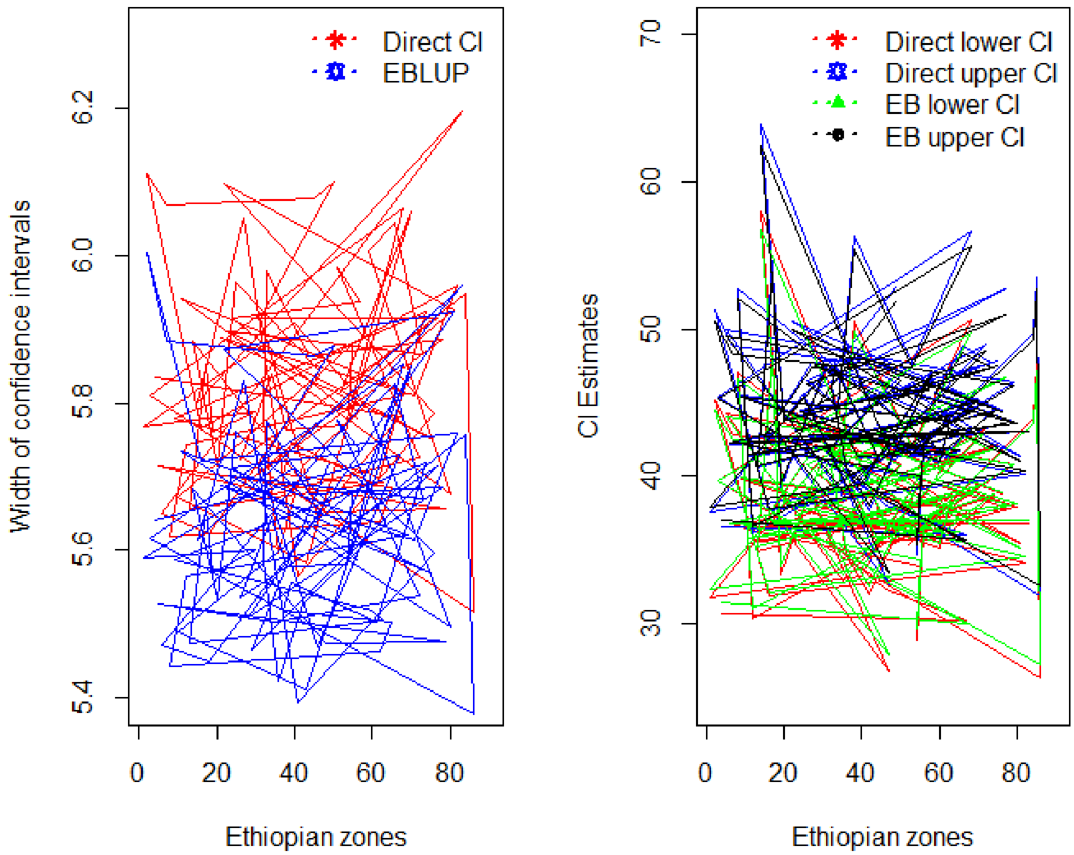

Figure 6.2: A plot illustrating the direct and EBLUP CIs and the width of CIs of percentage of food expenditure estimates versus Ethiopian zones in 2010/11

6.8 Head Count Ratio Estimates

The head count ratio is the measure of the proportion of households/individuals that are poor. It is the most commonly used method of measuring poverty because it is easy to understand and use.

6.8.1 Total Head Count Ratio Estimates

The measurement and analysis of poverty is crucial for understanding people's situations of well-being, and the factors determining their poverty situation. In addition to this, the outcomes of the analysis are often used to inform policy making as well as contributing to designing appropriate interventions, and assessing the effectiveness of on-going policies and strategies. Furthermore, addressing poverty has been an important component of the Millennium Development Goals (MDGs) as declared by the heads of states at the Millennium Summit in September 2000 that set out goals and targets to be met by the year 2015. The Ethiopian food poverty line and total poverty line in 2011 were 1985 and 3781 birr respectively (MoFED, 2012). These are the values used in the calculation of direct estimators of poverty proportion.

Table D.4 shows that zones where the proportion of the population under the total poverty line is smallest (below 5%) are Nifas silk (2.907%), Mekelle (3.071%) and Bahirdar (4.288%). On the other hand zones with higher poverty proportion (over 45%) are those situated in the Oromiya and SNNP regions. Many rural Ethiopian zones have a proportion greater than 35% of population with annual household expenditure below z=3781 birr. The zones with a high level of poverty prevalence (over 65%) are Wolayta (65.889%), Debub Omo (66.604%), Basketo (70.345%), Borena (65.872%), Awi (65.292%) and Alaba (75.43%). Furthermore, Nifas Silk (2.907%) has the smallest total poverty measure while Alaba (75.43%) has the highest poverty measures as shown in Table D.4 in the Appendix D.

In Table D.4, zones are divided into two groups:

- 37.21% of the zones with total poverty estimates that are lower than the national poverty level (which is 29.6%); and
- 62.79% of the zones with total poverty estimates that are higher than the national poverty level.

6.8.2 Food Head Count Ratio Estimates

The food poverty index measures the proportion of food-poor people that fall below the food poverty line. The achievement of food self-sufficiency is consistent with the MDG goal of eradicating extreme poverty or hunger. As for total poverty, the direct and EBLUP estimates with different model diagnostics are also computed for food poverty measures. As shown in Table D.3, Alaba (84.62%), Konso (88.16%), Burji (80.66%), Derashe (87.41%), Wolayta (82.56%), Konta (82.56%) and Awi (80.92%) have the highest number (above 80%) of households below the food poverty line whereas Bole (15.48%), Arada (17.29%), Mekelle (15.97%), Gigjiga 15.51%, Harari (19.33%) and Kolfe (19.95%) have the lowest number (less than 20%) of households below the food poverty line as shown in Table D.3 in the Appendix D.

In Table D.3, zones are divided into two groups:

- 29.07% of zones with food poverty estimates that are lower than the national food poverty level (which is 33.6%); and
- 70.03% of the zones with food poverty estimates that are higher than the national food poverty level.

The purpose of small area estimation is to produce reliable estimates of poverty and percentage of food expenditure measures for zones which only have small samples or no samples are available (Pfeffermann, 2013). Some of the zones in Ethiopia such as: Degahabor, Fik, Korahe, Gode, Warder, Afdera, Zone 2 and Zone 4 do not have direct survey estimates because they are non-sampled zones. Thus, we provide the poverty and percentage of food expenditure measures in Table D.1 in terms of the synthetic estimates. Synthetic poverty and percentage of food expenditure estimates for the zones that have not been sampled are estimated using information from the auxiliary variables. This is one of the advantages of using small area estimation.

Direct estimates of the percentage of food expenditure and poverty measures for Ethiopian zones do not have enough precision, due to the limited zonal sample sizes of the HCE 2010/11. For this reason, official zonal estimates of percentage of food expenditure and poverty measures at the zone level are not produced.

The HCE survey already provides estimates at the regional and national level. However, there is a high variability of the model-based percentage of food expenditure and poverty estimates for zones belonging to the same region. However, regional estimates usually mask variations at the zone level and so yield only partial information for planning and allocation of resources at the local level. Thus, conducting the analysis of poverty and percentage of food expenditure at regional level could mask important dissimilarities characterizing those territories. Regional estimates can also lead policy makers into the ‘ecological fallacy’ of assuming that the average value for a region applies to all zones within it when in fact there might be differences between zones.

As we map poverty and percentage of food expenditure at the zone level, geographic variability that was hidden at the national and regional level becomes apparent. This is clearly illustrated in the percentage of people's expenditure on food, food poverty and total poverty maps shown in Figures E.3, E.4 and E.5 in the Appendix E. These maps play an important role in communicating inequality in Ethiopia. They also show visually how the poverty and percentage of food expenditure varies across zones. The blue colour illustrates the higher the percentage of food expenditure and the higher the number of people under the poverty line, while red indicates the lesser poverty and percentage of food expenditure. Moreover, significant variation of small area estimates of poverty and percentage of food expenditure may be due to differences in agro-ecological resource endowment, education and health services (Deichmann, 1999).

6.9 Model Diagnostics

After fitting a model, it is crucial to determine the reliability of the model-based estimates before performing inference. If we see any violations, the resulting inference would be invalid, which leads to wrong conclusions. Thus, it is important to perform appropriate diagnostic procedures. Diagnostic procedures are used to check the validity of the model-based estimates versus direct survey estimates. An evaluation of the diagnostic measures confirms that there is reasonably close correspondence between the model-based estimates and the direct survey estimates. In this study, bias diagnostic, goodness of fit diagnostic and coverage diagnostic are used to check the validity of model based estimates. In addition, standard error and coefficient of variation are also used to check the reliability of our estimates.

Coverage Diagnostic

This method evaluates the validity of the CIs based on EBLUP estimates. It is used to measure the overlap between the 95% CIs generated by the direct and the model based estimates in small area estimation procedure (Brown et al., 2001). This test can be done as follows. Let X_1 and X_2 be two independent normal random variables having the same mean and variance $\sigma^2_{X_1}$ and $\sigma^2_{X_2}$ respectively. If $z(\alpha)$ is such that the probability that a standard normal variable takes values greater than $z(\alpha)$ is $\alpha/2$, then a sufficient condition for there to be probability of α that the two intervals $X_1 \pm z(\beta)\sigma_{X_1}$ and $X_2 \pm z(\beta)\sigma_{X_2}$ do not overlap is when

$$z(\beta) = z(\alpha)\left(1 + \frac{\sigma_{X_1}}{\sigma_{X_2}}\right)^{-1} \sqrt{1 + \frac{\sigma^2_{X_1}}{\sigma^2_{X_2}}},$$

where σ_{X_1} is the estimated standard error of the EBLUP estimate and σ_{X_2} is the estimated standard error of the direct estimate, $z(\alpha) = 1.96$. In our real data application the $z(\beta)$'s for food poverty, total poverty and percentage of food expenditure are 1.387302, 1.387346 and 1.387348 respectively. This confirms that the overlap proportion is 95%. In addition, Figure 6.2 shows that the CIs based on the EBLUP estimates lies within the CIs of the direct estimates.

Bias Diagnostic

The bias diagnostic is used to assess the deviation of the model based estimates from the direct survey estimates. If the relationship between the poverty measures and the auxiliary variables has not been misspecified, a linear relationship of the type y = x is expected between the direct survey estimates and the model-based estimates. To check for predictive bias in the model-based estimates, we plot model-based estimates against direct estimates. Then we test whether the regression line can be fitted to these points and if it is significantly different from the identity line (Brown et al., 2001). Figure 6.5 shows the bias scatter plot of the direct survey estimates against the model-based (EBLUP) estimates for each zone. From this figure, we can easily observe that all points lie around the line and the slope of all model outputs were also close to one for both poverty and percentage of food expenditure measures. These plots include all estimates without outlying estimates.

Goodness of Fit Diagnostic

The main aim of this method is to assess the unconditional bias of the EBLUP estimates. To do this we use the Wald statistic W which is given by the ratio of the weighted squared difference between the direct and EBLUP estimates to their variance. This will help us to test whether there is a difference between the direct estimator and the expected value of the EBLUP estimator provided that the sample sizes in the small areas are sufficient to justify the use of the central limit theorem. Under this consideration, W will have a χ^2 distribution with m degrees of freedom, where m is the number of sampled areas (Brown et al., 2001). In our real data application example, the smallest sample size is 48. The Wald statistics for food poverty, total poverty and percentage of food expenditure are 17.38, 9.99 and 6.12 respectively. From this we can conclude that there is no difference between the direct estimator and the expected value of the EBLUP estimator since the Wald statistics are less than the $\chi^2(86)$ value which is equal to 108.65.

Coefficient of Variation and Standard Error

In addition, Figure 6.3 (top) shows the scatter plots of the CVs of the EBLUP and direct estimates against the zones. It can be seen from the plots that the CVs from the model-based estimates are less than the CVs from the direct estimates. Figure 6.3 (bottom) shows that the plots of the standard errors versus zones. The scatter plots of the standard errors of the direct estimates are more dispersed than the model-based estimates. It is apparent that the standard errors of the direct estimates are larger than the standard errors of the EBLUP estimates due to small sample sizes (Pfeffermann, 2013). As shown by Tables D.2-D.4 model-based estimates of the poverty and percentage of food expenditure measures based on the FH model have lower estimated standard errors than the direct estimates. This indicates that the EBLUP estimates are more reliable than the direct estimates (Coondoo, Majumder and Chattopadhyay, 2011). In general, in terms of CV and standard errors, we observe that there is an overall clear gain of precision when using the EBLUP estimates based on the FH model instead of direct estimates.

Residuals and Standardized Residuals

Figure 6.4 plots the residuals versus the fitted values of the models. The residuals tend to be positive on the right part of the figure. This shows that the model is smoothing the larger values of the direct estimator. This indicates that direct domain estimates are not affected by outliers (Esteban et al., 2012). From the residual versus EBLUP plot we can see that a few anomalous observations are observed even though the values are almost evenly distributed around zero.

Note that the HCE data contains sampling weights. In Figure E.1 (bottom) of the Appendix E we have plotted residuals against sampling weights for both total and food expenditure to check for an indication of informative sampling in this application. The null pattern in the plots shows that there is no evidence of informative sampling (Molina, Nandram and Rao, 2014).

Additionally, Figure E.1 (top) in the Appendix E presents the index plots of residuals for total and food expenditure. Similarly, index plots of residuals for total poverty, food poverty and percentage of food expenditure are included in Figure E.2 of Appendix E. Both plots are approximately acceptable without any clear pattern.

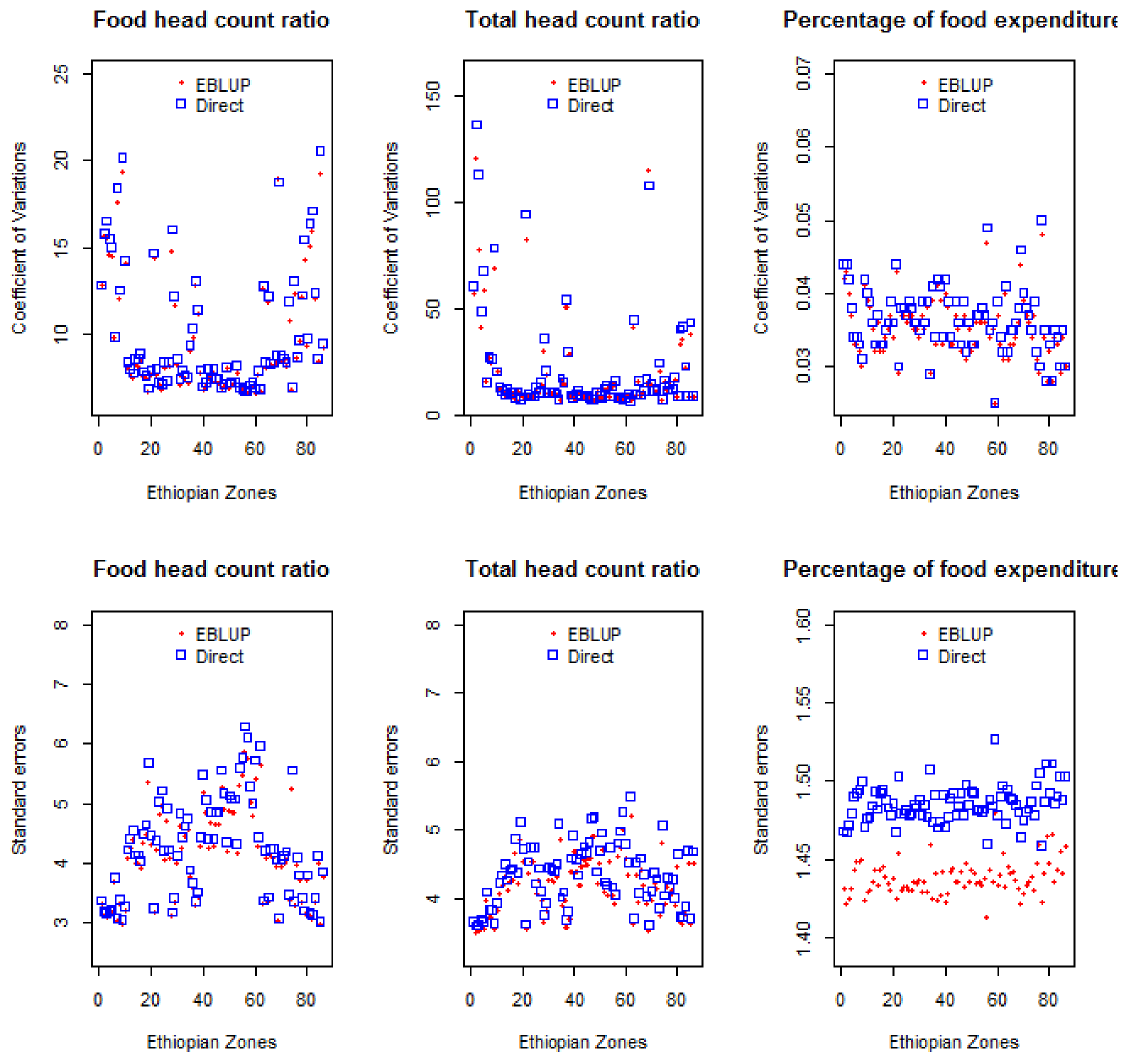

Figure 6.3: A plot illustrating the standard errors (bottom) and CVs (top) along the Ethiopian zones in 2010/11

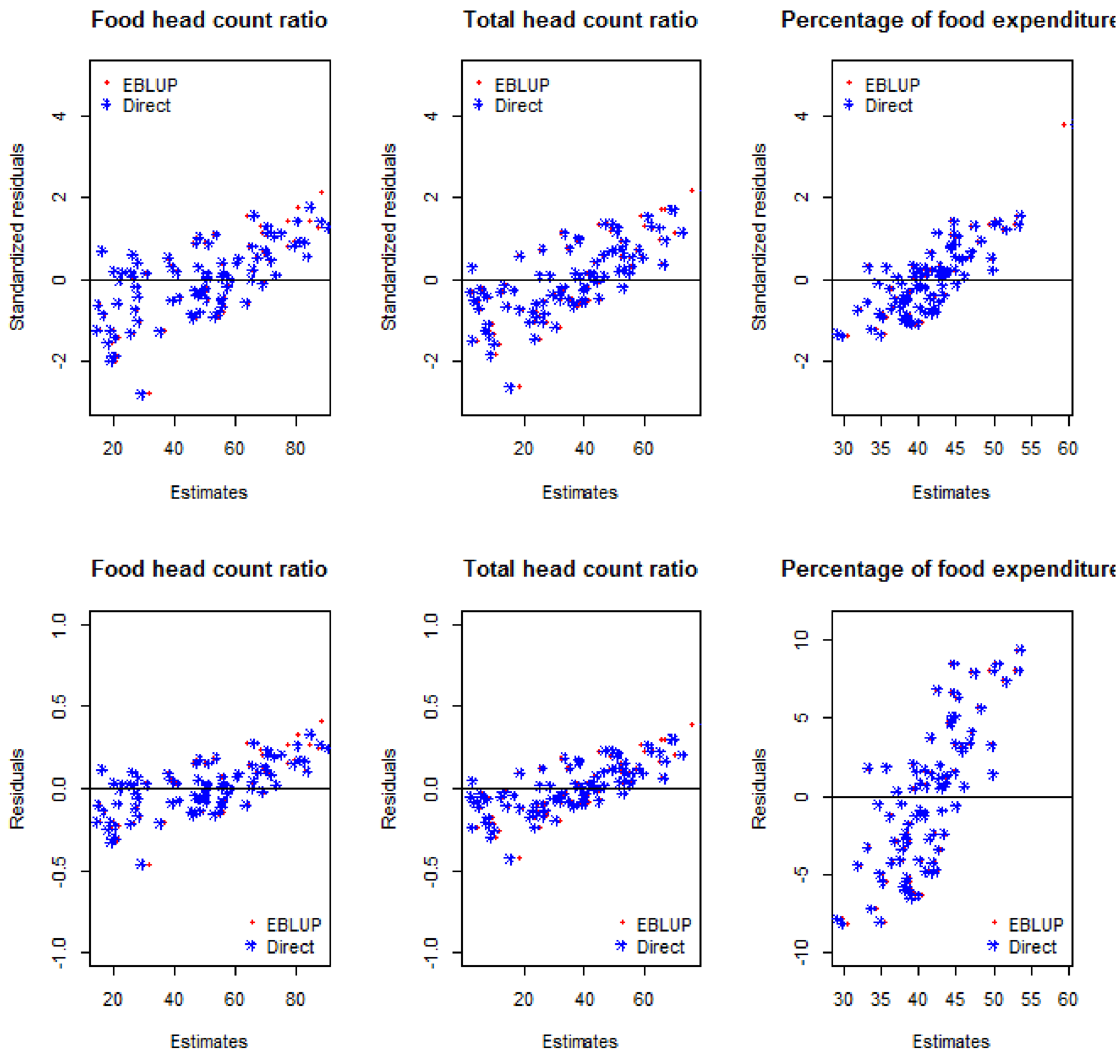

Figure 6.4: Standardized residuals (top); residuals (bottom) versus fitted values

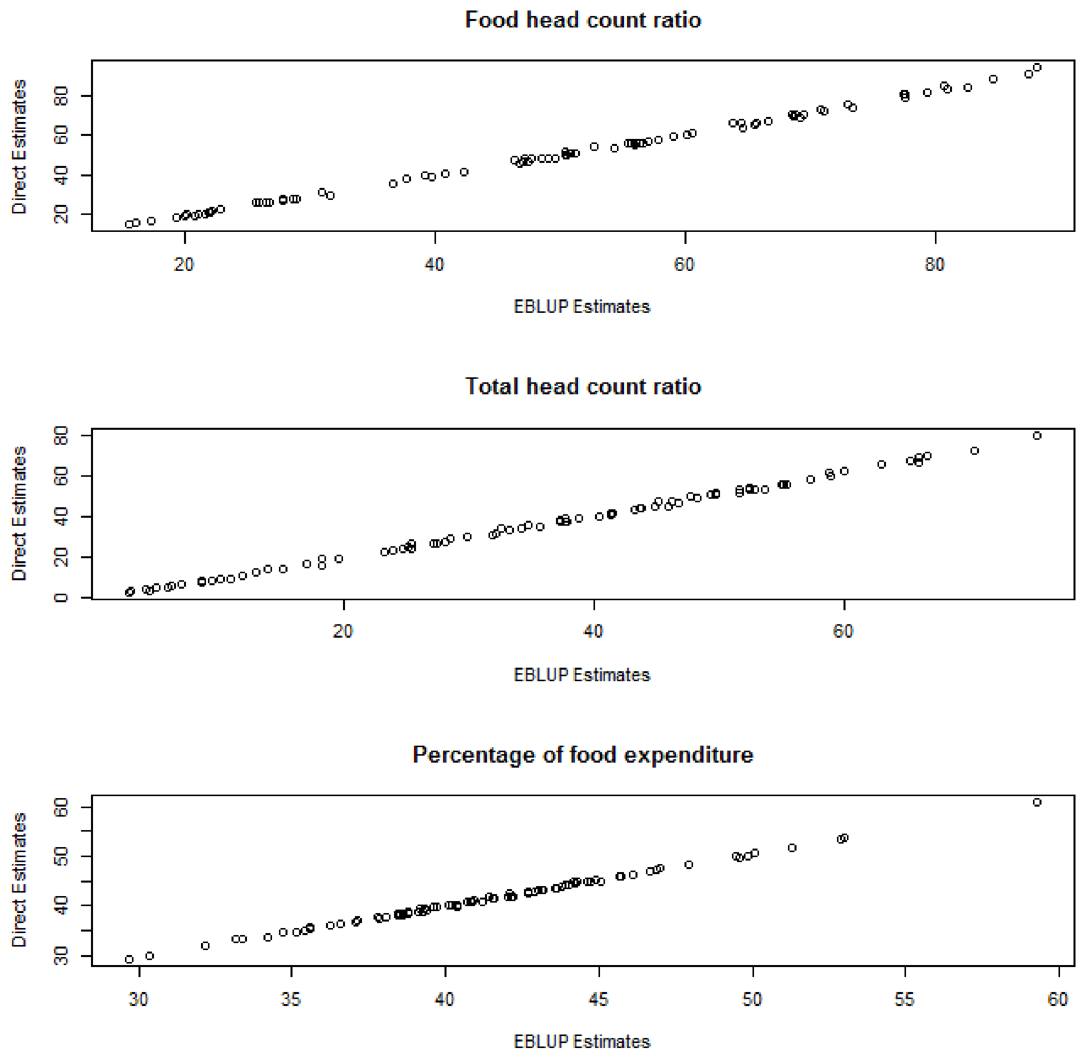

Figure 6.5: EBLUP estimates of food poverty measures (top), total poverty measures (middle) and percentage of food expenditure (bottom) against direct estimates for each zone

The use of the diagnostic measures confirms that the model-based zonal estimates are robust enough to provide reliable information for the guidance of policy, resource allocation, and the planning and evaluation of poverty and percentage of food expenditure measure programmes. The results of this study demonstrates the feasibility of producing small area poverty and percentage of food expenditure estimates based on the 2010/11 HCE sample survey.

6.10 The Links between Poverty and Spending on Food

The Pearson correlation coefficients and their corresponding p-values between food poverty and total poverty, food poverty and percentage of food expenditure, and total poverty and percentage of food expenditure are 0.853 (0.000), 0.043 (0.697) and 0.348 (0.001) respectively. This shows that there is a strong correlation between food and total poverty. Similarly, total poverty and percentage of food expenditure are positively correlated. People who live with limited income resources tend to be highly affected by food price increases (Deichmann, 1999). However, this relationship may not always hold. Food poverty and percentage of food expenditure are linearly uncorrelated (corr=0.043, p-value=0.697). According to The Economist (2013), Indians, for example, spend less on food than Russians do.

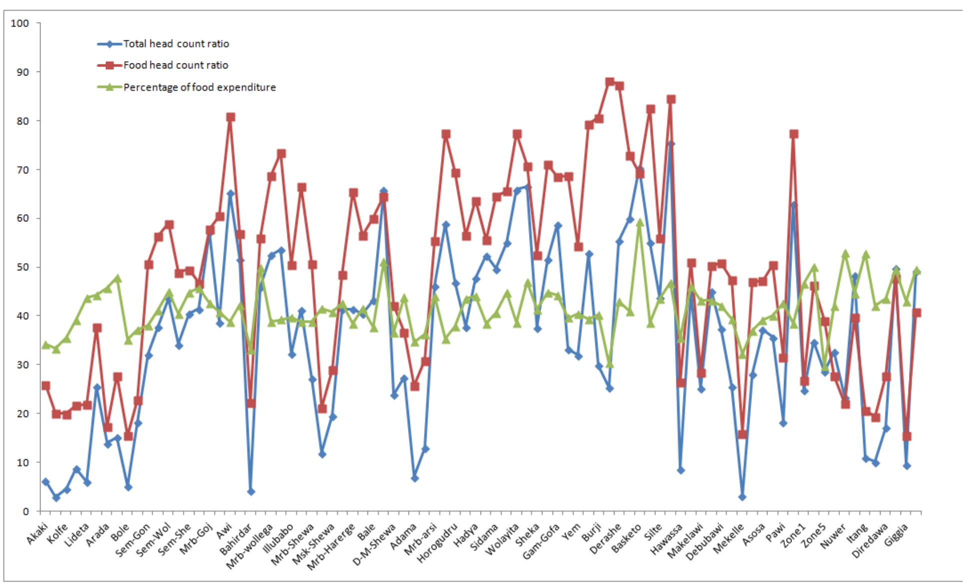

Figure 6.6: The scatter plot between poverty and spending on food. The points are connected with lines for visibility purposes

Chapter 7

Conclusions and Areas for Further Research

7.1 Conclusions

The findings of this study indicate that there is no single method which dominates all other methods in terms of MSE. However, in terms of bias and risk, the REML and FH methods perform consistently better than other methods for the normal random effects distribution. Also, for the Laplace random effects distribution, the FH method performs better than other methods. Furthermore, MSE_JY1 performed well in terms of RB and RR for the ML, REML and FH estimators for all variance patterns. When looking at RB and RR, MSE_JY1 performs even better than the existing MSE estimators based on the ML and REML methods. The difference however, gets smaller as the number of areas increases for all variance patterns. All estimators are very similar with regard to bias and risk for pattern (I). The best balance between bias, risk and MSE estimator are achieved when $\hat{A} = \hat{A}_{FH}$ and $\hat{A} = \hat{A}_{REML}$ for all variance patterns and the normal random effects distributions. Asymptotically (as $m \to \infty$), MSE, RB and RR based on all methods are identical for pattern (I) and (II), and the observed differences are due to the method of estimating A.

From the results and discussion, it becomes clear that MSE_JY1 out performs the rest for the ML, REML and FH methods. Therefore, we consider MSE_JY1 as a reasonable estimator for estimating the MSE in small area estimation. Our simulation results also suggest that the FH method performs better than its competitors for different sampling variance patterns and distribution of random effects. Thus, it is highly recommended to use MSE estimators based on the REML and FH methods for the normal random effects distribution. Whereas, we recommend the FH method for the Laplace random effects distribution.

In this study, we proposed CIs based on EBLUP and weighted estimators with fixed weights under the basic area level model. The proposed methods based on the weighted estimator with fixed weights gives an estimate when either the MSE or the estimated variance of the random effect, $\hat{A}$ are negative. In addition, our proposed methods (CIRao and CIJY) are area specific CIs. All of our methods have been numerically shown to be superior to the Cox (1975) and Prasad and Rao (1990) CIs in terms of the CPs of the intervals. These methods, while maintaining a bit higher ALs for small samples, have coverage which is better than the naive methods. All of these methods offer an easy framework and provide CIs that can easily be applied. This highlights the advantages of the proposed methods when compared to CIs based on resampling methods.

Furthermore, CIs based on resampling approach do not have sensible empirical plug-in estimators which are routinely used in place of the unknown parameters (Chatterjee and Lahiri, 2002). However, the advantages of CIs based on the Taylor series expansion methods (for example, Diao et al., 2013 and our methods) are computational. The unknown model parameters need to be estimated only once using either R or SAS softwares and then we can plug-in the estimates in the proposed MSEs and CIs.

The usefulness of our methods has also been shown through the application to the HCE survey data in Ethiopia to estimate poverty and percentage of food expenditure measures at the zone level. We derive zonal-level estimates of poverty and percentage of food expenditure measures by using small area estimation techniques to link data from the 2010/11 HCE and 2007 population and housing Census.

Additionally, the total and food poverty indices of Ethiopia are 29.6% and 33.6% respectively (MoFED, 2012). However, the zonal level total and food head count indices vary from 2.91% to 75.43% and 15.48% to 88.16% respectively. For instance, the total and food head count indices of the Southern Nations Nationalities and Peoples Regional State (SNNPRS) are 29.6% and 25.9% respectively (MoFED, 2012). However, the zonal level total and food head count indices of SNNPRS vary from 8.71% to 75.43% and 26.44% to 88.16% respectively. Similarly, the percentage of food expenditure varies from 29.679% to 59.30%. These results show that there is a considerable geographical variation in the poverty and percentage of food expenditure estimates for zones belonging to the same region. Thus, conducting the analysis of poverty as well as the percentage of food expenditure at regional level could mask important dissimilarities characterizing those territories.

The model-based methods have also been found to be very effective for developing zonal level estimates of poverty and percentage of food expenditure measures. For most of the zones the reduction in CV and standard errors are quite evident. The results demonstrate the feasibility for using small area estimation techniques to improve EBLUP estimates at the zone level for planning and programme evaluation.

Moreover, it is somehow surprising how large the number of zones have a proportion greater than 50% of population with annual consumption expenditure below the poverty line. Hence it would be desirable that the Ethiopian government implement polices to reduce poverty proportion in those zones in which poverty proportion is the highest. Furthermore, improved small area poverty and percentage of food expenditure estimates are very important for the government, policy makers, Non Governmental organisations (NGOs), and other concerned bodies. This is because before taking any remedial action, there is a need to identify zones in which the poverty and percentage of food expenditure measures are high. All the methods of estimating poverty and percentage of food expenditure are the recommended tools for monitoring progress towards MDG achievement.

7.2 Areas for Further Research

Considering this, further research could also be conducted to modify CIs for the linear combinations of small area means under the area level model. However, construction of such CIs is troublesome because of the correlation between small area means (Diao et al., 2013). Furthermore, construction of CIs for the difference of two small area means under the unit level (nested error regression) models is not yet done. In addition, we cannot construct CIs under the unit level model when the MSE is negative. This is another drawback of CIs based on EBLUP under the nested error regression model. Therefore it makes sense to extend interval estimation for weighted estimator with fixed weights under the unit level models.

Appendix A

Proof of Theorem 4.1.2

For any real $z > 0$,

$$P\left[\theta_i \in \hat{\theta}_i^{EB} \pm z \times \sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}\right] = 2\Phi(z) - 1 - \frac{z\phi(z)}{4}\eta + O(m^{-3/2})$$

where

$$\eta = (z^2+1)\frac{\psi_i^2}{A^2(A+\psi_i)^2}V(\hat{A}).$$

Proof. The proof of the Theorem about CIs based on the EBLUP estimator is given. The conditional distribution of θ_i given y_i is given by

$$\theta_i \mid y_i \sim N(\hat{\theta}_i^{EB}, g_{1i}(A)), \text{ where } \hat{\theta}_i^{EB} = (1-\hat{\gamma}_i)y_i + \hat{\gamma}_i \mathbf{x}_i'\hat{\beta}, \quad \tilde{\theta}_i^{B} = (1-\gamma_i)y_i + \gamma_i \mathbf{x}_i'\tilde{\beta}, \quad g_{1i}(A) = \frac{A\psi_i}{A+\psi_i}.$$

Consider the following expression as follows:

$$P\left[\theta_i \le \hat{\theta}_i^{EB} + z \times \sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}\right] = E[\Phi(z+F(z))].$$

The symbols $\Phi(z)$ and $\phi(z)$ represents the distribution function and the probability density function, respectively, of the $N(0,1)$ distribution. Thus, the CP of $I^{CEB}(\hat{A})$ is written as

$$\begin{aligned} P[\theta_i \in I^{CEB}(\hat{A})] &= P\left[-z+F(-z) < \frac{\theta_i - \hat{\theta}_i^{EB}}{\sqrt{g_{1i}(A)}} < z + F(z)\right] \\ &= E[\Phi(z+F(z)) - \Phi(-z+F(-z))] \end{aligned}$$

where $F(z) = \dfrac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})} - \sqrt{g_{1i}(A)}\right) + \hat{\theta}_i^{EB} - \tilde{\theta}_i^{B}}{\sqrt{g_{1i}(A)}}$.

Using the Taylor series expansion with an integral remainder term, $\Phi(z+F(z))$ is evaluated as

$$\begin{aligned} \Phi(z+F(z)) &= \Phi(z) + F(z)\phi(z) + \tfrac{1}{2}F^2(z)\phi'(z) + \tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2\phi''(x)dx \\ &= \Phi(z) + [F(z) - \tfrac{z}{2}F^2(z)]\phi(z) - \tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2(1-x^2)\phi(x)dx. \end{aligned}$$

Taking expectations on both sides reveals that

$$E(\Phi(z+F(z))) = \Phi(z) + E\left[F(z)\phi(z) - \tfrac{z}{2}F^2(z)\phi(z) - \tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2(1-x^2)\phi(x)dx\right].$$

Consider first that,

$$\hat{\theta}_i^{EB} - \tilde{\theta}_i^{B} = (1-\hat{\gamma}_i)y_i + \hat{\gamma}_i \mathbf{x}_i'\hat{\beta} - (1-\gamma_i)y_i - \gamma_i \mathbf{x}_i'\tilde{\beta}$$

$$= (\gamma_i - \hat{\gamma}_i)r_i + \gamma_i \mathbf{x}_i'(\hat{\beta}-\tilde{\beta}), \text{ where } r_i = (y_i - \mathbf{x}_i'\hat{\beta}).$$

Thus F(z) can be decomposed into three components as follows:

$$F(z) = \frac{z \times \left(\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})} - \sqrt{g_{1i}(A)}\right)}{\sqrt{g_{1i}(A)}} + \frac{(\gamma_i-\hat{\gamma}_i)r_i}{\sqrt{g_{1i}(A)}} + \frac{\gamma_i \mathbf{x}_i'(\hat{\beta}-\tilde{\beta})}{\sqrt{g_{1i}(A)}}$$

$$= S_1 + S_2 + S_3 \quad \text{say.}$$

Let us evaluate the expected value of each term as follows

$$EF(z) = E(S_1 + S_2 + S_3)$$

$$= E(S_1) + E(S_2) + E(S_3).$$

Applying the Taylor series expansion and taking expectations on both sides

$$ES_1 = z\tfrac{1}{2}\left(g_{1i}(A)\right)^{-1} E\left(\widehat{MSE}(\hat{\theta}_i^{EB}) - g_{1i}(A)\right)$$

$$-z\tfrac{1}{8}\left(g_{1i}(A)\right)^{-2} E\left(\widehat{MSE}(\hat{\theta}_i^{EB}) - g_{1i}(A)\right)^2$$

$$+z\tfrac{3}{16}\left(g_{1i}(A)\right)^{-1/2} E\int_{g_{1i}(A)}^{\widehat{MSE}(\hat{\theta}_i^{EB})} x^{-5/2}\left(\widehat{MSE}(\hat{\theta}_i^{EB}) - g_{1i}(A)\right)^2 dx.$$

Let us compute the expectations of each term

$$E\left(\widehat{MSE}(\hat{\theta}_i^{EB}) - g_{1i}(A)\right) = E(\widehat{MSE}(\hat{\theta}_i^{EB})) - E(g_{1i}(A)).$$

$E(\widehat{MSE}(\hat{\theta}_i^{EB})) = MSE(\hat{\theta}_i^{EB})$, since $E(\widehat{MSE}(\hat{\theta}_i^{EB}))$ is nearly a second order unbiased estimator of $MSE(\hat{\theta}_i^{EB})$.

Thus, using the Taylor series expansion, the second term can be expressed as

$$\widehat{MSE}(\hat{\theta}_i^{EB}) = MSE(\hat{\theta}_i^{EB}(A)) + MSE(\hat{\theta}_i^{EB}(A^*))^{(1)}(\hat{A}-A) + O(m^{-3/2})$$

$$E\left(\widehat{MSE}(\hat{\theta}_i^{EB}) - g_{1i}(A)\right)^2 = (g_{1i}^{(1)}(A^*))^2 E(\hat{A}-A)^2,$$

where $g_{1i}(A) = \frac{A\psi_i}{A+\psi_i}$, $g_{1i}^{(1)}(A) = \frac{\psi_i}{A+\psi_i} - \frac{A\psi_i}{(A+\psi_i)^2} = \frac{\psi_i^2}{(A+\psi_i)^2}$, $(g_{1i}^{(1)}(A))^2 = \frac{\psi_i^4}{(A+\psi_i)^4}$.

Substituting reveals that

$$E\left(\widehat{MSE}(\hat{\theta}_i^{EB}) - g_{1i}(A)\right)^2 = \frac{\psi_i^4}{(A+\psi_i)^4} E(\hat{A}-A)^2 + O(m^{-3/2}),$$

where, $E(\hat{A}-A)^2 = V(\hat{A}) + O_p(m^{-1})$, $V(\hat{A})$ can be the asymptotic variance of the FH, PR, ML and REML estimators of A.

Substituting reveals that

$$E\left(\widehat{MSE}(\hat{\theta}_i^{EB}) - g_{1i}(A)\right)^2 = \frac{\psi_i^4}{(A+\psi_i)^4}\frac{2}{m^2}\sum_{i=1}^m (A+\psi_i)^2 + O(m^{-1}).$$

The integral term can be simplified using a similar argument to Theorem 4.1.1. Thus,

$$E\int_{g_{1i}(A)}^{\widehat{MSE}(\hat{\theta}_i^{EB})} x^{-5/2}\left(\widehat{MSE}(\hat{\theta}_i^{EB}) - g_{1i}(A)\right)^2 dx = O(m^{-3/2}).$$

Let us compute S_2 as follows

$$S_2 = \frac{(\gamma_i-\hat{\gamma}_i)(y_i - \mathbf{x}_i'\hat{\beta})}{\sqrt{g_{1i}(A)}}.$$

By the Taylor series expansion,

$\gamma_i(\hat{A}) = \gamma_i(A) + \gamma_i^{(1)}(A^*)(\hat{A}-A)$ and $\hat{\beta} = \tilde{\beta}(A) + \tilde{\beta}^{(1)}(A^*)(\hat{A}-A)$, where $|A^*-A| \le |\hat{A}-A|$ and $\gamma_i^{(1)}(A^*)$

is the derivative of $\tilde{\beta}(\mathrm{A})$ evaluated at $\mathrm{A}=\mathrm{A}^*$, similarly $\tilde{\beta}^{(1)}(\mathrm{A}^*)$ is the derivative of $\gamma_i(\mathrm{A})$ evaluated at $\mathrm{A}=\mathrm{A}^*$. Thus, $\gamma_i(\hat{\mathrm{A}})-\gamma_i(\mathrm{A})=\gamma_i^{(1)}(\hat{\mathrm{A}}-\mathrm{A})$ and $\hat{\beta}-\tilde{\beta}(\mathrm{A})=\tilde{\beta}^{(1)}(\hat{\mathrm{A}}-\mathrm{A})$.

Substituting reveals that

$S_2=-g_{1i}^{-1/2}\gamma_i^{(1)}\left((\hat{A}-A)(y_i-\mathbf{x}_i'\tilde{\beta}(A))-(\hat{A}-A)\mathbf{x}_i'\tilde{\beta}^{(1)}(\hat{A}-A)\right)$.

Taking expectations on both sides

$$ES_2=-g_{1i}^{-1/2}E\left[(\hat{\gamma}_i-\gamma_i)(y_i-\mathbf{x}_i'\hat{\beta})\right]$$
$$=-g_{1i}^{-1/2}\gamma_i^{(1)}E\left[(\hat{A}-A)(y_i-\mathbf{x}_i'\tilde{\beta}(A))\right]+g_{1i}^{-1/2}\gamma_i^{(1)}E\left[(\hat{A}-A)\mathbf{x}_i'\tilde{\beta}^{(1)}(\hat{A}-A)\right].$$

Let the first part can be written as

$$S_2^*=-g_{1i}^{-1/2}\gamma_i^{(1)}E\left[(\hat{A}-A)(y_i-\mathbf{x}_i'\tilde{\beta}(A))\right]$$
$$=-g_{1i}^{-1/2}\gamma_i^{(1)}E\left[(\hat{A}-A)(y_i-\mathbf{x}_i'\beta)\right]+o(m^{-1}).$$

Using the Stein identity

$ES_2^*=-g_{1i}^{-1/2}\gamma_i^{(1)}(A+\psi_i)E(\frac{\partial\hat{A}}{\partial y_i})=0$, since $\frac{\partial\hat{A}}{\partial y_i}=O_P(m^{-1})$.

The expected value of S_3 can be evaluated as follows

$S_3=g_{1i}(A)^{-1/2}\gamma_i\mathbf{x}_i'\left(\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i}\right)^{-1}\sum_{i=1}^m\frac{\mathbf{x}_i}{A+\psi_i}e_i$.

since $\tilde{\beta}=\left(\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i}\right)^{-1}\sum_{i=1}^m\frac{\mathbf{x}_i}{A+\psi_i}y_i$

$E(S_3)=0$.

Combination of the above results, we have

$$EF(z)=E(S_1)+E(S_2)+E(S_3)$$
$$=E(S_1)+0+0$$
$$=\frac{z}{2}\left(g_{1i}(A)\right)^{-1}(g_{2i}(A)+g_{3i}(A))-\frac{z}{8}\left(g_{1i}(A)\right)^{-2}\frac{\psi_i^4}{(A+\psi_i)^4}V(\hat{A})+O(m^{-3/2})$$
$$=\frac{z}{2}\left(\frac{A\psi_i}{A+\psi_i}\right)^{-1}\left(\frac{\psi_i^2}{(A+\psi_i)^2}\mathbf{x}_i'(\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{A+\psi_i})^{-1}\mathbf{x}_i+\frac{\psi_i^2}{(A+\psi_i)^3}V(\hat{A})\right)$$
$$-\frac{z}{8}\left(\frac{A\psi_i}{A+\psi_i}\right)^{-2}\frac{\psi_i^4}{(A+\psi_i)^4}V(\hat{A})+O(m^{-3/2}).$$

$$EF(z)^2=E(S_1+S_2+S_3)^2$$
$$=E(S_1^2)+E(S_2^2)+E(S_3^2)+2E(S_1S_2)+2E(S_1S_3)+2E(S_2S_3).$$

Compute all the expectations values turn by turn as follows.

Let us evaluate S_1S_2 using the Taylor series expansion as follows

$$S_1S_2=\frac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}-\sqrt{g_{1i}(A)}\right)(\gamma_i-\hat{\gamma}_i)(y_i-\mathbf{x}_i'\hat{\beta})}{g_{1i}(A)}$$
$$=g_{1i}^{-1}(A)\frac{-z}{2}\left(\widehat{MSE}(\hat{\theta}_i^{EB})-g_{1i}(A)\right)(\gamma_i-\hat{\gamma}_i)(y_i-\mathbf{x}_i'\beta)$$

$$= -g_{1i}^{-1}(A)\tfrac{z}{2}g_{1i}^{(1)}(A)\gamma_i^{(1)}(A)\left((\hat{A}-A)^2(y_i-\mathbf{x}_i'\tilde{\beta}(A))-(\hat{A}-A)^3\mathbf{x}_i'\tilde{\beta}^{(1)}(A)\right).$$

Let $c = -g_{1i}^{-1}(A)\frac{z}{2}g_{1i}^{(1)}(A)\gamma_i^{(1)}(A)$ and taking expectations on both sides gives that

$ES_1S_2 = cE\left((\hat{A}-A)^2(y_i-\mathbf{x}_i'\tilde{\beta}(A))\right)-cE\left((\hat{A}-A)^3\mathbf{x}_i'\tilde{\beta}^{(1)}(A)\right) = r_{11}+r_{12}$ say.

$$Er_{11} = cE\left((\hat{A}-A)^2(y_i-\mathbf{x}_i'\tilde{\beta}(A))\right)$$
$$= cE\left((\hat{A}-A)^2(y_i-\mathbf{x}_i'\beta)\right)+o(m^{-1}).$$

Using the Stein identity

$$Er_{11} = c(A+\psi_i)E\left(\frac{\partial}{\partial y_i}(\hat{A}-A)^2\right)$$
$$= c(A+\psi_i)E\left(2(\hat{A}-A)\frac{\partial\hat{A}}{\partial y_i}\right),\text{ but } \frac{\partial\hat{A}}{\partial y_i} = O_P(m^{-1})$$
$$Er_{12} = cE\left((\hat{A}-A)^3\mathbf{x}_i'\tilde{\beta}^{(1)}(A)\right)$$
$$= c\mathbf{x}_i'\tilde{\beta}^{(1)}(A)E(\hat{A}-A)^3.$$

Since $E(\hat{A}-A)^{2S} = O(m^{-S})$, (Prasad and Rao, 1990), thus $E(\hat{A}-A)^3 = O(m^{-3/2})$.

Let us evaluate S_2S_3 as using the Taylor series expansion as follows

$$S_2S_3 = g_{1i}^{-1}(y_i-\mathbf{x}_i'\hat{\beta})\gamma_i\mathbf{x}_i'(\hat{\beta}-\tilde{\beta})(\gamma_i-\hat{\gamma}_i)$$
$$= -g_{1i}^{-1}(y_i-\mathbf{x}_i'\tilde{\beta}(A))\gamma_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)\gamma_i^{(1)}(A)(\hat{A}-A)$$
$$+ g_{1i}^{-1}\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)\gamma_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)\gamma_i^{(1)}(A)(\hat{A}-A)$$
$$= r_{21}+r_{22}.$$

Taking expectations on both sides gives

$$Er_{21} = -g_{1i}^{-1}\gamma_i^{(1)}(A)\gamma_iE\left((y_i-\mathbf{x}_i'\tilde{\beta}(A))\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)^2\right)$$
$$= -g_{1i}^{-1}\gamma_i^{(1)}(A)\gamma_iE\left((y_i-\mathbf{x}_i'\beta)\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)^2\right)+o(m^{-1}).$$

Using the Stein identity

$$Er_{21} = -g_{1i}^{-1}\gamma_i^{(1)}(A)\gamma_i(A+\psi_i)E\left(\frac{\partial}{\partial y_i}\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)^2\right\}\right)+o(m^{-1}).$$

Note that

$$\frac{\partial}{\partial y_i}\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)^2\right\}$$
$$= \frac{\partial}{\partial y_i}\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)\right\}(\hat{A}-A)^2+2\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)\frac{\partial\hat{A}}{\partial y_i},$$

where

$$\frac{\partial}{\partial y_i}\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)\right\} = \frac{\partial}{\partial A}\mathbf{x}_i'\left\{\frac{\partial}{\partial y_i}\tilde{\beta}(A)\right\}$$
$$= \frac{\partial}{\partial A}\left\{\frac{\tilde{h}_{ii}}{A+\psi_i}\right\}+O(m^{-1}),$$

where $\tilde{h}_{ii} = \mathbf{x}_i'\left\{\sum_{j=1}^m(A+\psi_j)^{-1}\mathbf{x}_j\mathbf{x}_j'\right\}^{-1}$ is $O(m^{-1})$, $\frac{\partial\hat{A}}{\partial y_i} = O_P(m^{-1})$ and $\frac{\partial}{\partial y_i}(\hat{\beta}(\hat{A})) = O_P(m^{-1})$ (see Kubokawa, 2010, page 14; Datta et al., 2011, page 20).

$$\begin{aligned} Er_{22} &= g_{1i}^{-1}\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)\gamma_i\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A)\gamma_i^{(1)}(A)(\hat{A}-A) \\ &= g_{1i}^{-1}\gamma_i\gamma_i^{(1)}(A)\left\{\mathbf{x}_i'\tilde{\beta}^{(1)}(A)\right\}^2 E(\hat{A}-A)^3 = O(m^{-3/2}). \end{aligned}$$

Thus, $E(S_2S_3) = O(m^{-3/2})$.

Let us evaluate $E(S_1S_3)$ as follows.

$$S_1S_3 = \frac{z\times\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}-\sqrt{g_{1i}(A)}\right)\gamma_i\mathbf{x}_i'(\hat{\beta}-\tilde{\beta})}{g_{1i}(A)}.$$

Using the Taylor series expansion we have

$$= g_{1i}^{-1}z\gamma_i g_{1i}^{(1)}\mathbf{x}_i'\tilde{\beta}^{(1)}(\hat{A}-A)^2.$$

Taking expectations

$$\begin{aligned} E(S_1S_3) &= g_{1i}^{-1}z\gamma_i g_{1i}^{(1)}\mathbf{x}_i'\tilde{\beta}^{(1)}E(\hat{A}-A)^2 \\ &= g_{1i}^{-1}z\gamma_i g_{1i}^{(1)}\mathbf{x}_i'\tilde{\beta}^{(1)}Var(\hat{A}). \end{aligned}$$

We know that,

$$\begin{aligned} S_1^2 &= \left(\frac{z\left(\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB})}-\sqrt{g_{1i}(A)}\right)}{\sqrt{g_{1i}(A)}}\right)^2 \\ &= \frac{z^2}{4}\left(g_{1i}^{-1}(A)(\widehat{MSE}(\hat{\theta}_i^{EB})-g_{1i}(A))\right)^2. \end{aligned}$$

Using the Taylor series expansion, we have

$$S_1^2 = \frac{z^2}{4}\left(g_{1i}^{-1}(A)(\widehat{MSE}(\hat{\theta}_i^{EB})-g_{1i}(A))\right)^2.$$

Taking expectations on both sides

$$\begin{aligned} ES_1^2 &= \frac{z^2}{4}E\left(g_{1i}^{-1}(A)(g_{1i}^{1}(A)(\hat{A}-A))\right)^2 + O(m^{-3/2}) \\ &= \frac{z^2}{4}g_{1i}^{-2}(A)(g_{1i}^{1}(A))^2E(\hat{A}-A)^2 + O(m^{-3/2}) \\ &= \frac{z^2\psi_i^2}{4A^2(A+\psi_i)^2}V(\hat{A}) + O(m^{-3/2}). \end{aligned}$$

Let us evaluate $S_2^2 = g_{1i}^{-1}(\gamma_i-\hat{\gamma}_i)^2(y_i-\mathbf{x}_i'\hat{\beta})^2$.

By the Taylor series expansion,

$$\begin{aligned} S_2^2 &= g_{1i}^{-1}(\gamma_i-\hat{\gamma}_i)^2(y_i-\mathbf{x}_i'\hat{\beta})^2 \\ &= g_{1i}^{-1}(\gamma_i^{(1)}(A))^2(y_i-\mathbf{x}_i'\tilde{\beta}(A))^2(\hat{A}-A)^2 - 2g_{1i}^{-1}(\gamma_i^{(1)}(A))^2\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(y_i-\mathbf{x}_i'\tilde{\beta}(A))(\hat{A}-A)^3 \\ &\quad + (\mathbf{x}_i'\tilde{\beta}^{(1)}(A))^2(\hat{A}-A)^4 \\ &= r_{31}+r_{32}+r_{33}. \end{aligned}$$

Taking expectations on both sides

$$Er_{31} = g_{1i}^{-1}(\gamma_i^{(1)}(A))^2E\left\{(y_i-\mathbf{x}_i'\tilde{\beta}(A))^2(\hat{A}-A)^2\right\}$$

$$= g_{1i}^{-1}(\gamma_i^{(1)}(A))^2 E\left\{(y_i - \mathbf{x}_i'\beta)^2(\hat{A} - A)^2\right\} + o(m^{-1}).$$

Using the Stein identity

$$E[(\hat{A} - A)^2(y_i - \mathbf{x}_i'\beta)^2] = (A + \psi_i)E\left[\frac{\partial}{\partial y_i}\left\{(\hat{A} - A)^2(y_i - \mathbf{x}_i'\beta)\right\}\right]$$

$$= (A + \psi_i)E\left[2(\hat{A} - A)(y_i - \mathbf{x}_i'\beta)\frac{\partial \hat{A}}{\partial y_i} + (\hat{A} - A)^2\frac{\partial}{\partial y_i}\left\{(y_i - \mathbf{x}_i'\beta)\right\}\right] + o(m^{-1})$$

$$= (A + \psi_i)E(\hat{A} - A)^2 + o(m^{-1}).$$

$$Er_{31} = g_{1i}^{-1}(\gamma_i^{(1)}(A))^2(A + \psi_i)E(\hat{A} - A)^2 + o(m^{-1}).$$

$$Er_{32} = 2g_{1i}^{-1}(\gamma_i^{(1)}(A))^2\mathbf{x}_i'\tilde{\beta}^{(1)}(A)E\left\{(y_i - \mathbf{x}_i'\beta)(\hat{A} - A)^3\right\} + o(m^{-1})$$

$$= 6g_{1i}^{-1}(\gamma_i^{(1)}(A))^2\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(A + \psi_i)E(\hat{A} - A)^2\frac{\partial \hat{A}}{\partial y_i} + o(m^{-1}).$$

$$Er_{33} = (\mathbf{x}_i'\tilde{\beta}^{(1)}(A))^2E(\hat{A} - A)^4 = O(m^{-2}), \text{ since } E(\hat{A} - A)^{2S} = O(m^{-S}).$$

Substituting reveals that

$$ES_2^2 = Er_{31} + Er_{32} + Er_{33}$$

$$= g_{1i}^{-1}(\gamma_i^{(1)})^2(A + \psi_i)E(\hat{A} - A)^2 + O(m^{-2})$$

$= \frac{\psi_i}{A(A+\psi_i)^2}V(\hat{A}) + O(m^{-2})$, since $(\gamma_i^{(1)})^2 = \frac{\psi_i^2}{(A+\psi_i)^4}$ and $E(\hat{A} - A)^2 = V(\hat{A}) + o_p(m^{-1})$.

Using this fact, we obtain

$$ES_2^2 = \frac{\psi_i}{A(A+\psi_i)^2}V(\hat{A}) + O(m^{-2}).$$

Similarly, we have

$E(S_3^2) = \frac{\psi_i}{A(A+\psi_i)}\mathbf{x}_i'\left(\sum_{j=1}^m (A + \psi_j)^{-1}\mathbf{x}_j\mathbf{x}_j'\right)^{-1}\mathbf{x}_i$, where $Var(e_i) = \psi_i$

$E(S_3^3) = O(m^{-2})$.

Note that the remainder term can be simplified using a similar argument to Theorem 4.1.1. Finally, combination and simplification of all the above expressions gives the CP of $I^{CEB}(\hat{A})$ as

$$P[\theta_i \in I^{CEB}(\hat{A})] = 2\Phi(z) - 1 - \frac{z\phi(z)}{4}\eta + O(m^{-3/2})$$

where $\eta = (z^2 + 1)\frac{\psi_i^2}{A^2(A+\psi_i)^2}V(\hat{A})$. □

Appendix B

Proof of Theorem 4.2.1

$$\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}) = \widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB}) + \widehat{\mathrm{MSE}}(\hat{\theta}_j^{EB}) - \frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\sum_{i=1}^{m}\frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j.$$

Proof. $\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}) = \mathrm{Var}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}) + \left\{\mathrm{Bias}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB})\right\}^2$
where
$\mathrm{Var}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}) = \mathrm{Var}(\hat{\theta}_i^{EB}) + \mathrm{Var}(\hat{\theta}_j^{EB}) - 2\mathrm{Cov}(\hat{\theta}_i^{EB}, \hat{\theta}_j^{EB})$
and
$\mathrm{Bias}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}) = \mathrm{Bias}(\hat{\theta}_i^{EB}) - \mathrm{Bias}(\hat{\theta}_j^{EB})$
$\left\{\mathrm{Bias}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB})\right\}^2 = \left\{\mathrm{Bias}(\hat{\theta}_i^{EB}) - \mathrm{Bias}(\hat{\theta}_j^{EB})\right\}^2$

$$= \mathrm{Bias}^2(\hat{\theta}_i^{EB}) + \mathrm{Bias}^2(\hat{\theta}_j^{EB}) - 2\mathrm{Bias}(\hat{\theta}_i^{EB})\mathrm{Bias}(\hat{\theta}_j^{EB}).$$

Substituting reveals that

$$\begin{aligned}\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}) &= \mathrm{Var}(\hat{\theta}_i^{EB}) + \mathrm{Var}(\hat{\theta}_j^{EB}) - 2\mathrm{Cov}(\hat{\theta}_i^{EB}, \hat{\theta}_j^{EB}) \\ &\quad + \mathrm{Bias}^2(\hat{\theta}_i^{EB}) + \mathrm{Bias}^2(\hat{\theta}_j^{EB}) - 2\mathrm{Bias}(\hat{\theta}_i^{EB})\mathrm{Bias}(\hat{\theta}_j^{EB})\end{aligned}$$

$$= \mathrm{Var}(\hat{\theta}_i^{EB}) + \mathrm{Bias}^2(\hat{\theta}_i^{EB}) + \mathrm{Var}(\hat{\theta}_j^{EB}) + \mathrm{Bias}^2(\hat{\theta}_j^{EB})$$
$$-2\left\{\mathrm{Cov}(\hat{\theta}_i^{EB}, \hat{\theta}_j^{EB}) + \mathrm{Bias}(\hat{\theta}_i^{EB})\mathrm{Bias}(\hat{\theta}_j^{EB})\right\}.$$

$$\begin{aligned}\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}) &= \widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB}) + \widehat{\mathrm{MSE}}(\hat{\theta}_j^{EB}) \\ &\quad -2\left\{\mathrm{Cov}(\hat{\theta}_i^{EB}, \hat{\theta}_j^{EB}) + \mathrm{Bias}(\hat{\theta}_i^{EB})\mathrm{Bias}(\hat{\theta}_j^{EB})\right\}.\end{aligned}$$

By definition the bias of $\hat{\theta}_i^{EB}$ is given by

$\text{Bias}(\hat{\theta}_i^{EB}) = E(\hat{\theta}_i^{EB}) - \theta_i$.

Since $\hat{\theta}_i^{EB}$ is an unbiased estimator of θ_i, then the $\text{Bias}(\hat{\theta}_i^{EB}) = 0$.

Thus,

$$\widehat{\text{MSE}}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}) = \widehat{\text{MSE}}(\hat{\theta}_i^{EB}) + \widehat{\text{MSE}}(\hat{\theta}_j^{EB}) - 2\text{Cov}(\hat{\theta}_i^{EB}, \hat{\theta}_j^{EB}).$$

The covariance of $\hat{\theta}_i^{EB}$ and $\hat{\theta}_j^{EB}$ is defined as

$$\begin{aligned}
\text{Cov}(\hat{\theta}_i^{EB}, \hat{\theta}_j^{EB}) &= E\left[(\hat{\theta}_i^{EB} - E(\hat{\theta}_i^{EB}))(\hat{\theta}_j^{EB} - E(\hat{\theta}_j^{EB}))\right] \\
&= E[\hat{\theta}_i^{EB}\hat{\theta}_j^{EB}] - E[\hat{\theta}_i^{EB}]E[\hat{\theta}_i^{EB}] \\
&= \frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\sum_{i=1}^m \frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j.
\end{aligned}$$

□

Appendix C

Proof of Theorem 4.2.3

For any $z > 0$

$$P\left[\theta_i - \theta_j \in I^{CEBD}(\hat{A})\right] = P\left[\theta_i - \theta_j \in \hat{\theta}_i^{EB} - \hat{\theta}_j^{EB} \pm z \times \sqrt{\widehat{MSE}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB})}\right]$$
$$= 2\Phi(z) - 1 - \tfrac{z\phi(z)}{4}\eta_D + O(m^{-3/2})$$

where

$$\eta_D = (z^2+1)(\tfrac{A\psi_i}{A+\psi_i} + \tfrac{A\psi_j}{A+\psi_j})^{-2}\left(\tfrac{\psi_i^2}{(A+\psi_i)^2} + \tfrac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A}).$$

Proof. The CP of $I^{CEB}(\hat{A})$ is written as

$$P\left[\theta_i - \theta_j \in I^{CEB}(\hat{A})\right] = P\left[-z + F(-z) < \tfrac{\theta_i - \theta_j - (\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB})}{\sqrt{g_{1i}(A) + g_{2j}(A}} < z + F(z)\right]$$
$$= E[\Phi(z+F(z)) - \Phi(-z+F(-z))]$$

where $F(z) = \frac{z \times \left(\sqrt{\widehat{MSE}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB})} - \sqrt{g_{1i}(A)+g_{1j}(A)}\right) + \hat{\theta}_i^{EB} - \tilde{\theta}_i^{B} - (\hat{\theta}_j^{EB} - \tilde{\theta}_j^{B})}{\sqrt{g_{1i}(A)+g_{1j}(A)}}$,
$g_{1i}(A) = \frac{A\psi_i}{A+\psi_i}$, $g_{1j}(A) = \frac{A\psi_j}{A+\psi_j}$.

Using the Taylor series expansion with an integral remainder term, $\Phi(z+F(z))$ is evaluated as

$$\Phi(z+F(z)) = \Phi(z) + F(z)\phi(z) + \tfrac{1}{2}F^2(z)\phi'(z) + \tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2\phi''(x)dx$$
$$= \Phi(z) + [F(z) - \tfrac{z}{2}F^2(z)]\phi(z) - \tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2(1-x^2)\phi(x)dx.$$

Now observe that

$$E(\Phi(z+F(z))) = \Phi(z) + E\Big[F(z)\phi(z) - \tfrac{z}{2}F^2(z)\phi(z)$$
$$- \tfrac{1}{2}\int_z^{z+F(z)}(z+F(z)-x)^2(1-x^2)\phi(x)dx\Big].$$

In a similar manner to a one population case, we have

$$\hat{\theta}_i^{EB} - \tilde{\theta}_i^{B} - \left\{\hat{\theta}_j^{EB} - \tilde{\theta}_j^{B}\right\}$$
$$= (1-\hat{\gamma}_i)y_i + \hat{\gamma}_i\mathbf{x}_i'\hat{\beta} - (1-\gamma_i)y_i - \gamma_i\mathbf{x}_i'\tilde{\beta} - \left\{(1-\hat{\gamma}_j)y_j + \hat{\gamma}_j\mathbf{x}_j'\hat{\beta} - (1-\gamma_j)y_j - \gamma_j\mathbf{x}_j'\tilde{\beta}\right\}$$

$$= (\gamma_i - \hat{\gamma}_i)r_i + \gamma_i \mathbf{x}_i'(\hat{\beta} - \tilde{\beta}) - (\gamma_j - \hat{\gamma}_j)r_j - \gamma_j \mathbf{x}_j'(\hat{\beta} - \tilde{\beta})$$

where $r_i = (y_i - \mathbf{x}_i'\hat{\beta})$ and $r_j = (y_j - \mathbf{x}_j'\hat{\beta})$.

Thus F(z) can be decomposed into three components as follows:

$$F(z) = \frac{z \times \left(\sqrt{\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB})} - \sqrt{g_{1i}(A) + g_{1j}(A)} \right)}{\sqrt{g_{1i}(A) + g_{1j}(A)}} + \frac{(\gamma_i - \hat{\gamma}_i)r_i - (\gamma_j - \hat{\gamma}_j)r_j}{\sqrt{g_{1i}(A) + g_{1j}(A)}} + \frac{\gamma_i \mathbf{x}_i'(\hat{\beta} - \tilde{\beta}) - \gamma_j \mathbf{x}_j'(\hat{\beta} - \tilde{\beta})}{\sqrt{g_{1i}(A) + g_{1j}(A)}}$$

$$= Q_1 + Q_2 + Q_3 \quad \text{say.}$$

We obtained the expected values of Q_1 and Q_1^2 given above

$$EQ_1 = (g_{1i} + g_{1j})^{-1} \frac{z}{2} \left(g_{2i}(A) + g_{3i}(A) + g_{2j}(A) + g_{3j}(A) \right)$$

$$-(g_{1i} + g_{1j})^{-1} \frac{z}{2} \left(\frac{2\psi_i \psi_j}{(\hat{A} + \psi_i)(\hat{A} + \psi_j)} \mathbf{x}_i' \left\{ \sum_{i=1}^{m} \frac{\mathbf{x}_i \mathbf{x}_i'}{\hat{A} + \psi_i} \right\}^{-1} \mathbf{x}_j \right)$$

$$-(g_{1i} + g_{1j})^{-2} \frac{z^2}{8} \left(\frac{\psi_i^2}{(A + \psi_i)^2} + \frac{\psi_j^2}{(A + \psi_j)^2} \right)^2 V(\hat{A}) + O(m^{-3/2}).$$

$$E(Q_1^2) = (g_{1i} + g_{1j})^{-2} z^2 E \left(\widehat{\mathrm{MSE}}(\hat{\theta}_i^{EB} - \hat{\theta}_j^{EB}) - (g_{1i}(A) + g_{1j}(A)) \right)^2$$

$$= \left(\frac{A\psi_i}{(A + \psi_i)} + \frac{A\psi_j}{(A + \psi_j)} \right)^{-2} z^2 \left(\frac{\psi_i^2}{(A + \psi_i)^2} + \frac{\psi_j^2}{(A + \psi_j)^2} \right)^2 V(\hat{A}) + O(m^{-3/2}).$$

Let us now evaluate Q_2^2 as follows:

$$Q_2^2 = \left(\frac{(\gamma_i - \hat{\gamma}_i)r_i - (\gamma_j - \hat{\gamma}_j)r_j}{\sqrt{g_{1i}(A) + g_{1j}(A)}} \right)^2$$

$$= (g_{1i} + g_{1j})^{-1} \left((\gamma_i - \hat{\gamma}_i)r_i - (\gamma_j - \hat{\gamma}_j)r_j \right)^2$$

$$= (g_{1i} + g_{1j})^{-1} \left\{ ((\gamma_i - \hat{\gamma}_i)r_i)^2 + \left((\gamma_j - \hat{\gamma}_j)r_j \right)^2 - 2(\gamma_i - \hat{\gamma}_i)r_i(\gamma_j - \hat{\gamma}_j)r_j \right\}$$

$$= Q_{11} + Q_{12} + Q_{13} \text{ Say.}$$

Let us evaluate each term turn by turn as follows:

$$Q_{11} = (g_{1i} + g_{1j})^{-1} (\gamma_i - \hat{\gamma}_i)^2 (y_i - \mathbf{x}_i'\hat{\beta})^2.$$

Using the Taylor series expansion,

$$Q_{11} = (g_{1i} + g_{1j})^{-1} (\gamma_i^{(1)}(A))^2 (\hat{A} - A)^2 (y_i - \mathbf{x}_i'\tilde{\beta} - \mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A} - A))^2$$

$$= (g_{1i} + g_{1j})^{-1} (\gamma_i^{(1)}(A))^2 (\hat{A} - A)^2 (y_i - \mathbf{x}_i'\tilde{\beta})^2$$

$$-2(g_{1i} + g_{1j})^{-1} (\gamma_i^{(1)}(A))^2 (\hat{A} - A)^2 \mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A} - A)(y_i - \mathbf{x}_i'\tilde{\beta})$$

$$+ (g_{1i} + g_{1j})^{-1} (\gamma_i^{(1)}(A))^2 (\hat{A} - A)^2 \left(\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A} - A) \right)^2.$$

We evaluated this part for the case of one population mean, thus

$$EQ_{11} = (g_{1i} + g_{1j})^{-1} (\gamma_i^{(1)})^2 (A + \psi_i) E(\hat{A} - A)^2$$

$$EQ_{11} = (g_{1i} + g_{1j})^{-1} (\gamma_i^{(1)})^2 (A + \psi_i) V(\hat{A}).$$

Similarly,

$EQ_{12} = (g_{1i}+g_{1j})^{-1}(\gamma_j^{(1)})^2(A+\psi_j)V(\hat{A})$

$Q_{13} = -2(g_{1i}+g_{1j})^{-1}(\gamma_i-\hat{\gamma}_i)(y_i-\mathbf{x}_i'\hat{\beta})(\gamma_j-\hat{\gamma}_j)(r_j-\mathbf{x}_j'\hat{\beta})$.

Using the Taylor series expansion we have,

$Q_{13} = -2(g_{1i}+g_{1j})^{-1}(\gamma_i^{(1)}(A)(\hat{A}-A)(y_i-\mathbf{x}_i'\tilde{\beta}-\mathbf{x}_i'\tilde{\beta}^{(1)}(A)(\hat{A}-A))(\gamma_j^{(1)}(A)(\hat{A}-A)(y_j-\mathbf{x}_j'\tilde{\beta}-\mathbf{x}_j'\tilde{\beta}^{(1)}(A)(\hat{A}-A))$.

Using the Stein identity, we have

$Q_{13} = -2(g_{1i}+g_{1j})^{-1}(A+\psi_i)\gamma_i^{(1)}\frac{\partial\hat{A}}{\partial y_i}(A+\psi_j)\gamma_j^{(1)}\frac{\partial\hat{A}}{\partial y_j}$, since $\frac{\partial\hat{A}}{\partial y_i} = O_P(m^{-1})$.

This is therefore,

$$\begin{aligned}
Q_2^2 &= Q_{11}+Q_{12}+Q_{13}\\
&= (g_{1i}+g_{1j})^{-1}\left((\gamma_i^{(1)})^2(A+\psi_i)+(\gamma_j^{(1)})^2(A+\psi_j)\right)V(\hat{A})\\
&= \left(\tfrac{A\psi_i}{(A+\psi_i)}+\tfrac{A\psi_j}{(A+\psi_j)}\right)^{-1}\left(\tfrac{\psi_i^2}{(A+\psi_i)^3}+\tfrac{\psi_j^2}{(A+\psi_j)^3}\right)V(\hat{A}).
\end{aligned}$$

The expected value of Q_3^2 can be given as follows

$$\begin{aligned}
EQ_3^2 &= E\left(\frac{\gamma_i\mathbf{x}_i'(\hat{\beta}-\tilde{\beta})-\gamma_j\mathbf{x}_j'(\hat{\beta}-\tilde{\beta})}{\sqrt{g_{1i}(A)+g_{1j}(A)}}\right)^2\\
&= (g_{1i}+g_{1j})^{-1}\Big[\gamma_i^2\mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i\\
&+\gamma_j^2\mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i\\
&-2\,\gamma_i\gamma_j\mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_j\Big]\\
&= \left(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j}\right)^{-1}\Big[\frac{\psi_i^2}{(A+\psi_i)^2}\mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i\\
&+\frac{\psi_j^2}{(A+\psi_j)^2}\mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u v_u'\right\}^{-1}\mathbf{x}_i\\
&-\ 2\frac{\psi_i}{(A+\psi_i)}\frac{\psi_j}{(A+\psi_j)}\mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_j\Big].
\end{aligned}$$

Combining all the above results reveal that

$$\begin{aligned}
EF(z) &= (g_{1i}+g_{1j})^{-1}\frac{z}{2}\Big(g_{2i}(A)+g_{3i}(A)+g_{2j}(A)+g_{3j}(A)\\
&-\frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\textstyle\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j\Big)\\
&-(g_{1i}+g_{1j})^{-2}\frac{z}{8}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2V(\hat{A})+O(m^{-3/2}).
\end{aligned}$$

$$\begin{aligned}
E(F(z)-F(-z)) = 2\frac{z}{2}\left(\frac{A\psi_i}{(A+\psi_i)}+\frac{A\psi_j}{(A+\psi_j)}\right)^{-1}\Bigg[&\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2V(\hat{A})\\
&+\frac{\psi_i^2}{(A+\psi_i)^2}\mathbf{x}_i'\left\{\textstyle\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i
\end{aligned}$$

$$+\frac{\psi_j^2}{(A+\psi_j)^2}\mathbf{x}_i'\left\{\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$$

$$\left.-\ \frac{2\psi_i\psi_j}{(\hat{A}+\psi_i)(\hat{A}+\psi_j)}\mathbf{x}_i'\left\{\sum_{i=1}^m\frac{\mathbf{x}_i\mathbf{x}_i'}{\hat{A}+\psi_i}\right\}^{-1}\mathbf{x}_j\right]$$

$$-\left(\frac{A\psi_i}{(A+\psi_i)}+\frac{A\psi_j}{(A+\psi_j)}\right)^{-2}\frac{z}{4}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2V(\hat{A})$$

$$+\ O(m^{-3/2}).$$

$$E[F(z)^2]=EQ_1^2+EQ_2^2+EQ_3^2$$

$$=\left(\frac{A\psi_i}{(A+\psi_i)}+\frac{A\psi_j}{(A+\psi_j)}\right)^{-2}\frac{z^2}{4}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2V(\hat{A})$$

$$+\left(\frac{A\psi_i}{(A+\psi_i)}+\frac{A\psi_j}{(A+\psi_j)}\right)^{-1}\left[\left(\frac{\psi_i^2}{(A+\psi_i)^3}+\frac{\psi_j^2}{(A+\psi_j)^3}\right)V(\hat{A})\right.$$

$$+\ \frac{\psi_i^2}{(A+\psi_i)^2}\mathbf{x}_i'\left\{\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$$

$$+\frac{\psi_j^2}{(A+\psi_j)^2}\mathbf{x}_i'\left\{\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$$

$$\left.-2\frac{\psi_i}{(A+\psi_i)}\frac{\psi_j}{(A+\psi_j)}\mathbf{x}_i'\left\{\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_j\right]$$

$$+\ O(m^{-3/2}).$$

$$E[F(z)^2+F(-z)^2]$$

$$=2\left(\frac{A\psi_i}{(A+\psi_i)}+\frac{A\psi_j}{(A+\psi_j)}\right)^{-2}\frac{z^2}{4}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2V(\hat{A})$$

$$+2\left(\frac{A\psi_i}{(A+\psi_i)}+\frac{A\psi_j}{(A+\psi_j)}\right)^{-1}\left[\left(\frac{\psi_i^2}{(A+\psi_i)^3}+\frac{\psi_j^2}{(A+\psi_j)^3}\right)V(\hat{A})\right.$$

$$+\ \frac{\psi_i^2}{(A+\psi_i)^2}\mathbf{x}_i'\left\{\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$$

$$+\frac{\psi_j^2}{(A+\psi_j)^2}\mathbf{x}_i'\left\{\sum_{u=1}^m(A+\psi_u)^{-1}x\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_i$$

$$\left.-2\frac{\psi_i}{(A+\psi_i)}\frac{\psi_j}{(A+\psi_j)}\mathbf{x}_i'\left\{\sum_{u=1}^m(A+\psi_u)^{-1}\mathbf{x}_u\mathbf{x}_u'\right\}^{-1}\mathbf{x}_j\right].$$

$$E\left[F(z)-F(-z)-\frac{z}{2}(F(z)^2+F(-z)^2)\right]$$

$$=-2\frac{z}{2}\left(\frac{A\psi_i}{(A+\psi_i)}+\frac{A\psi_j}{(A+\psi_j)}\right)^{-2}\frac{z^2}{4}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2V(\hat{A})$$

$$-\left(\frac{A\psi_i}{(A+\psi_i)}+\frac{A\psi_j}{(A+\psi_j)}\right)^{-2}\frac{z}{4}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2V(\hat{A})$$

$$=-\frac{z\phi(z)}{4}(z^2+1)\left(\frac{A\psi_i}{A+\psi_i}+\frac{A\psi_j}{A+\psi_j}\right)^{-2}\left(\frac{\psi_i^2}{(A+\psi_i)^2}+\frac{\psi_j^2}{(A+\psi_j)^2}\right)^2V(\hat{A})$$

$$+O(m^{-3/2}).$$

Note that the remainder and the integral terms can be simplified using a similar argument to Theorem

4.1.1.

Combination of all the above results gives as

$$P\left[\theta_i - \theta_j \in I^{CEB}(\hat{A})\right] = 2\Phi(z) - 1 - \frac{z\phi(z)}{4}\eta + O(m^{-3/2})$$

where

$$\eta = (z^2+1)\left(\frac{A\psi_i}{A+\psi_i} + \frac{A\psi_j}{A+\psi_j}\right)^{-2}\left(\frac{\psi_i^2}{(A+\psi_i)^2} + \frac{\psi_j^2}{(A+\psi_j)^2}\right)^2 V(\hat{A}). \qquad \square$$

Appendix D

Tables

Table D.1: Synthetic estimates of total poverty, food poverty and percentage of food expenditure measures and associated standard errors and normal theory 95% CIs in 2011: FH results

Zones	Total head count index			Food head count index		
	Synthetic	Est.mse	Std.error	Synthetic	Est.mse	Std.error
Degahabur	44.8853	0.021853	0.147826	58.2195	0.023806	0.154291
Fik	45.3120	0.021837	0.147773	61.1400	0.023787	0.154231
Korahe	30.3939	0.021845	0.147799	42.5948	0.023796	0.15426
Gode	24.5796	0.021836	0.147771	35.8634	0.023787	0.154229
Warder	31.1624	0.02183	0.147751	43.2562	0.023779	0.154206
Afedera	33.4567	0.021839	0.14778	43.4725	0.02379	0.154239
Zone 2	31.6207	0.021843	0.147792	40.9691	0.023794	0.154252
Zone 4	30.0685	0.021855	0.147835	38.0579	0.023808	0.15430

Zones	Percentage of food expenditure				
	Synthetic	Est.mse	Std.error	Lower Int	Upper Int
Degahabur	41.2002	85.6548	9.2550	23.0604	59.3399
Fik	41.4927	58.2048	7.6292	26.5394	56.4459
Korahe	42.1633	72.1804	8.4959	25.5113	58.8153
Gode	41.5729	58.006	7.6162	26.6452	56.5005
Warder	42.1526	47.2688	6.8752	28.6771	55.6280
Afedera	41.5493	62.4139	7.9002	26.0648	57.0338
Zone 2	41.5294	68.6582	8.2860	25.2888	57.7700
Zone 4	41.7799	90.5468	9.5156	23.1293	60.4305

Table D.2: Direct and EBLUP percentage of food expenditure estimates and associated standard errors, CVs and 95% CIs in 2011: FH results

Zones	N	Estimates			Std Error		Coef. Variation		CI estimates	
		Direct	EBLUP	Diff	SE-Dire	SE-EB	CV-Dire	CV-EBLUP	Lower I	Upper I
Akaki	379	33.68	34.181	0.501	1.468	1.431	0.044	0.042	31.377	36.985
Nifas	380	33.14	33.368	0.228	1.467	1.421	0.044	0.043	30.582	36.153
Kolfe	379	35.20	35.584	0.384	1.472	1.425	0.042	0.040	32.791	38.377
Gulele	381	38.69	39.111	0.421	1.479	1.431	0.038	0.037	36.305	41.916
Lideta	371	43.87	43.772	-0.098	1.490	1.443	0.034	0.033	40.944	46.600
Kirkos	377	44.58	44.215	-0.365	1.492	1.448	0.034	0.033	41.377	47.054
Arada	363	45.90	45.674	-0.226	1.494	1.448	0.033	0.032	42.836	48.512
Addis	375	48.34	47.930	-0.410	1.500	1.450	0.031	0.030	45.088	50.772
Bole	366	34.80	35.145	0.345	1.471	1.424	0.042	0.041	32.353	37.937
Yeka	370	37.13	37.108	-0.022	1.476	1.429	0.040	0.039	34.307	39.909
Sem-Gon	883	37.82	38.057	0.237	1.477	1.430	0.039	0.038	35.255	40.860
Deb-Gon	460	40.88	41.219	0.339	1.484	1.435	0.036	0.035	38.408	44.031
Sem-Wol	388	45.03	45.073	0.043	1.493	1.443	0.033	0.032	42.246	47.901
Deb-Wol	860	40.07	40.359	0.289	1.482	1.433	0.037	0.036	37.550	43.168
Sem-She	539	44.85	44.74	-0.110	1.492	1.443	0.033	0.032	41.911	47.569
Msrk-Goj	436	45.90	45.694	-0.206	1.494	1.445	0.033	0.032	42.862	48.525
Mrb-Goj	352	42.85	42.707	-0.143	1.488	1.439	0.035	0.034	39.888	45.527
Waghmr	279	40.79	40.881	0.090	1.483	1.434	0.036	0.035	38.069	43.692
Awi	194	38.41	38.777	0.367	1.478	1.430	0.039	0.037	35.974	41.580
Orom-Zo	204	41.88	42.181	0.301	1.486	1.438	0.036	0.034	39.361	45.000
Bahirdar	383	33.27	33.145	-0.126	1.467	1.425	0.044	0.043	30.352	35.937
Argoba	84	49.96	49.855	-0.105	1.503	1.454	0.030	0.029	47.005	52.705
Mrb-wollega	292	38.58	38.788	0.208	1.479	1.430	0.038	0.037	35.985	41.590
Msk-Wollega	132	38.82	39.255	0.435	1.479	1.432	0.038	0.037	36.449	42.061
Illubabo	368	39.61	39.699	0.089	1.481	1.432	0.037	0.036	36.892	42.506
Jimma	208	38.37	38.755	0.385	1.478	1.430	0.039	0.037	35.953	41.557
Mrb-Shewa	367	38.67	38.797	0.127	1.479	1.430	0.038	0.037	35.994	41.600
Sem-Shewa	303	41.63	41.529	-0.101	1.485	1.436	0.036	0.035	38.715	44.344
Msk-Shewa	579	40.76	40.834	0.074	1.483	1.434	0.036	0.035	38.023	43.645
Arsi	488	42.75	42.665	-0.085	1.488	1.437	0.035	0.034	39.848	45.483
Mrb-Harerge	279	38.04	38.449	0.409	1.478	1.429	0.039	0.037	35.648	41.250
Msk-Harerge	284	41.41	41.605	0.195	1.485	1.435	0.036	0.035	38.792	44.417
Bale	315	37.73	37.764	0.034	1.477	1.429	0.039	0.038	34.964	40.565
Borena	108	51.80	51.262	-0.538	1.507	1.459	0.029	0.029	48.402	54.122
D-M-Shewa	292	36.29	36.586	0.296	1.474	1.425	0.041	0.039	33.793	39.379
Guji	279	44.22	43.880	-0.340	1.491	1.441	0.034	0.033	41.055	46.704
Adama	384	34.66	34.697	0.037	1.470	1.424	0.042	0.041	31.905	37.489
Jimma-Liyu	384	36.17	36.26	0.090	1.474	1.427	0.041	0.039	33.464	39.056

Mrb-arsi	324	44.36	44.025	-0.335	1.491	1.442	0.034	0.033	41.199	46.851
Kellem-wol	168	34.85	35.409	0.559	1.471	1.423	0.042	0.040	32.620	38.199
Horogudru	196	37.53	37.818	0.288	1.477	1.428	0.039	0.038	35.018	40.617
Gurage	328	43.39	43.56	0.170	1.489	1.442	0.034	0.033	40.734	46.386
Hadya	256	44.78	44.170	-0.610	1.492	1.444	0.033	0.033	41.341	47.000
Kembata	158	38.29	38.458	0.168	1.478	1.435	0.039	0.037	35.646	41.270
Sidama	392	40.80	40.679	-0.121	1.484	1.435	0.036	0.035	37.867	43.492
Gedeo	227	44.99	44.747	-0.243	1.492	1.442	0.033	0.032	41.921	47.573
Wolayita	256	38.42	38.616	0.195	1.478	1.432	0.039	0.037	35.809	41.422
Deb-Omo	204	47.57	46.999	-0.571	1.498	1.447	0.032	0.031	44.164	49.834
Sheka	123	41.69	41.421	-0.269	1.485	1.436	0.036	0.035	38.607	44.235
Kefa	132	45.35	44.894	-0.456	1.493	1.443	0.033	0.032	42.065	47.723
Gam-Gofa	390	44.75	44.273	-0.477	1.492	1.441	0.033	0.033	41.447	47.098
Benc-Maj	267	39.74	39.628	-0.112	1.481	1.434	0.037	0.036	36.818	42.439
Yem	48	39.94	40.389	0.449	1.482	1.435	0.037	0.036	37.576	43.201
Amaro	58	39.37	39.335	-0.035	1.480	1.433	0.038	0.036	36.526	42.145
Burji	48	40.13	40.193	0.063	1.482	1.438	0.037	0.036	37.375	43.011
Konso	60	29.76	30.323	0.563	1.460	1.413	0.049	0.047	27.554	33.092
Derashe	60	42.69	42.933	0.243	1.488	1.443	0.035	0.034	40.104	45.761
Dawro	160	40.98	40.933	-0.047	1.484	1.435	0.036	0.035	38.120	43.746
Basketo	59	61.01	59.300	-1.710	1.527	1.480	0.025	0.025	56.399	62.202
Konta	60	38.19	38.611	0.421	1.478	1.433	0.039	0.037	35.802	41.420
Silte	196	43.64	43.585	-0.055	1.490	1.440	0.034	0.033	40.763	46.407
Alaba	60	47.16	46.861	-0.299	1.497	1.454	0.032	0.031	44.012	49.710
Hawassa	383	35.72	35.591	-0.129	1.473	1.432	0.041	0.040	32.783	38.398
Sem-Mrb	256	46.14	46.095	-0.045	1.495	1.445	0.0320	0.0310	43.262	48.928
Makelawi	584	43.16	43.117	-0.043	1.489	1.441	0.035	0.033	40.293	45.942
Misrakawi	380	43.13	43.195	0.065	1.488	1.442	0.035	0.033	40.369	46.021
Debubawi	460	41.67	42.019	0.349	1.485	1.437	0.036	0.034	39.202	44.836
Mirabawi	232	38.93	39.388	0.458	1.480	1.431	0.038	0.036	36.584	42.193
Mekelle	378	31.85	32.155	0.305	1.464	1.421	0.046	0.044	29.37	34.940
Metekel	285	36.84	37.041	0.201	1.475	1.428	0.040	0.039	34.242	39.839
Asosa	625	39.28	39.134	-0.146	1.480	1.433	0.038	0.037	36.325	41.943
Kemashi	200	40.14	40.079	-0.061	1.482	1.436	0.037	0.036	37.264	42.894
Pawi	72	42.37	42.708	0.338	1.487	1.439	0.035	0.034	39.889	45.528
Maokomo	147	38.32	38.506	0.185	1.478	1.430	0.039	0.037	35.702	41.309
Zone1	768	46.89	46.642	-0.248	1.497	1.447	0.032	0.031	43.806	49.478
Zone3	160	50.68	50.062	-0.618	1.505	1.459	0.030	0.029	47.203	52.921
Zone5	207	29.14	29.679	0.539	1.459	1.422	0.050	0.048	26.892	32.465
Agnwak	208	42.59	42.102	-0.488	1.487	1.441	0.035	0.034	39.277	44.927
Nuwer	876	53.69	53.000	-0.690	1.511	1.465	0.028	0.028	50.129	55.872
Mejenger	399	44.96	44.593	-0.367	1.492	1.447	0.033	0.032	41.757	47.429
Itang	64	53.47	52.878	-0.592	1.511	1.466	0.028	0.028	50.004	55.752

Harari	669	41.98	42.150	0.170	1.486	1.436	0.035	0.034	39.335	44.965
Diredawa	668	43.69	43.601	-0.089	1.490	1.443	0.034	0.033	40.773	46.428
Shinle	446	49.83	49.593	-0.237	1.503	1.455	0.030	0.029	46.741	52.444
Giggia	825	43.05	43.011	-0.039	1.488	1.441	0.035	0.034	40.186	45.837
Liben	448	50.08	49.490	-0.590	1.503	1.458	0.030	0.030	46.633	52.348

Table D.3: Direct and EBLUP food poverty head count ratio estimates and associated standard errors and CVs in 2011: FH results

Zones	N	Estimates			Std Error		Coef. Variation	
		Direct	EBLUP	Diff	SE-Dire	SE-EB	CV-Dire	CV-EBLUP
Akaki	379	26.39	25.93	-0.46	3.377	3.32	12.80	12.81
Nifas	380	20.21	20.08	-0.13	3.19	3.13	15.79	15.6
Kolfe	379	19.13	19.95	0.82	3.159	3.10	16.51	15.55
Gulele	381	20.66	21.63	0.97	3.204	3.14	15.51	14.54
Lideta	371	21.52	21.91	0.39	3.229	3.17	15.00	14.47
Kirkos	377	38.21	37.71	-0.50	3.765	3.68	9.85	9.76
Arada	363	16.75	17.29	0.54	3.09	3.04	18.45	17.58
Addis	375	27.23	27.79	0.56	3.403	3.33	12.5	11.99
Bole	366	15.10	15.48	0.38	3.044	2.99	20.16	19.34
Yeka	370	22.97	22.85	-0.12	3.272	3.21	14.25	14.04
Sem-Gon	883	50.82	50.79	-0.03	4.228	4.09	8.32	8.06
Deb-Gon	460	55.6	56.27	0.67	4.419	4.26	7.95	7.57
Sem-Wol	388	58.98	58.99	0.01	4.558	4.39	7.73	7.44
Deb-Wol	860	48.58	48.96	0.38	4.142	4.01	8.53	8.19
Sem-She	539	48.58	49.5	0.92	4.142	4.01	8.53	8.11
Msrk-Goj	436	45.85	46.77	0.92	4.039	3.92	8.81	8.38
Mrb-Goj	352	57.53	57.79	0.26	4.498	4.33	7.82	7.50
Waghmr	279	61.26	60.52	-0.74	4.655	4.47	7.60	7.39
Awi	194	82.94	80.92	-2.02	5.683	5.36	6.85	6.63
Orom-Zo	204	57.06	56.94	-0.12	4.479	4.32	7.85	7.59
Bahirdar	383	22.16	22.18	0.02	3.248	3.19	14.66	14.39
Argoba	84	54.88	55.94	1.06	4.39	4.24	8.00	7.58
Mrb-wollega	292	69.89	68.66	-1.23	5.04	4.81	7.21	7.00
Msk-Wollega	132	73.63	73.41	-0.22	5.216	4.97	7.08	6.77
Illubabo	368	50.24	50.49	0.25	4.206	4.07	8.37	8.06
Jimma	208	67.44	66.56	-0.88	4.927	4.71	7.31	7.08
Mrb-Shewa	367	50.75	50.65	-0.10	4.226	4.09	8.33	8.07
Sem-Shewa	303	19.84	21.08	1.24	3.179	3.12	16.03	14.8
Msk-Shewa	579	28.15	28.94	0.79	3.432	3.36	12.19	11.6
Arsi	488	48.08	48.50	0.42	4.123	3.99	8.58	8.23
Mrb-Harerge	279	65.47	65.47	0.00	4.838	4.63	7.39	7.08
Msk-Harerge	284	55.95	56.55	0.60	4.433	4.27	7.92	7.56
Bale	315	60.64	60.05	-0.59	4.628	4.45	7.63	7.41
Borena	108	63.77	64.63	0.86	4.763	4.58	7.47	7.09
D-M-Shewa	292	41.77	42.21	0.44	3.891	3.78	9.31	8.96
Guji	279	35.5	36.63	1.13	3.672	3.58	10.34	9.78
Adama	384	25.8	25.67	-0.13	3.359	3.29	13.02	12.82
Jimma-Liyu	384	31.02	30.89	-0.13	3.524	3.44	11.36	11.15

Mrb-arsi	324	55.93	55.39	-0.54	4.432	4.28	7.92	7.72
Kellem-wol	168	79.24	77.53	-1.71	5.492	5.20	6.93	6.71
Horogudru	196	70.56	69.41	-1.15	5.071	4.84	7.19	6.97
Gurage	328	55.71	56.57	0.86	4.423	4.27	7.94	7.55
Hadya	256	66.26	63.73	-2.53	4.874	4.67	7.36	7.33
Kembata	158	55.71	55.63	-0.08	4.423	4.28	7.94	7.70
Sidama	392	65.77	64.47	-1.30	4.852	4.65	7.38	7.21
Gedeo	227	66.07	65.68	-0.39	4.866	4.66	7.36	7.09
Wolayita	256	80.55	77.49	-3.06	5.559	5.27	6.9	6.80
Deb-Omo	204	72.76	70.8	-1.96	5.174	4.92	7.11	6.95
Sheka	123	54.01	52.61	-1.40	4.355	4.20	8.06	7.99
Kefa	132	71.95	71.11	-0.84	5.136	4.89	7.14	6.88
Gam-Gofa	390	70.78	68.49	-2.29	5.081	4.84	7.18	7.07
Benc-Maj	267	70.78	68.74	-2.04	5.081	4.85	7.18	7.06
Yem	48	53.15	54.29	1.14	4.32	4.18	8.13	7.70
Amaro	58	81.31	79.28	-2.03	5.598	5.30	6.88	6.68
Burji	48	84.76	80.66	-4.10	5.779	5.47	6.82	6.78
Konso	60	94.09	88.16	-5.93	6.297	5.86	6.69	6.65
Derashe	60	90.74	87.41	-3.33	6.106	5.75	6.73	6.57
Dawro	160	75.19	73.01	-2.18	5.291	5.03	7.04	6.89
Basketo	59	69.06	69.23	0.17	5.001	4.80	7.24	6.94
Konta	60	83.84	82.56	-1.28	5.73	5.42	6.83	6.57
Silte	196	56.23	56.02	-0.21	4.444	4.28	7.90	7.65
Alaba	60	88.10	84.62	-3.48	5.959	5.64	6.76	6.66
Hawassa	383	26.51	26.44	-0.07	3.381	3.32	12.75	12.56
Sem-Mrb	256	50.56	51.14	0.58	4.218	4.08	8.34	7.98
Makelawi	584	28.22	28.54	0.32	3.434	3.36	12.17	11.78
Misrakawi	380	51.39	50.31	-1.08	4.251	4.12	8.27	8.19
Debubawi	460	51.00	50.82	-0.18	4.235	4.10	8.30	8.07
Mirabawi	232	46.36	47.38	1.02	4.058	3.94	8.75	8.31
Mekelle	378	16.41	15.97	-0.44	3.081	3.03	18.77	18.98
Metekel	285	46.73	47.03	0.30	4.072	3.95	8.71	8.40
Asosa	625	48.36	47.18	-1.18	4.134	4.01	8.55	8.50
Kemashi	200	50.13	50.44	0.31	4.202	4.07	8.38	8.08
Pawi	72	29.31	31.53	2.22	3.469	3.39	11.84	10.76
Maokomo	147	80.58	77.52	-3.06	5.56	5.26	6.90	6.79
Zone1	768	25.79	26.77	0.98	3.359	3.29	13.02	12.29
Zone3	160	47.28	46.26	-1.02	4.093	3.98	8.66	8.60
Zone5	207	39.42	39.12	-0.30	3.807	3.73	9.66	9.53
Agnwak	208	28.07	27.75	-0.32	3.43	3.36	12.22	12.11
Nuwer	876	20.76	22.03	1.27	3.207	3.15	15.45	14.31
Mejenger	399	39.24	39.73	0.49	3.801	3.71	9.69	9.33
Itang	64	19.37	20.7	1.33	3.166	3.12	16.34	15.05

Harari	669	18.33	19.33	1.00	3.136	3.08	17.11	15.92
Diredawa	668	27.63	27.77	0.14	3.416	3.35	12.36	12.05
Shinle	446	48.06	47.71	-0.35	4.123	4	8.58	8.39
Giggia	825	14.75	15.51	0.76	3.034	2.98	20.57	19.25
Liben	448	41.01	40.84	-0.17	3.864	3.77	9.42	9.23

Table D.4: Direct and EBLUP total poverty head count ratio and associated standard errors and CVs in 2011: FH results

Zones		Estimates			Std Error		Coef. Variation	
	N	Direct	EBLUP	Diff	SE-Dire	SE-EB	CV-Dire	CV-EBLUP
Akaki	379	6.05	6.256	0.206	3.589	3.666	57.365	60.601
Nifas	380	2.64	2.907	0.267	3.508	3.599	120.681	136.321
Kolfe	379	3.20	4.546	1.346	3.519	3.610	77.39	112.811
Gulele	381	7.53	8.719	1.189	3.597	3.696	41.249	49.081
Lideta	371	5.38	6.027	0.647	3.56	3.653	59.07	67.898
Kirkos	377	26.36	25.486	-0.874	3.976	4.094	15.60	15.533
Arada	363	13.77	13.929	0.159	3.720	3.823	26.708	27.767
Addis	375	14.50	15.168	0.668	3.728	3.839	24.579	26.473
Bole	366	4.630	5.118	0.488	3.545	3.638	69.259	78.578
Yeka	370	18.89	18.264	-0.626	3.814	3.931	20.884	20.812
Sem-Gon	883	31.84	32.193	0.353	4.073	4.218	12.653	13.248
Deb-Gon	460	37.04	37.79	0.750	4.178	4.339	11.055	11.715
Sem-Wol	388	43.72	43.634	-0.086	4.321	4.500	9.903	10.292
Deb-Wol	860	33.98	34.208	0.228	4.114	4.267	12.027	12.559
Sem-She	539	40.14	40.473	0.333	4.247	4.413	10.493	10.994
Msrk-Goj	436	41.24	41.515	0.275	4.268	4.439	10.281	10.764
Mrb-Goj	352	58.49	57.241	-1.249	4.651	4.876	8.125	8.336
Waghmr	279	38.81	38.753	-0.057	4.215	4.381	10.878	11.289
Awi	194	67.10	65.297	-1.803	4.853	5.110	7.432	7.615
Orom-Zo	204	52.84	51.669	-1.171	4.529	4.728	8.766	8.948
Bahirdar	383	3.84	4.288	0.448	3.537	3.622	82.487	94.335
Argoba	84	45.26	45.907	0.647	4.36	4.537	9.499	10.025
Mrb-wollega	292	53.42	52.499	-0.921	4.534	4.743	8.636	8.879
Msk-Wollega	132	53.34	53.64	0.300	4.537	4.741	8.459	8.889
Illubabo	368	31.67	32.258	0.588	4.066	4.214	12.604	13.307
Jimma	208	41.05	41.338	0.288	4.263	4.435	10.312	10.804
Mrb-Shewa	367	26.47	27.137	0.667	3.96	4.097	14.592	15.477
Sem-Shewa	303	10.41	11.947	1.537	3.648	3.754	30.533	36.063
Msk-Shewa	579	18.82	19.597	0.777	3.808	3.930	19.431	20.881
Arsi	488	41.64	41.424	-0.216	4.273	4.449	10.315	10.685
Mrb-Harerge	279	40.68	41.406	0.726	4.255	4.426	10.276	10.880
Msk-Harerge	284	39.75	40.476	0.726	4.234	4.404	10.460	11.078
Bale	315	43.45	43.244	-0.206	4.316	4.493	9.980	10.341
Borena	108	66.57	65.872	-0.698	4.855	5.095	7.370	7.654
D-M-Shewa	292	23.00	23.908	0.908	3.889	4.020	16.265	17.480
Guji	279	26.22	27.434	1.214	3.956	4.091	14.418	15.603
Adama	384	6.75	7.070	0.320	3.584	3.68	50.702	54.523
Jimma-Liyu	384	12.68	12.985	0.305	3.694	3.801	28.446	29.975

Mrb-arsi	324	47.09	46.193	-0.897	4.396	4.583	9.517	9.732
Kellem-wol	168	59.78	58.857	-0.923	4.681	4.910	7.954	8.214
Horogudru	196	46.95	46.813	-0.137	4.391	4.579	9.381	9.753
Gurage	328	37.01	37.859	0.849	4.185	4.338	11.054	11.722
Hadya	256	49.85	47.753	-2.097	4.460	4.652	9.341	9.332
Kembata	158	53.91	52.431	-1.479	4.568	4.756	8.712	8.822
Sidama	392	50.85	49.723	-1.127	4.479	4.678	9.007	9.199
Gedeo	227	56.01	55.102	-0.908	4.591	4.811	8.332	8.589
Wolayita	256	69.17	65.889	-3.281	4.911	5.167	7.453	7.471
Deb-Omo	204	69.89	66.604	-3.286	4.913	5.188	7.376	7.423
Sheka	123	39.00	37.69	-1.310	4.218	4.386	11.191	11.245
Kefa	132	51.91	51.584	-0.326	4.502	4.705	8.727	9.063
Gam-Gofa	390	61.49	58.788	-2.702	4.717	4.956	8.024	8.060
Benc-Maj	267	32.88	33.245	0.365	4.096	4.242	12.321	12.902
Yem	48	30.56	31.988	1.428	4.051	4.189	12.663	13.707
Amaro	58	53.24	52.845	-0.395	4.538	4.739	8.587	8.901
Burji	48	29.50	29.951	0.451	4.035	4.165	13.472	14.118
Konso	60	24.37	25.409	1.039	3.917	4.050	15.415	16.620
Derashe	60	55.99	55.458	-0.532	4.614	4.81	8.321	8.591
Dawro	160	62.32	59.97	-2.350	4.741	4.978	7.905	7.988
Basketo	59	72.65	70.345	-2.305	5.017	5.266	7.132	7.249
Konta	60	55.31	55.045	-0.265	4.594	4.792	8.346	8.664
Silte	196	44.37	43.746	-0.624	4.333	4.516	9.906	10.177
Alaba	60	80.19	75.430	-4.760	5.212	5.486	6.910	6.842
Hawassa	383	8.310	8.705	0.395	3.626	3.712	41.654	44.665
Sem-Mrb	256	44.82	44.909	0.089	4.347	4.527	9.680	10.100
Makelawi	584	25.26	25.133	-0.127	3.941	4.070	15.682	16.112
Misrakawi	380	47.17	45.148	-2.022	4.407	4.585	9.761	9.720
Debubawi	460	37.55	37.333	-0.217	4.192	4.351	11.229	11.588
Mirabawi	232	23.84	25.49	1.650	3.908	4.039	15.332	16.941
Mekelle	378	3.350	3.071	-0.279	3.526	3.613	114.824	107.845
Metekel	285	27.15	28.138	0.988	3.977	4.112	14.134	15.145
Asosa	625	38.63	37.257	-1.373	4.216	4.377	11.317	11.330
Kemashi	200	35.04	35.667	0.627	4.145	4.292	11.622	12.249
Pawi	72	15.59	18.324	2.734	3.749	3.862	20.461	24.769
Maokomo	147	65.3	62.973	-2.327	4.812	5.060	7.642	7.748
Zone1	768	24.10	24.778	0.678	3.915	4.044	15.800	16.782
Zone3	160	35.71	34.728	-0.982	4.166	4.308	11.995	12.064
Zone5	207	28.71	28.591	-0.119	4.035	4.147	14.113	14.444
Agnwak	208	34.07	32.621	-1.449	4.126	4.270	12.648	12.532
Nuwer	876	22.15	23.317	1.167	3.888	4.002	16.673	18.066
Mejenger	399	49.32	48.274	-1.046	4.460	4.639	9.238	9.405
Itang	64	9.21	10.994	1.784	3.641	3.730	33.116	40.498

Harari	669	8.87	10.170	1.300	3.618	3.723	35.577	41.971
Diredawa	668	16.87	17.137	0.267	3.777	3.888	22.042	23.049
Shinle	446	51.77	49.793	-1.977	4.509	4.701	9.056	9.080
Giggia	825	8.500	9.542	1.042	3.619	3.715	37.924	43.711
Liben	448	51.08	49.270	-1.810	4.504	4.683	9.141	9.169

Appendix E

Figures

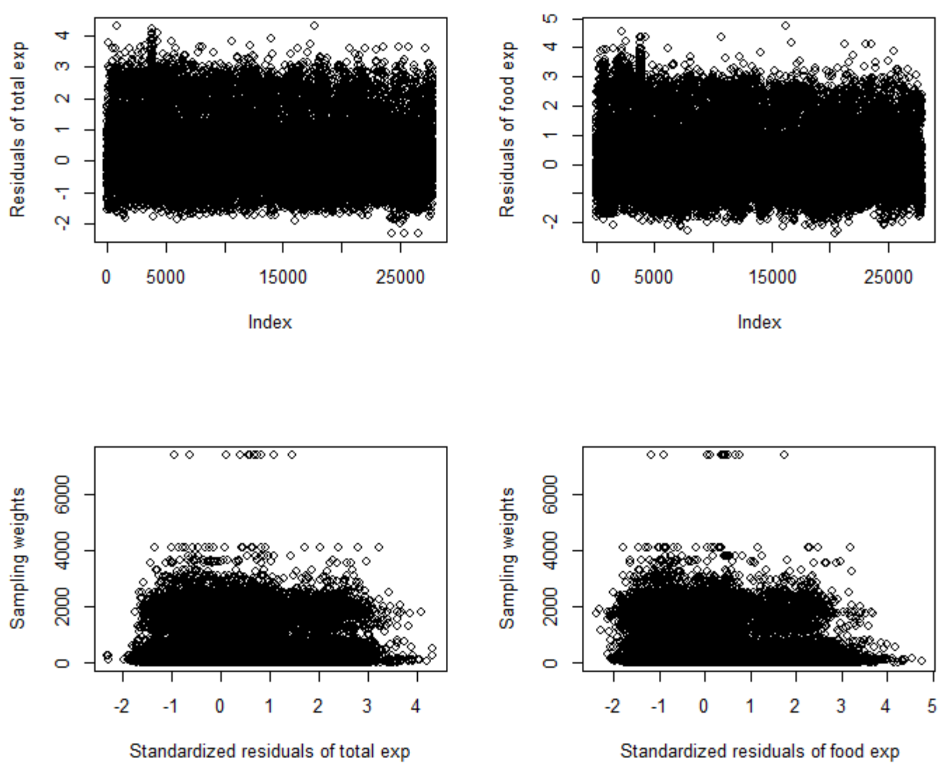

Figure E.1: Index plots of residuals (top) and residuals against sampling weights (bottom)

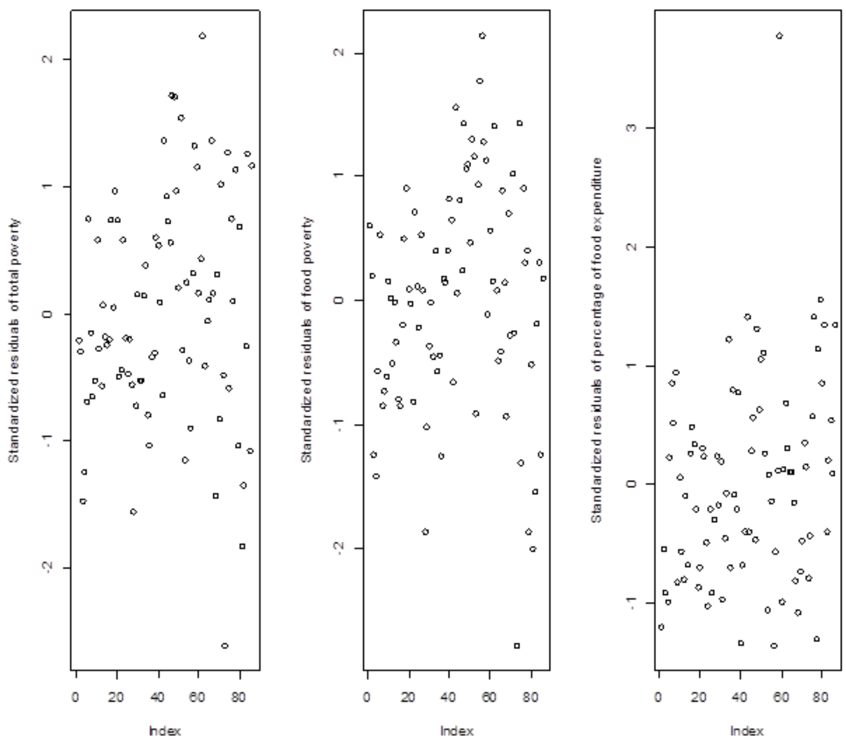

Figure E.2: Index plots of residuals for total poverty (left), food poverty (middle) and percentage of food expenditure (right).

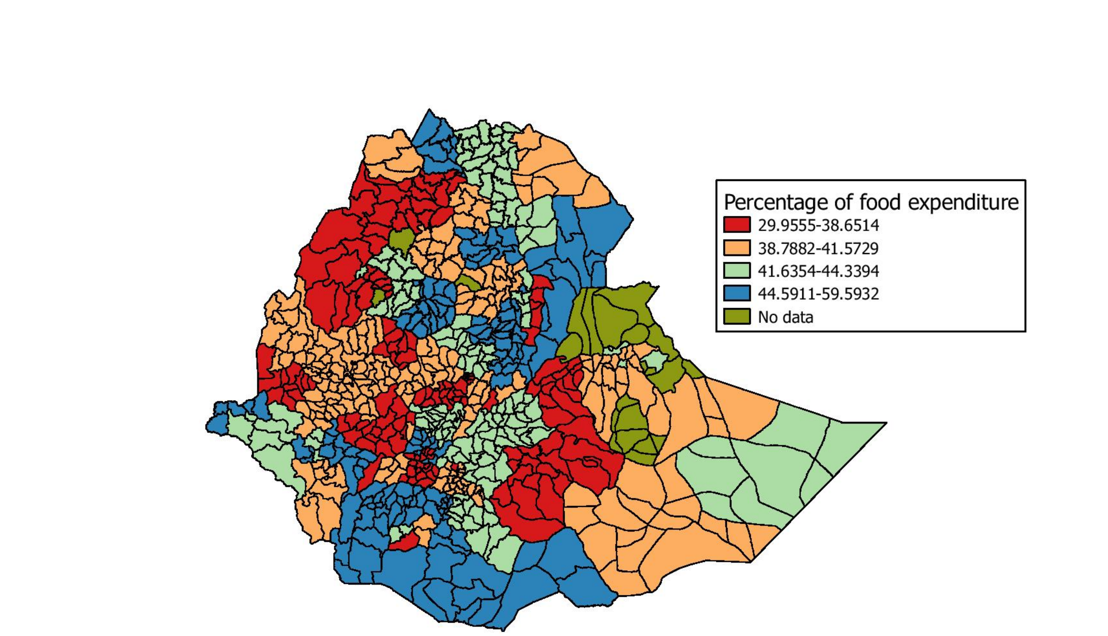

Figure E.3: A map of Ethiopia showing the food costs as a percentage of expenditure

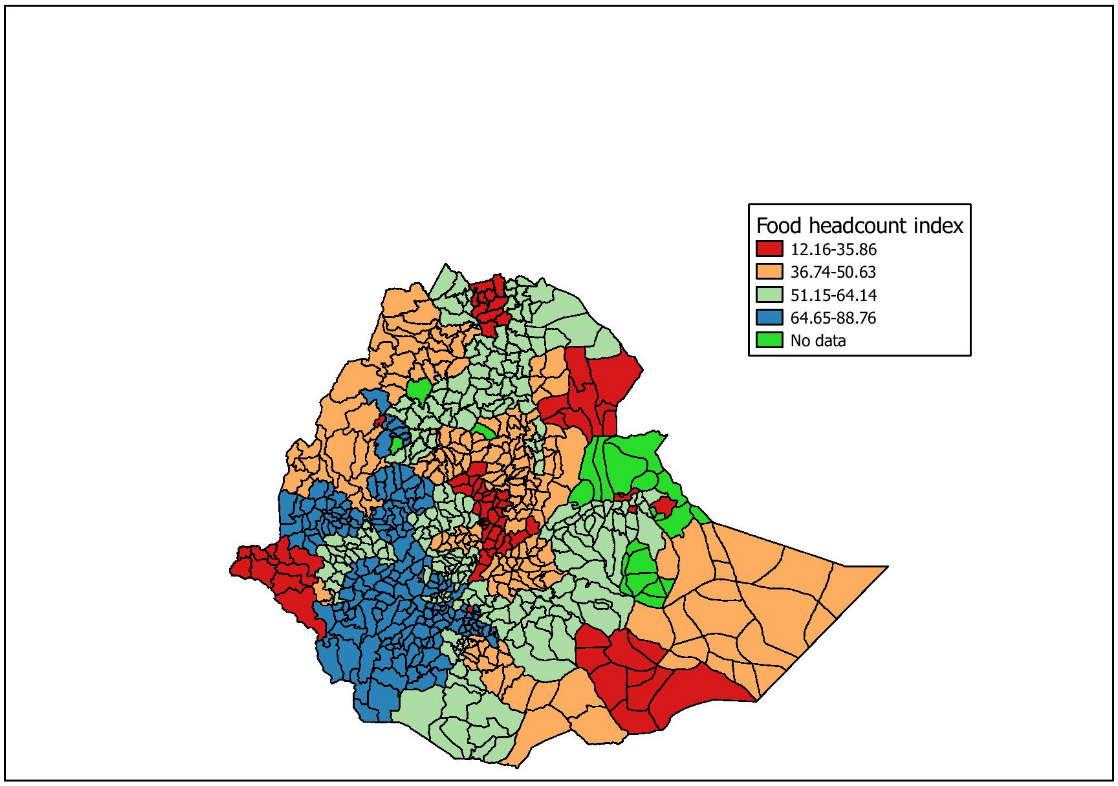

Figure E.4: A map of Ethiopia based on the food head count ratio

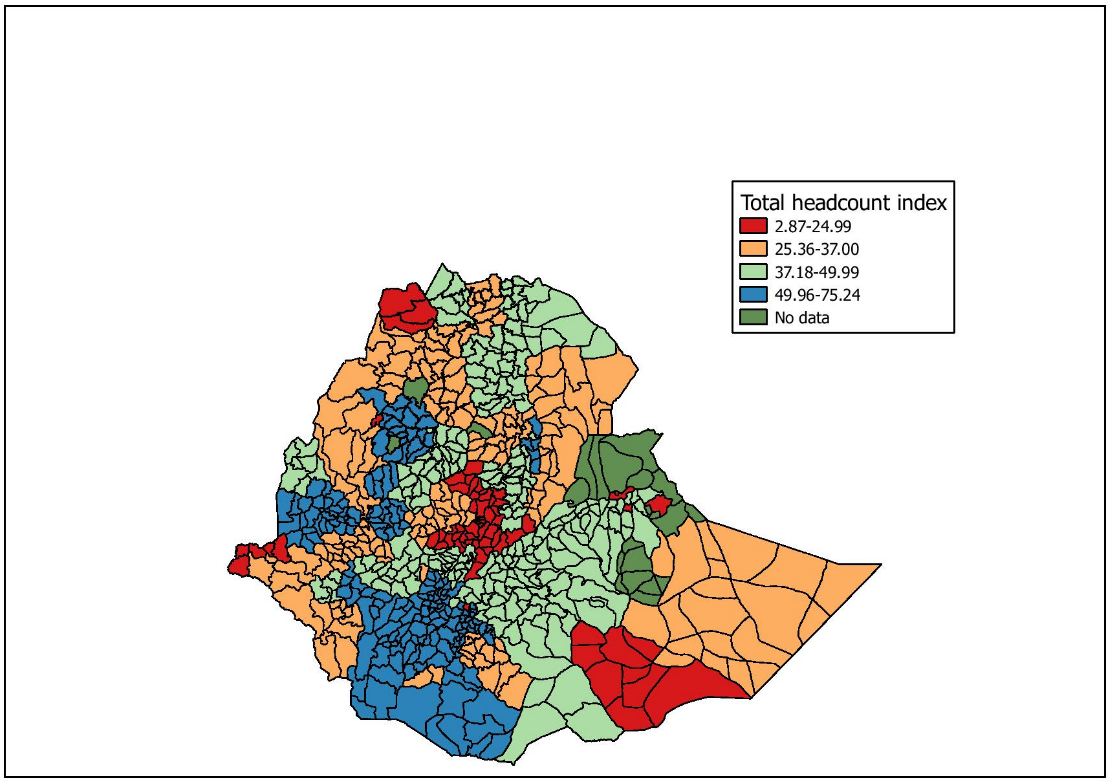

Figure E.5: A map of Ethiopia based on the total head count ratio

Bibliography

Ananya, R. (2007) *Empirical and Hierarchical Bayesian Methods with Applications to Small Area Estimation*. PhD Dissertation, University of Florida.

Basu, R., Ghosh, J.K. and Mukerjee, R. (2003) Empirical Bayes prediction intervals in a normal regression model: higher order asymptotics. *Statistics and Probability Letters*, 63, 197-203.

Battese, G.E., Harter, R.M. 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.

Boonstra, J.H. and Buelens, B. (2011) *Model based estimation*. Statistics Netherlands, The Hague/Heerlen.

Brown, G., Chambers, R., Heady, P., Heasman, D. (2001) Evaluation of small area estimation methods - An application to unemployment estimates from the UK LFS, in *Proceedings of Statistics Canada Symposium 2001: Achieving Data Quality in a Statistical Agency: A Methodological Perspective*, Statistics Canada.

Brown, H. and Prescott, R. (2006) *Applied mixed models in medicine*. Second Edition, John Wiley and Sons Ltd, England.

Butar, F. and Lahiri, P. (2003) On measures of uncertainty of empirical Bayes small area estimators. *Journal of Statistical Planning and Inference*, 112 , 63-76.

Carlin, B.P. and Louis, T.A. (1996) *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman and Hall, London, 399 pages.

Carr, J., Lee, D., Scaife, A. and Hayes, I. (2014) *Food statistics pocketbook 2013-in year update.* Department of Environment, Food and Rural Affairs, UK.

Chandra, H. and Sud, U.C. (2012) Small area estimation for zero-inflated data. *Communications in Statistics - Simulation and Computation*, 41, 632-643.

Chatterjee, S. and Lahiri, P. (2002) *Parametric bootstrap confidence interval in small area estimation problems*. Unpublished report; University of Minnesota and University of Nebraska-Lincoln.

Chatterjee, S., Lahiri, P. and Li, H. (2008) Parametric bootstrap approximation to the distribution of EBLUP and related prediction intervals in linear mixed models. *The Annals of Statistics*, 36, 1221-1245.

Citro, C.F. and Kalton, G. (2000) *Small area income and poverty estimates*. National Academy Press, Washington D.C., 271 pages.

Coondoo, D., Majumder, A. and Chattopadhyay, S. (2011) District-level poverty estimation: a proposed method. *Journal of Applied Statistics*, 38, 2327-2343.

Costa, A., Satorra, A. and Ventura, E. (2006) Improving small area estimation by combining surveys: New perspectives in regional statistics. *SORT*, 30, 101-122.

Cox, D. R. (1975) Prediction intervals and empirical Bayes confidence intervals. In *Perspectives in probability and statistics*, Papers in honor of M.S. Bartlett (ed. J. Gani), 47-55, Academic Press, London.

CSA (2012) The Ethiopian household consumption expenditure survey 2010/11 analytical report. *Statistical Bulletin*, I, Addis Ababa, Ethiopia.

Das, K., Jiang, J. and Rao, J.N.K. (2004) Mean squared error of empirical predictor. *The Annals of Statistics*, 32, 818-840.

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

Datta, G.S. and Ghosh, M. (2012) Small area shrinkage estimation. *Statistical Science*, 27, 95-114.

Datta, G.S., Ghosh, M., Smith, D.D. and Lahiri, P. (2002) On an asymptotic theory of conditional and unconditional coverage probabilities of Empirical Bayes confidence intervals. *Scandinavian Journal of Statistics*, 29, 139-152.

Datta, G.S., Kubokawa, T., Molina, I. and Rao, J.N.K. (2011) Estimation of mean squared error of model-based small area estimators. *Test*, 20, 367-388.

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.

Datta, G.S., Rao, J.N.K. and Smith, D.D. (2002) *On measures of uncertainty of small area estimators in the Fay-Herriot model*. Technical Report, University of Georgia, Athens.

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.

Deichmann, U. (1999) *Geographic aspects of inequality and poverty*. Avilable form: http://www.worldbank.org/poverty/inequalty index.htm. [Accessed: 21st October 2014].

Diao, L., Smith, D.D, Datta, G.S., Maiti, T. and Opsomer, J.D. (2013) Accurate confidence interval estimation of small area parameters under the Fay-Herriot model. *Scandinavian Journal of Statistics*, 41, 497-515.

Esteban, M.D., Morales, D., Perez, A. and Santamaria, L. (2012) Small area estimation of poverty proportions under area-level time models. *Computational Statistics and Data Analysis*, 56, 2840-2855.

Fay, R. E. and Herriot, R. A. (1979) Estimates of income for small places: An application of James-Stein procedure to census data. *Journal of the American Statistical Association*, 74, 269-277.

Federal Democratic Republic of Ethiopia Population and Census Commission (2008) *Sum-*

mary and statistical report of the 2007 population and housing census. Addis Ababa, Ethiopia.

Foster, J., Greer, J. and Thorbecke, E. (1984) A class of decomposable poverty measures. *Econometrica*, 52, 761-766.

Ghosh, M. and Rao, J.N.K. (1994) Small area estimation: An appraisal. *Statistical Science*, 9, 55-76.

Goh, C. and Virola, R.A. (2005) Estimation of local poverty in the Philippines. *National Statistical Coordination Board in collaboration with the World Bank*.

Ha, N.S., Lahiri, P. and Parsons, V. (2011) *Methods and results for small area estimation using smoking data from the 2008 National Health Interview Survey*. Section on Survey Research Methods, Toledo Rd, Hyattsville, LeFrak Hall College Park.

Hall, P., Lee, S. M. and Young, G. A. (2000) Importance of interpolation when constructing double-bootstrap confidence intervals. *Journal of the Royal Statistical Society*, 62, 479-491.

Hall, P. and Maiti, T. (2006) On parametric bootstrap methods for small area prediction. *Journal of the Royal Statistical Society B*, 68, Part 2, 221-238.

Harville, D.A. (1991) Comment on "That BLUP is a good thing: The estimation of random effects", by G.K.Robinson, *Statistical Science* 6, 35-39.

Hawala, S. and Lahiri, P. (2010) *Variance modeling in the U.S. small area income and poverty estimates program for the American community survey.* Section on Survey Methods-JSM, U.S. Census Bureau, United States.

Henderson, C.R. (1950) Estimation of genetic parameters. *Annals of Mathematical Statistics*, 21, 309-310.

Henderson, C.R. (1953) Estimation of variance and covariance components. *Biometrics*, 9, 226-252.

Henderson, C. R. (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31, 423-447.

Household consumption expenditure on food (2014). Avilable form: www.destatis.de/EN/FactsFigures/CountriesRegions/InternationalStatistics/Topic/Tables/BasicDataHouseholdExpFood.html;jsessionid=6EEB86E5F09DF.565D716BF41FDAB9614.cae1. [Accessed: 16th October 2014].

Hulting, F. L. and Harville, D. A. (1991) Some Bayesian and non-Bayesian procedures for the analysis of comparative experiments and for small-area estimation: computational aspects, frequentist properties and relationships. *Journal of the American Statistical Association*, 86, 557-568.

Isidro, M.C. (2010) *Intercensal updating of small area estimates*. PhD Thesis, Massey University, Palmerston North, New Zealand.

Jiang, J. and Lahiri, P. (2006) Mixed model prediction and small area estimation. *Test*, 15, 1-96.

Jiang, J. and Tang, E. (2011) The best EBLUP in the Fay and Herriot model. *Annals of the Institute of Statistical Mathematics*, 63, 1123-1140.

Jiang, J. and Zhang, W. (2002) Distribution free prediction intervals in mixed linear models. *Statistica Sinica*, 12, 537-553.

Kackar, R. and Harville, D.A. (1981) Unbiasedness of two-stage estimation and prediction procedures for mixed linear models. *Communications in Statistics - Theory and Methods*, 10, 1249-1261.

Kackar, R. 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.

Kim, J.K. and Rao, J.N.K. (2012) Combining data from two independent surveys: A model assisted approach. *Biometrika*, 99, 85-100.

Kubokawa, T. (2011) On measuring uncertainty of small area estimators with higher order accuracy. *Journal of Japan Statistical Society*, 41, 93-119.

Kubokawa, T. and Nagashima, B. (2012) Parametric bootstrap methods for bias correction in linear mixed models. *Journal of Multivariate Analysis*, 106, 1-16.

Kurnia, A. and Notodiputro, K.A. (2006) *EB-EBLUP MSE estimator on small area estimation with application to BPS data*. Department of Statistics and CESPO, Indonesia.

Lee, S., Davis, W.W., Nguyen, H.A., Mcneel, T.S., Brick, J.M., and Flores-Cervantes, I. (2007) *Examining trends and averages using combined cross sectional survey data from multiple years*. California Health Interview Survey Methodology Paper: UCLA Center for Health Policy Research.

Li, H. (2006) *An application of parametric bootstrap method in small area estimation problem*. Proceedings of the American Statistical Association: Section on Survey Research Methods, University of Maryland, United States.

Li, H. (2007) *Small area estimation: An empirical best linear unbiased predictor approach*. PhD dissertation, University of Maryland, United States.

Li, H. and Lahiri, P. (2010) An adjusted maximum likelihood method for solving small area estimation problems. *Journal of Multivariate Analysis*, 101, 882-892.

Militino, A.F., Goicoa, T. and Ugarte, M.D. (2012) Estimating the percentage of food expenditure in small areas using bias-corrected P-spline based estimators. *Computational Statistics and Data Analysis*, 56, 2934-2948.

MoFED (2012) *Ethiopia's progress towards eradicating poverty: an interim report on poverty analysis study*. Development Planning and Research Directorate, Addis Ababa, Ethiopia.

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

Molina, I. and Marhuenda, Y. (2013) *Small area estimation*. Package 'sae'. Avilable form: http://cran.rproject.org/web/packages/sae/sae.pdf. [Accessed: 15th July 2013].

Molina, I. and Rao, J. N. K. (2010) Small area estimation of poverty indicators. *The Canadian Journal of Statistics*, 38, 369-385.

Morris, C.N. (1983a) *Parametric Empirical Bayes Confidence Intervals, in Scientific Inference, Data Analysis and Robustness*. Academic Press, New York, 25-50.

Morris, C.N. (1983b) Parametric Empirical Bayes inference: theory and applications (with discussion). *Journal of the American Statistical Association*, 78, 47-65.

Morrison, R.M. (2014) *Selected charts 2014, Ag and food statistics: charting the essentials.* Economic Research Service, USDA, Adminstrative Publication No. (Ag-067) 12pp.

Nandram, B. (1999) An Empirical Bayes prediction interval for the finite population mean of small area. *Statistica Sinica*, 9, 325-343.

Nankervis, J. C. (2005) Computational algorithms for double bootstrap confidence Intervals. *Computational Statistics and Data Analysis*, 49, 461-475.

Nissinen, K. (2009) *Linear mixed models from unit level panel and rotating panel data*. University of Jyvaskyla, Department of Mathematics and Statistics, Report, 117.

Pfeffermann, D. (2002) Small area estimation-new developments and directions. *International Statistical Review*, 70, 125-143.

Pfeffermann, D. (2013) New important developments in small area estimation. *Statistical Science*, 28, 40-68.

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.

Pryseley, A., Tchonlafi, C., Verbeke, G. and Molenberghs, G. (2011) Estimating negative variance components from Gaussian and non-Gaussian data: A mixed models approach. *Computational Statistics and Data Analysis*, 55, 1071-1085.

Rao, J.N.K. (2001) EB and EBLUP in small area estimation in S. E. Ahmed and N. Reid

(Eds.); Empirical Bayes and Likelihood Inference, Lecture Notes in Statistics 148, New York: Springer, 33-43.

Rao, J.N.K. (2003a) *Small Area Estimation*. John Wiley and Sons, Inc., New York, 327 pages.

Rao, J.N.K. (2003b) Some new developments in small area estimation. *Journal of the Iranian Statistical Society*, 2, 145-169.

Rao, J.N.K. (2012) Small area estimation: methods and applications. In *Seminar presented on applications of small area estimation techniques in the social sciences*, Iberoamerican University.

Ray, B.D, Mazhari, H.K., Passah, P.M. and Pandey, M.C. (2000) *Population poverty and environment in north east India*. Commercial Block, New Delhi, India, 412 pages.

Rivest, L.P. and Belmonte, E. (2000) A Conditional Mean Squared Error of Small Area Estimators. *Survey Methodology*, 26, 79-90.

Rivest, L.P. and Vandal, N. (2003) Mean squared error estimation for small areas when the small area variances are estimated: *proceedings of the International Conference on Recent Advances in Survey Sampling*, Carleton University.

Saei, A. and Chambers, R. (2003) *Small area estimation: A review of methods based on the application of mixed models*. Southampton Statistical Sciences Research Institute working paper, Highfield, Southampton, United Kingdom.

Schenker, N. and Raghunathan, T.E. (2007) Combining information from multiple surveys to enhance estimation of measures of health. *Statistics in Medicine*, 26, 1802-1811.

Searle, S.R., Casella, G. and McCulloch, C.E. (1992) *Variance Components*. John Wiley and Sons, Inc., Hoboken, New Jersey, 524 pages.

Stein, C. (1981) Estimation of the mean of a multivariate normal distribution, *Annals of Statistics*, 9, 1135-1151.

Tanton, R. (2007) SpatialMSM: The Australian spatial microsimulation model. The 1st General Conference of the International Microsimulation Association, Vienna, Austria 20-22 August 2007.

Tarozzi, A. and Deaton, A. (2007) *Using census and survey data to estimate poverty and inequality for small areas*. Duke University, Durham, Princeton University, Princeton.

The Economist (2013) *Daily chart: Thought for food.* Avilable form: http://www.economist.com/blogs/graphicdetail/2013/03/daily-chart-5. [Accessed: 25^{th} September 2014].

World Bank (1998) *Using disaggregated poverty maps to plan sectoral investments*. PREM notes.

World Bank (2010) *Small area estimation of poverty in rural Bhutan*. Technical report jointly prepared by National Statistics Bureau of Bhutan and the World Bank.