2010

Adaptive inference and design for multistage surveys

Loai Mahmoud Awad Al-Zou'bi
University of Wollongong

Adaptive Inference and Design for Multistage Surveys

A thesis submitted in fulfilment of the requirements for the award of the degree of

DOCTOR OF PHILOSOPHY

from

University of Wollongong

by

Loai Mahmoud Awad Al-Zou'bi

B.Sc. Mathematics, M.Sc. Statistics

School of Mathematics and Applied Statistics

Wollongong 2522, NSW, Australia

November 2010

Dedicated to

My Parents

My Wife

Abstract

Two-stage sampling usually leads to higher variances for estimators of means and regression coefficients, because of intra-class homogeneity. This thesis will develop and evaluate adaptive strategies for designing and analyzing two-stage surveys, where sample data will be used to determine the appropriate way of allowing for intraclass correlation.

The approach to analysis will be based on fitting a linear regression model to estimate means and regression coefficients. One method for allowing for clustering in fitting a linear regression model is to use a linear mixed model with two levels. If the estimated intra-class correlation is close to zero, it may be acceptable to ignore clustering and use a single level model. This thesis will evaluate an adaptive approach for estimating the variances of estimated regression coefficients. The strategy is based on testing the null hypothesis that the random effect variance component is zero. If this hypothesis is not rejected the estimated variances of estimated regression coefficients are extracted from the one-level linear model. Otherwise, the estimated variance

is based on the linear mixed model, or, alternatively the Huber-White robust variance estimator is used.

Another adaptive strategy based on assessing the estimated design effect due to clustering is also evaluated. This is based on testing the null hypothesis that the random effect variance component is zero and at the same time comparing the estimated design effect to a predetermined cutoff value. If the null hypothesis is rejected and the estimated design effect is more than the predetermined cutoff value the estimated variances of estimated regression coefficients are extracted from the linear mixed model, or, alternatively the Huber-White robust variance estimator is used. Otherwise, the estimated variance is based on the one-level linear model. This approach is found to be nearly identical in practice to the adaptive approach based on just testing the null hypothesis that the random effect variance component is zero.

This adaptive strategy for estimation will be developed based on a two-level linear model assuming normality. It will be evaluated by simulation using normal data, with equal and unequal numbers of observations per cluster, and also using log-normal data, to assess the robustness of the approach to non-normality. The simulations indicate that extreme designs with 5 or less PSUs and many observations per cluster should be avoided. For these extreme designs, most methods perform poorly, including the adaptive methods and the linear mixed model, due to the difficulty of appropriately defining the degrees of freedom for this model. Apart from these extreme designs, the adaptive strategy is found to perform acceptably well, resulting in simpler analysis and slightly shorter confidence intervals.

The use of a pilot survey to estimate the intraclass correlation will also be considered. The pilot estimate of this parameter can be used to estimate the optimal within-PSU sample size for the main survey. The best design based on a "cost-adjusted design effect" and the estimated variance of the estimated regression coefficients will be considered.

An upper cutoff should be placed on the sample size to be selected from each PSU, to allow for the possibility of an under-estimate of the intraclass correlation from the pilot data. The optimal value of this cutoff is found to be between 10 and 50 depending on the pilot sample sizes.

Some results are also obtained on appropriate sample sizes of PSUs and units in the pilot study.

Certification

I, Loai Mahmoud Awad Al-Zou'bi, declare that this thesis, submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Mathematics and Applied Statistics, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.

Loai Mahmoud Awad Al-Zou'bi

29 November 2010

Acknowledgements

All praise are due to ***ALLAH***, the most magnificent, the most beneficent. I thank Him for his guidance and protection through my life. It is with His guidance and help that I have been able to complete this huge task.

I express my deep sense of gratitude to my supervisor Senior Lecturer Robert Clark, for his great effort and valuable guidance in helping me throughout my study. He has taught me a great deal of interesting inferential statistics. He gave me the opportunity to explore challenging research problems and he has been a constant source of guidance and inspiration.

I express my gratitude to Professor David Steel for his comments and suggestions on the final form of this work. I also express my appreciation to staff in the School of Mathematics and Applied Statistics for their help and support.

Finally, I owe my deepest thanks and gratitude to my parents, my wife, our families; the gift of unbounded love and support has no equal.

Thanks to my friends at the University of Wollongong, in Jordan, in

Australia and in United Arab Emirates for their help and support.

Thank you Dr. Raed Alzghool, Ghaith Al-Zou'bi, Dr. Maen Al-Hawari, Mohammad Al-Kadiri, Dr. Ala Maghayreh, Akram Matarneh, Ahamad Aloqaily, and many many anonymous.

List of Conferences and Publications

- Adaptive inference for multi-stage survey data, accepted for publication in the Communications in Statistics, Simulations and Computations (2010);
- Adaptive modeling for complex survey data, Australian Statistical Conference, Melbourne, 2008.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Cluster and Multistage Surveys

Two-stage sampling designs are used in many surveys of social, health, economic and demographic topics. Final population units are grouped into primary sampling units (PSUs). The first stage of selection is a sample of PSUs and the second stage is a sample of units within selected PSUs. For example, PSUs and units could be schools and students, or households and people, or geographic areas and households (see for example (see for example Cochran, 1977; Kish, 1965)).

Two-stage sampling is typically used because

- There is no sampling frame of final units, but a frame of PSUs (e.g. a list of suburbs) is available.
- Cost; for example it is much cheaper to draw a two-stage sample of 100 students from 10 schools than draw a simple random sample of

100 students, as those students might be dispersed over 100 schools (Snijders, 2001).

- Within-group correlations may be of interest in their own right. For instance, the correlation between values for students in the same school might be of interest.

A complication of two-stage sampling is that values of a variable of interest may tend to be more similar for units from the same PSU than for units from different PSUs. The intraclass correlation (ICC), ρ, is a measure of the association between the observations for members of the same PSU. It also describes the PSU homogeneity (Hansen et al., 1953, Chapter 6). If the intraclass correlation is non-zero, the clustered nature of the design should be reflected in the analysis procedure. One way of doing this is by fitting a multilevel model (MLM) (Goldstein, 2003, Chapter 1).

In practice the intraclass correlation is often quite small. For example, if units within PSUs are no more homogenous than units over all PSUs, then the intraclass correlation is zero. On the other hand, if units from the same PSU have equal values then the intraclass correlation is 1. The intraclass correlation may take a negative value, but in practice it is generally positive. If each PSU in the population contains M units, the smallest possible value of ρ is $-1/(M-1)$. This occurs when the population is finite with high heterogeneity within PSUs, and zero variance between PSU means (Hansen et al., 1953, p.260, show this for repeated probability sampling from a fixed finite population).

In this thesis we will focus on modeling two-stage survey data. In the case of equal number of observations in each PSU, ρ is usually less than 0.1 when PSUs are geographic areas and final units are households in these areas (Verma et al., 1980). When PSUs are households and final units are people in households it is usually between 0 and 0.2 (Clark and Steel, 2002).

Variances of estimators obtained from two-stage samples are often higher than those from a simple random sample of the same size. Kish (1965, Chapter 5) defined the design effect as the ratio of the design variance (the variance over repeated probability sampling from a finite population) under the sampling technique used, to the variance assuming simple random sampling with the same sample size.

If the number of PSUs is large, each PSU contains M units, and the sample size in each PSU is equal to m, then the design effect for the sample mean is given by

$$deff = 1 + (m-1)\rho. \tag{1.1}$$

When PSUs have unequal sample sizes, the deff is not expressible in terms of ρ. Some design effect approximations have been suggested, one of these is

$$deff = 1 + (\bar{m} - 1)\rho, \tag{1.2}$$

where $\bar{m}$ stands for the average PSU sample size (Kish, 1965, p.162).

The optimal value of m can be chosen using simple cost models. Hansen et al. (1953, p.272) and Kish (1965, p.268, Equation 8.3.5) defined a simple cost model for two-stage sampling as

$$C = C_0 + cC_1 + nC_2 \tag{1.3}$$

where C is the total cost, c is number of PSUs in the sample, n is the total sample size, C_0 is the fixed cost, C_1 is the cost of including a new PSU in the sample, and C_2 is the average cost of including an extra unit in the sample. Hansen et al. (1953, p.286) showed that the optimal PSU sample size that minimizes the variance of the sample mean subject to fixed total cost is

$$m_{opt} = \sqrt{\frac{C_1}{C_2}\frac{1-\rho}{\rho}} \tag{1.4}$$

In practice ρ would have to be estimated, sometimes from a pilot survey, in which case the estimator of ρ could be quite imprecise (Ukoumunne, 2002). In the balanced data case, that is when all PSUs have the same number of sample observations, m, Equation (1.3) can be rewritten as $C = C_0 + cC_1 + mcC_2$, therefore $c = (C - C_0)/(C_1 + mC_2)$. Hence, the optimal value of c is

$$c_{opt} = \frac{C - C_0}{C_1 + m_{opt}C_2} \tag{1.5}$$

1.2 Multilevel Analysis of Clustered Data

One way of allowing for correlations between values for units between PSUs is to fit a multilevel model. Multilevel models are a generalization of regression models. Let y_{ij} be a dependent variable of interest, and $\mathbf{x}_{ij}$ a vector of covariates for unit j in PSU i. The two-level linear mixed model (LMM) (Goldstein, 2003, Chapter 2) is given by

$$y_{ij} = \boldsymbol{\beta}'\mathbf{x}_{ij} + b_i + e_{ij}, \quad i = 1, 2, \ldots, c, \quad j = 1, 2, \ldots, m_i, \tag{1.6}$$

where c denotes the number of PSUs in the sample, m_i denotes the number of observations selected in PSU i, $\boldsymbol{\beta}$ is the vector of unknown regression coefficients, $b_i \sim N(0, \sigma_b^2)$ is a PSU specific random effect, and e_{ij} is assumed to be $N(0, \sigma_e^2)$. Therefore $y_{ij} \sim N(\boldsymbol{\beta}'\mathbf{x}_{ij}, \sigma_b^2 + \sigma_e^2)$, with variance $\sigma_y^2 = \sigma_b^2 + \sigma_e^2$. Variances of regression coefficient estimates can be estimated by either standard likelihood theory (West et al., 2007), or by using the robust Huber-White estimator (Huber, 1967; White, 1982). Maximum likelihood or restricted maximum likelihood methods can be used to estimate the model parameters.

The sampler is assumed to know the values of the design variable; hence the sampling design can be ignored (Sugden and Smith, 1984). Unequal selection probabilities are often used in the sampling designs that lie behind the sample selection, at least in some stages of the selection procedure. The use of OLS estimators or other estimators that ignore the sampling design can bring in large bias and therefore mislead the inference when these probabilities are related to dependent variable values (Pfeffermann and Sverchkov, 1999). In this thesis, it is assumed that the sampling design is ignorable (Sugden and Smith, 1984) so that a simple LMM can be applied to the sample. The issues associated with the effect of more complex sampling designs on multilevel models are discussed by Pfeffermann et al. (1998).

1.3 Adaptive Procedures for Analyzing Two-Stage Survey Data

There are number of possible approaches for estimating regression coefficients and their variances when the intraclass correlation (ρ) is thought to be small or has been estimated as a small value. One approach is to fit a linear mixed model regardless. Another is to fit a linear model assuming independent observations, i.e. ρ=0. However, if the sample design is relatively clustered, that is a large number of final units are selected from each PSU, the estimated variances resulting from a linear mixed model can be much larger than those obtained from a linear model assuming independent observations, leading to wider confidence intervals. Moreover, a linear mixed model is more complicated to fit and explain than a simple linear model, so the latter is preferable provided it does not give misleading inference. This thesis will explore a third alternative: an adaptive strategy based on testing the null hypothesis that the PSU-level variance component, σ_b^2, is zero. If the null hypothesis is not rejected we use the linear model for estimating the variances of the estimated regression coefficients $\hat{\boldsymbol{\beta}}$. On the other hand, if the null hypothesis is rejected we use the estimated variance for $\hat{\boldsymbol{\beta}}$ either using the standard likelihood theory variance estimator for the LMM or the Huber-White method.

This strategy is explained in Figure 1.1, where $\widehat{var}_{LM}(\hat{\boldsymbol{\beta}})$ is the estimator

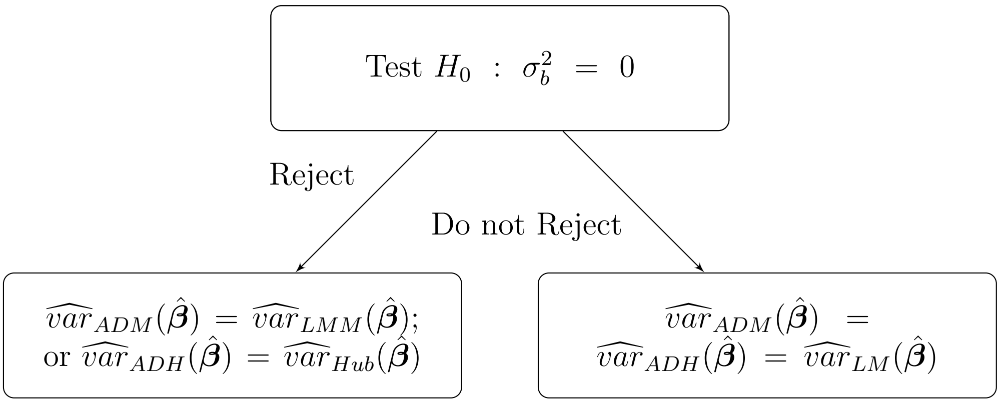


Figure 1.1: Flowchart explaining the adaptive procedure relying on testing $H_0 : \sigma_b^2 = 0$ using LMM-REML variance estimator or Huber-White variance estimator as an alternative

of $var(\hat{\boldsymbol{\beta}})$ using the LM strategy, $\widehat{var}_{LMM}(\hat{\boldsymbol{\beta}})$ is the estimator of $var(\hat{\boldsymbol{\beta}})$ using the LMM strategy, $\widehat{var}_{ADM}(\hat{\boldsymbol{\beta}})$ is the adaptive estimator based on the LMM variance estimator as an alternative and $\widehat{var}_{ADH}(\hat{\boldsymbol{\beta}})$ is the adaptive estimator based on the robust Huber-White variance estimator as an alternative.

Another possible strategy is to use the LMM variance estimator or the robust Huber-White variance estimator as an alternative when $H_0 : \sigma_b^2 = 0$ is rejected and the estimated design effect of the estimated regression coefficient, $\widehat{deff}(\hat{\boldsymbol{\beta}})$, is larger than some cutoff (d). This might be a good approach because the linear model could still be a reasonable approximation even when H_0 is rejected, because of the small estimate of the intraclass correlation ρ. Several cutoff points are considered later in this thesis. Figure 1.2 explains this adaptive strategy.

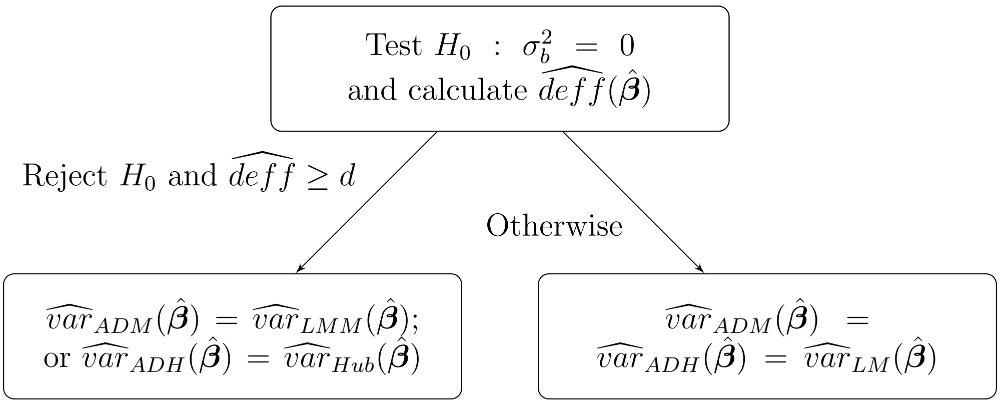


Figure 1.2: Flowchart showing the adaptive procedure based on testing $H_0 : \sigma_b^2 = 0$ and comparing $\widehat{deff}$ to a predetermined cutoff (d), using LMM-REML variance estimator or Huber-White variance estimator as an alternative

1.4 Adaptive Design based on a Pilot survey

A pilot survey is a small study designed to test survey procedures and possibly obtain data to guide sample design, prior to conducting the full survey. It also can help the researcher to address the inadequacy in the proposed design and avoid problems in the large scale studies (Lancaster et al., 2004). For example, Niser (2010) conducted a pilot survey to understand how the field of "study abroad" of Higher Education Institutions in the six New England states of the USA is organized. The contributions of 195 institutions were examined. He used websites, publications and telephone interviews to collect the information for his study. This study revealed that most of the institutions offered study abroad programs. It also revealed that providers played an important role in the broad programs offered to students from different institutions.

The use of a pilot survey to estimate the intraclass correlation ρ is considered in this thesis, assuming the intercept-only model. An estimator of ρ can be substituted in Equation (1.4) to give a within-PSU sample size for the main survey. Because ρ appears in the denominator of (1.4), a small estimated value of ρ might lead to a very large PSU sample size being calculated, which could lead to very high variances from the main survey. Besides, the estimate of ρ is often 0 in multilevel models, which happens often because of small variance across PSU-level units (Muthén and Satorra, 1995). To deal with these possibilities, m will be truncated if it is greater than a cutoff, A. The value of m will also be truncated below to be greater than or equal to 2, to ensure that we can estimate the intraclass correlation ρ. A range of values of the cutoff A will be evaluated by simulation. A range of values of the pilot sample sizes of PSUs (c_p) and units per PSU (m_p) will also be evaluated. Figure 1.3 illustrates the approach.

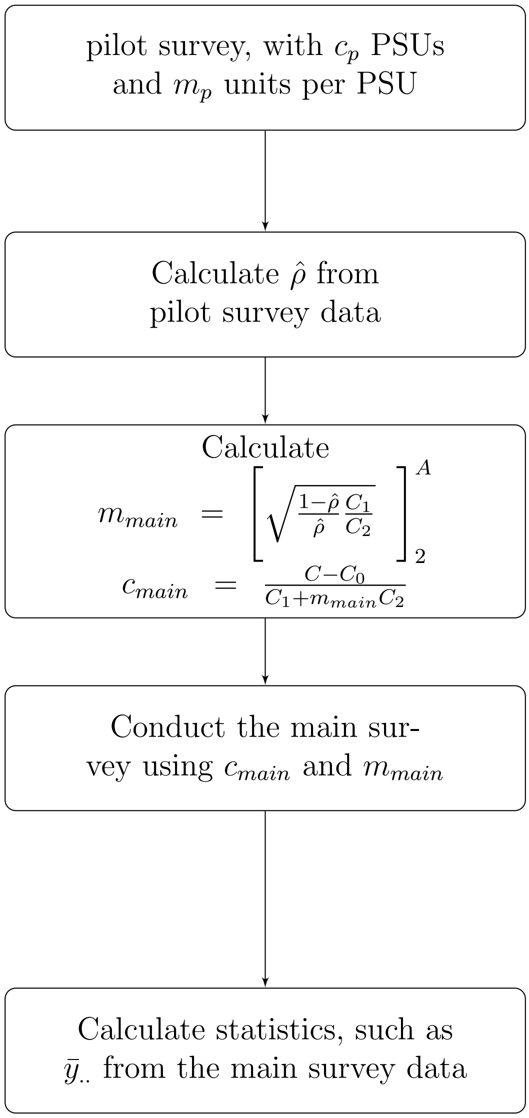


Figure 1.3: Flowchart explaining an adaptive procedure based on a pilot survey

1.5 Outline of Thesis

This thesis is divided into seven chapters.

In Chapter 2, a summary of literature relevant to the thesis will be given. Topics will include linear mixed models, cluster and multistage sampling and the limited literature available on adaptive analysis of survey data.

Chapter 3 will consider two adaptive strategies. Both of them rely on the idea of testing the variance component σ_b^2 in model (1.6). In the first adaptive strategy, if we reject $H_0 : \sigma_b^2 = 0$, we use the LMM estimators of $var(\hat{\beta})$. On the other hand, if we accept H_0, then we assume that $\sigma_b^2 = 0$ and we fit the standard linear model with independent errors. The second adaptive strategy is using the robust Huber-White estimator $\widehat{var}_{Hub}(\hat{\beta})$ is used instead of $\widehat{var}_{LMM}(\hat{\beta})$ when H_0 is rejected. The two strategies are summarized in Figure 1.1. The adaptive strategies will be evaluated in a simulation study of normally distributed data from balanced and unbalanced designs.

The linear mixed model assumes that data are normally distributed. Chapter 4 will evaluate whether the adaptive procedures evaluated in Chapter 3 with simulated normal data are robust to this assumption. This will be done by simulating log-normal data with varying degrees of skewness.

The adaptive procedures of Chapter 3 are based on using the linear model whenever $H_0 : \sigma_b^2 = 0$ is retained. It is possible that $H_0 : \sigma_b^2 = 0$ is rejected but that ρ is still relatively small, so that a linear model may still be a reasonable model. Chapter 5 will evaluate a strategy to deal with this possibility. The LM estimators of $var(\hat{\beta})$ will be used when $H_0 : \sigma_b^2 = 0$ is not rejected or $\widehat{deff} < d$, where d is a cutoff value. If H_0 is rejected and $\widehat{deff} \geq d$, the LMM variance estimators or alternatively the Huber-White variance estimators will be used. This approach is summarized in Figure 1.2. Several cutoff values, d, will be evaluated using simulated normal data.

Chapter 6 will develop approaches for using a pilot survey to estimate ρ, and hence to derive the best m and c for the main survey, as described in Section 1.4. Approaches will be evaluated by simulating pilot data, calculating c_{opt} and m_{opt} based on the pilot data, and then simulating main survey data using these values. The simulation will assume model (1.6) including the assumption of normality. Conclusions will be drawn on appropriate values for m and c for the pilot survey, and for a maximum value A for m in the main survey.

Finally, in Chapter 7 we will state conclusions and suggest directions for future research.

The Appendices contain derivations of some equations as well as some extra tables and the simulation programs.

Chapter 2

Review of Relevant Literature

2.1 List of notations

Symbol	Definition
y_{ij}	j^{th} observation in PSU i
$\mathbf{Y}$	complete set of observations in all PSUs
$\mathbf{x}_i$	vector of covariates
$\mathbf{X}$	the $n \times p$ matrix of explanatory variables
p	number of regressors
$\boldsymbol{\beta}$	vector of unknown regression coefficients
$\mathbf{b}_i$	vector of random coefficients
$\mathbf{e}_{ij}$	error or residual term
c	number of PSUs
m_i	number of observations in PSU i in the unbalanced design
m	number of observations per PSU in the balanced design
n	total number of observations in all PSUs
σ_b^2	random-effect variance component
σ_e^2	error term variance component

$\mathbf{V}$	block diagonal variance-covariance matrix of the complete set of observations in all PSUs, with diagonal elements $\mathbf{V}_i$
$\mathbf{V}_i$	diagonal elements of $\mathbf{V}$, $\mathbf{V}_i = \sigma_b^2 \mathbf{J}_{m_i} + \sigma_e^2 \mathbf{I}_{m_i}$
$\hat{\mathbf{V}}_i$	estimate of $\mathbf{V}_i$
$\lvert\mathbf{V}\rvert$	the determinant of the variance-covariance matrix $\mathbf{V}$
$\mathbf{J}_{m_i}$	$m_i \times m_i$ matrix where all entries are 1
$\mathbf{I}_{m_i}$	$m_i \times m_i$ identity matrix
MSE	mean square error within PSUs
MSA	mean square among PSUs
$\bar{y}_{i.}$	the sample mean for PSU i
ρ	intraclass correlation (ICC)
m_{main}	number of observations per PSU in the main survey of the pilot survey
c_{main}	number of sample PSUs in the main survey of the pilot survey
$deff$	design effect
n_e	effective sample size
C	total cost
C_1	cost of including a new PSU in the sample
C_2	average cost of including an extra element in the sample
C_f	the total cost
$\bar{y}_w$	REML estimate of β in the unbalanced data case, $\sum_{i=1}^{c} \frac{m_i \bar{y}_{i.}}{\hat{\lambda}_i} \Big/ \sum_{i=1}^{c} \frac{m_i}{\hat{\lambda}_i}$
λ_i	variance reciprocal of the mean for PSU i, $\lambda_i = \frac{m_i}{\sigma_e^2 + m_i \sigma_b^2} = (var(\bar{y}_{i.}))^{-1}$
$\hat{\lambda}_i$	estimate of λ_i
Λ	likelihood ratio test
B	number of PSUs in the population

M	population size for each PSU, in the case where all PSUs are of the same size.
c_p	number of PSUs in the pilot survey
m_p	number of observation per PSU in the pilot survey
A	maximum number of observations per PSU
m_{opt}	optimal PSU sample size
c_{opt}	optimal number of PSUs in the sample
μ	mean of the normal distribution in Chapter 4
σ^2	variance of the normal distribution in Chapter 4
ℓ_R	restricted log-likelihood function
ℓ_M	log-likelihood function

2.2 The Two-Level Linear Mixed Model

2.2.1 The Model

Let $\mathbf{X}$ be the $n \times p$ design matrix, which is assumed to be of rank p, and $\mathbf{Y} = (\mathbf{y}'_1, \ldots, \mathbf{y}'_c)'$ be the complete set of $n = \sum_{i=1}^{c} m_i$ observations in the c groups, where $\mathbf{y}_i = (y_{i1}, \ldots, y_{im_i})'$ is the observed vector for the i^{th} PSU. Model (1.6) can also be written as

$$\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{V}), \tag{2.1}$$

where $\mathbf{V}$ is a block diagonal matrix, $\mathbf{V} = diag(\mathbf{V}_i, i = 1, \ldots, c)$, and

$$\mathbf{V}_i = \sigma_b^2 \mathbf{J}_{m_i} + \sigma_e^2 \mathbf{I}_{m_i}, \tag{2.2}$$

where $\mathbf{J}_{m_i}$ is an $m_i \times m_i$ matrix with all entries equal to 1, and $\mathbf{I}_{m_i}$ is the $m_i \times m_i$ identity matrix. $\boldsymbol{\beta}$ is the vector of unknown regression coefficients.

A simple special case of model (1.6) is the intercept-only model, this model includes just a grand mean parameter, it is defined by setting x_{ij} to 1 for all i, j:

$$y_{ij} = \beta + b_i + e_{ij}, \quad i = 1, 2, \ldots, c, \quad j = 1, 2, \ldots, m_i, \tag{2.3}$$

where c denotes number of the sample PSUs, m_i denotes the number of units selected in PSU i, $b_i \sim N(0, \sigma_b^2)$ is a PSU specific random effect and $b_i s$ are independent and identically distributed (iid), and e_{ij} is assumed to be $N(0, \sigma_e^2)$. The parameters σ_b^2 and σ_e^2 are the between- and within-PSUs variance components. This model will be used in the simulation studies in Chapters 3-6.

Observations for different units from the same PSU are correlated. It is assumed that b_i is uncorrelated with e_{ij}, and that b_i and $b_{i'}$ for $i \neq i'$ are uncorrelated. Therefore,

$$\begin{aligned} V(y_{ij}) &= V(b_i) + V(e_{ij}) = \sigma_b^2 + \sigma_e^2, \\ Cov(y_{ij}, y_{ij'}) &= V(b_i) = \sigma_b^2 \text{ for } j \neq j', \text{ and} \\ Cov(y_{ij}, y_{i'j}) &= 0 \text{ for } i \neq i'. \end{aligned} \tag{2.4}$$

(Rao, 1997).

Assuming balanced data design, with $i = 1, \ldots, c$ and $(j \neq j') = 1, \ldots, m$, Rao (1997) defined the intraclass correlation as

$$\rho = \frac{Cov(y_{ij}, y_{ij'})}{\sqrt{V(y_{ij})V(y_{ij'})}}. \tag{2.5}$$

Therefore, substituting (2.4) into (2.5), we obtain

$$\rho = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2}. \tag{2.6}$$

Notice that under model (2.3), the intraclass correlation is always greater than or equal to 0.

Given estimates $\hat{\sigma}_b^2$ and $\hat{\sigma}_e^2$, an estimator for ρ is

$$\hat{\rho} = \frac{\hat{\sigma}_b^2}{\hat{\sigma}_b^2 + \hat{\sigma}_e^2}. \tag{2.7}$$

2.2.2 Likelihood Theory Estimation of Model Parameters

The variance components σ_b^2 and σ_e^2 are generally not known, and are usually estimated by Restricted Maximum Likelihood (REML), giving estimates $\hat{\mathbf{V}}_i$ of $\mathbf{V}_i$.

REML was first introduced by Patterson and Thompson (1971) as a modification of Maximum Likelihood. The REML method is often presented as a technique based on maximization of the likelihood of a set of linear combinations of the elements of the response variable $\mathbf{y}$, say $\mathbf{k}'\mathbf{y}$, where $\mathbf{k}$ is chosen so that $\mathbf{k}'\mathbf{y}$ is free of fixed effects. One of the attractive aspects of REML is that it takes into account the degrees of freedom used up by the estimation of the fixed effects (Diggle et al., 1994, Chapter 4). There is also no loss of information about the variance components when the inference is derived from $\mathbf{k}'\mathbf{y}$ rather than $\mathbf{y}$.

The restricted log-likelihood function is given by West et al. (2007, p.28) by the Equation

$$\begin{aligned} \ell_R &= -\tfrac{1}{2}\big[(n-1)log(2\pi) + log|\mathbf{V}| + log|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}| \\ &\quad +\mathbf{Y}'\mathbf{V}^{-1}\{\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\}\mathbf{V}^{-1}\mathbf{Y}\big], \end{aligned} \tag{2.8}$$

where $\mathbf{V} = diag(\mathbf{V}_i)$ and $\mathbf{V}_i$ are given by (2.2). Maximizing (2.8) with respect to σ_b^2 and σ_e^2 gives the REML estimates of these parameters. The REML estimate of $\hat{\boldsymbol{\beta}}$ is given by

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{Y} \\ &= (\textstyle\sum_{i=1}^{c} \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{x}_i)^{-1}\sum_{i=1}^{c} \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{y}_i. \end{aligned} \tag{2.9}$$

In the intercept-only model, the REML estimates are defined by the following system of equations:

$$\begin{aligned} \frac{n-c}{\hat{\sigma}_e^2} + \sum_{i=1}^{c}\frac{\hat{\lambda}_i}{m_i} - \frac{\sum_{i=1}^{c}\frac{\hat{\lambda}_i^2}{m_i}}{\sum_{i=1}^{c}\hat{\lambda}_i} &= \frac{(n-c)MSE}{\hat{\sigma}_e^4} + \sum_{i=1}^{c}\frac{\hat{\lambda}_i^2}{m_i}(\bar{y}_{i.} - \hat{\beta})^2 \\ \sum_{i=1}^{c}\hat{\lambda}_i - \frac{\sum_{i=1}^{c}\hat{\lambda}_i^2}{\sum_{i=1}^{c}\hat{\lambda}_i} &= \sum_{i=1}^{c}\hat{\lambda}_i^2(\bar{y}_{i.} - \hat{\beta})^2 \\ \hat{\beta} &= \frac{\sum_{i=1}^{c}\hat{\lambda}_i\bar{y}_{i.}}{\sum_{i=1}^{c}\hat{\lambda}_i}, \end{aligned} \tag{2.10}$$

(Sahai and Ojeda, 2005, p.106), where $\bar{y}_{i.}$ is the mean of PSU i and

$$MSE = \frac{1}{n-c}\sum_{i=1}^{c}\sum_{j=1}^{m_i}(y_{ij} - \bar{y}_{i.})^2,$$

and

$$\lambda_i = \frac{m_i}{\sigma_e^2 + m_i\sigma_b^2} = (var(\bar{y}_{i.}))^{-1},$$

is the variance reciprocal of the mean of PSU i, and

$$\hat{\lambda}_i = \frac{m_i}{\hat{\sigma}_e^2 + m_i \hat{\sigma}_b^2},$$

is the estimate of λ_i. Equations in (2.10) must be solved numerically with respect to $\hat{\sigma}_b^2$ and $\hat{\sigma}_e^2$. In the balanced data case ($m_i = m$ for all i), the REML estimates have a simpler form. Let $MSA = \frac{m}{c-1}\sum_{i=1}^{c}(\bar{y}_{i.} - \bar{y}_{..})^2$, the system of equations (2.10) becomes

$$\begin{aligned}
\hat{\sigma_e^2} &= min\left(MSE, \frac{n-c}{n-1}MSE + \frac{c-1}{n-1}MSA\right); \\
\hat{\sigma_b^2} &= \frac{1}{m} max(MSA - MSE, 0); \\
\hat{\beta} &= \bar{y}_{..}.
\end{aligned}$$

(Sahai and Ojeda, 2005, p.40).

2.2.3 Likelihood Theory Estimation of $var(\hat{\boldsymbol{\beta}})$

In this section we discuss the variances of the estimated regression coefficients and their estimators. The estimated variance of the REML $\hat{\boldsymbol{\beta}}$ is given by

$$\begin{aligned}
\widehat{var}(\hat{\boldsymbol{\beta}}) &= \left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1} \\
&= \left(\sum_{i=1}^{c} \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{x}_i\right)^{-1},
\end{aligned} \tag{2.11}$$

where $\hat{\mathbf{V}}_i = \hat{\sigma}_b^2\mathbf{J}_{m_i} + \hat{\sigma}_e^2\mathbf{I}_{m_i}$. For the intercept-only model given by (2.3), in the unbalanced data case, this simplifies to

$$\widehat{var}(\hat{\beta}) = \left\{\sum_{i=1}^{c} \frac{m_i}{\hat{\sigma}_e^2 + m_i\hat{\sigma}_b^2}\right\}^{-1} = (\sum_{i=1}^{c}\hat{\lambda}_i)^{-1}, \tag{2.12}$$

Proof:

$$\begin{aligned}
\widehat{var}(\hat{\beta}) &= \frac{1}{(\sum_{i=1}^{c}\hat{\lambda}_i)^2}\Big[\sum_{i=1}^{c}\widehat{var}(\hat{\lambda}_i\bar{y}_{i.})\Big] \\
&= \frac{1}{(\sum_{i=1}^{c}\hat{\lambda}_i)^2}\Big[\sum_{i=1}^{c}\hat{\lambda}_i^2\widehat{var}(\bar{y}_{i.})\Big] \\
&= \frac{1}{(\sum_{i=1}^{c}\hat{\lambda}_i)^2}\Big[\sum_{i=1}^{c}\hat{\lambda}_i^2\Big\{\frac{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}{m_i}\Big\}\Big] \\
&= \frac{1}{(\sum_{i=1}^{c}\hat{\lambda}_i)^2}\Big[\sum_{i=1}^{c}\Big\{\hat{\lambda}_i^2.\frac{1}{\hat{\lambda}_i}\Big\}\Big] \\
&= \frac{1}{(\sum_{i=1}^{c}\hat{\lambda}_i)^2}\Big(\sum_{i=1}^{c}\hat{\lambda}_i\Big) \\
&= \frac{1}{\sum_{i=1}^{c}\hat{\lambda}_i} \\
&= \Big(\sum_{i=1}^{c}\hat{\lambda}_i\Big)^{-1}.
\end{aligned}$$

In the balanced data case, where $m_i = m$, the variance estimator simplifies further to

$$\widehat{var}(\hat{\beta}) = \frac{1}{c}\Big[\hat{\sigma}_b^2 + \frac{\hat{\sigma}_e^2}{m}\Big]. \tag{2.13}$$

A confidence interval for β could be constructed using the Equation

$$(1-\alpha)100\%CI = \hat{\beta} \pm t_{(df,1-\frac{\alpha}{2})}\sqrt{\widehat{var}(\hat{\beta})}. \tag{2.14}$$

However, it is not clear how the degrees of freedom in (2.14) should be defined for mixed models. Faes et al. (2009) suggested the following approximate confidence interval for the mixed models based on a scaled t-distribution:

$$(1-\alpha)100\%CI = \hat{\beta} \pm \delta^{-1}t_{(\nu,1-\frac{\alpha}{2})}\sqrt{\widehat{var}(\hat{\beta})}, \tag{2.15}$$

where

$$\begin{aligned}
\delta &= \sqrt{\frac{\nu}{(\nu-2)\hat{V}(T)}};\\
\nu &= \sum_{i=1}^{c} \frac{m_i}{1+(m_i-1)\hat{\rho}} - 1;\\
\hat{V}(T) &= 1 + \left(\frac{\hat{\beta}^2}{4(\widehat{var}(\hat{\beta}))^3} \widehat{var}[\widehat{var}(\hat{\beta})] \right),
\end{aligned} \tag{2.16}$$

with $\widehat{var}(\hat{\beta})$ defined in (2.11) and $T = \frac{\hat{\beta}}{\sqrt{\widehat{var}(\hat{\beta})}}$. The scale factor, δ was chosen so that the first two moments of δt agreed with the moments of $t_{\nu-1}$. Faes et al. (2009) did not specify how $V(T)$ or $\widehat{var}[\widehat{var}(\hat{\beta})]$ should be estimated; we will use a parametric bootstrap (see Subsection 3.4.3 for details). Other approaches have been suggested, see for example Satterthwaite (1941) and Kenward and Roger (1997). The method of Faes et al. (2009) has the advantage that it extends naturally to non-Gaussian model, unlike the other approaches.

2.2.4 Bootstrap Approaches

Although in complex survey data there are many methods to estimate the variance and calculate confidence intervals of nonlinear statistics such as regression coefficients, these methods are often awkward or do not broaden to complex designs or nonlinear estimators. Resampling methods such as the bootstrap, the jackknife and balanced repeated replication naturally deal with complex statistics and designs.

Rao and Wu (1988) considered extensions of the *iid* bootstrap to complex survey data of a nonlinear statistics. They applied the bootstrap method to two-stage cluster sampling with equal probabilities at both stages and without replacement to estimate the variances of the estimated regression coefficients. They found that this method is extendable to general sampling designs, such as stratified cluster sampling in which the clusters are sampled with replacement, stratified simple random sampling without replacement, unequal probability sampling without replacement, and two-stage cluster sampling with equal probabilities and without replacement.

Rao and Wu (1988) divided the population into B PSUs with M_i elements each and assumed that the population size is unknown. A simple random sample of c PSUs is selected without replacement, with m_i elements from the M_i elements in each population PSU chosen without replacement. To estimate the variance of $\hat{\beta}$, c PSUs from the c sample PSUs are selected with replacement, then m_i elements are drawn with replacement from from the m_i elements in each selected PSU.

Sitter (1992) extended existing bootstrap with replacement and without replacement to more complex designs including stratified sampling and two-stage cluster sampling. The proposed resampling method was based on resampling a smaller number, c' of the c sample PSUs selected from the B population PSUs without replacement. This step is repeated independently $h = c(1 - c'/c)/(c'(1 - c/B))$ times. Then from the m_i elements in each resampled PSU i a number of within-PSU elements, $1 \leq m_i' < m_i$, are resampled without replacement. This step is also repeated independently $[m_i(1-m_i'/m_i)/m_i'(1-m_i/M_i)](B/hc')$ times. The variance of statistics such as $\hat{\beta}$ is estimated by repeating the procedure a large number of times.

2.2.5 Huber-White Estimator of $var(\hat{\boldsymbol{\beta}})$

Liang and Zeger (1986) suggested the generalized estimation equation (GEE) approach as an alternative to the ML and REML approaches for modeling longitudinal and cross-sectional data. The GEE approach to linear modeling of clustered data can use either ordinary least squares (OLS) or generalized least squares (GLS).

The OLS estimator for $\boldsymbol{\beta}$ is defined by

$$\hat{\boldsymbol{\beta}}_{ols} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}. \tag{2.17}$$

The estimator $\hat{\boldsymbol{\beta}}_{ols}$, when the observations from different PSUs are uncorrelated but the same PSU observations are correlated with common intraclass correlation ρ, is unbiased (Scott and Holt, 1982) with variance equal to

$$var(\hat{\boldsymbol{\beta}}_{ols}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}. \tag{2.18}$$

In general, $\mathbf{V}$ is not known and it can be estimated by $\hat{\mathbf{V}}$, therefore the estimated variance for $\hat{\boldsymbol{\beta}}_{ols}$ is defined by

$$\widehat{var}(\hat{\boldsymbol{\beta}}_{ols}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}. \tag{2.19}$$

The estimator $\widehat{var}(\hat{\boldsymbol{\beta}})$ in (2.11) will be approximately unbiased provided that the variance model (2.2) is correct. If this is not the case, $\widehat{var}(\hat{\boldsymbol{\beta}})$ will be biased and inference will be incorrect. An alternative to ML or REML estimates of $var(\hat{\boldsymbol{\beta}})$ is the robust variance estimate approach described by Liang and Zeger (1986), in the context of modeling longitudinal data using generalized estimating equations (GEE). This approach can be applied to the analysis of data collected using PSUs, where observations within PSUs might be correlated and the observations in different PSUs are independent.

This approach can be referred to as robust or Huber-White variance estimation (Huber, 1967; White, 1982). It will be used as an alternative approach to estimating $var(\hat{\boldsymbol{\beta}})$ in this thesis. The method yields asymptotically consistent covariance matrix estimates even if the variances and covariances assumed in model (1.6) are incorrect. It is still necessary to assume that observations from different PSUs are independent.

In Equation (2.11) in Subsection 2.2.3, the variance of $\hat{\boldsymbol{\beta}}$ was estimated by substituting REML estimates of σ_b^2 and σ_e^2 into $\mathbf{V}_i$. An alternative estimator of $\mathbf{V}_i$ is $\hat{\mathbf{V}}_i^{Hub} = \hat{\mathbf{e}}_i\hat{\mathbf{e}}_i'$, where $\hat{\mathbf{e}}_i = \mathbf{y}_i - \mathbf{x}_i'\hat{\boldsymbol{\beta}}$. $\hat{\mathbf{V}}_i^{Hub}$ is approximately unbiased for $\mathbf{V}_i$ even if (2.2) does not apply.

$$\begin{aligned} E\Big(\hat{\mathbf{V}}_i^{Hub}\Big) &= E(\hat{\mathbf{e}}_i\hat{\mathbf{e}}_i') \\ &\approx E[(\mathbf{y}_i - \mathbf{x}_i'\boldsymbol{\beta})(\mathbf{y}_i - \mathbf{x}_i'\boldsymbol{\beta})'] \\ &= \mathbf{V}_i. \end{aligned} \tag{2.20}$$

Note that

$$\begin{aligned} var(\hat{\boldsymbol{\beta}}) &= var\big(\big(\textstyle\sum_{i=1}^c \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{x}_i\big)^{-1}\big(\textstyle\sum_{i=1}^c \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{y}_i\big)\big) \\ &\approx \big(\textstyle\sum_{i=1}^c \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{x}_i\big)^{-1}\big(\textstyle\sum_{i=1}^c \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{V}_i\hat{\mathbf{V}}_i^{-1}\mathbf{x}_i\big) \\ &\quad \big(\textstyle\sum_{i=1}^c \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{x}_i\big)^{-1}. \end{aligned} \tag{2.21}$$

One way to construct a robust estimator of $var(\hat{\boldsymbol{\beta}})$ is to substitute the robust estimator $\hat{\mathbf{V}}_i^{Hub}$ in (2.21) as follows (Liang and Zeger, 1986),

$$\begin{aligned} \widehat{var}_{Hub}(\hat{\boldsymbol{\beta}}) &= \Big(\textstyle\sum_{i=1}^c \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{x}_i\Big)^{-1}\Big(\textstyle\sum_{i=1}^c \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\hat{\mathbf{V}}_i^{Hub}\hat{\mathbf{V}}_i^{-1}\mathbf{x}_i\Big) \\ &\quad \Big(\textstyle\sum_{i=1}^c \mathbf{x}_i'\hat{\mathbf{V}}_i^{-1}\mathbf{x}_i\Big)^{-1}. \end{aligned} \tag{2.22}$$

When there is only an intercept in the model ($\mathbf{x}_{ij}$=1), (2.22) becomes

$$\widehat{var}_{Hub}(\hat{\beta}) = \frac{\sum_{i=1}^c \hat{\lambda}_i^2(\bar{y}_{i.} - \hat{\beta})^2}{(\sum_{i=1}^c \hat{\lambda}_i)^2}. \tag{2.23}$$

Proof: See Appendix A.

In the balanced data case, (i.e. $m_i = m$), from Equation (2.23) and since $\hat{\lambda}_i$ is constant this estimator becomes

$$\widehat{var}_{Hub}(\hat{\beta}) = \frac{1}{c(c-1)}\sum_{i=1}^{c}(\bar{y}_{i.} - \bar{y}_{..})^2. \tag{2.24}$$

Exact confidence intervals can then be calculated using (2.15) with degrees of freedom equal to c-1 (MacKinnon and White, 1985).

2.3 Testing $H_0 : \sigma_b^2 = 0$ in the Linear Mixed Model

A hypothesis of particular interest in model (1.6) is whether σ_b^2 is zero. If the null hypothesis $H_0 : \sigma_b^2 = 0$ is retained, then there is no significant correlation within PSUs. Two methods to test the hypothesis $H_0 : \sigma_b^2 = 0$ will now be described: the t-test and the restricted-likelihood ratio test.

2.3.1 t-Test

One approach to test $H_0 : \sigma_b^2 = 0$ vs $H_1 : \sigma_b^2 > 0$ is a t-test approach. This approach is the default of the statistical software SPSS (SPSS, 2007). Assuming the intercept-only model for the balanced design with $m_i = m$, the variance of $\hat{\sigma}_b^2$ can be approximated by

$$var(\hat{\sigma}_b^2) = \frac{2}{c-1}\left(\sigma_b^2 + \frac{\sigma_e^2}{m}\right)^2 + \frac{2}{m^2(n-c)}\,\sigma_e^4, \tag{2.25}$$

(Rao, 1997) when the probability that $\hat{\sigma}_b^2 = 0$ is small. (This would be a poor approximation if σ_b^2 is small or zero). Following Berkhof and Snijders (2001), the t-test statistic is the ratio of the restricted maximum likelihood estimator $\hat{\sigma}_b^2$ to its estimated standard deviation $\widehat{se}(\hat{\sigma}_b^2) = (\widehat{var}(\hat{\sigma}_b^2))^{\frac{1}{2}}$; it is given by

$$t \;=\; \frac{\hat{\sigma}_b^2}{\widehat{se}(\hat{\sigma}_b^2),} \tag{2.26}$$

where

$$\widehat{var}(\hat{\sigma}_b^2) \quad = \quad \frac{2}{c-1}\Big(\hat{\sigma}_b^2 + \frac{\hat{\sigma}_e^2}{m}\Big)^2 + \frac{2}{m^2(n-c)}\hat{\sigma}_e^4. \tag{2.27}$$

The null hypothesis $H_0 : \sigma_b^2 = 0$ is rejected if $t > t_{n-1,\alpha}$, where n is the sample size and α is the significance level. This approach is based on an assumption that $t \sim t_{n-1}$ when H_0 is true. However, it is easy to see that this is not justified. For the intercept-only model (2.3) with $m_i = m$, the maximum likelihood estimator for σ_b^2 is given by

$$\hat{\sigma}_b^2 \quad = \quad \frac{1}{m}\Big\{\Big(1 - \frac{1}{c}\Big)MSA - MSE\Big\}, \tag{2.28}$$

provided that this estimator is positive and 0 otherwise. The probability that $\hat{\sigma}_b^2$=0 tends to 0.5 under H_0 for large c and m (Berkhof and Snijders, 2001). When H_0 is not true, the approximate distribution of $\hat{\sigma}_b^2$ is $\frac{\sigma_b^2}{c}\chi^2_{c-1}$ with standard error $\sqrt{2(c-1)}\sigma_b^2/c$, for large m and fixed values of c, σ_b^2 and $var(\hat{\sigma}_b^2)$. Hence the t-test statistic would be expected to give flawed inference for testing that $H_0 : \sigma_b^2 = 0$.

2.3.2 Restricted Likelihood Ratio Test (RLRT)

A better option is to use REML estimators to derive the likelihood ratio test (LRT) statistic for testing $H_0 : \sigma_b^2 = 0$.

The problem of testing $H_0 : \sigma_b^2 = 0$ using the likelihood ratio test is discussed by Self and Liang (1987) using ML estimators for the variance components. Self and Liang (1987) allowed the true parameter values to be on the boundary of the parameter space, and showed that the large sample distribution of the likelihood ratio test is a mixture of χ^2 distributions under nonstandard conditions assuming that response variables are *iid*. This assumption does not generally hold in linear mixed models, at least under the alternative hypothesis. Stram and Lee (1994) used the results of Self and Liang (1987) to prove that the asymptotic distribution of the likelihood ratio test for testing $H_0 : \sigma_b^2 = 0$ has an asymptotic 50:50 mixture of χ^2 with 0 and 1 degrees of freedom under H_0 rather than the classical single χ^2 if the data are *iid* under the null and alternative hypotheses. (χ_0^2 is defined to be the identically zero distribution.) This is because the chance of obtaining a negative estimate of σ_b^2 under the null hypothesis is 50% and the chance of obtaining a positive estimate is 50% as well. However, negative values of $\hat{\sigma}_b^2$ are not permitted and are therefore corrected to 0. When this happens, the chance of getting zero $\hat{\sigma}_b^2$ is approximately 50% (LaHuis and Ferguson, 2009).

From (2.8), the restricted likelihood ratio test is given by

$$\begin{aligned} \Lambda &= -2\, log(RLRT) \qquad (2.29) \\ &= 2 \overset{MAX}{H_A} \;\ell_R(\boldsymbol{\beta}, \sigma_b^2, \sigma_e^2) - 2 \overset{MAX}{H_0} \;\ell_R(\boldsymbol{\beta}, \sigma_b^2, \sigma_e^2). \end{aligned}$$

In the intercept-only model case (2.3) assuming balanced data, Visscher

(2006) gave the REML-based likelihood ratio test (RLRT) as

$$\Lambda = \begin{cases} (n-1)\,log\Big(\frac{n-c}{n-1} + \frac{c-1}{n-1}\,F\Big) - (c-1)\,log(F) & if \quad F > 1 \\ 0 & if \quad F \leq 1. \end{cases} \tag{2.30}$$

Derivation of 2.30: See Appendix A

The large sample distribution of the likelihood ratio Λ is a 50:50 mixture of χ^2 distribution with 0 and 1 degrees of freedom as the parameter values fall on the boundary of the parameter space (Self and Liang, 1987).

In the unbalanced data case, with the intercept-only model, the RLRT is

$$\begin{aligned} \Lambda &= -2\Big(\overset{MAX}{H_0}\ \ell_R - \overset{MAX}{H_A}\ \ell_R\Big) \\ &= ln(n) + (n-1)\,ln(MSE_0) + \frac{\sum_{i=1}^{c} m_i(\bar{y}_{i.} - \bar{y}_{..})^2}{MSE_0} \\ &\quad -(n-c)ln(MSE_A) - \sum_{i=1}^{c} ln(\hat{\eta}_i) - ln\Big(\sum_{i=1}^{c}(\hat{\lambda}_i)\Big) \\ &\quad - \sum_{i=1}^{c} \hat{\lambda}_i(\bar{y}_{i.} - \hat{\beta})^2, \end{aligned} \tag{2.31}$$

where $MSE_0 = \frac{1}{n-1}\sum_{i=1}^{c}\sum_{j=1}^{m_i}(y_{ij} - y_{..})^2$ is the mean squared error under the null hypothesis, $\sigma_b^2 = 0$ and $MSE_A = \hat{\sigma}_e^2$ is the mean squared error under the alternative hypothesis, $\sigma_b^2 > 0$ and $\eta_i = \sigma_e^2 + m_i\sigma_b^2$.

Derivation of 2.31: See Appendix A

2.4 Adaptive Procedures

2.4.1 Review of Longford (2008)

Longford (2008) has investigated the advantages of estimators based on selected models, assuming a two-stage sampling design and c PSUs with m_i observations from each. He used the one-way analysis of variance model, $y_{ij} = \beta_i + b_i + e_{ij},\ i = 1, \ldots, c;\ j = 1, \ldots, m_i$, where β_i is the mean of PSU i.

Two alternative sub-models were considered;

- Model A: no restrictions on β_i;

- Model B: the group means are all equal, $\beta_i = \beta,\ i = 1, \ldots, c$.

Longford (2008) was interested in estimation of β_i for each $i = 1,\ \ldots,\ c$, but for simplicity just β_1 was discussed.

For estimating β_i two estimators were considered $\hat{\beta}_{Ai} = \bar{y}_i$ under model A or $\hat{\beta}_{Bi} = \hat{\beta} = \bar{y}_{..}$ under model B.

The mean squared error $MSE = E[(\hat{\beta}_i - \beta_i)^2]$ of the alternative estimators of β_i were compared. The mean squared errors for $\hat{\beta}_i$ and $\hat{\beta}$ were $MSE(\hat{\beta}_{Ai}) = \frac{\sigma^2}{m_i}$ and $MSE(\hat{\beta}_{Bi}) = \frac{\sigma^2}{n} + (\beta_i - \beta)^2$. Longford recommended using whichever of $\hat{\beta}_{Ai}$ or $\hat{\beta}_{Bi}$ had lower MSE. This results in the following estimator of β_i:

$$\begin{cases} \bar{y}_{i.} & if\ (\beta_i - \beta)^2 > \sigma^2\ g_i \\ \bar{y}_{..} & otherwise, \end{cases} \tag{2.32}$$

where

$$g_i = \frac{1}{m_i} - \frac{1}{n}.$$

In practice (2.32) could not be used because β_i-β is unknown, but Longford used (2.32) to motivate several estimators which can be applied in practice.

One example was to estimate β_i using $\bar{y}_{..}$ when m_i was small, and using $\bar{y}_{i.}$ when m_i was large.

As an alternative estimator to (2.32), Longford (2008) considered the convex combination

$$\tilde{\beta}_i = (1 - t_i)\,\hat{\beta}_i + t_i\,\hat{\beta}, \tag{2.33}$$

where t_i is set to minimize $MSE(\tilde{\beta}_i; \beta_i) = E[(\tilde{\beta}_i - \beta_i)^2]$. The value of t_i that minimizes the MSE is

$$t_i^* \quad = \quad \frac{g_i}{g_i + \frac{(\beta_i - \beta)^2}{\sigma^2}}. \tag{2.34}$$

The "ideal synthetic estimator" is then

$$\tilde{\beta}_i(t_i^*) = (1 - t_i^*)\,\hat{\beta}_i + t_i^*\,\hat{\beta}. \tag{2.35}$$

In practice (2.35) can not be calculated as $(\beta_i - \beta)$ is unknown.

Assuming σ^2 is known, one approach would be to estimate t_i^* using

$$\hat{t}_i = \frac{g_i\,\sigma^2}{g_i\,\sigma^2 + (\hat{\beta}_i - \hat{\beta})^2}. \tag{2.36}$$

2.4.2 Model Averaging

Model averaging is an alternative to model selection. In model selection the best model is selected and used for estimating the model parameters. Model selection calculations are simple as they rely on a single model. On the other hand, model selection ignores model uncertainty and can give therefore over-optimistic inference. Model averaging combines models together and calculate the estimates as weighted averages. It requires more calculations but provides better estimates (Madigan and Ridgeway, 2003). Bayesian model averaging (BMA) is a widely used approach to model averaging approach in many fields, including medicine, meteorology and management sciences (Li and Shi, 2010).

Sorenson and Gianola (2002) define the following terms

$$\begin{aligned}
\Psi &= \text{parameter or future data point,} \\
\mathbf{y} &= \text{data,} \\
D &= \{D_1, D_2, \ldots, D_k\} \;\text{ set of models,} \\
p(D_r) &= \text{prior probability of model } r, r = 1, \ldots, k, \\
p(D_r|\mathbf{y}) &= \text{posterior probability of model } r.
\end{aligned}$$

It is commonly assumed that models are assigned equal prior probabilities, although this is not always true the case (Posada and Buckley, 2004).

The posterior distribution of Ψ in the usual Bayesian approach is given by

$$p(\Psi|\mathbf{y}, D_r) = \frac{p(\mathbf{y}|\Psi, D_r)p(\Psi|D_r)}{p(\mathbf{y}|D_r)}. \tag{2.37}$$

The posterior distribution Equation (2.37) shows the case in which, if the model is true, inferences are conditional on D_r. The idea of Bayesian model averaging, in contrast, is to find the average of the posterior distribution of models, resulting in the following Equation

$$p(\Psi|\mathbf{y}) \;=\; \sum_{r=1}^{k} p(\Psi|\mathbf{y}, D_\mathbf{r})\mathbf{p}(D_\mathbf{r}|\mathbf{y}). \tag{2.38}$$

It is attractive to use BMA, but two real challenges have arisen. The first is how to select the set of models $D_1, \ldots, D_k$. For computational reasons, it is preferable not to use too many models particularly if each model involves complex structure. One approach is to only use the models that operate well according to some criteria such as the Akaike information criteria (AIC) (Akaike, 1974) or Bayesian information criteria (BIC) (Kass and Wasserman, 1995).

Another problem is how to calculate the marginal model likelihood according to the likelihood of every model,

$$p(\Psi|D_r) \;=\; \int p(\Psi|\theta_r, D_r)p(\theta_r|D_r)d\theta_r, \tag{2.39}$$

where θ_r is the vector of parameters in model D_r.

Adaptive confidence intervals calculated in the model selection criterion do not incorporate the model selection uncertainty, and so may not have the correct coverage rates. In this thesis we will evaluate the extent of this problem by simulation. Estimates of the variances of regression coefficients could be done based on model averaging of the linear and the linear mixed model rather than selecting between them (see for example Hoeting et al., 1999; Yuan and Yang, 2005). This approach will not be developed in this thesis, because one of the objectives is to simplify the modeling process when the intraclass correlation is small.

2.5 Cluster and Multistage Sampling

2.5.1 Introduction

In multistage sampling, the population is divided into groups called primary sampling units (PSUs). A random sample from each selected PSU is then selected. If all units within each selected PSU are selected then two-stage sampling is called cluster sampling. Multistage sampling may employ more than two stages of selection. For example, in order to select a sample of local voters in New South Wales in Australia, a random sample of post codes could be surveyed. Then a sample of city blocks could be chosen within selected post codes. Then within each of these blocks a random sample of households could be selected.

One reason why two-stage sampling is used is to reduce cost with face-to-face interviewing (Lehtonen and Pahkinen, 1994). Although the variability of estimates is increased if two-stage sampling is used, it enables surveys to be completed faster with less cost. For example, in the first stage a sample of areas could be chosen; in the second stage a sample of respondents within those areas is selected (Tate and Hudgens, 2007). This can reduce travel and other administrative costs.

Two-stage sampling can be used when there is a list of all PSUs in the population but not of all units. Therefore, one might obtain a random sample of PSUs and then take a census or a sample within the selected PSUs (Hansen and Hurwitz, 1951).

Using a two-stage sample rather than simple random sample of the same size will increase the variance of estimates. The design effect is used to measure the increase in variance that happened when two-stage sampling is used. It is defined as the ratio of the variance of a statistic $\hat{\boldsymbol{\beta}}$ under a two-stage sampling design, $var_d(\hat{\boldsymbol{\beta}})$, to the variance of the statistic calculated under the simple random sampling design of the same sample size n (Kish, 1965, Chapter 5). If the sample PSUs are of equal sizes, m, then the design effect is given by (1.1) in the intercept-only model. If the sample PSUs have different sizes, one approximation of the design effect is given by Equation (1.2).

Under the intercept-only model (2.3), in the unbalanced case

$$
\begin{aligned}
deff(\bar{y}_{..}) &= 1 + \left(\frac{\sum m_i^2}{n} - 1\right)\rho \\
&= 1 + (\bar{m}(1 + c_m^2) - 1)\rho,
\end{aligned}
$$

where c_m is the coefficient of variation of the within PSU sample sizes. Hence provided the within PSU sample size do not vary considerably, c_m^2 will be small and (1.2) will provide a reasonable approximation.

2.5.2 Ignorable Two-Stage Sampling

This thesis assumes ignorable sampling, in the sense that multilevel models can be estimated from sample data without explicitly allowing for the sample design. This subsection reviews the concept of ignorability.

Sugden and Smith (1984) modeled the selection procedure by a sample selection method which relies on the design variables z and may rely on the response variables y and a vector of parameters θ. This design can be written as

$$p(s|y, z; \theta), \quad s \in \Omega, \tag{2.40}$$

where Ω is the set of feasible samples.

Sugden and Smith (1984) investigated ignorability conditions based on designs which depend on the design variables only, given partial information on the design. Such designs can be written as

$$p(s|z), \quad s \in \Omega, \tag{2.41}$$

They defined $d_s = D_s(z)$ to be data derived from knowledge of selection procedure (2.40) and from values of the available probabilities of selection $(s, p(s))$, as well as any values or functions of z. The fundamental condition for ignoring the sampling design given the design information is that $p(s|z) = p(s|d_s)$ for z such that $d_s = D_s(z)$.

Models fitted using data from a simple random sample are generally approximately unbiased for the model that would be estimated from the full population. If the sample design is more complex, the sample model could be biased for the population model (Pfeffermann, 1993).

Pfeffermann (1993) assumed that the population consists of N units and a vector of measurements (y_i, z_i) is linked with every unit i where (y_i, z_i) are independent draws and have a bivariate normal $BN(\mu, \Sigma)$. The aim was to estimate $\mu_y = E(Y)$, where Y is the variable of interest with values $Y_i,\ i = 1, \ldots, N$, from a sample s selected by a probability sampling method. If simple random sampling is used then $\bar{Y}_s$ is an unbiased estimator of μ_y, and it fulfils other optimal properties. It is obvious that inference can ignore the sampling design in this case. However, if probability proportional to z_i, with replacement, is used, then ignoring the sampling design can be misleading, and $\bar{Y}_s$ may be biased for μ_y.

The ignorability of the sampling design depends on the model and the parameters of interest as well as the sample design and the information available about the design. If all design variables are incorporated in the regressor variables in the regression model, then the sampling design is ignorable for estimating the regression coefficients. It is not ignorable for estimating the unconditional mean and variance for the regression dependent variables, if the values of the design variables are only known for units in the sample.

2.5.3 Cost-Variance Modeling and Optimal Design

Two-stage sampling normally leads to estimates having a higher variance than simple random sampling with the same sample size. Therefore, when its effect on reducing the unit cost is more than the increase of the unit variance, two-stage sampling is recommended. Increasing the within-PSU sample size increases both the cost and the variance (Kish, 1965, p.263). Even small values of intraclass correlation lead to a significant increase in variance when the average PSU sample size is large (Gao and Smith, 1998). multi-stage sampling, assuming equal sized PSUs with equal sample size, and simple random sample at both stages:

$$C = C_0 + cC_1 + nC_2, \tag{2.42}$$

Hansen et al. (1953, p.271) stated: "We shall assume, for the particular illustrative sample survey under consideration, that on the basis of prior experience and experimental work we have estimated that $C_2 = \$1$". Whereas C_1 is often not easy to estimate since it includes interviewer travel costs. The fixed costs C_0 do not affect the optimal design.

Kalsbeek et al. (1981) stated that "We believe that the ideal cost model has the following three characteristics. First, it must realistically represent the way in which costs are incurred in an actual survey operation. Second, the formulation should be simple enough so that the optimum solution is tractable. Third, unit costs which constitute the parameters of the cost model should be sufficiently straightforward in interpretation so that they can be easily understood by operations staff to develop useful estimates for calculating optimum allocations. The influence of clustering the sample on costs and variances generally is opposed; it reduces the costs and increases the variances. The economic design of a multistage sample requires the sampling statistician to estimate and balance these influences."

The approximate optimal number of sample PSUs and sample PSU sizes are given by

$$\begin{aligned} m_{opt} &= \sqrt{\tfrac{C_1}{C_2}\tfrac{1-\rho}{\rho}}; \\ c_{opt} &= \tfrac{C-C_0}{C_1+m_{opt}C_2}. \end{aligned}$$

In the discussion so far, it has been assumed that simple random sampling of PSUs and of elements within PSUs is used. In practice probability proportional to size (PPS) selection of PSUs may be preferable (Hansen and Hurwitz, 1943). In the PPS method, the probability of selecting a PSU varies according to the PSU size: the larger the PSU size is the greater the probability of selection will be, up to a maximum of 1. The PPS approach can increase the precision for a given sample size by targeting the sample towards large units that affect population estimates more. With suitable redefinition of the variance components similar results can be obtained for PPS.

Chapter 3

Adaptive Estimators Based on Testing the Variance Component in a Multilevel Model

3.1 Introduction

In multistage sampling, sample units are selected in stages. The target population is divided into primary sample units in the first stage. Sampling units are then subsampled from these PSUs. Further selection is made within each unit. It is used in many surveys of social, health, economic and demographic topics. It is a very flexible technique since many aspects of the design can be controlled, including the number of stages (eg PPS or equal probability, systematic or simple random sampling) of selection or the number of units and the number of units selected for each stage. In this thesis, we are going to consider two-stage samples.

Data from final units within the same PSU may be correlated. One way of analyzing this kind of data is with multilevel models. Multilevel models are a generalization of regression models. Goldstein (2003, Chapter 2) defined the two-level linear mixed model (LMM) by Equation (1.6).

The intraclass correlation (ρ) is a measure of the association between the regression residuals for members of the same PSU. It also expresses the between-PSU variance, σ_b^2 as the proportion of the sum of the between- and the within-PSU variance components (Commenges and Jacqmin, 1994), as described in Equation (2.7)

The intraclass correlation, ρ is quite small in many cases. For instance, it is zero if units within PSUs are homogeneous. The highest possible value of ρ is 1. This is true when values are equal for units from the same PSU (Kish, 1965, Chapter 5). The smallest value of the intraclass correlation is $\frac{-1}{M-1}$ when all PSUs contain M units, but this is rare. Model (2.3) implies that ρ is greater than or equal to 0. The intraclass correlation tends to be positive in typical two-stage surveys. Even small intraclass correlations can have a large effect on the variance, it within-PSU PSU sample sizes are large. In general, when geographic areas are PSUs and household are the final units, the intraclass correlation is less than 0.1 (Verma et al., 1980). If households are PSUs and people in these households are the final units it is usually between 0 and 0.2 (Clark and Steel, 2002).

Regression coefficients, $\boldsymbol{\beta}$, and the variances of their estimates, $var(\boldsymbol{\beta})$, can be estimated using many possible procedures when the intraclass correlation is considered small. The linear model with independent observations is one possible procedure. The linear mixed model is another approach that can be used. An alternative is to use an adaptive approach based on testing the null hypothesis that the PSU-level variance component, σ_b^2, is zero. Accepting the null hypothesis, the linear model will be used for estimating $var(\boldsymbol{\beta})$. If the null hypothesis is rejected, the linear mixed model or the robust Huber-White variance estimator will be used for estimating the variance of the regression coefficient estimates.

This chapter is divided into four sections. Section 3.2 will describe the adaptive strategies. A simulation study of the adaptive and other methods will be described in Section 3.4. In Section 3.5 we will draw conclusions.

3.2 Adaptive Strategies

In this Chapter two adaptive strategies will be considered based on the intercept-only model. Both of them rely on the idea of testing the variance component σ_b^2 in model (1.6). In the first adaptive strategy, if we reject $H_0 : \sigma_b^2 = 0$, we use the LMM estimators of $var(\hat{\boldsymbol{\beta}})$ defined in Equation (2.11). On the other hand, if we accept H_0, then we assume that $\sigma_b^2 = 0$ and we fit the standard linear model with independent errors. This strategy is explained in Figure 1.1.

In the unbalanced case, $\hat{\boldsymbol{\beta}}_{LMM}$ (from the linear mixed model) will depend on $\hat{\lambda}_i$ and therefore on m_i and $\hat{\rho}$. But in the balanced case $\hat{\boldsymbol{\beta}}_{LMM}$ does not depend on $\hat{\rho}$ irrespective of its value as the values of $\hat{\lambda}_i$ are all equal and cancel out.

The second adaptive strategy, is identical, except that the robust Huber-White estimator $\widehat{var}_{Hub}(\hat{\boldsymbol{\beta}})$ is used instead of $\widehat{var}_{LMM}(\hat{\boldsymbol{\beta}})$ when H_0 is rejected.

The two adaptive strategies (ADM) and (ADH) are defined as

$$\widehat{var}_{ADM}(\hat{\boldsymbol{\beta}}) = \begin{cases} \widehat{var}_{LMM}(\hat{\boldsymbol{\beta}}) & \text{if} \quad H_0 \quad \text{is} \quad \text{not} \quad \text{retained} \\ \widehat{var}_{LM}(\hat{\boldsymbol{\beta}}) & \text{if} \quad H_0 \quad \text{is} \quad \text{retained,} \end{cases} \tag{3.1}$$

$$\widehat{var}_{ADH}(\hat{\boldsymbol{\beta}}) = \begin{cases} \widehat{var}_{Hub}(\hat{\boldsymbol{\beta}}) & \text{if} \quad H_0 \quad \text{is} \quad \text{not} \quad \text{retained} \\ \widehat{var}_{LM}(\hat{\boldsymbol{\beta}}) & \text{if} \quad H_0 \quad \text{is} \quad \text{retained.} \end{cases} \tag{3.2}$$

The Huber-White variance estimator is approximately but not exactly unbiased. For the intercept-only model, it is straightforward to show that

$$\frac{E\big(\widehat{var}_{Hub}(\hat{\beta})\big)}{var(\hat{\beta})} = \frac{(\sum_{i=1}^{c}\lambda_i)^2 - \sum_{i=1}^{c}(\lambda_i^2)}{(\sum_{i=1}^{c}\lambda_i)^2}. \tag{3.3}$$

Derivation: See Appendix B

where $\hat{\beta}$ and $var(\hat{\beta})$ are given by (2.10) and (2.23), respectively. Hence a bias-adjusted estimator is given by dividing (2.23) by the right hand side of (3.3), giving:

$$\widehat{var}_{Hub}(\hat{\beta}) = \frac{1}{(\sum_{i=1}^{c}\hat{\lambda}_i)^2 - \sum_{i=1}^{c}\hat{\lambda}_i^2}\sum_{i=1}^{c}\hat{\lambda}_i^2(\bar{y}_{i.} - \hat{\beta})^2, \tag{3.4}$$

The LMM 90% confidence intervals for β are given by

$$(1-\alpha)100\%CI = \hat{\beta} \pm \delta^{-1} t_{(df,1-\frac{\alpha}{2})} \sqrt{\widehat{var}(\hat{\beta})}, \tag{3.5}$$

where $\delta = \sqrt{\frac{\nu}{(\nu-2)\hat{V}(T)}}$, $\alpha = 0.1$ and the degrees of freedom (df) are defined to be:

$$df = \begin{cases} n-1 & \text{using LM Est.} \\ \nu - 1 & \text{using LMM Est.} \\ c-1 & \text{using Huber-White Est..} \end{cases} \tag{3.6}$$

Degrees of freedom for adaptive strategies (ADM) and (ADH) are defined as

$$df_{ADM} = \begin{cases} n-1 & \text{if } H_0 \text{ not rejected} \\ \nu - 1 & \text{if } H_0 \text{ rejected;} \end{cases} \tag{3.7}$$

$$df_{ADH} = \begin{cases} n-1 & \text{if } H_0 \text{ not rejected} \\ c - 1 & \text{if } H_0 \text{ rejected,} \end{cases} \tag{3.8}$$

where ν represents the effective sample size, with $\hat{\nu} = \frac{n}{\widehat{deff}(\hat{\beta})}$. The effective sample size is the ratio of the sample size to the design effect of the $\hat{\beta}$. The degrees of freedom for the linear mixed model are only an approximation (Faes et al., 2009). However, the degrees of freedom of the linear model and Huber-White are exact (MacKinnon and White, 1985). See Subsections 2.2.3 and 2.2.5 for further discussion of the LMM and Huber-White variance estimators and confidence intervals.

The advantage of the adaptive strategy is that we use the simple linear model to derive variance estimators, unless there is strong evidence against

$H_0 : \sigma_b^2 = 0$. This has the benefit of simplifying the model and may also give tighter confidence intervals.

The adaptive confidence intervals may not have the correct coverage rates as they might not incorporate the model selection uncertainty. The extent of this problem will be evaluated by simulation. An alternative approach would be to fit both the LM and LMM and base estimates and inference on model averaging of these two models (see for example Hoeting et al., 1999; Yuan and Yang, 2005). This approach will not be developed in this thesis, because one of the objectives is to simplify the modeling process when the intraclass correlation is small.

3.3 Type 1 and Type 2 Errors of LM and Huber-White Approaches

The choice between the Huber-White and LM estimators of $var(\hat{\beta})$ can be considered as a tradeoff of type 1 and type 2 error. In this context, type 1 error means using $\widehat{var}_{Hub}(\hat{\beta})$ when σ_b^2=0 and type 2 error means $\widehat{var}_{LM}(\hat{\beta})$ when $\sigma_b^2 > 0$. This section derives a result on the mean squared errors of the two approaches reflecting the type 1 and type 2 errors. For simplicity a balanced design and an intercept-only model are assumed, and only the Huber-White and LM approaches are compared.

When σ_b^2=0, we know that $\widehat{var}_{LM}(\hat{\beta}) = \frac{1}{n}s^2$, where $s^2 = \sum_{i=1}^{c}\sum_{j=1}^{m}(y_{ij} -$

$\bar{y}_{..})^2$ which is distributed as a $\sigma_e^2\chi^2_{(n-1)}/(n-1)$ (Hocking, 1996, Theorem 3.1). Therefore,

$$\begin{aligned} E\big(\widehat{var}_{LM}(\hat{\beta})\big) &= (n-1)\frac{\sigma_e^2}{n(n-1)} = \frac{\sigma_e^2}{n} \\ var\big(\widehat{var}_{LM}(\hat{\beta})\big) &= 2(n-1)\frac{\sigma_e^4}{n^2(n-1)^2} = \frac{2\sigma_e^4}{n^2(n-1)}. \end{aligned} \tag{3.9}$$

In the general case when $\sigma_b^2 > 0$, the estimated variance of $\hat{\beta}$ is given by

$$\begin{aligned} \widehat{var}_{LM}(\hat{\beta}) &= \frac{1}{n}s^2 = \frac{1}{n(n-1)}\sum_{i=1}^{c}\sum_{j=1}^{m}(y_{ij}-\bar{y}_{..})^2 \\ &= \frac{1}{n(n-1)}\sum_{i=1}^{c}\sum_{j=1}^{m}[(y_{ij}-\bar{y}_{i.})-(\bar{y}_{i.}-\bar{y}_{..})]^2 \\ &= \frac{1}{n(n-1)}\Big[\sum_{i=1}^{c}\sum_{j=1}^{m}(y_{ij}-\bar{y}_{i.})^2+\sum_{i=1}^{c}\sum_{j=1}^{m}(\bar{y}_{i.}-\bar{y}_{..})^2\Big] \\ &= \frac{1}{n(n-1)}[SSE+SSA]. \end{aligned} \tag{3.10}$$

But SSE and SSA are stochastically independent and have $\sigma_e^2\chi^2_{(n-c)}$ and $(\sigma_e^2+m\sigma_b^2)\chi^2_{(c-1)}$ distributions, respectively (Sahai and Ojeda, 2004, Theorems 2.3.2 and 2.3.3). Hence,

$$\begin{aligned} E\big(\widehat{var}_{LM}(\hat{\beta})\big) &= \frac{1}{n(n-1)}(E(SSE)+E(SSA)) \\ &= \frac{1}{n(n-1)}((n-c)\sigma_e^2+(c-1)(\sigma_e^2+m\sigma_b^2)) \\ var\big(\widehat{var}_{LM}(\hat{\beta})\big) &= \frac{1}{n^2(n-1)^2}(var(SSE)+var(SSA)) \\ &= \frac{2}{n^2(n-1)^2}\big[(n-c)\sigma_e^4+(c-1)(\sigma_e^2+m\sigma_b^2)^2\big] \end{aligned}$$

Therefore, the mean squared error for $\widehat{var}_{LM}(\hat{\beta})$ when σ_b^2=0, MSE_{LM0}, is derived as

$$\begin{aligned} MSE_{LM0} &= \big[E\big(\widehat{var}_{LM}(\hat{\beta})\big) - var_{LM}(\hat{\beta})\big]^2 + var\big(\widehat{var}_{LM}(\hat{\beta})\big) \\ &= \big[E\big(\frac{s^2}{n}\big) - \frac{\sigma_e^2}{n}\big]^2 + \frac{2\sigma_e^4}{n^2(n-1)} = \frac{2\sigma_e^4}{n^2(n-1)}. \end{aligned} \tag{3.11}$$

The mean squared error for $\widehat{var}_{LM}(\hat{\beta})$ when σ_b^2 >0, MSE_{LMG}, is derived as

$$\begin{aligned} MSE_{LMG} &= \Big[\frac{1}{n(n-1)}\big((n-c)\sigma_e^2 + (c-1)(\sigma_e^2 + m\sigma_b^2)\big) - \frac{1}{n}(\sigma_e^2 + m\sigma_b^2)\Big]^2 \\ &\quad + \frac{2}{n^2(n-1)^2}\big((n-c)\sigma_e^4 + (c-1)(\sigma_e^2 + m\sigma_b^2)^2\big) \\ &= \frac{1}{n^2}\Big[\bigg(\frac{n-c}{n-1}\bigg)\sigma_e^2 + \bigg(\frac{c-1}{n-1} - 1\bigg)(\sigma_e^2 + m\sigma_b^2)\Big]^2 \\ &\quad + \frac{2}{n^2(n-1)^2}\big((n-c)\sigma_e^4 + (c-1)(\sigma_e^2 + m\sigma_b^2)^2\big) \\ &= \frac{1}{n^2}\Big[\frac{n-c}{n-1}\sigma_e^2 - \bigg(\frac{n-c}{n-1}\bigg)(\sigma_e^2 + m\sigma_b^2)\Big]^2 \\ &\quad + \frac{2}{n^2(n-1)^2}\bigg((n-c)\sigma_e^4 + (c-1)(\sigma_e^2 + m\sigma_b^2)^2\bigg) \\ &= \frac{1}{n^2}\Big[\frac{m(n-c)}{n-1}\sigma_b^2\Big]^2 + \frac{2}{n^2(n-1)^2}\big((n-c)\sigma_e^4 \\ &\quad + (c-1)(\sigma_e^2 + m\sigma_b^2)^2\big). \end{aligned} \tag{3.12}$$

But $\sigma_e^2 = (1-\rho)\sigma_y^2$ and $\sigma_b^2 = \rho\sigma_y^2$, so that $\sigma_e^2 + m\sigma_b^2 = (1 + (m-1)\rho)\sigma_y^2$.

Therefore, substituting these terms into (3.12), we have

$$\begin{aligned} MSE_{LMG} &= \frac{1}{n^2(n-1)^2}\Big[m^2(n-c)^2\rho^2 + 2(n-c)(1-\rho)^2 \\ &\quad +2(c-1)(1+(m-1)\rho)^2\Big]\sigma_y^4 \end{aligned} \tag{3.13}$$

The Huber-White variance estimator of $\hat{\beta}$ is given by

$$\widehat{var}_{Hub}(\hat{\beta}) = \frac{1}{c(c-1)}\sum_{i=1}^{c}(\bar{y}_{i.} - \bar{y}_{..})^2. \tag{3.14}$$

But $\bar{y}_{i.}$ is normally distributed with mean 0 and variance $(\sigma_e^2 + m\sigma_b^2)/m$; hence, $\widehat{var}_{Hub}(\hat{\beta})$ has a $\frac{(\sigma_e^2+m\sigma_b^2)/m}{c(c-1)}\chi^2_{(c-1)}$ distribution. Therefore,

$$\begin{aligned} E\big(\widehat{var}_{Hub}(\hat{\beta})\big) &= (c-1)\frac{\frac{\sigma_e^2+m\sigma_b^2}{m}}{c(c-1)} = \frac{\sigma_e^2+m\sigma_b^2}{n} \\ var\big(\widehat{var}_{Hub}(\hat{\beta})\big) &= 2(c-1)\frac{(\sigma_e^2+m\sigma_b^2)^2}{m^2c^2(c-1)^2} = \frac{2(\sigma_e^2+m\sigma_b^2)^2}{n^2(c-1)}. \end{aligned} \tag{3.15}$$

Therefore, the mean squared error for $\widehat{var}_{Hub}(\hat{\beta})$, MSE_H, is derived as

$$\begin{aligned} MSE_H &= \big[E\big(\widehat{var}_{Hub}(\hat{\beta})\big) - var_{Hub}(\hat{\beta})\big]^2 + var\big(\widehat{var}_{Hub}(\hat{\beta})\big) \\ &= \frac{2(\sigma_e^2+m\sigma_b^2)^2}{n^2(c-1)} \\ &= \frac{2(1+(m-1)\rho)^2}{n^2(c-1)}\sigma_y^4. \end{aligned} \tag{3.16}$$

When σ_b^2=0, this reduces to MSE_{H0}, where

$$MSE_{H0} = \frac{2\sigma_e^4}{n^2(c-1)}. \tag{3.17}$$

Comparing MSE_{H0} to MSE_{LM0}, it is obvious that MSE_{LM0} is always less than MSE_{H0}. MSE_{LM0} is m times smaller than MSE_{H0} when n and c are large.

By manipulating (3.13) and (3.16), it is clear that $MSE_H < MSE_{LM}$ when

$$\begin{aligned}&\tfrac{m^2c^2(m-1)^2\rho^2+2c(m-1)(1-\rho)^2+2(c-1)(1+(m-1)\rho)^2}{(mc-1)^2}\\&-\tfrac{2(1+(m-1)\rho)^2}{(c-1)}>0.\end{aligned} \tag{3.18}$$

The left hand side of (3.18) is a quadratic and can be rewritten as:

$$\begin{aligned}&m^2c^2(m-1)^2(c-1)\rho^2+2c(c-1)(m-1)(1-\rho)^2\\&+2(c-1)^2(1+(m-1)\rho)^2-2(mc-1)^2(1+(m-1)\rho)^2>0.\end{aligned} \tag{3.19}$$

Simplifying this inequality, we have

$$\begin{aligned}&2(c-1)^2+2c(m-1)(c-1)-2(mc-1)^2+(4(m-1)(c-1)^2\\&-4c(m-1)(c-1)-4(m-1)(mc-1)^2)\rho+(2c(m-1)(c-1)\\&+2(m-1)^2(c-1)^2+c^2m^2(m-1)^2(c-1)\\&-2(m-1)^2(mc-1)^2)\rho^2>0\end{aligned}$$

Expanding the constant term and the coefficients of ρ and ρ^2, this inequality simplifies to:

$$\begin{aligned}&-2c(m-1)(mc-1)-4c(m-1)(mc-2m+1)\rho\\&+c(m-1)(2(2m^2-4m+1)-mc(3m^2-3m-2)+m^2c^2(m-1))\rho^2>0\end{aligned}$$

Dividing by $c(m-1)$, we get

$$\begin{aligned}&-2(mc-1)-4(mc-2m+1)\rho+(2(2m^2-4m+1)\\&-mc(3m^2-3m-2)+m^2c^2(m-1))\rho^2>0\end{aligned} \tag{3.20}$$

Setting the left hand side of (3.20) to zero, we obtain the following roots of this quadratic equation:

$$\rho_1 = \tfrac{2(mc-2m+1)+\sqrt{4(mc-2m+1)^2+2(mc-1)(m^2c^2(m-1)+2(2m^2-4m+1)-mc(3m^2-3m-2))}}{m^2c^2(m-1)+2(2m^2-4m+1)-mc(3m^2-3m-2)}$$

$$\rho_2 = \frac{2(mc-2m+1)-\sqrt{4(mc-2m+1)^2+2(mc-1)(m^2c^2(m-1)+2(2m^2-4m+1)-mc(3m^2-3m-2))}}{m^2c^2(m-1)+2(2m^2-4m+1)-mc(3m^2-3m-2)}$$

But $2m^2-4m+1$ is positive if and only if $m \geq \frac{2+\sqrt{2}}{2}$ and $3m^2-3m-2$ is positive if and only if $m \geq \frac{32+\sqrt{33}}{6}$. It follows that $2m^2-4m+1$ and $3m^2-3m-2$ are positive if and only if $m \geq 2$. Therefore, $\rho_1 \geq 0$ and $\rho_2 \leq 0$ whenever $m \geq 2$. In practice, m would almost always be greater than or equal 2. It follows that the Huber-White variance estimator has lower MSE than the LM variance estimator when $\rho \geq \rho_1$, since ρ would almost always be greater than or equal to 0.

Figure 3.1 shows the values of m, c and ρ such that the Huber-White variance estimator (MSE_H) has lower mean squared error than the LM variance estimator (MSE_{LM}) for different values of ρ, m and c. The Figure shows that for $\rho = 0$, MSE_H is larger than MSE_{LM} for all values of m and c, except when m=1 for all values of c. When m=1, the two estimators have equal mean squared error for all ρ and c; this is clear from (3.21).

For ρ=0.01, the Huber-White variance estimator does better than the LM variance estimator for values of $m > 20$ with $c > 17$. In case of ρ=0.025, the region such that the Huber-White variance estimator is better than the LM variance estimator becomes larger. $MSE_H < MSE_{LM}$ for values of $m > 8$ with $c > 8$.

For $\rho = 0.05$, the Huber-White variance estimator has lower mean squared

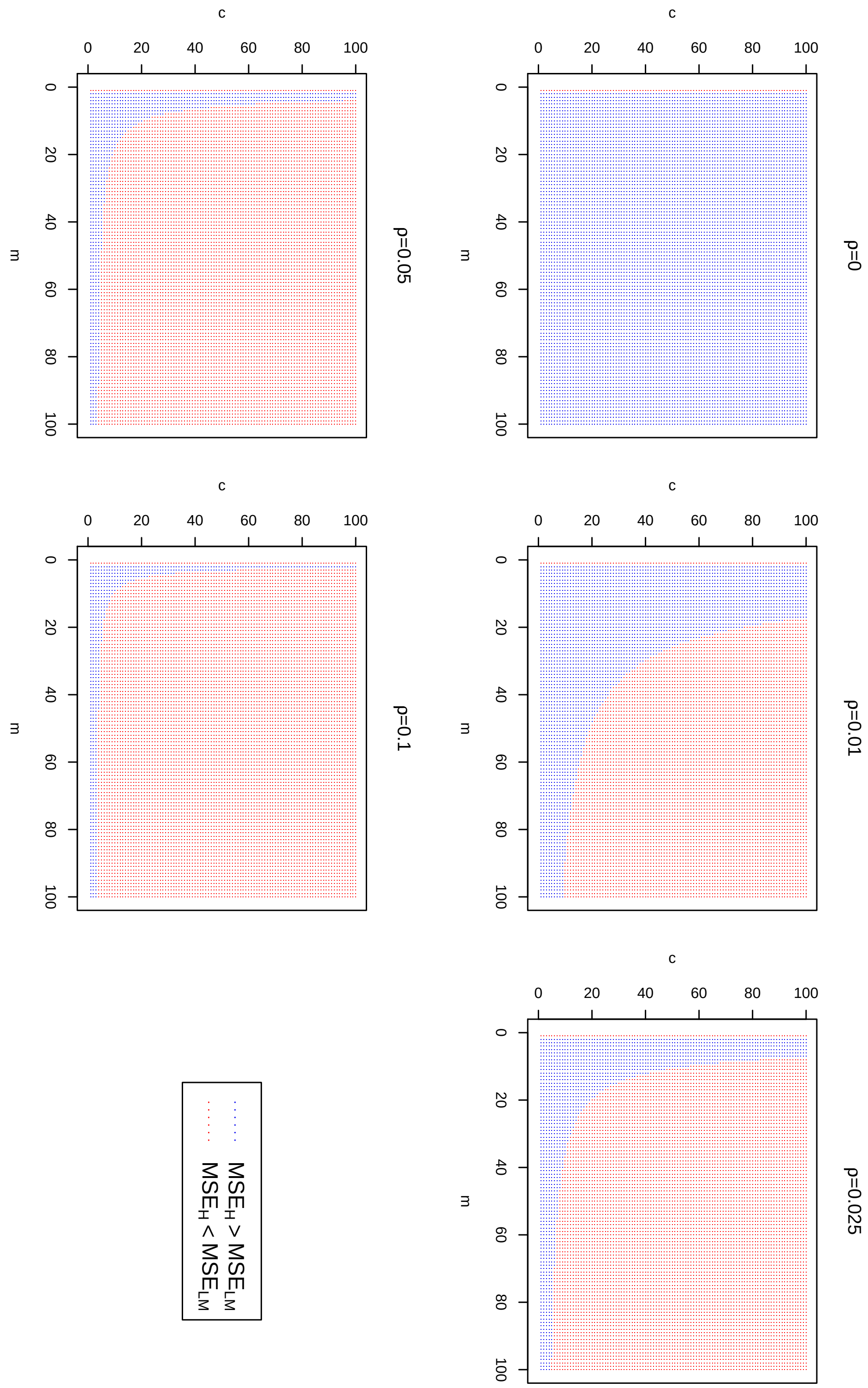


Figure 3.1: The values of m and c such that $MSE_H < MSE_{LM}$

error than the LM variance estimator for values of $c \geq 4$ with $m \geq 4$. Finally, in case of $\rho = 0.1$ the region such that the Huber-White variance estimator has lower mean squared error than the LM variance estimator is bounded by values of $c \geq 3$ with $m \geq 2$.

The Figure shows that as the value of ρ increases the region such that Huber-White variance estimator has lower mean squared error than LM variance estimator increases.

For large m and c, Equation (3.20) can be rewritten as:

$$-2c - 4c\rho + (4m - 3m^2c + m^2c^2)\rho^2 > 0 \tag{3.21}$$

For large m and c, the quadratic term in (3.21) is dominated by m^2c^2; therefore, (3.21) reduces to

$$-2 - 4\rho + m^2c\rho^2 = 0 \tag{3.22}$$

Setting the left hand side of (3.22) to 0, we obtain the following roots:

$$\rho_1 = \frac{2 + \sqrt{2(2 + m^2c)}}{m^2c},$$

$$\rho_2 = \frac{2 - \sqrt{2(2 + m^2c)}}{m^2c}$$

It is clear that $\rho_1 > 0$ and $\rho_2 < 0$. Hence Huber-White variance estimator does better than LM variance estimator whenever $\rho \geq \rho_1$. We can further approximate ρ_1 to be

$$\rho_1 \approx \frac{2}{m^2c} + \sqrt{\frac{2}{m^2c}} \approx \sqrt{\frac{2}{m^2c}}$$

So the Huber-White variance estimator generally does better than the LM variance estimator when $\rho \geq \sqrt{\frac{2}{m^2 c}}$. This lower bound on ρ tends to 0 as c and particularly m increase.

3.4 Simulation Study

3.4.1 Design of Simulation Study

A simulation study was conducted to compare the adaptive and non-adaptive methods for estimating $var(\hat{\beta})$. Data were generated from the normal distribution, with $m_i = m$ and an intercept only model (2.3). The values of ρ, m and c were varied. 1000 samples were generated in each case. The values of σ_b^2 and σ_e^2 were set to $\frac{\rho}{1-\rho}$ and 1 respectively, to ensure that the intraclass correlation was ρ.

For each sample the estimated regression coefficients $\hat{\beta}$ and the estimators of $var(\hat{\beta})$ were calculated for the LMM and LM models using the *lme*4 and *lm* packages (Pinheiro and Bates, 2000) in the R statistical environment (R Development Core Team, 2007). The true variance of $\hat{\beta}$ was determined by calculating the variance over all 1000 simulations.

The hypothesis $H_0 : \sigma_b^2 = 0$ was tested as described in Subsection 2.3.2. The two adaptive strategies ADM and ADH are defined by (3.1) and (3.2).

90% confidence intervals were calculated for the LMM method using the method of Faes et al. (2009) as described in Subsection 2.2.3. Huber-White confidence intervals were calculated as discussed in Subsection 2.2.5, and the adaptive confidence intervals were calculated as discussed in Section 3.2. The hope is that the adaptive procedures give shorter confidence intervals as they will use the LM when H_0 is not rejected and for small sample sizes these cases still have $\hat{\rho}$ away from zero. As the sample size increases, H_0 will only be not rejected when $\hat{\rho}$ is close to zero.

Two methods were evaluated for testing $H_0 : \sigma_b^2 = 0$: a t-test (as described in Subsection 2.3.1) and the restricted likelihood ratio test (as described in Subsection 2.3.2).

The values of ρ, m and c were varied. The parameter ρ was varied over a range of values of 0, 0.01, 0.025, 0.05 and 0.1; c was varied over 2, 5, 10 and 25; and m was varied over 2, 5, 10, 15, 25 and 50. So the design effects varied from 1 to 5.9.

The estimated regression coefficients $\hat{\beta}$ and the estimators of $var(\hat{\beta})$ were calculated for the LMM and LM models using the $lme4$ and lm packages (Pinheiro and Bates, 2000) in the R statistical environment (R Development Core Team, 2007). The t-test for $H_0 : \sigma_b^2 = 0$ was applied by coding Equation (2.26) in R.

3.4.2 Simulation Results on Testing $H_0 : \sigma_b^2 = 0$

This subsection will summarize the performance of the t-test and the RLRT for testing $H_0 : \sigma_b^2 = 0$. Results for the intercept-only model with equal-sized PSUs will be used, with ρ=0 and 0.05.

Table 3.1: Non-coverage of testing $H_0 : \sigma_b^2 = 0$ using RLRT and t-test with ρ=0 and $\rho = 0.05$.

PSUs	Obser vations	P(Reject $H_0 : \sigma_b^2 = 0$) when			
		$\rho = 0$		$\rho = 0.05$	
c	m	t-test (%)	RLRT (%)	t-test (%)	RLRT (%)
2	2	0.0	10.5	0.0	10.2
2	5	0.0	5.8	0.0	8.6
2	10	0.0	5.2	0.0	11.5
2	15	0.0	5.0	0.0	11.7
2	25	0.0	3.8	0.0	19.4
2	50	0.0	4.5	0.0	29.2
5	2	0.9	10.5	1.0	10.9
5	5	0.0	8.0	0.1	13.1
5	10	0.0	7.5	0.3	22.7
5	15	0.0	6.4	1.3	27.8
5	25	0.0	7.4	4.4	43.6
5	50	0.0	7.2	16.8	62.7
10	2	11.1	10.1	14.7	11.2
10	5	11.3	8.5	23.9	19.7
10	10	11.1	6.3	37.0	32.2
10	15	10.7	8.6	52.8	47.4
10	25	9.3	9.7	71.6	66.1
10	50	9.7	6.7	90.7	89.3
25	2	14.4	10.9	20.9	14.4
25	5	12.9	9.4	37.9	30.0
25	10	10.2	7.6	58.2	57.4
25	15	12.3	10.9	77.8	73.0
25	25	10.2	8.7	93.3	92.2
25	50	16.0	8.0	99.9	99.3

Table 3.1 shows the probability of rejecting $H_0 : \sigma_b^2 = 0$ based on the t-

test using the derived standard error defined by Equation (2.27) as well as the rejection probability based on the restricted likelihood ratio test using two different values of ρ of 0 and 0.05. We expect that the probability of rejecting H_0 should be close to 0.1 when ρ=0, while the probability of rejecting H_0 should be as high as possible when $\rho > 0$.

The t-test performed very poorly as the proportions of samples where $H_0 : \sigma_b^2 = 0$ is rejected were very small, in general, when there where small number of PSUs (5 or less) for both values of ρ. Proportions of samples where $H_0 : \sigma_b^2 = 0$ is rejected were close to the nominal rate when ρ=0 and unacceptably high when ρ=0.05. For example: This is an important finding, because the t-test is the method used by the SPSS statistical software.

The RLRT proportions of samples where $H_0 : \sigma_b^2 = 0$ was rejected were closer to the nominal rate when ρ=0. For ρ=0.05, the proportions of samples where H_0 is rejected were much better than the t-test when there small numbers of sample PSUs. For example:

- When c=2 and m=5, the proportion of sample where H_0 is rejected were 5.8% when ρ=0 and 8.6% when ρ=0.05.
- When c=5 and m=15, the proportion of sample where H_0 is rejected were 6.4% when ρ=0 and 27.8% when ρ=0.05.
- When c=10 and m=25, the proportion of sample where H_0 is rejected

were 9.7% when ρ=0 and 66.1% when ρ=0.05.

- When c=25 and m=2, the proportion of sample where H_0 is rejected were 10.9% when ρ=0 and 14.4% when ρ=0.05.

Therefore, the RLRT method used in this thesis.

3.4.3 Simulation Results on Adaptive Confidence Intervals for β for Balanced Data

A simulation study based on equal sized PSUs, $m_i = m$, and an intercept only model was conducted to compare the adaptive and non-adaptive methods for estimating $var(\hat{\beta})$. Data were generated from the intercept-only model (2.3). The values of ρ, m and c were varied. 1000 samples were generated in each case. In this study we used the parametric bootstrap to estimate $V(T)$ because the scale parameter δ relies on $V(T)$ (see Equation 2.16) and Faes et al. (2009) did not specify how $V(T)$ can be estimated.

To apply the parametric bootstrap method to estimate $var(T)$, 100 samples were generated from the intercept-only model (2.3) with variances $\hat{\sigma}_b^2$ and $\hat{\sigma}_e^2$. For each sample, we estimated β and $var(\hat{\beta})$ to find the value of $T = \frac{\hat{\beta}}{\sqrt{\widehat{var}(\hat{\beta})}}$. The variance of the 100 values of T was calculated and used to estimate $V(T)$.

Another way to estimate $var(T)$ is to estimate $\widehat{var}[\widehat{var}(\hat{\beta})]$, and then substitute into (2.16), but Faes et al. (2009) also did not specify how to estimate this parameter, therefore we have tried to do that using the parametric bootstrap. The same procedure above is used, but now we estimated $var(\hat{\beta})$ from the fitted model and then calculated the variance of the 100 estimated values of $var(\hat{\beta})$. Then $\widehat{var}(T)$ was calculated by coding Equation (2.16) in R. The LMM non-coverage rates were very small, specially for small number of sample PSUs (5 or less).

In the end the method of estimating $V(T)$ by calculating the variance of the 100 estimated values of T performed better than the method uses $\widehat{var}[\widehat{var}(\hat{\beta})]$ to estimate $var(T)$. Therefore, the first was used in the simulation studies in Chapters 3-5 in the balanced design.

The hypothesis $H_0 : \sigma_b^2 = 0$ was tested as described in Subsection 2.3.2 using the restricted likelihood ratio test defined in Equation (2.30). The two adaptive strategies ADM and ADH were as defined in Section 3.2. 90% confidence intervals were calculated for the LMM method using the method of Faes et al. (2009) as described in Subsection 2.2.3. Huber-White confidence intervals were calculated as discussed in Subsection 2.2.5, and the adaptive confidence intervals were calculated as discussed in Section 3.2.

Tables 3.2 - 3.4 show the ratio of the mean estimated variance of $\hat{\beta}$, $E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$, using the four strategies of estimation (ADM, ADH, LMM and Huber) with values of ρ of 0, 0.025 and 0.1. In all tables we used $\beta = 0$ and significance level $\alpha = 0.1$ for testing $\sigma_b^2 = 0$. The tables show the non-coverage rates of 90% confidence intervals of β and the average lengths of these confidence intervals. The proportion of samples where $H_0 : \sigma_b^2 = 0$ was rejected are also shown. Results on non-coverage and 90% confidence intervals average length are shown in graphical form in Figures 3.2 - 3.4. In the graphs we also include the LM strategy of estimation so that the effect of completely ignoring the clustered nature of the data can be examined.

The variance estimators were generally approximately unbiased, as all ratios are approximately 1. There were some exceptions. The first was the variance estimator using the LMM strategy; it tended to be biased when there were 10 or less sample PSUs with approximately all numbers of observations per PSU for ρ=0. For 0.025, it tended to biased when there were 2 sample PSUs with 25 or less observations per PSU and when there were 5 sample PSUs with 2 and 15 observations per PSU. It tended to be biased when there were 2 sample PSUs with 5 observations or less per PSU in case of $\rho = 0.1$, as well. In case of ρ=0, it also tended to be biased when there were 5 PSUs with all numbers of observations per PSU. The other exception was the ADM and the ADH variance estimators, they tended to be biased when there were 2 sample PSUs with 2, 5, 15 and 50 observations per PSU when ρ=0.

Non-coverage rates for confidence intervals for β were close to the nominal rate of 10% when $\rho = 0$ for all methods.

Table 3.2: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT with ρ=0.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.290	1.290	1.553	1.183	8.4	8.4	11.6	7.8	11.2	5.676	3.059	6.330	5.429
2	5	1.283	1.283	1.517	1.042	9.2	9.0	9.1	10.3	6.3	1.243	1.564	1.327	3.142
2	10	1.259	1.259	1.523	1.055	8.9	8.9	9.2	10.7	5.1	0.862	1.053	0.952	2.270
2	15	1.179	1.179	1.412	0.927	9.8	9.8	10.1	10.4	3.8	0.683	0.794	0.749	1.772
2	25	1.165	1.165	1.419	0.976	10.8	10.8	11.3	8.9	4.2	0.528	0.621	0.584	1.442
2	50	1.318	1.318	1.581	1.087	7.9	7.9	9.4	9.5	5.5	0.389	0.477	0.426	1.015
5	2	1.074	1.074	1.183	0.986	9.4	9.2	10.2	9.4	9.9	1.173	1.181	1.190	1.255
5	5	1.163	1.163	1.288	1.057	9.3	9.3	10.0	8.5	7.6	0.716	0.721	0.732	0.801
5	10	1.152	1.152	1.282	1.044	8.0	8.0	8.8	9.5	6.7	0.500	0.505	0.513	0.569
5	15	1.133	1.133	1.259	1.017	9.2	9.2	10.1	10.1	7.9	0.412	0.417	0.423	0.465
5	25	1.124	1.124	1.234	0.999	9.4	9.4	10.0	10.4	7.9	0.317	0.321	0.324	0.360
5	50	1.157	1.157	1.294	1.059	8.0	7.9	8.5	8.7	7.0	0.224	0.226	0.232	0.258
10	2	1.036	1.036	1.103	0.976	10.9	10.8	10.9	11.0	11.4	0.788	0.787	0.793	0.794
10	5	1.148	1.148	1.221	1.071	8.1	8.1	9.1	9.3	8.0	0.489	0.490	0.492	0.503
10	10	1.033	1.033	1.095	0.972	10.8	11.0	10.9	10.6	8.5	0.347	0.348	0.348	0.361
10	15	1.205	1.205	1.282	1.116	7.2	7.1	8.9	9.1	8.6	0.282	0.283	0.285	0.292
10	25	1.203	1.203	1.268	1.103	7.3	7.3	7.9	9.4	8.3	0.219	0.219	0.219	0.225
10	50	1.137	1.137	1.209	1.045	9.7	9.7	10.0	10.7	8.0	0.154	0.154	0.155	0.159
25	2	0.950	0.950	0.994	0.920	11.4	11.4	11.7	12.3	10.1	0.483	0.483	0.483	0.482
25	5	0.956	0.956	0.963	0.913	10.6	10.7	11.9	11.1	8.1	0.303	0.302	0.298	0.302
25	10	1.038	1.038	1.051	0.984	10.6	10.6	11.1	11.5	8.7	0.214	0.214	0.212	0.214
25	15	1.013	1.013	1.025	0.961	9.5	9.4	10.8	10.9	6.8	0.174	0.173	0.171	0.173
25	25	1.123	1.123	1.137	1.069	8.4	8.4	8.8	9.0	9.1	0.135	0.135	0.134	0.136
25	50	1.045	1.046	1.068	0.971	9.4	9.4	10.3	10.3	8.8	0.096	0.096	0.095	0.095

Table 3.3: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT with ρ=0.025.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.189	1.189	1.440	1.074	8.2	8.2	12.0	8.9	9.9	4.934	2.977	5.552	5.299
2	5	1.182	1.182	1.415	1.012	11.4	10.9	11.1	12.6	7.5	1.330	1.693	1.437	3.330
2	10	1.139	1.139	1.356	0.985	11.4	11.4	12.9	11.6	7.5	0.915	1.211	1.009	2.413
2	15	1.005	1.005	1.227	0.947	15.5	15.5	13.9	8.5	8.6	0.780	1.034	0.880	2.167
2	25	1.074	1.074	1.273	1.058	18.6	18.6	16.1	8.7	14.4	0.710	1.054	0.800	1.911
2	50	0.886	0.886	1.056	0.911	23.4	23.4	19.9	9.3	16.9	0.548	0.835	0.623	1.469
5	2	1.123	1.123	1.245	1.061	9.6	9.6	8.5	9.2	11.9	1.224	1.230	1.254	1.323
5	5	0.986	0.986	1.102	0.941	10.8	10.9	11.2	10.8	11.2	0.751	0.759	0.780	0.854
5	10	0.967	0.967	1.093	0.975	12.4	12.1	12.1	8.7	14.7	0.543	0.551	0.572	0.644
5	15	1.080	1.080	1.233	1.118	11.3	11.3	10.1	8.8	17.1	0.457	0.465	0.487	0.549
5	25	0.895	0.895	1.007	0.945	15.8	15.8	13.6	10.7	24.7	0.383	0.392	0.409	0.462
5	50	0.812	0.812	0.888	0.861	19.3	19.3	16.3	11.6	40.2	0.320	0.330	0.339	0.376
10	2	1.044	1.044	1.125	1.008	9.0	9.1	8.2	9.5	11.9	0.809	0.812	0.819	0.830
10	5	0.953	0.953	1.016	0.948	11.6	11.6	12.0	10.2	12.5	0.507	0.507	0.514	0.537
10	10	1.045	1.045	1.121	1.062	11.0	10.9	11.5	9.6	18.6	0.370	0.371	0.378	0.398
10	15	0.935	0.935	1.008	0.975	13.1	13.0	11.4	10.5	25.8	0.318	0.318	0.327	0.344
10	25	1.002	1.002	1.072	1.061	11.6	11.5	11.4	8.6	40.5	0.269	0.270	0.279	0.292
10	50	0.995	0.995	1.029	1.030	14.1	13.3	12.8	10.3	67.1	0.231	0.232	0.237	0.243
25	2	1.018	1.018	1.066	0.997	9.6	9.7	9.9	9.7	12.7	0.489	0.489	0.490	0.491
25	5	1.007	1.007	1.022	1.017	10.5	10.6	10.7	10.2	17.2	0.316	0.314	0.314	0.323
25	10	0.991	0.991	1.008	1.034	11.7	11.6	12.1	10.1	29.0	0.232	0.231	0.232	0.242
25	15	0.990	0.990	1.009	1.038	11.3	11.3	11.0	9.6	42.3	0.199	0.197	0.200	0.206
25	25	1.049	1.049	1.063	1.087	10.8	11.1	10.8	9.7	66.0	0.171	0.169	0.172	0.175
25	50	0.947	0.948	0.953	0.954	11.0	10.7	10.9	10.7	92.7	0.144	0.144	0.145	0.145

Table 3.4: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT with ρ=0.1.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.096	1.096	1.322	1.016	9.4	9.4	13.8	11.0	10.2	5.112	3.145	5.791	5.538
2	5	1.102	1.102	1.335	1.103	15.7	15.3	13.8	9.4	13.4	1.607	2.238	1.775	4.127
2	10	0.887	0.887	1.055	0.900	20.3	20.2	17.7	10.0	15.2	1.119	1.748	1.266	3.141
2	15	0.883	0.883	1.037	0.929	25.5	25.5	19.5	9.4	18.7	1.067	1.711	1.216	2.930
2	25	0.949	0.949	1.058	0.996	28.7	28.6	21.9	7.1	29.1	1.130	1.920	1.249	2.812
2	50	0.895	0.895	0.945	0.920	29.1	28.6	23.9	11.2	45.6	1.281	2.126	1.353	2.622
5	2	1.063	1.063	1.189	1.062	11.0	10.9	10.5	9.7	16.6	1.294	1.310	1.342	1.434
5	5	0.929	0.929	1.042	0.962	13.4	13.1	13.2	10.6	20.5	0.837	0.851	0.882	0.984
5	10	1.010	1.010	1.111	1.070	14.0	13.7	12.9	10.1	37.7	0.711	0.727	0.755	0.828
5	15	1.043	1.043	1.111	1.095	15.2	15.0	11.8	9.0	53.3	0.692	0.709	0.726	0.783
5	25	0.966	0.966	0.999	0.993	15.4	14.9	13.5	10.4	67.1	0.643	0.652	0.662	0.694
5	50	0.959	0.959	0.968	0.967	12.6	11.2	11.9	9.8	85.2	0.644	0.632	0.650	0.646
10	2	0.978	0.978	1.060	0.979	9.8	9.9	9.8	9.5	16.2	0.838	0.841	0.856	0.874
10	5	0.983	0.983	1.051	1.043	12.3	12.1	12.4	10.0	33.6	0.584	0.586	0.601	0.634
10	10	0.974	0.974	1.025	1.028	13.3	13.0	11.6	8.8	57.3	0.485	0.487	0.501	0.519
10	15	0.984	0.984	1.011	1.011	12.2	11.9	11.8	10.5	73.7	0.455	0.453	0.465	0.470
10	25	1.009	1.009	1.017	1.018	11.3	11.1	10.5	9.8	89.4	0.436	0.434	0.439	0.440
10	50	0.993	0.993	0.994	0.994	10.1	9.6	10.0	9.3	98.4	0.415	0.408	0.416	0.408
25	2	0.971	0.971	1.021	0.983	10.5	10.3	10.7	10.2	21.9	0.522	0.523	0.526	0.533
25	5	0.985	0.985	1.001	1.027	12.1	11.9	12.3	10.1	53.5	0.369	0.366	0.371	0.380
25	10	0.974	0.974	0.979	0.989	11.2	11.4	10.8	10.0	86.5	0.308	0.306	0.309	0.310
25	15	0.982	0.982	0.983	0.985	9.9	10.0	9.9	9.9	96.3	0.288	0.285	0.288	0.285
25	25	0.974	0.974	0.974	0.974	9.8	9.4	9.8	9.4	99.9	0.264	0.263	0.264	0.263
25	50	1.054	1.054	1.054	1.054	9.0	8.7	9.0	8.7	100.0	0.244	0.244	0.244	0.244

Figure 3.2: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0

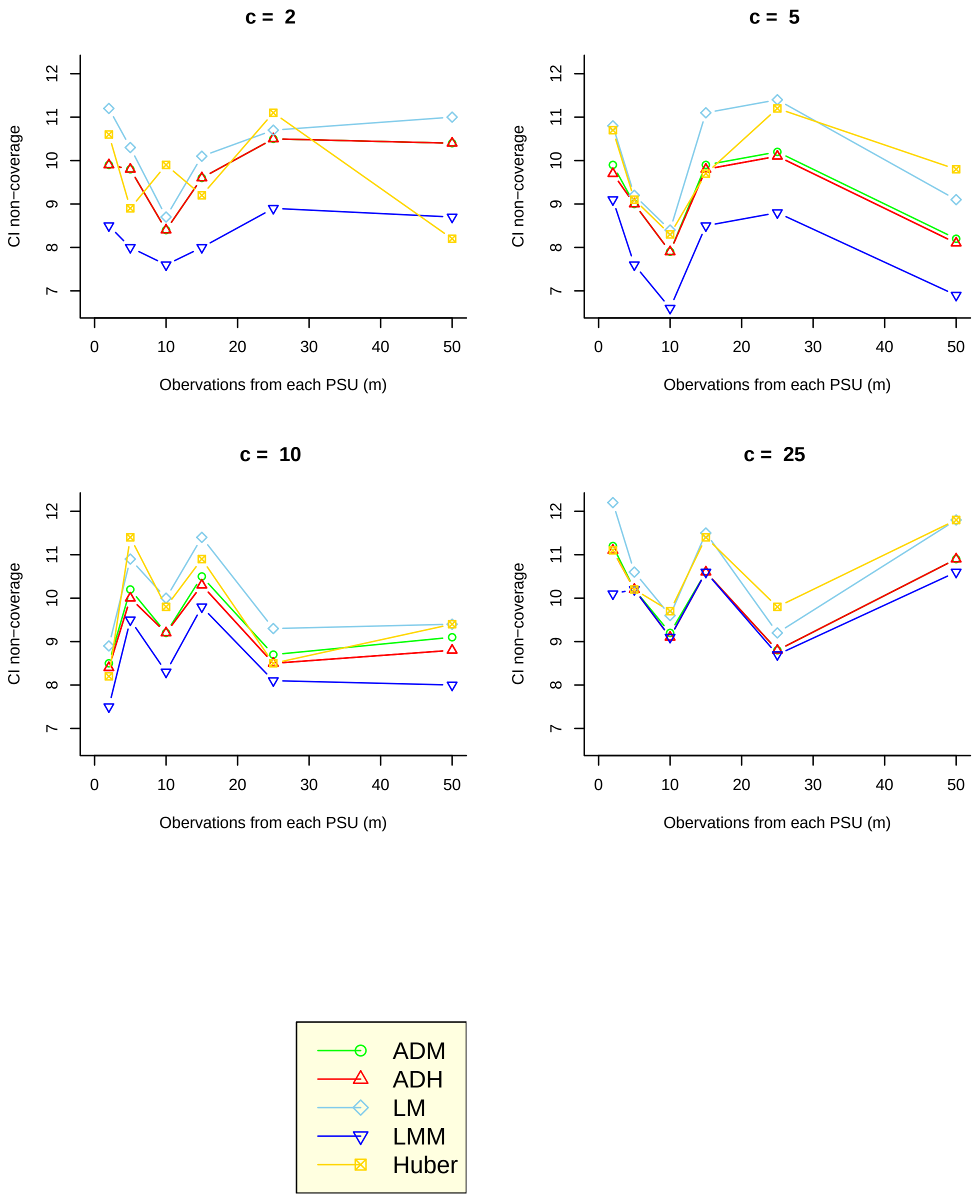


Figure 3.3: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.025

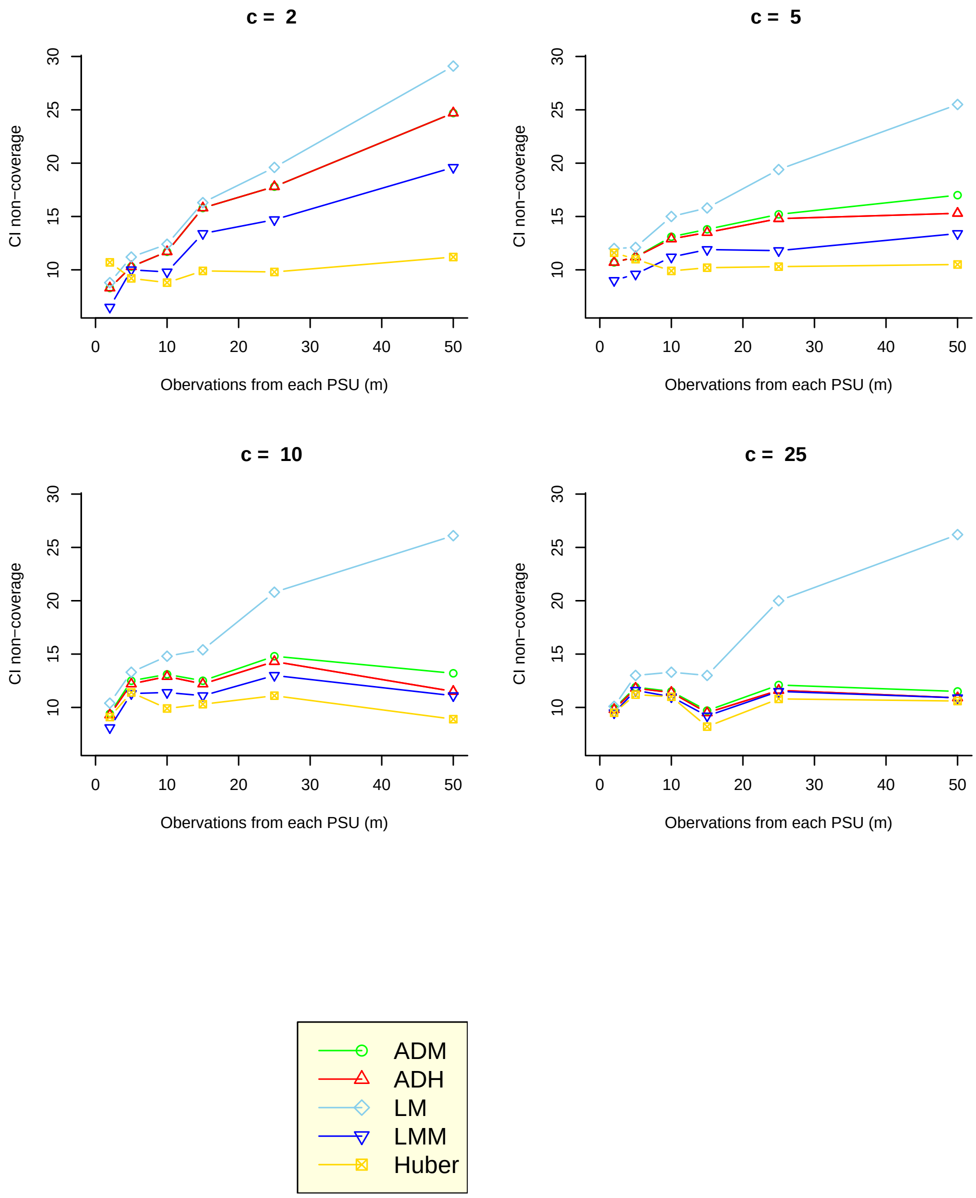


Figure 3.4: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.1

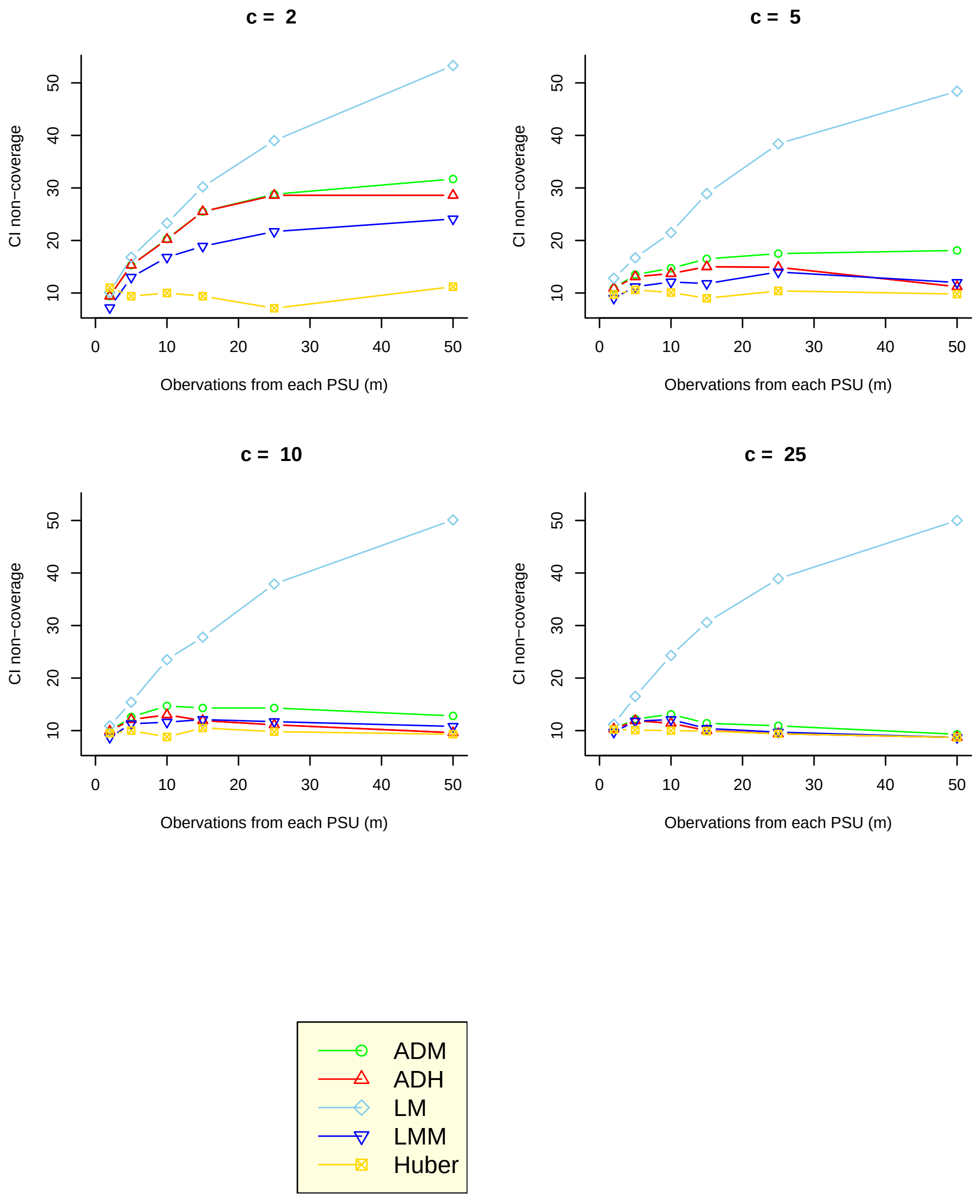


Figure 3.5: Confidence interval average lengths using different variance estimation methods and for various values of m and c, ρ=0

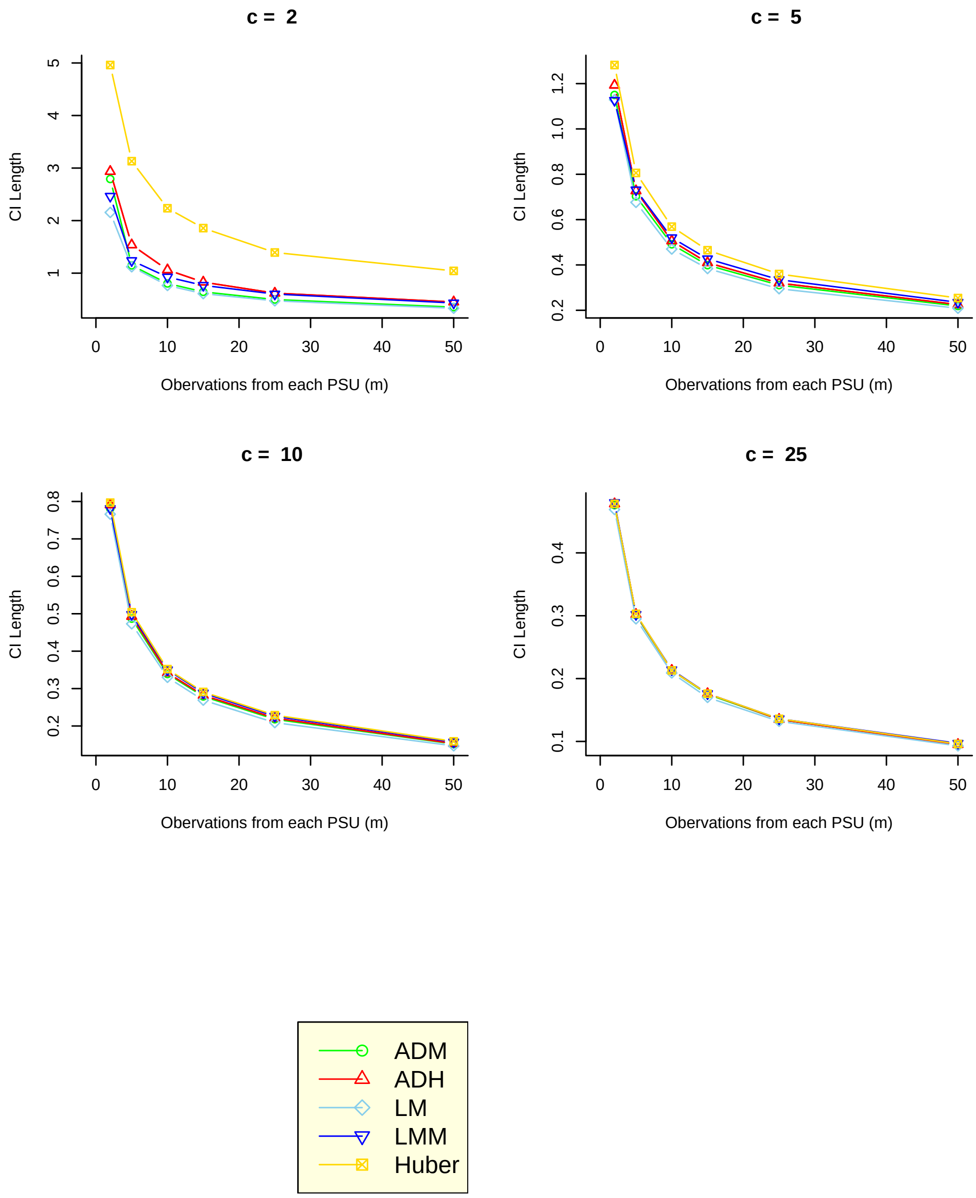


Figure 3.6: Confidence interval average lengths using different variance estimation methods and for various values of m and c, ρ=0.025

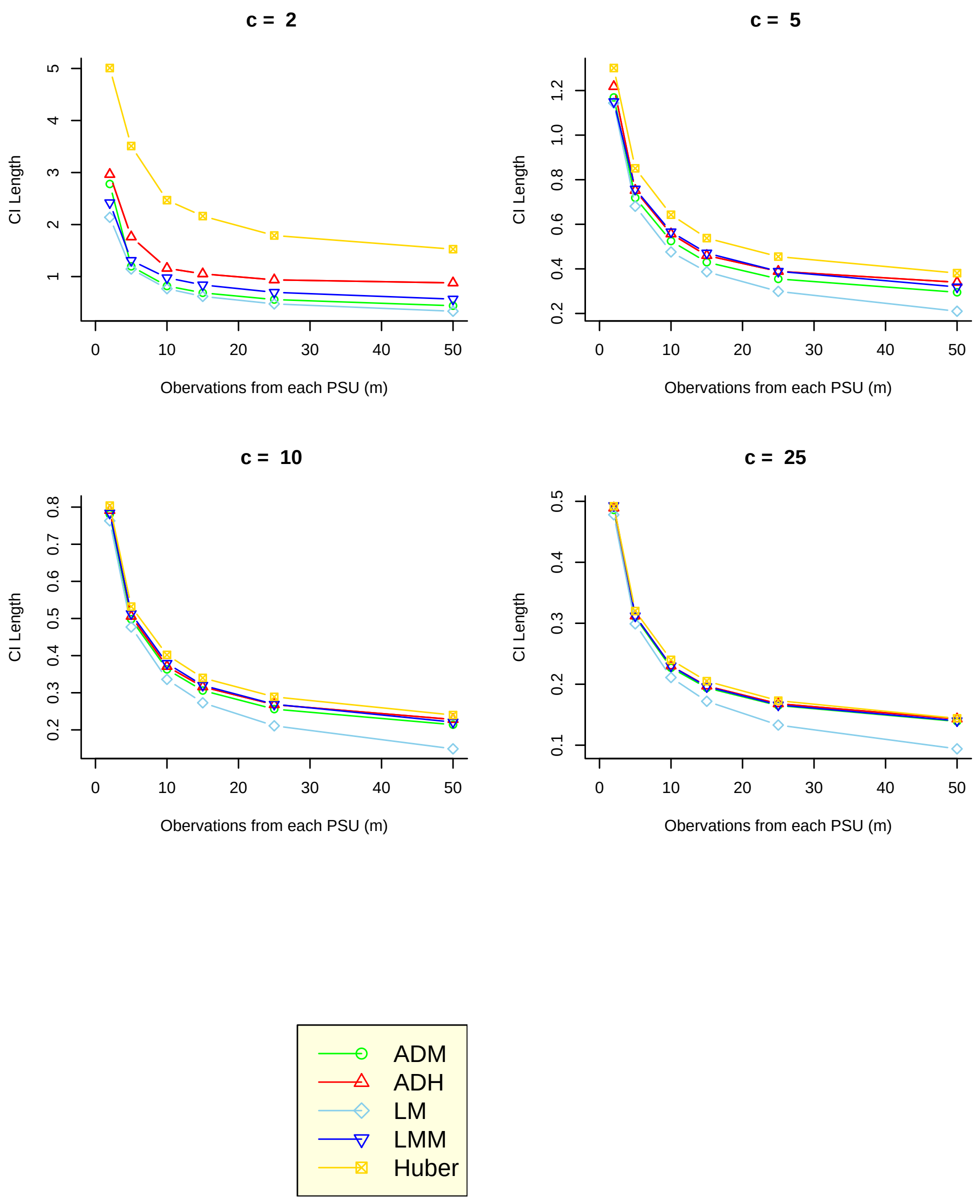
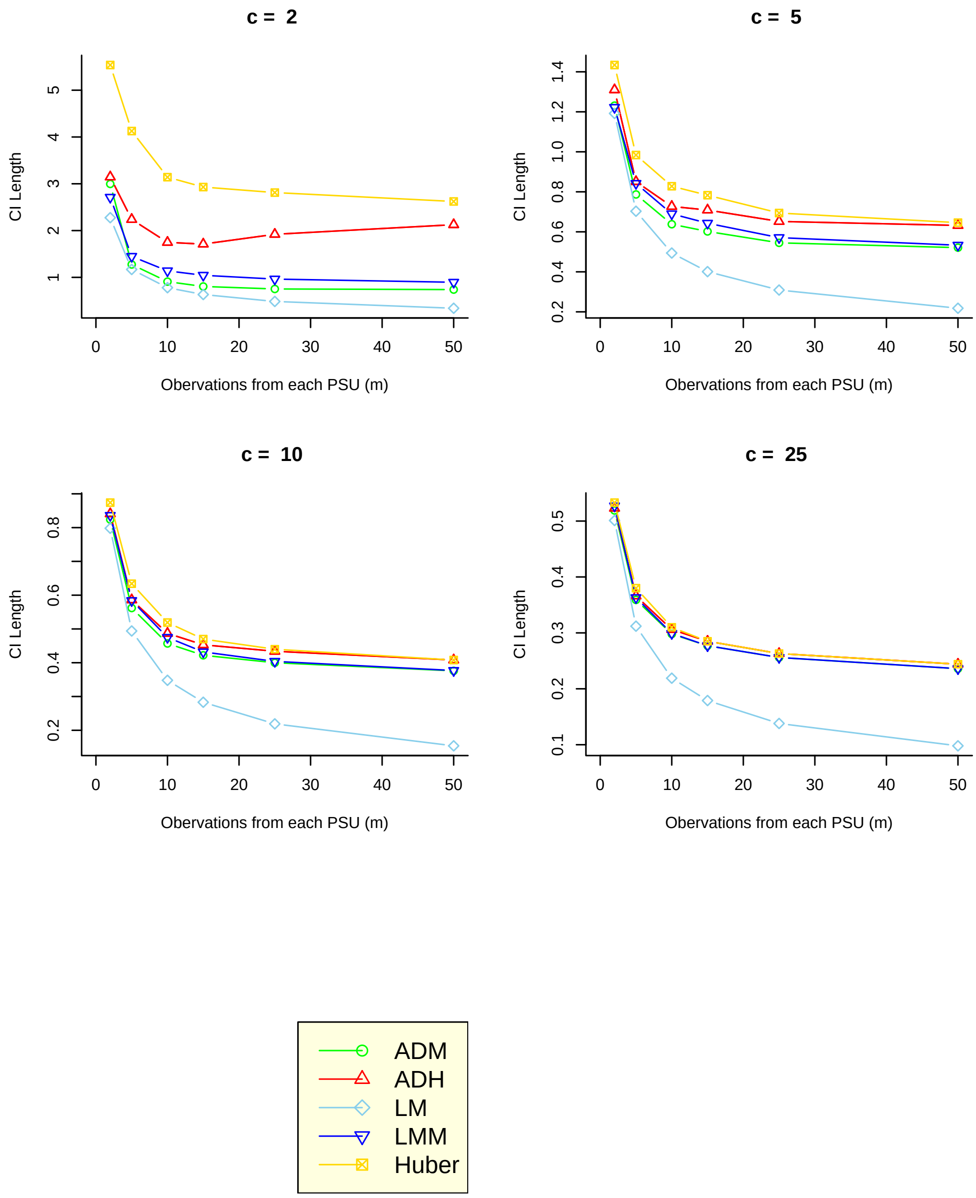


Figure 3.7: Confidence interval average lengths using different variance estimation methods and for various values of m and c, ρ=0.1



For $\rho \neq 0$, Huber non-coverage was close to 10% in all cases. For $\rho \neq 0$, non-coverage rates for confidence intervals for β were generally close to the nominal rate of 10% for all other methods of estimation. There were some exceptions. The first was the non-coverage rates for the LMM, ADM and ADH strategies; they tended to be much higher when there were small number of sample PSUs (10 or less) with 10 or more observations per PSU when ρ=0.025, in general. They also tended to be high when ρ=0.1, when there were 2 and 5 sample PSUs with approximately 5 or more observations per PSU. The ADM and ADH non-coverage rates tended to be high when there were 10 sample PSUs with 10 or more observations per PSU. This may be because of the difficulty in determining the appropriate degrees of freedom in the LMM case.

For ρ =0.1, the LMM non-coverage rates were high when c was small (10 or less) and m was large (5 or more), in general.

The ADH average lengths of confidence intervals for β were almost always shorter than the Huber average lengths of confidence intervals for β. When there were 2 sample PSUs it was very clear that ADH average lengths of confidence intervals for β were much shorter than Huber average lengths of confidence intervals for β, with orders 40-60% when ρ=0 and 0.025, and 30-70% when ρ=0.1. When there were 5 sample PSUs, the average lengths for the ADH were shorter with order of 15-25% for ρ=0 and 0.025, and 30-70%

when there were 25 or less observations per PSU when ρ=0.1. When there were 10 sample PSUs, the ADH average lengths were shorter for ρ=0.025 when there were 10-25 observations per PSU with order of about 15% and when there were 5 and 10 observations per PSU for ρ=0.1 of order 10-15%. There were no clear difference when there were more than 10 observations per PSU.

Figure 3.2 shows that LM non-coverage was close to 10% when $\rho = 0$. It was very high otherwise as shown by Figures 3.3 and 3.4. Hence, the use of LM without at least checking $H_0 : \sigma_b^2 = 0$ is not a viable strategy.

Figures 3.5 - 3.7 show the average lengths of confidence intervals for β using the LM strategy were the shortest, however this strategy is not viable because of its high non-coverage when $\rho \neq 0$. The Huber based approach gave the widest intervals in general. The ADM average lengths of confidence intervals for β were almost always shorter than the LMM average lengths of confidence intervals for β. When there were 2 sample PSUs it was very clear that ADM average lengths of confidence intervals for β were much shorter than LMM average lengths of confidence intervals for β, with orders 7-15% when ρ=0 and 0.025, and 10-20% when ρ=0.1. There were no clear difference otherwise. For example:

- in case of c=2 and m=2 and ρ=0, ADM and ADH average lengths of confidence intervals for β were 5.676 and 3.059, respectively, while the average lengths of confidence intervals for β of LMM and Huber were 6.330 and 5.429, respectively.

- in case of c=10 and m=5 and ρ=0.025, ADM and ADH average lengths of confidence intervals for β were 0.507 and 0.507, respectively, while the average lengths of confidence intervals for β of LMM and Huber were 0.514 and 0.537, respectively.

- in case of c=25 and m=15 and ρ=0.1, ADM and ADH average lengths of confidence intervals for β were 0.288 and 0.285, respectively, while the average lengths of confidence intervals for β of LMM and Huber were 0.288 and 0.285, respectively.

3.4.4 Simulation Results on Adaptive Confidence Intervals for β for Unbalanced Data

A simulation study was conducted to compare the adaptive and non-adaptive methods for estimating $var(\hat{\beta})$ using PSUs with unequal sample sizes. Data were generated from model (2.3), with different PSU sizes, m_i. The value of ρ was varied over a range of values of 0, 0.025 and 0.1. The number of PSUs, c, was also varied over a range of values of 2, 5, 10, 25 and 50. m_i generated randomly from uniform distribution. The average number of observations per PSU, $\bar{m}$ was varied to be 3, 10 and 25 to be consistent with the balanced data case. For this purpose three cases were used. In case 1, the number of observations was generated to be an integer between 2 and 4 with average equal to 3 observations per PSU. In case 2, this number varied from 5 to 15, with average equal to 10. Finally, in case 3, the average was 25, with m_i varying between 15 and 35. 1000 samples were generated in each case. The hypothesis $H_0 : \sigma_b^2 = 0$ was tested as described in Subsection 2.3.2 using the restricted likelihood ratio test defined in Equation (2.31).

Tables 3.5 - 3.7 show the results for the unbalanced data case. They show the ratio of the mean estimated variance of $\hat{\beta}$, $E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$, using the four strategies of estimation (ADM, ADH, LMM and Huber) with values of ρ of 0, 0.025 and 0.1. In all tables we used $\beta = 0$ and significance level $\alpha = 0.1$ for testing $\sigma_b^2 = 0$. The tables show the non-coverage rates of 90% confidence intervals for β and the average lengths of these confidence intervals. The proportion of samples where $H_0 : \sigma_b^2 = 0$ was rejected are also shown.

The variance estimators were generally approximately unbiased as most ratios were close to 1. There were some exceptions. The first was the LMM, ADM and ADH variance estimators, which tended to be biased when there were 2 sample PSUs with all average numbers of observations per PSU and when there 5 sample PSUs with 10 or less average number of observations per PSU for ρ=0. For ρ=0.025, it tended to be biased when c was 2 with $\bar{m}$ was (10 or less) and when there were 5 sample PSU with $\bar{m}$ was 3. For

ρ=0.1, it tended to be biased when c was 2 with $\bar{m}$ was 3 only.

Non-coverage rates for β were close to the nominal rate of 10% when ρ=0 for all methods except for the LMM method. The LMM non-coverage rates were a bit smaller than the nominal rate when c=2 with all average numbers of observations per PSU. The LMM non-coverage was good when there were 5 or more PSUs.

For $\rho \neq 0$, Huber non-coverage rate was close to 10% in all cases.

For ρ=0.025, the LMM and ADM non-coverage rates were much higher than the nominal rate when there were (25 or less) sample PSUs with average number of observations per PSU was large 25. The ADH non-coverage rate was higher than the nominal rate when there were 2 sample PSUs with $\bar{m}$=25. In case of ρ=0.1, the LMM and ADM non-coverage rates were much higher than the nominal rate when c $\leq$10 and $\bar{m}$=10 or 25, and when c=50 with $\bar{m}$=3. The ADH non-coverage rate was about the same as the nominal rate in most cases except when c=5 with all values of $\bar{m}$ for ρ=0, when c=2 and 5 with $\bar{m}$=25 and 3, respectively when ρ=0.025 and when c=2 with $\bar{m}$=10 and 25, c=5 with $\bar{m}$=25 and when c=50 with $\bar{m}$=3 in case of ρ=0.1.

The ADM average lengths of confidence intervals for β were similar to the LMM average lengths of confidence intervals for β for $c \geq 5$ with all average numbers of observations per PSU for all values of ρ. When c=2, the ADM average lengths were about 6-12% shorter. The ADH average lengths of

Table 3.5: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	Lrt	ADM	ADH	LMM	Hub
2	3	1.380	1.380	1.493	1.099	8.3	8.2	7.6	10.8	19.1	2.254	2.916	2.402	4.385
2	10	1.391	1.391	1.460	1.018	8.9	8.9	8.1	10.2	15.9	0.938	1.512	0.966	2.332
2	25	1.310	1.310	1.401	0.953	8.2	8.1	6.9	9.9	13.5	0.542	0.873	0.562	1.433
5	3	1.240	1.240	1.253	1.037	8.2	7.7	7.9	10.9	27.1	1.021	1.069	1.027	1.066
5	10	1.213	1.213	1.217	0.972	8.2	7.7	8.2	8.9	24.8	0.526	0.567	0.527	0.568
5	25	1.184	1.184	1.189	0.973	8.1	7.6	8.1	9.3	23.0	0.326	0.353	0.327	0.365
10	3	1.164	1.164	1.165	1.038	8.7	8.4	8.7	10.2	25.5	0.648	0.660	0.649	0.648
10	10	1.129	1.129	1.129	0.973	8.4	8.0	8.4	10.1	21.3	0.354	0.363	0.354	0.354
10	25	1.206	1.206	1.206	1.053	8.1	8.1	8.1	10.1	23.2	0.223	0.230	0.223	0.226
25	3	1.050	1.050	1.050	1.001	10.5	10.4	10.5	10.0	15.4	0.395	0.397	0.395	0.394
25	10	1.126	1.126	1.126	1.058	8.6	8.5	8.6	8.5	13.0	0.215	0.216	0.215	0.215
25	25	1.125	1.125	1.125	1.050	8.5	8.5	8.5	10.3	11.6	0.136	0.136	0.136	0.135
50	3	0.992	0.992	0.992	0.970	10.2	10.2	10.2	10.3	6.8	0.273	0.273	0.273	0.273
50	10	1.042	1.042	1.042	1.015	9.1	9.0	9.1	9.9	6.6	0.149	0.150	0.149	0.150
50	25	1.027	1.027	1.027	0.990	9.5	9.5	9.5	9.7	10.9	0.095	0.095	0.095	0.094

Table 3.6: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0.025.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	Lrt	ADM	ADH	LMM	Hub
2	3	1.343	1.343	1.454	1.109	9.3	9.2	8.2	10.1	19.7	2.339	3.044	2.502	4.625
2	10	1.291	1.291	1.365	1.031	11.0	10.8	9.6	9.8	19.8	1.017	1.761	1.054	2.602
2	25	1.141	1.141	1.195	0.982	15.1	14.9	13.9	8.9	27.0	0.637	1.347	0.657	1.890
5	3	1.220	1.220	1.233	1.046	7.8	7.4	7.7	9.1	30.9	1.048	1.103	1.055	1.109
5	10	1.096	1.097	1.102	0.950	10.3	8.8	10.1	9.2	34.8	0.560	0.619	0.561	0.631
5	25	1.008	1.009	1.011	0.939	13.7	11.6	13.6	10.7	49.1	0.383	0.446	0.384	0.460
10	3	1.137	1.137	1.138	1.037	10.1	9.8	10.1	11.5	30.1	0.665	0.678	0.665	0.671
10	10	1.069	1.070	1.069	1.007	10.1	9.5	10.1	10.0	38.4	0.381	0.397	0.381	0.401
10	25	1.028	1.028	1.028	1.007	12.0	10.0	12.0	10.6	61.0	0.268	0.287	0.268	0.291
25	3	1.011	1.011	1.011	0.991	10.2	10.2	10.2	10.3	20.4	0.404	0.406	0.404	0.409
25	10	1.057	1.057	1.057	1.075	10.1	9.8	10.1	9.4	38.9	0.233	0.237	0.233	0.243
25	25	0.973	0.973	0.973	0.992	12.1	11.6	12.1	10.5	71.9	0.166	0.171	0.166	0.174
50	3	0.952	0.952	0.952	0.964	11.3	11.3	11.3	10.6	11.9	0.279	0.279	0.279	0.283
50	10	0.936	0.936	0.936	0.991	11.4	11.3	11.4	9.8	35.8	0.162	0.163	0.162	0.169
50	25	1.008	1.008	1.008	1.018	10.2	9.8	10.2	9.5	88.6	0.119	0.121	0.119	0.122

Table 3.7: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0.1.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	Lrt	ADM	ADH	LMM	Hub
2	3	1.258	1.258	1.356	1.100	10.3	10.2	8.6	10.5	24.2	2.648	3.570	2.817	5.150
2	10	1.145	1.145	1.188	1.025	16.3	15.5	15.3	9.6	32.7	1.285	2.609	1.321	3.375
2	25	1.028	1.028	1.050	0.987	22.3	20.3	21.2	9.1	45.8	0.917	2.386	0.935	2.845
5	3	1.149	1.149	1.165	1.047	10.7	10.3	10.4	9.8	39.5	1.132	1.205	1.142	1.232
5	10	0.986	0.986	0.989	0.943	13.9	11.0	13.9	10.3	59.4	0.685	0.796	0.687	0.813
5	25	0.934	0.934	0.934	0.926	18.7	12.1	18.7	11.4	81.3	0.557	0.684	0.557	0.692
10	3	1.073	1.073	1.074	1.029	10.9	10.4	10.9	9.8	44.3	0.723	0.744	0.724	0.749
10	10	1.055	1.055	1.055	1.052	12.4	10.6	12.4	10.2	76.7	0.483	0.519	0.483	0.525
10	25	0.972	0.973	0.972	0.973	14.8	11.0	14.8	10.9	94.6	0.400	0.438	0.400	0.439
25	3	0.963	0.963	0.963	0.980	10.4	10.2	10.3	9.5	43.1	0.441	0.445	0.441	0.455
25	10	1.087	1.087	1.087	1.093	10.3	9.0	10.3	8.5	91.3	0.306	0.315	0.306	0.317
25	25	0.998	0.998	0.998	0.998	10.7	10.0	10.7	10.0	99.8	0.256	0.264	0.256	0.264
50	3	0.916	0.916	0.916	0.962	12.5	12.5	12.5	10.9	38.3	0.305	0.306	0.305	0.316
50	10	0.988	0.988	0.988	0.990	11.0	10.5	11.0	10.3	98.1	0.217	0.221	0.217	0.221
50	25	1.008	1.008	1.008	1.008	9.8	9.4	9.8	9.4	100.0	0.182	0.185	0.182	0.185

confidence intervals for β were similar to the Huber average lengths of confidence intervals for β for all sample PSUs with all values of m and ρ except when c=2, as the ADH average lengths were shorter than the Huber average lengths of order about 30-65%.

The proportions of samples where $H_0 : \sigma_b^2 = 0$ is rejected were generally much higher than 10% when ρ=0, and was a very high 27% when c=5 and $\bar{m}$=3. They were much higher than for the balanced data case. This might be because the PSU sizes in the unbalanced design case have a wide range, for example; in case of $\bar{m}$=25, the PSU sizes vary between 15 and 35. Or this might be because of the distribution of the RLRT. It was assumed that the distribution is a 50:50 mixture of χ_0^2 and χ_1^2 following Chernoff (1954) in the balanced and unbalanced designs. The 50:50 mixture of χ^2 distribution of the likelihood ratio test might not perform well in the unbalanced designs because the response can not be divided into identically distributed sub-vectors as in Stram and Lee (1994). This approximation may not be a very good approximation in the unbalanced designs if the response is divided into small or moderate number of sub-vectors, even if the responses are independent (Scheipl et al., 2007).

3.5 Conclusions

i. Adaptive confidence intervals can perform poorly in designs with few sample PSUs and large sample sizes in each PSU. In these designs, even a small intraclass correlation will substantially inflate the variance of the mean, however the PSU-level variance component is unlikely to be statistically significant even if the intraclass correlation is as high as 0.1. As a result, when the number of PSUs (c) is 2 or 5, and the number of observations per PSU (m or $\bar{m}$) is 25 or more both of the adaptive estimators have higher than desirable non-coverage when the intraclass correlation is non-zero, of the order of 15-20%. It appears that for these extreme designs, clustering must be allowed for in variance estimates, even if it is not statistically significant.

ii. In comparing the Linear Mixed Model (LMM) with the adaptive version (ADM), we find that:

- Both the LMM and ADM approaches have close to nominal non-coverage, except for extreme designs of the kind discussed in i. For these designs, the adaptive and non-adaptive LMM methods both have high non-coverage. In the case of the adaptive method, this is presumably because there is not much power to detect the PSU-level variance component, even when it is substantial. For the non-adaptive LMM, the problem seems to be that the LMM confidence intervals are not exact and do not do well for small sample sizes.

- The ADM confidence intervals are noticeably narrower (10-20%) than the LMM for c equal to 2 and 5, but there is not much to choose between ADM and LMM for c=10 or more.

iii. In comparing the robust Huber-White approach with the adaptive version (ADH), we find that:

- The Huber approach has close to nominal non-coverage in all cases. So does the ADH approach, except for the extreme designs mentioned in i.

- The Huber method gives wide confidence intervals when c is small (2 or 5) with order of 10-80% eventhough the non-coverage is close to the nominal 10%. This is because the degrees of freedom for this method is equal to (c-1). ADH has much narrower confidence intervals (10-80%) , because its degrees of freedom are equal to (n-1) rather than (c-1) if the PSU-level variance component is not significant.

iv. This leads to the following recommendations:

- Designs with fewer than 10 PSUs, and a large sample size in each PSU should be avoided, even if the intraclass correlation is believed to be low. Hence, we recommend ignoring clustering if the PSU-level variance effect is insignificant.

Chapter 4

Robustness of Adaptive Estimators based on Linear Mixed Models to Non-Normality

4.1 Introduction

In Chapter 3, the methods were based on fitting a linear mixed model. Data were assumed to be normally distributed. In this chapter, the purpose is to see if these methods still work well if the assumption of normality is not justified. For this purpose, the same methods applied in Chapter 3 will be applied to data that are log-normal rather than normal.

Log-normal distributions play a very important role in many sciences including ecology and biology (Ott, 1995). A random variable Y is said to have a log-normal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ if the natural logarithm of Y, $X = ln(Y)$, follows a normal distribution with mean μ and standard deviation σ (i.e. $X \sim N(\mu, \sigma^2)$) (Kapadia et al., 2005). Therefore, it is equivalent to $Y = e^X$ where X is normally distributed with mean μ and standard deviation σ. The log-normal distribution is a continuous distribution which is typically used to model right-skewed variables. The log-normal distribution is useful for many intrinsically positive variables, for example residential property prices (Zabel, 1999) and household income (Longford and Pittau, 2006), and organisms size and number of species in biology (Krishnamoorthy and Mathew, 2003).

The log-normal distribution will be denoted in this thesis by $LN(\mu, \sigma^2)$. Crow and Shimizo (1988) defined the probability density function (*pdf*) of $Y \sim LN(\mu, \sigma^2)$ by

$$f(y) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma y} exp\Big[-\frac{(ln(y)-\mu)^2}{2\sigma^2} \Big] & y > 0, \\ 0 & y \leq 0. \end{cases} \tag{4.1}$$

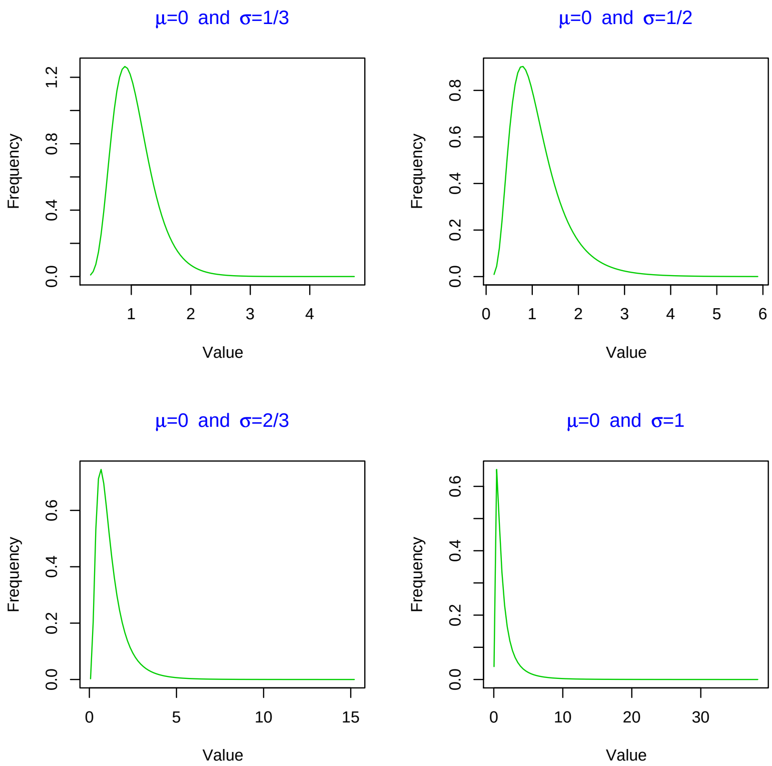
Figure 4.1 shows the log-normal probability density function with different values of σ - this parameter controls the skewness of Y.

Crow and Shimizo (1988) noted that the mean ($E(Y)$) and the variance ($Var(Y)$) of the log-normal random variable Y as

$$\begin{aligned} E(Y) &= \exp(\mu + \frac{1}{2}\sigma^2); \\ Var(Y) &= \exp(2\mu + \sigma^2)\{\exp(\sigma^2) - 1.\} \end{aligned} \tag{4.2}$$

This chapter is divided into four sections. Sections 4.2 and 4.3 describe the simulation studies conducted to evaluate the adaptive methods described

Figure 4.1: Probability density function of the Log-normal distribution plotted for a sample of size 10,000



in Chapter 3 when data are log-normal rather than normal. Section 4.2 describes simulation of a balanced design and Section 4.3 covers unbalanced designs. In Section 4.4, we will state the conclusions of this chapter.

4.2 Simulation Study of Log-Normal Data in a Balanced Two-Stage Design

A simulation study was performed to compare the adaptive and non-adaptive methods utilized in Chapter 3 for estimating $var(\hat{\beta})$ and associated confidence intervals where data are log-normal rather than normal. This study was based on equal sample sizes within PSUs. Data were generated from the intercept only model (2.3) assuming that b_i and e_{ij} are normally distributed with zero mean and variances equal to $\sigma_b^2 = \frac{\rho}{1-\rho}\sigma^2$ and $\sigma_e^2 = \sigma^2$, respectively, where σ was $\frac{1}{3}$, $\frac{1}{2}$ or $\frac{2}{3}$. Then the Equation $Y = e^X$ was applied to generate log-normally distributed values.

The five procedures for estimating the variance of $\hat{\beta}$ used in Chapter 3 were used in this simulation as well. These strategies are the linear model strategy (LM), the linear mixed model strategy (LMM), the robust Huber-White variance estimator strategy (Hub) and the two adaptive strategies, the LMM based and the Huber based adaptive strategies. 1000 samples were generated in each case. All methods used in this chapter were identical to those used in Chapter 3. The values of ρ, c, m and σ were varied. The parameter ρ was varied over a range of values of 0 and 0.025. The number of PSUs, c, was varied over a range of values of 2, 5, 10 and 25 and the PSU sample size was varied over a range of values of 2, 5, 10, 15, 25 and 50.

For each sample, the estimated regression coefficients $\hat{\beta}$ and the estimators of $var(\hat{\beta})$ were calculated for the LMM and LM models using the $lme4$ and lm packages (Pinheiro and Bates, 2000) in the R statistical environment (R Development Core Team, 2007). The true variance of $\hat{\beta}$ was determined by calculating the variance over all 1000 simulations.

The hypothesis $H_0 : \sigma_b^2 = 0$ was tested as described in Section 2.3.2 using the restricted likelihood ratio test defined in equations (2.30) and (2.31). The two adaptive strategies ADM and ADH were as defined in Section 3.2. 90% confidence intervals for β were calculated for the LMM method using the method of Faes et al. (2009) as described in Subsection 2.2.3. Huber-White confidence intervals for β were calculated as discussed in Subsection 2.2.5, and the adaptive confidence intervals for β were calculated as discussed in Section 3.2. The approaches are applied to Y but the intraclass correlation, ρ, applies to X.

The results for the simulation study using several log-normal distributions with two values of σ ($\frac{1}{3}$ and $\frac{2}{3}$), and two values of ρ (0 and 0.025) are shown in Tables 4.1 - 4.4. At this section we assumed that the PSUs have the same number of observations, that is $m_i = m$, for all i=1, 2, ..., c. Results for other values of ρ and σ are shown in Appendix C.

As in Chapter 3, the ratio of the estimated variance to the true variance of $\hat{\beta}$, $E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$, was calculated. The tables also include the non-coverage rates for β as well as the average lengths of the 90% confidence intervals for β. The restricted likelihood ratio test probabilities of rejecting $H_0 : \sigma_b^2 = 0$ are included in these tables as well. Four strategies of estimation are included in the tables, ADM, ADH, LMM and Hub. The LM strategy of estimation is not shown, because Chapter 3 showed that this method was

Table 4.1: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT with ρ=0, using Log-Normal data with level of skewness $\sigma = \frac{1}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.138	1.138	1.319	0.953	9.5	9.5	12.9	10.9	10.7	1.599	1.016	1.813	1.768
2	5	1.265	1.265	1.545	1.092	10.0	9.8	10.1	10.6	6.2	0.456	0.555	0.495	1.170
2	10	1.266	1.266	1.540	1.079	11.0	11.0	11.0	9.0	5.5	0.315	0.383	0.349	0.831
2	15	1.089	1.089	1.340	0.900	11.1	11.1	11.1	9.3	3.4	0.240	0.276	0.268	0.653
2	25	1.224	1.224	1.491	1.008	9.7	9.7	10.4	8.2	4.2	0.191	0.227	0.211	0.514
2	50	1.280	1.281	1.572	1.114	8.7	8.7	9.5	10.2	5.1	0.139	0.169	0.156	0.380
5	2	1.211	1.211	1.333	1.129	9.9	9.9	10.2	12.1	11.2	0.432	0.436	0.440	0.464
5	5	1.087	1.087	1.192	0.964	8.9	8.9	10.1	9.7	5.9	0.253	0.254	0.257	0.281
5	10	1.045	1.045	1.164	0.940	10.6	10.5	10.7	10.1	6.6	0.180	0.182	0.185	0.204
5	15	1.114	1.114	1.242	0.983	9.5	9.5	10.2	9.8	5.2	0.145	0.146	0.149	0.164
5	25	1.013	1.013	1.119	0.909	10.5	10.5	12.0	11.1	8.4	0.115	0.117	0.118	0.131
5	50	1.152	1.152	1.284	1.026	8.8	8.8	8.4	10.0	6.7	0.081	0.082	0.084	0.092
10	2	1.059	1.059	1.126	0.984	10.2	10.1	10.7	11.3	9.6	0.283	0.285	0.283	0.285
10	5	1.009	1.009	1.061	0.938	10.9	10.9	12.8	11.5	8.8	0.178	0.178	0.177	0.182
10	10	1.043	1.043	1.101	0.975	9.4	9.3	10.2	10.6	8.6	0.125	0.125	0.125	0.129
10	15	1.059	1.059	1.112	0.972	9.6	9.6	10.3	10.1	8.7	0.102	0.102	0.102	0.105
10	25	1.127	1.127	1.192	1.030	9.0	9.0	9.8	9.6	9.5	0.079	0.079	0.079	0.081
10	50	1.015	1.015	1.077	0.944	12.2	12.2	12.5	12.2	8.4	0.056	0.056	0.056	0.058
25	2	0.988	0.988	1.030	0.952	11.9	12.0	12.3	11.9	9.4	0.174	0.174	0.173	0.173
25	5	1.126	1.127	1.135	1.074	8.1	8.1	8.1	8.6	10.1	0.110	0.110	0.109	0.110
25	10	1.004	1.004	1.017	0.955	10.2	10.2	11.5	10.8	7.5	0.077	0.077	0.077	0.077
25	15	1.152	1.152	1.164	1.087	8.5	8.8	9.3	9.3	10.4	0.064	0.064	0.063	0.063
25	25	1.030	1.030	1.043	0.987	9.7	9.7	10.0	10.0	9.3	0.049	0.049	0.049	0.049
25	50	1.047	1.047	1.079	1.006	10.2	10.3	10.1	9.9	9.0	0.035	0.035	0.035	0.035

Table 4.2: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT with ρ=0.025, using Log-Normal data with level of skewness $\sigma = \frac{1}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.147	1.147	1.349	1.002	11.9	12.0	14.0	10.4	10.7	1.846	1.009	2.027	1.795
2	5	1.204	1.204	1.489	1.133	11.4	10.8	11.1	8.8	7.7	0.466	0.594	0.516	1.266
2	10	1.019	1.019	1.277	0.929	13.0	13.0	13.3	9.9	6.2	0.315	0.394	0.358	0.872
2	15	1.056	1.056	1.294	1.001	15.1	15.1	14.1	8.1	9.0	0.277	0.374	0.314	0.770
2	25	1.070	1.070	1.309	1.065	18.2	18.2	14.3	9.3	11.7	0.237	0.339	0.272	0.657
2	50	1.033	1.033	1.208	1.062	22.9	22.9	19.4	9.1	18.2	0.205	0.317	0.230	0.544
5	2	1.101	1.101	1.199	1.013	11.1	11.1	11.4	11.3	11.0	0.431	0.434	0.436	0.463
5	5	1.122	1.122	1.254	1.075	10.1	10.0	10.3	10.4	10.3	0.265	0.267	0.273	0.302
5	10	1.057	1.057	1.191	1.054	12.8	12.7	11.1	8.4	14.3	0.192	0.194	0.201	0.226
5	15	0.953	0.953	1.086	0.972	14.3	14.1	13.7	10.6	15.4	0.160	0.163	0.169	0.192
5	25	1.002	1.002	1.135	1.058	12.3	12.1	12.0	9.0	23.3	0.135	0.138	0.145	0.163
5	50	0.931	0.932	1.018	0.987	17.8	17.3	14.3	10.0	38.9	0.114	0.117	0.121	0.133
10	2	1.074	1.074	1.145	1.021	10.3	10.1	10.7	9.8	12.5	0.281	0.281	0.281	0.285
10	5	1.002	1.002	1.076	0.992	11.7	11.7	12.2	11.3	12.3	0.180	0.180	0.183	0.190
10	10	0.982	0.982	1.059	1.011	12.3	12.0	12.1	10.3	19.2	0.133	0.133	0.136	0.144
10	15	0.963	0.963	1.046	1.014	12.2	12.2	12.6	9.9	25.5	0.113	0.114	0.117	0.124
10	25	0.990	0.990	1.057	1.044	12.8	12.8	11.8	9.7	39.3	0.096	0.096	0.099	0.104
10	50	0.963	0.963	1.002	1.003	12.9	12.7	11.0	9.1	61.6	0.080	0.080	0.082	0.084
25	2	1.016	1.016	1.063	0.997	11.9	12.0	12.7	12.6	12.6	0.172	0.173	0.172	0.173
25	5	0.982	0.982	0.996	0.987	10.7	10.6	11.9	9.7	16.4	0.112	0.112	0.111	0.115
25	10	0.964	0.964	0.975	0.991	10.9	11.0	11.5	10.1	29.8	0.083	0.083	0.083	0.086
25	15	0.876	0.876	0.892	0.916	12.0	12.2	11.9	10.7	40.4	0.071	0.070	0.071	0.073
25	25	0.991	0.991	1.005	1.024	11.2	11.7	11.3	10.3	63.5	0.060	0.060	0.061	0.062
25	50	1.079	1.079	1.084	1.087	9.0	8.6	8.7	8.1	91.9	0.051	0.051	0.051	0.051

Table 4.3: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT with ρ=0, using Log-Normal data with level of skewness $\sigma = \frac{2}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.136	1.136	1.273	0.976	13.3	13.3	16.0	11.6	10.2	3.469	2.295	3.745	4.171
2	5	1.120	1.120	1.360	0.989	15.3	15.2	14.8	10.7	4.7	1.067	1.267	1.161	2.878
2	10	1.136	1.136	1.425	0.995	12.8	12.8	13.2	11.2	3.6	0.751	0.857	0.841	2.060
2	15	1.171	1.171	1.448	1.007	12.3	12.3	13.7	10.3	4.9	0.617	0.732	0.691	1.674
2	25	1.201	1.201	1.492	1.049	10.9	10.9	10.6	9.9	4.0	0.471	0.553	0.527	1.328
2	50	1.246	1.246	1.530	1.080	9.3	9.3	9.9	9.1	5.8	0.352	0.433	0.393	0.967
5	2	1.107	1.107	1.190	1.036	16.0	15.5	15.6	13.5	11.8	1.025	1.032	1.026	1.113
5	5	1.122	1.122	1.214	1.016	11.7	11.6	12.2	10.8	6.8	0.635	0.639	0.638	0.713
5	10	1.011	1.011	1.105	0.903	12.8	12.8	13.6	12.2	6.1	0.450	0.452	0.454	0.509
5	15	1.061	1.061	1.179	0.985	11.7	11.6	12.1	11.1	5.4	0.371	0.373	0.380	0.432
5	25	1.081	1.081	1.205	0.994	11.0	11.0	11.4	12.6	7.4	0.294	0.297	0.303	0.339
5	50	1.102	1.102	1.226	0.987	9.6	9.6	9.8	9.9	5.5	0.204	0.206	0.210	0.234
10	2	1.033	1.033	1.092	0.980	12.2	12.2	12.5	12.5	10.0	0.691	0.694	0.691	0.706
10	5	1.050	1.050	1.104	0.992	11.0	11.0	12.1	11.2	5.6	0.445	0.445	0.443	0.463
10	10	1.067	1.067	1.123	0.990	10.9	10.7	12.1	10.8	5.7	0.316	0.316	0.314	0.327
10	15	1.065	1.065	1.127	0.993	10.2	10.2	10.0	10.4	8.5	0.259	0.260	0.261	0.269
10	25	1.110	1.110	1.176	1.027	8.4	8.4	9.6	9.6	7.9	0.202	0.202	0.202	0.209
10	50	1.097	1.097	1.158	0.991	9.9	9.9	9.8	10.6	5.7	0.142	0.142	0.142	0.146
25	2	1.079	1.079	1.116	1.040	11.2	11.3	11.9	11.8	10.7	0.437	0.437	0.433	0.436
25	5	0.990	0.990	1.001	0.949	11.1	11.3	11.8	11.6	9.3	0.283	0.282	0.279	0.282
25	10	1.070	1.070	1.080	1.013	9.8	9.7	10.5	10.0	8.3	0.199	0.199	0.196	0.199
25	15	1.026	1.026	1.038	0.980	10.6	10.7	11.9	10.8	9.8	0.164	0.164	0.162	0.164
25	25	1.057	1.057	1.069	1.009	8.9	9.0	9.0	10.0	8.5	0.126	0.126	0.124	0.126
25	50	1.012	1.012	1.044	0.969	11.0	10.8	11.8	12.0	8.6	0.089	0.089	0.089	0.090

Table 4.4: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT with ρ=0.025, using Log-Normal data with level of skewness $\sigma = \frac{2}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.133	1.133	1.323	1.046	15.1	15.2	17.0	12.1	11.4	3.765	2.322	4.140	4.386
2	5	1.062	1.062	1.323	1.027	16.0	15.6	15.5	10.9	6.7	1.123	1.381	1.244	3.093
2	10	1.045	1.045	1.270	0.908	15.3	15.3	14.8	10.8	6.2	0.779	0.969	0.861	2.130
2	15	0.924	0.924	1.167	0.875	17.3	17.2	16.4	8.5	5.8	0.634	0.776	0.724	1.833
2	25	1.039	1.039	1.305	1.058	17.7	17.7	15.7	10.0	10.8	0.581	0.805	0.667	1.633
2	50	0.884	0.884	1.064	0.900	23.1	23.1	20.6	11.7	16.5	0.475	0.721	0.540	1.280
5	2	1.121	1.121	1.213	1.072	14.3	14.2	14.4	14.2	11.4	1.065	1.073	1.074	1.169
5	5	1.101	1.101	1.242	1.088	13.0	12.8	13.6	11.9	9.5	0.636	0.641	0.658	0.743
5	10	1.009	1.009	1.153	1.016	11.9	11.8	12.1	11.5	10.6	0.467	0.472	0.490	0.557
5	15	1.039	1.040	1.184	1.069	13.6	13.5	13.3	10.4	13.7	0.404	0.410	0.428	0.489
5	25	1.003	1.004	1.136	1.055	14.8	14.6	13.7	11.3	23.1	0.344	0.352	0.370	0.413
5	50	0.992	0.992	1.098	1.060	17.4	17.0	14.1	10.8	35.0	0.279	0.287	0.298	0.333
10	2	1.111	1.111	1.170	1.054	11.1	11.0	10.8	10.9	9.7	0.699	0.700	0.697	0.713
10	5	1.000	1.000	1.058	0.974	13.7	13.6	14.4	14.2	10.5	0.443	0.444	0.444	0.466
10	10	1.043	1.043	1.122	1.069	11.7	11.5	10.7	9.8	15.8	0.331	0.331	0.337	0.358
10	15	0.986	0.986	1.072	1.034	12.7	12.7	11.4	9.2	21.0	0.284	0.285	0.294	0.311
10	25	0.952	0.952	1.020	1.002	14.1	13.6	13.0	9.8	32.7	0.235	0.235	0.242	0.255
10	50	0.894	0.895	0.943	0.943	14.1	14.7	13.0	11.7	54.3	0.195	0.194	0.202	0.208
25	2	1.039	1.039	1.082	1.014	12.2	12.3	12.2	12.9	12.4	0.438	0.438	0.435	0.439
25	5	1.036	1.036	1.048	1.039	10.6	10.4	11.2	9.9	13.8	0.285	0.285	0.282	0.292
25	10	0.937	0.937	0.951	0.959	12.2	12.1	13.3	11.8	23.5	0.212	0.210	0.210	0.218
25	15	0.908	0.908	0.927	0.953	12.7	12.9	13.6	11.7	33.2	0.177	0.176	0.177	0.184
25	25	0.904	0.904	0.921	0.941	12.5	12.6	12.2	10.7	55.6	0.150	0.148	0.150	0.154
25	50	1.039	1.039	1.052	1.056	10.5	10.5	10.2	9.6	83.9	0.124	0.123	0.125	0.125

inadequate even for normal data.

Similar to the Huber variance ratios calculated in Chapter 3, Huber variance ratios were as close to 1 for all values of ρ and σ.

The LMM variance estimators have approximately the same bias as the LMM variance estimators in Chapter 3, when ρ=0 regardless of the skewness level, σ. These biases were large when there were 5 or less sample PSUs. They became smaller for larger numbers of sample PSUs (10 or more).

For ρ=0.025, the LMM variance estimators for log-normal data have smaller bias than the LMM variance estimators for normal data when there were 2 sample PSUs for $\sigma = \frac{1}{3}$. These biases were larger in the case of $\sigma = \frac{2}{3}$ when c=2. The biases were approximately the same for other numbers of sample PSUs (5 or more).

For ρ=0 and ρ=0.025, the ADM and ADH variance estimators for the log-normal data with both skewness levels have approximately the same bias as the ADM and ADH variance estimators in Chapter 3. The biases were large when there was 2 sample PSUs and small otherwise. For the log-normal as well as the normal data, the biases for the ADM variance estimator were smaller than the biases of the LMM variance estimator, regardless the intraclass correlation value, ρ, and the skewness level, σ.

For ρ=0 and $\sigma = \frac{1}{3}$, non-coverage rates of confidence intervals of β using the LMM, ADM and ADH methods were close to the nominal rate (10%) as well as the LMM, ADM and ADH non-coverage rates in Chapter 3. For ρ=0 and $\sigma = \frac{2}{3}$, the LMM, ADM and ADH non-coverage rates differed significantly from 10% when there were 2 or 5 sample PSUs with small number of observations per PSU (15 or less). For ρ=0.025, the LMM, ADM and ADH non-coverage rates differed appreciably from 10% for all values of c and m, in general, for both skewness levels. For all values of ρ and σ the Huber non-coverage rates were close to the nominal rate, as in Chapter 3.

For both values of ρ, the LMM non-coverage rates when $\sigma = \frac{1}{3}$ were closer to the nominal rate than the LMM non-coverage rates when $\sigma = \frac{2}{3}$ for small sample PSUs (5 or less). The average length of the 90% confidence intervals for β using all methods of estimation were obviously shorter when $\sigma = \frac{1}{3}$ than when $\sigma = \frac{2}{3}$ for all values of c.

The 90% confidence intervals for β were much shorter using the log-normal data than those calculated using the normal data in Chapter 3, because $Var(Y) = exp(2\mu + \sigma^2)\{exp(\sigma^2) - 1\}$ less than the variance of the simulated data in Chapter 3.

For ρ=0 with $\sigma = \frac{1}{3}$ and $\sigma = \frac{2}{3}$, when there were 2 sample PSUs the average length of the 90% confidence intervals for β using the ADM method was 15-25% shorter than the LMM method. The ADM were 10-15% shorter when there were 5 sample PSUs with large number of observations per PSU (15 or more). The average length of the 90% ADH confidence intervals for β

were 65-80% shorter than the Huber when there were 2 sample PSUs, 15-30% shorter when there were 5 sample PSUs, and 10-15% shorter when there were 10 sample PSUs with 5 or more observations per PSU. Differences between the adaptive and non-adaptive confidence interval lengths were negligible in all other cases.

For ρ=0.025 with $\sigma = \frac{1}{3}$ and $\sigma = \frac{2}{3}$, when there were 2 sample PSUs the average lengths of the 90% ADM confidence intervals for β were 10-25% shorter than the LMM method. The average lengths of the 90% ADH confidence intervals for β were 65-85% shorter than the Huber when there were 2 sample PSUs, and 10-20% shorter when there were 5 sample PSUs. There were no relevant differences, otherwise.

The proportions of samples where $H_0 : \sigma_b^2 = 0$ was rejected were similar to those in Chapter 3.

4.3 Simulation Study of Log-normal Data in an Unbalanced Two-Stage Design

A simulation study was conducted to compare the adaptive and non-adaptive methods for estimating $var(\hat{\beta})$ using PSUs with unequal sample sizes and log-normal data. Tables 4.5 - 4.8 show the results of the simulation study. Log-normal data were generated in the same way as described in Section 4.2. Data were generated assuming unequal sample within PSU sizes, m_i.

Results for other values of ρ and σ are shown in Appendix C. Two values of ρ were used, 0 and 0.025. The number of sample PSUs, c, was also varied over a range of values of 2, 5, 10, 25 and 50. The value of σ was varied over $\frac{1}{3}$ and $\frac{2}{3}$. Results for other values of ρ and σ are shown in Appendix C. The average number of observations per PSU, $\bar{m}$ was set to 3, 10 and 25. The three cases used for this purpose are explained in Subsection 3.4.4. The hypothesis $H_0 : \sigma_b^2 = 0$ was tested as described in Subsection 2.3.2 using the restricted likelihood ratio test defined in Equation (2.31). In all tables we used $\beta = 0$ and significance level $\alpha = 0.1$ for testing $\sigma_b^2 = 0$. The tables show the non-coverage rates of 90% confidence intervals for β and the average lengths of these confidence intervals. The proportion of samples where $H_0 : \sigma_b^2 = 0$ was rejected are also shown. The tables show the ratio of the mean estimated variance of $\hat{\beta}$, $E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$, using the four strategies of estimation (ADM, ADH, LMM and Huber) with values of ρ of 0 and 0.025 and skewness levels of $\sigma = \frac{1}{3}$ and $\sigma = \frac{2}{3}$.

For all values of ρ and σ, the average length of the 90% ADM confidence intervals for β was 10-15% shorter than the LMM confidence intervals for β when there were 2 sample PSUs with all values of $\bar{m}$. For all values of ρ and σ, the average length of the 90% ADH confidence intervals for β was much shorter than the Huber (50-65%) when there were 2 sample PSUs with all values of $\bar{m}$. For ρ=0 and $\sigma = \frac{2}{3}$, the average length of the 90% ADH confid-

Table 4.5: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0, $\sigma = \frac{1}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	3	1.357	1.357	1.475	1.025	9.3	9.2	8.3	10.8	16.3	0.730	0.919	0.781	1.435
2	10	1.297	1.297	1.379	0.943	9.2	8.9	8.4	8.5	13.3	0.327	0.493	0.339	0.827
2	25	1.477	1.477	1.564	1.094	8.3	8.2	7.3	10.4	15.0	0.200	0.335	0.206	0.532
5	3	1.238	1.238	1.255	1.073	9.2	8.9	9.2	9.6	25.1	0.357	0.373	0.360	0.381
5	10	1.289	1.289	1.294	1.041	8.1	7.2	8.1	9.6	23.1	0.188	0.201	0.188	0.204
5	25	1.286	1.286	1.290	1.043	8.8	8.1	8.6	10.0	23.3	0.118	0.128	0.118	0.130
10	3	1.137	1.137	1.138	1.001	8.0	8.0	8.0	9.6	23.0	0.235	0.238	0.235	0.233
10	10	1.094	1.094	1.094	0.952	9.9	9.4	9.9	11.7	20.7	0.128	0.131	0.128	0.129
10	25	1.075	1.075	1.075	0.936	9.8	9.3	9.8	9.5	19.7	0.080	0.082	0.080	0.081
25	3	1.184	1.184	1.184	1.113	7.4	7.3	7.4	8.5	13.0	0.143	0.143	0.143	0.141
25	10	1.000	1.000	1.000	0.937	10.0	10.0	10.0	11.0	13.8	0.078	0.078	0.078	0.078
25	25	1.087	1.087	1.087	1.005	8.9	8.7	8.9	9.1	11.1	0.049	0.049	0.049	0.049
50	3	1.035	1.035	1.035	1.016	10.6	10.6	10.6	10.8	5.7	0.098	0.098	0.098	0.098
50	10	0.995	0.995	0.995	0.961	9.6	9.5	9.6	9.8	5.3	0.054	0.054	0.054	0.054
50	25	1.064	1.064	1.064	1.021	8.4	8.3	8.4	9.1	9.4	0.034	0.034	0.034	0.034

Table 4.6: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0.025, $\sigma = \frac{1}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	3	1.325	1.325	1.437	1.023	9.8	9.4	8.7	11.0	17.4	0.745	0.954	0.795	1.456
2	10	1.086	1.086	1.161	0.853	12.7	12.7	11.5	9.9	18.3	0.342	0.572	0.356	0.893
2	25	1.285	1.285	1.348	1.088	13.9	13.5	12.2	10.0	23.7	0.223	0.451	0.230	0.645
5	3	1.183	1.183	1.198	1.039	9.3	8.6	9.0	10.0	29.3	0.361	0.380	0.364	0.388
5	10	1.158	1.158	1.166	1.023	11.0	10.2	11.0	10.6	34.6	0.199	0.220	0.199	0.227
5	25	1.071	1.071	1.073	1.001	12.5	10.3	12.4	9.2	50.1	0.135	0.158	0.135	0.163
10	3	1.098	1.098	1.101	1.000	8.4	8.3	8.4	9.6	26.1	0.237	0.241	0.237	0.239
10	10	1.041	1.041	1.041	0.975	10.6	9.4	10.6	9.3	37.3	0.135	0.141	0.135	0.142
10	25	0.994	0.994	0.994	0.974	12.3	10.2	12.3	10.5	58.1	0.094	0.100	0.094	0.102
25	3	1.102	1.102	1.102	1.068	9.0	9.0	9.0	9.0	17.3	0.144	0.145	0.144	0.145
25	10	0.908	0.908	0.908	0.916	11.3	11.1	11.3	10.4	35.5	0.083	0.084	0.083	0.086
25	25	1.016	1.016	1.016	1.034	11.3	11.1	11.3	9.6	68.9	0.059	0.061	0.059	0.062
50	3	0.976	0.976	0.976	0.988	11.9	11.9	11.9	11.6	10.7	0.099	0.099	0.099	0.100
50	10	0.914	0.914	0.914	0.967	11.4	11.4	11.4	9.8	33.3	0.057	0.058	0.057	0.060
50	25	0.996	0.996	0.996	1.007	10.3	10.0	10.3	9.7	84.7	0.042	0.042	0.042	0.043

Table 4.7: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0, $\sigma = \frac{2}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	3	1.272	1.272	1.392	0.992	13.9	13.5	12.6	11.7	15.0	1.665	2.050	1.789	3.453
2	10	1.261	1.261	1.356	0.949	12.1	11.6	10.7	9.3	13.4	0.801	1.204	0.834	2.092
2	25	1.445	1.445	1.518	1.071	9.8	9.7	9.3	9.5	15.9	0.505	0.860	0.520	1.365
5	3	1.217	1.217	1.236	1.087	13.0	12.4	12.5	12.0	23.3	0.874	0.911	0.882	0.951
5	10	1.281	1.281	1.285	1.045	9.9	8.9	9.9	10.7	23.7	0.474	0.509	0.475	0.521
5	25	1.272	1.272	1.275	1.041	9.5	9.0	9.5	10.3	22.6	0.300	0.324	0.300	0.334
10	3	1.137	1.137	1.139	1.014	10.4	10.3	10.4	12.0	20.6	0.583	0.591	0.584	0.584
10	10	1.089	1.089	1.089	0.948	11.1	10.6	11.1	12.2	21.9	0.325	0.333	0.325	0.328
10	25	1.082	1.082	1.082	0.949	8.9	8.6	8.9	10.1	19.4	0.204	0.209	0.204	0.209
25	3	1.196	1.196	1.196	1.132	7.8	7.7	7.8	8.7	11.1	0.363	0.364	0.363	0.361
25	10	1.008	1.008	1.008	0.946	10.0	10.0	10.0	10.6	12.7	0.200	0.201	0.200	0.200
25	25	1.107	1.107	1.107	1.022	8.7	8.4	8.7	9.0	12.1	0.127	0.127	0.127	0.125
50	3	1.040	1.040	1.040	1.020	10.6	10.6	10.6	11.1	5.9	0.249	0.249	0.249	0.249
50	10	0.999	0.999	0.999	0.964	9.8	9.8	9.8	10.8	5.5	0.139	0.139	0.139	0.138
50	25	1.068	1.068	1.068	1.025	9.1	9.1	9.1	9.5	9.8	0.088	0.089	0.088	0.088

Table 4.8: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0.025, $\sigma = \frac{2}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	3	1.231	1.231	1.364	0.993	14.9	14.4	13.4	12.8	15.2	1.686	2.079	1.821	3.498
2	10	1.085	1.085	1.169	0.872	16.1	15.5	14.6	11.1	17.3	0.830	1.363	0.869	2.236
2	25	1.274	1.274	1.341	1.070	14.9	14.2	13.8	10.1	24.0	0.551	1.113	0.568	1.594
5	3	1.174	1.174	1.196	1.061	12.8	12.4	12.5	12.8	27.2	0.882	0.925	0.891	0.963
5	10	1.156	1.156	1.163	1.020	12.0	11.4	12.0	11.9	32.9	0.496	0.545	0.497	0.566
5	25	1.070	1.070	1.072	0.988	13.2	11.1	13.0	10.4	47.6	0.337	0.390	0.337	0.403
10	3	1.103	1.103	1.105	1.008	10.5	10.2	10.5	11.1	24.7	0.589	0.599	0.589	0.597
10	10	1.044	1.044	1.044	0.972	11.3	10.9	11.3	11.1	36.5	0.341	0.356	0.341	0.358
10	25	1.001	1.001	1.001	0.977	12.5	11.0	12.5	11.3	53.2	0.233	0.248	0.233	0.253
25	3	1.127	1.127	1.127	1.089	8.6	8.5	8.6	9.3	16.5	0.366	0.368	0.366	0.368
25	10	0.929	0.929	0.929	0.937	12.0	12.0	12.0	10.6	29.9	0.209	0.212	0.209	0.217
25	25	1.013	1.014	1.013	1.037	10.6	10.0	10.6	8.8	60.6	0.147	0.150	0.147	0.154
50	3	0.979	0.979	0.979	0.988	11.6	11.6	11.6	11.4	9.8	0.250	0.250	0.250	0.254
50	10	0.921	0.921	0.921	0.968	10.9	10.8	10.9	9.6	28.7	0.146	0.147	0.146	0.152
50	25	1.009	1.009	1.009	1.026	10.3	10.2	10.3	9.9	77.1	0.104	0.106	0.104	0.107

ence intervals for β was shorter than the Huber (5-8%) when there were 5 sample PSUs with $\bar{m}$=2 and 25. For ρ=0.025 and $\sigma = \frac{1}{3}$, the average length of the 90% ADH confidence intervals for β was shorter than the Huber (about 6%) when there were 5 sample PSUs with $\bar{m} \geq 10$. For ρ=0.025 and $\sigma = \frac{2}{3}$, the average length of the 90% ADH confidence intervals for β was shorter than the Huber (5-8%) when there were 5 sample PSUs with all values of $\bar{m}$ and when c=50 with $\bar{m}$=10. There were no relevant differences, otherwise.

The proportions of samples where $H_0 : \sigma_b^2 = 0$ was rejected were relatively smaller for log-normal data than for normal data, regardless the value of the intraclass correlation, ρ and the skewness level, σ.

The Huber non-coverage rates were close to the nominal rate (10%) for all values of ρ and σ, as in Chapter 3.

For ρ=0, the LMM, ADM and ADH non-coverage rates were close to the nominal rate for both values of σ, as in Chapter 3, except when there were small number of sample PSUs (5 or less) with 2 observations per PSU.

For ρ=0.025 with $\sigma = \frac{1}{3}$, the non-coverage rates of the LMM, ADM and ADH confidence intervals were close to 10%, except when there were 2 sample PSUs with 10 or more observations per PSU, as in Chapter 3. The LMM and ADM non-coverage rates were higher than the nominal rate when there were 5 and 2 sample PSUs with average number of observations per PSU equal to 25. When $\sigma = \frac{2}{3}$, the non-coverage rates of the LMM, ADM and

ADH confidence intervals were close to 10%, except when there was a small number of sample PSUs (5 or less) with all all values of $\bar{m}$.

Similar to Chapter 3 all variance estimators for ρ=0.025 and both values of σ were approximately unbiased as all variance ratios were approximately 1, except that the LMM, ADM and ADH variance estimators tended to be biased when there were small numbers of sample PSUs (5 or less) with all average numbers of observations per PSU.

The proportions of samples where $H_0 : \sigma_b^2 = 0$ was rejected were higher than the nominal rate (10%), but they were lower for normal data in Chapter 3 for both values of ρ. Possible reasons why these proportions are higher than 10% are discussed in Subsection 3.4.4.

4.4 Conclusion

- Huber variance estimators were unbiased regardless of σ and ρ.
- Huber has close to the nominal non-coverage in all cases.
- For ρ=0 with both values of σ, LMM variance estimators have similar biases to Chapter 3.
- LMM variance estimators have smaller bias than in Chapter 3 when c=2, $\sigma = \frac{1}{3}$, and ρ=0.025.

- LMM variance estimators have larger bias than in Chapter 3 when c=2, $\sigma = \frac{2}{3}$, and ρ=0.025.

- ADM and ADH variance estimators have approximately the same bias as in Chapter 3, regardless the values of ρ and σ.

- When $c \leq 5$ and for all values of ρ and σ, ADM variance estimators have smaller biases than the LMM variance estimators, similar to what was in Chapter 3.

- In the unbalanced data designs, LMM, ADM and ADH variance estimators tended to be biased when $c \leq 5$ for all $\bar{m}$.

- LMM, ADM and ADH non-coverage rates were

  - close to the nominal rate when ρ=0 and $\sigma = \frac{1}{3}$.

  - significantly larger than the nominal rate when $c \leq 5$ with $m \leq 15$ when ρ=0 and $\sigma = \frac{2}{3}$, in the balanced data design. They were larger than 10%, in the unbalanced data designs when $c \leq 5$ with $\bar{m} = 2$.

  - were considerably different from 10% for all values of c, m and σ when $\rho = 0.025$, in the balanced data design. They were close to the nominal rate except when c=2 with $\bar{m} \geq 10$

- Log-normal data with $\sigma = \frac{1}{3}$ gave shorter confidence intervals than log-normal data with $\sigma = \frac{2}{3}$ in all cases.

- In comparing adaptive and non adaptive confidence intervals when ρ=0 and for both values of σ, in the balanced data designs:

  - ADM was 15-25% shorter than the LMM when c=2, 10-15% shorter when c=5 with $m \geq 15$.

  - ADH was 65-80% shorter than the Huber when c=2, 15-30% shorter when c=5 and 10-15% shorter when c=10 with $m \geq 5$.

- In comparing adaptive and non adaptive confidence intervals when ρ=0.025 and for both values of σ, in the balanced data designs:

  - ADM was 10-25% shorter than the LMM when c=2, 10-15% shorter when c=5 with $m \geq 15$.

  - ADH was 65-85% shorter than the Huber when c=2 and 10-20% shorter when c=5.

- In the unbalanced data designs

  - the ADM confidence intervals were 10-15% shorter than the LMM confidence intervals when c=2.

  - the ADH confidence intervals were 50-65% shorter than the Huber confidence intervals when c=2, for both values of ρ and σ. They were 5-8% when c=5, in general.

- Proportions of samples where $H_0 : \sigma_b^2 = 0$ is rejected were similar to those in Chapter 3 in the balanced data designs and relatively smaller in the unbalanced data designs.

Chapter 5

A Modified Adaptive Strategy based on the Estimated Design Effect

5.1 Introduction

The design effect is the ratio of the design variance of a statistic, $\hat{\boldsymbol{\beta}}$, to the variance under simple random sampling with the same sample size (Kish, 1965, p.162). For two-stage sampling with equal probability of selection at both stages, it can be approximated by $deff = 1 + (\bar{m} - 1)\rho$, where $\bar{m}$ is the average number of observations per sample PSU. One way of estimating the design effect is

$$\widehat{deff} = 1 + (\bar{m} - 1)\hat{\rho},$$

where $\hat{\rho}$ is obtained from a REML fit of the linear mixed model (2.3).

In Chapter 3, the adaptive strategies were defined based on the linear mixed model. Clustering was allowed for in the estimation of $var(\hat{\boldsymbol{\beta}})$ only if the PSU-level variance component (σ_b^2) was statistically significant. However, it is possible that the estimated intraclass correlation could be quite small, even if σ_b^2 is significant. In this case, it may still be preferable to ignore clustering when estimating $var(\hat{\boldsymbol{\beta}})$. Then ρ could be large and $\bar{m}$ is small so design effect is small. This chapter evaluates adaptive strategies along these lines.

In this chapter the adaptive strategies are based on the linear mixed model and normal data as in Chapter 3. The new adaptive strategies will be defined based on testing the null hypothesis $H_0 : \sigma_b^2 = 0$ and on comparing the estimated design effect to a cutoff value, d. If we reject the null hypothesis and, at the same time the estimated design effect $\widehat{deff}$ is larger than the cutoff point, d, the variance estimators are extracted from the linear mixed model or are estimated using the robust Huber-White variance estimator. Otherwise, the variance estimators are extracted from the linear model. Several cutoff points were evaluated. The flowchart in Figure 5.1 summarizes the two adaptive estimators of $var(\hat{\boldsymbol{\beta}})$: $\widehat{var}_{ADM}(\hat{\boldsymbol{\beta}})$ and $\widehat{var}_{ADH}(\hat{\boldsymbol{\beta}})$.

The two adaptive strategies (ADM) and (ADH) are defined as

$$\widehat{var}_{ADM}(\hat{\boldsymbol{\beta}}) = \begin{cases} \widehat{var}_{LMM}(\hat{\boldsymbol{\beta}}) & \text{if } H_0 \text{ is not retained,} \\ & \text{and } \widehat{deff} \geq d \\ \widehat{var}_{LM}(\hat{\boldsymbol{\beta}}) & \text{otherwise.} \end{cases} \tag{5.1}$$

$$\widehat{var}_{ADH}(\hat{\boldsymbol{\beta}}) = \begin{cases} \widehat{var}_{Hub}(\hat{\boldsymbol{\beta}}) & \text{if } H_0 \text{ is not retained,} \\ & \text{and } \widehat{deff} \geq d \\ \widehat{var}_{LM}(\hat{\boldsymbol{\beta}}) & \text{otherwise.} \end{cases} \tag{5.2}$$

This chapter is divided into three sections. Section 5.2 will evaluate the adaptive and other methods using cutoff values of d of 1.05 and 1.5 by simulation using balanced and unbalanced data cases. In Section 5.3 we will draw conclusions.

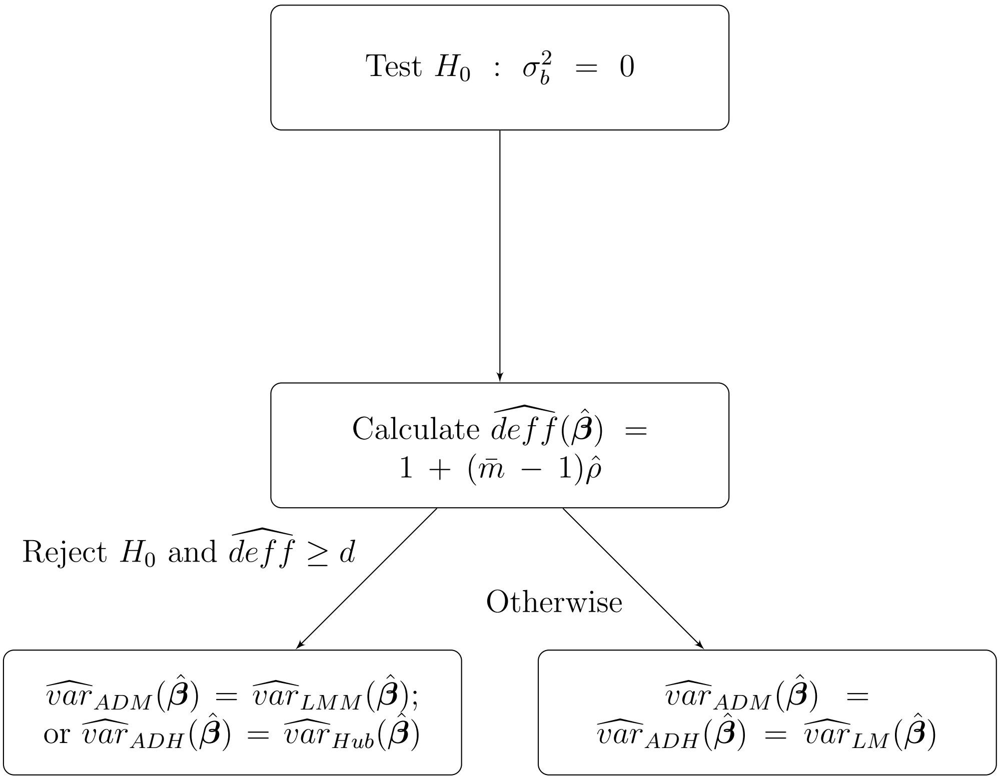


Figure 5.1: Flowchart showing the adaptive procedure based on testing $H_0 : \sigma_b^2 = 0$ and comparing $\widehat{deff}$ to a predetermined cutoff (d), using LMM-REML variance estimator or Huber-White variance estimator as an alternative

5.2 Simulation Study

The simulation study conducted in this chapter takes the balanced and unbalanced designs cases into consideration. In this study data were generated from intercept-only model defined in Equation (2.3). $H_0 : \sigma_b^2 = 0$ is tested using the restricted likelihood ratio test (2.30). The intraclass correlation (ρ) is estimated using Equation (2.7). The estimated intraclass correlation is then used to estimate the design effect for $\hat{\beta}$ defined by Equation (1.1). The estimated design effect with the RLRT, simultaneously, are used to define the adaptive strategies in equations (5.1) and (5.2). Values of d of 1.05 and 1.5 were used.

5.2.1 Simulation Study Using Balanced Data case

Similar to the simulation study conducted in Chapter 3, a simulation study was conducted in this chapter to compare the adaptive and non-adaptive methods for estimating $var(\hat{\beta})$ based on testing whether the PSU-level variance component is zero and at the same time estimating the design effect and comparing it to a cutoff value, d. The simulation study aimed to compare the effect of using the estimated design effect on the adaptive methods for estimating $var(\hat{\beta})$. The intercept only model (2.3) was used to generate the data for the simulation study, with $m_i = m$. The values of ρ, c and m were varied. The parameter ρ was varied over a range of values of 0, 0.025, 0.05

and 0.1; c takes the values 2, 5, 10 and 25; and m takes the range of values of 2, 5, 10, 15, 25 and 50. 1000 samples were generated in each case.

For each sample the estimated regression coefficients ($\hat{\beta}$) and the estimators of $var(\hat{\beta})$ were calculated for the LMM and LM models using the $lme4$ and lm packages (Pinheiro and Bates, 2000) in the R statistical environment (R Development Core Team, 2007). The Huber-White variance estimator of $\hat{\beta}$ was calculated as well by coding Equation (2.24) in R. The true variance of $\hat{\beta}$ was determined by calculating the variance over all 1000 simulations.

The two adaptive strategies ADM and ADH were as defined in Section 5.1. 90% confidence intervals of $\hat{\beta}$ were calculated for the LMM method using the method of Faes et al. (2009) as described in Subsection 2.2.3. Huber confidence intervals of $\hat{\beta}$ were calculated as discussed in Subsection 2.2.5, and the adaptive confidence intervals of $\hat{\beta}$ were calculated as discussed in Section 3.2.

Tables 5.1 - 5.4 show the ratio of the mean estimated variance of $\hat{\beta}$ using the four strategies of estimation (ADM, ADH, LMM and Huber) with values of ρ of 0 and 0.025, and values of d of 1.05 and 1.5. Results for other values of ρ and d are shown in Appendix D. In all tables we used $\beta = 0$ and significance level $\alpha = 0.1$ for testing $H_0 : \sigma_b^2 = 0$ and at the same time checking if $\widehat{deff} \geq d$. The tables show the non-coverage rates of 90% confidence intervals of β and the average lengths of these confidence intervals.

Table 5.1: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $\widehat{deff} \geq 1.05$ with ρ=0, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.05| Rej $H_0]$	$p[\widehat{deff}$ > 1.05 & Rej $H_0]$	$E(\widehat{deff})$	$p[\widehat{deff}$ > 1.05]	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.290	1.290	1.553	1.183	8.4	8.4	11.6	7.8	11.2	100.0	11.2	1.241	43.2	5.676	3.059	6.330	5.429
2	5	1.283	1.283	1.517	1.042	9.2	9.0	9.1	10.3	6.3	100.0	6.3	1.309	25.1	1.243	1.564	1.327	3.142
2	10	1.259	1.259	1.523	1.055	8.9	8.9	9.2	10.7	5.1	100.0	5.1	1.386	26.4	0.862	1.053	0.952	2.270
2	15	1.179	1.179	1.412	0.927	9.8	9.8	10.1	10.4	3.8	100.0	3.8	1.365	24.8	0.683	0.794	0.749	1.772
2	25	1.165	1.165	1.419	0.976	10.8	10.8	11.3	8.9	4.2	100.0	4.2	1.407	24.8	0.528	0.621	0.584	1.442
2	50	1.318	1.318	1.581	1.087	7.9	7.9	9.4	9.5	5.5	100.0	5.5	1.493	24.5	0.389	0.477	0.426	1.015
5	2	1.074	1.074	1.183	0.986	9.4	9.2	10.2	9.4	9.9	100.0	9.9	1.165	44.2	1.173	1.181	1.190	1.255
5	5	1.163	1.163	1.288	1.057	9.3	9.3	10.0	8.5	7.6	100.0	7.6	1.207	27.3	0.716	0.721	0.732	0.801
5	10	1.152	1.152	1.282	1.044	8.0	8.0	8.8	9.5	6.7	100.0	6.7	1.223	26.2	0.500	0.505	0.513	0.569
5	15	1.133	1.133	1.259	1.017	9.2	9.2	10.1	10.1	7.9	100.0	7.9	1.242	26.4	0.412	0.417	0.423	0.465
5	25	1.124	1.124	1.234	0.999	9.4	9.4	10.0	10.4	7.9	100.0	7.9	1.242	25.2	0.317	0.321	0.324	0.360
5	50	1.157	1.157	1.294	1.059	8.0	7.9	8.5	8.7	7.0	100.0	7.0	1.257	26.8	0.224	0.226	0.232	0.258
10	2	1.036	1.036	1.103	0.976	10.9	10.8	10.9	11.0	11.4	100.0	11.4	1.124	42.6	0.788	0.787	0.793	0.794
10	5	1.148	1.148	1.221	1.071	8.1	8.1	9.1	9.3	8.0	100.0	8.0	1.133	23.2	0.489	0.490	0.492	0.503
10	10	1.033	1.033	1.095	0.972	10.8	11.0	10.9	10.6	8.5	100.0	8.5	1.145	21.7	0.347	0.348	0.348	0.361
10	15	1.205	1.205	1.282	1.116	7.2	7.1	8.9	9.1	8.6	100.0	8.6	1.151	23.2	0.282	0.283	0.285	0.292
10	25	1.203	1.203	1.268	1.103	7.3	7.3	7.9	9.4	8.3	100.0	8.3	1.147	20.0	0.219	0.219	0.219	0.225
10	50	1.137	1.137	1.209	1.045	9.7	9.7	10.0	10.7	8.0	100.0	8.0	1.149	21.7	0.154	0.154	0.155	0.159
25	2	0.950	0.950	0.994	0.920	11.4	11.4	11.7	12.3	10.1	100.0	10.1	1.082	42.0	0.483	0.483	0.483	0.482
25	5	0.956	0.956	0.963	0.913	10.6	10.7	11.9	11.1	8.1	100.0	8.1	1.050	10.6	0.303	0.302	0.298	0.302
25	10	1.038	1.038	1.051	0.984	10.6	10.6	11.1	11.5	8.7	100.0	8.7	1.058	12.2	0.214	0.214	0.212	0.214
25	15	1.013	1.013	1.025	0.961	9.5	9.4	10.8	10.9	6.8	100.0	6.8	1.049	10.2	0.174	0.173	0.171	0.173
25	25	1.123	1.123	1.137	1.069	8.4	8.4	8.8	9.0	9.1	100.0	9.1	1.068	12.8	0.135	0.135	0.134	0.136
25	50	1.045	1.046	1.068	0.971	9.4	9.4	10.3	10.3	8.8	100.0	8.8	1.076	16.0	0.096	0.096	0.095	0.095

Table 5.2: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $\widehat{deff} \geq 1.05$ with ρ=0.025, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff} >1.05 \mid \text{Rej } H_0]$	$p[\widehat{deff} > 1.05\ \&\ \text{Rej } H_0]$	$E(\widehat{deff})$	$p[\widehat{deff} > 1.05]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.189	1.189	1.440	1.074	8.2	8.2	12.0	8.9	9.9	100.0	9.9	1.234	42.5	4.934	2.977	5.552	5.299
2	5	1.182	1.182	1.415	1.012	11.4	10.9	11.1	12.6	7.5	100.0	7.5	1.381	29.5	1.330	1.693	1.437	3.330
2	10	1.139	1.139	1.356	0.985	11.4	11.4	12.9	11.6	7.5	100.0	7.5	1.475	27.9	0.915	1.211	1.009	2.413
2	15	1.005	1.005	1.227	0.947	15.5	15.5	13.9	8.5	8.6	100.0	8.6	1.630	33.2	0.780	1.034	0.880	2.167
2	25	1.074	1.074	1.273	1.058	18.6	18.6	16.1	8.7	14.4	100.0	14.4	1.985	38.6	0.710	1.054	0.800	1.911
2	50	0.886	0.886	1.056	0.911	23.4	23.4	19.9	9.3	16.9	100.0	16.9	2.291	43.7	0.548	0.835	0.623	1.469
5	2	1.123	1.123	1.245	1.061	9.6	9.6	8.5	9.2	11.9	100.0	11.9	1.190	48.7	1.224	1.230	1.254	1.323
5	5	0.986	0.986	1.102	0.941	10.8	10.9	11.2	10.8	11.2	100.0	11.2	1.276	33.2	0.751	0.759	0.780	0.854
5	10	0.967	0.967	1.093	0.975	12.4	12.1	12.1	8.7	14.7	100.0	14.7	1.398	39.9	0.543	0.551	0.572	0.644
5	15	1.080	1.080	1.233	1.118	11.3	11.3	10.1	8.8	17.1	100.0	17.1	1.487	44.2	0.457	0.465	0.487	0.549
5	25	0.895	0.895	1.007	0.945	15.8	15.8	13.6	10.7	24.7	100.0	24.7	1.691	51.6	0.383	0.392	0.409	0.462
5	50	0.812	0.812	0.888	0.861	19.3	19.3	16.3	11.6	40.2	100.0	40.2	2.185	66.1	0.320	0.330	0.339	0.376
10	2	1.044	1.044	1.125	1.008	9.0	9.1	8.2	9.5	11.9	100.0	11.9	1.143	49.6	0.809	0.812	0.819	0.830
10	5	0.953	0.953	1.016	0.948	11.6	11.6	12.0	10.2	12.5	100.0	12.5	1.179	28.4	0.507	0.507	0.514	0.537
10	10	1.045	1.045	1.121	1.062	11.0	10.9	11.5	9.6	18.6	100.0	18.6	1.275	36.4	0.370	0.371	0.378	0.398
10	15	0.935	0.935	1.008	0.975	13.1	13.0	11.4	10.5	25.8	100.0	25.8	1.400	45.2	0.318	0.318	0.327	0.344
10	25	1.002	1.002	1.072	1.061	11.6	11.5	11.4	8.6	40.5	100.0	40.5	1.639	61.2	0.269	0.270	0.279	0.292
10	50	0.995	0.995	1.029	1.030	14.1	13.3	12.8	10.3	67.1	100.0	67.1	2.231	81.8	0.231	0.232	0.237	0.243
25	2	1.018	1.018	1.066	0.997	9.6	9.7	9.9	9.7	12.7	100.0	12.7	1.092	45.2	0.489	0.489	0.490	0.491
25	5	1.007	1.007	1.022	1.017	10.5	10.6	10.7	10.2	17.2	100.0	17.2	1.107	22.3	0.316	0.314	0.314	0.323
25	10	0.991	0.991	1.008	1.034	11.7	11.6	12.1	10.1	29.0	100.0	29.0	1.202	34.7	0.232	0.231	0.232	0.242
25	15	0.990	0.990	1.009	1.038	11.3	11.3	11.0	9.6	42.3	100.0	42.3	1.309	49.2	0.199	0.197	0.200	0.206
25	25	1.049	1.049	1.063	1.087	10.8	11.1	10.8	9.7	66.0	100.0	66.0	1.580	71.8	0.171	0.169	0.172	0.175
25	50	0.947	0.948	0.953	0.954	11.0	10.7	10.9	10.7	92.7	100.0	92.7	2.214	96.2	0.144	0.144	0.145	0.145

Table 5.3: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $\widehat{deff} \geq 1.5$ with ρ=0, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.5$\vert$ Rej $H_0]$	$p[\widehat{deff}$ > 1.5 & Rej $H_0]$	$E(\widehat{deff})$	$p[\widehat{deff}$ > 1.5$]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.290	1.290	1.553	1.183	8.4	8.4	11.6	7.8	11.2	100.0	11.2	1.241	24.3	5.676	3.059	6.330	5.429
2	5	1.283	1.283	1.517	1.042	9.2	9.0	9.1	10.3	6.3	100.0	6.3	1.309	21.7	1.243	1.564	1.327	3.142
2	10	1.259	1.259	1.523	1.055	8.9	8.9	9.2	10.7	5.1	100.0	5.1	1.386	23.2	0.862	1.053	0.952	2.270
2	15	1.179	1.179	1.412	0.927	9.8	9.8	10.1	10.4	3.8	100.0	3.8	1.365	20.9	0.683	0.794	0.749	1.772
2	25	1.165	1.165	1.419	0.976	10.8	10.8	11.3	8.9	4.2	100.0	4.2	1.407	21.9	0.528	0.621	0.584	1.442
2	50	1.318	1.318	1.581	1.087	7.9	7.9	9.4	9.5	5.5	100.0	5.5	1.493	22.2	0.389	0.477	0.426	1.015
5	2	1.074	1.074	1.183	0.986	9.4	9.2	10.2	9.4	9.9	100.0	9.9	1.165	12.0	1.173	1.181	1.190	1.255
5	5	1.163	1.163	1.288	1.057	9.3	9.3	10.0	8.5	7.6	100.0	7.6	1.207	18.8	0.716	0.721	0.732	0.801
5	10	1.152	1.152	1.282	1.044	8.0	8.0	8.8	9.5	6.7	100.0	6.7	1.223	19.3	0.500	0.505	0.513	0.569
5	15	1.133	1.133	1.259	1.017	9.2	9.2	10.1	10.1	7.9	100.0	7.9	1.242	20.3	0.412	0.417	0.423	0.465
5	25	1.124	1.124	1.234	0.999	9.4	9.4	10.0	10.4	7.9	100.0	7.9	1.242	19.2	0.317	0.321	0.324	0.360
5	50	1.157	1.157	1.294	1.059	8.0	7.9	8.5	8.7	7.0	100.0	7.0	1.257	20.6	0.224	0.226	0.232	0.258
10	2	1.009	1.009	1.103	0.976	11.2	11.2	10.9	11.0	11.4	49.1	5.6	1.124	5.6	0.775	0.774	0.793	0.794
10	5	1.148	1.148	1.221	1.071	8.1	8.1	9.1	9.3	8.0	100.0	8.0	1.133	11.3	0.489	0.490	0.492	0.503
10	10	1.033	1.033	1.095	0.972	10.8	11.0	10.9	10.6	8.5	100.0	8.5	1.145	14.1	0.347	0.348	0.348	0.361
10	15	1.205	1.205	1.282	1.116	7.2	7.1	8.9	9.1	8.6	100.0	8.6	1.151	14.6	0.282	0.283	0.285	0.292
10	25	1.203	1.203	1.268	1.103	7.3	7.3	7.9	9.4	8.3	100.0	8.3	1.147	14.5	0.219	0.219	0.219	0.225
10	50	1.137	1.137	1.209	1.045	9.7	9.7	10.0	10.7	8.0	100.0	8.0	1.149	14.7	0.154	0.154	0.155	0.159
25	2	0.920	0.920	0.994	0.920	11.6	11.6	11.7	12.3	10.1	5.0	0.5	1.082	0.5	0.475	0.475	0.483	0.482
25	5	0.938	0.938	0.963	0.913	11.0	11.0	11.9	11.1	8.1	44.4	3.6	1.050	3.6	0.299	0.299	0.298	0.302
25	10	1.019	1.019	1.051	0.984	10.7	10.7	11.1	11.5	8.7	47.1	4.1	1.058	4.1	0.212	0.212	0.212	0.214
25	15	0.998	0.998	1.025	0.961	9.6	9.6	10.8	10.9	6.8	50.0	3.4	1.049	3.4	0.172	0.172	0.171	0.173
25	25	1.105	1.105	1.137	1.069	8.7	8.6	8.8	9.0	9.1	59.3	5.4	1.068	5.4	0.134	0.134	0.134	0.136
25	50	1.034	1.034	1.068	0.971	9.6	9.6	10.3	10.3	8.8	70.5	6.2	1.076	6.2	0.095	0.095	0.095	0.095

Table 5.4: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $\widehat{deff} \geq 1.5$ with ρ=0.025, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff} > 1.5 \mid$ Rej $H_0]$	$p[\widehat{deff} > 1.5$ & Rej $H_0]$	$E(\widehat{deff})$	$p[\widehat{deff} > 1.5]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.189	1.189	1.440	1.074	8.2	8.2	12.0	8.9	9.9	100.0	9.9	1.234	24.4	4.934	2.977	5.552	5.299
2	5	1.182	1.182	1.415	1.012	11.4	10.9	11.1	12.6	7.5	100.0	7.5	1.381	25.0	1.330	1.693	1.437	3.330
2	10	1.139	1.139	1.356	0.985	11.4	11.4	12.9	11.6	7.5	100.0	7.5	1.475	24.7	0.915	1.211	1.009	2.413
2	15	1.005	1.005	1.227	0.947	15.5	15.5	13.9	8.5	8.6	100.0	8.6	1.630	29.1	0.780	1.034	0.880	2.167
2	25	1.074	1.074	1.273	1.058	18.6	18.6	16.1	8.7	14.4	100.0	14.4	1.985	35.8	0.710	1.054	0.800	1.911
2	50	0.886	0.886	1.056	0.911	23.4	23.4	19.9	9.3	16.9	100.0	16.9	2.291	40.6	0.548	0.835	0.623	1.469
5	2	1.123	1.123	1.245	1.061	9.6	9.6	8.5	9.2	11.9	100.0	11.9	1.190	14.7	1.224	1.230	1.254	1.323
5	5	0.986	0.986	1.102	0.941	10.8	10.9	11.2	10.8	11.2	100.0	11.2	1.276	24.8	0.751	0.759	0.780	0.854
5	10	0.967	0.967	1.093	0.975	12.4	12.1	12.1	8.7	14.7	100.0	14.7	1.398	31.5	0.543	0.551	0.572	0.644
5	15	1.080	1.080	1.233	1.118	11.3	11.3	10.1	8.8	17.1	100.0	17.1	1.487	35.6	0.457	0.465	0.487	0.549
5	25	0.895	0.895	1.007	0.945	15.8	15.8	13.6	10.7	24.7	100.0	24.7	1.691	44.6	0.383	0.392	0.409	0.462
5	50	0.812	0.812	0.888	0.861	19.3	19.3	16.3	11.6	40.2	100.0	40.2	2.185	60.5	0.320	0.330	0.339	0.376
10	2	1.015	1.015	1.125	1.008	9.4	9.5	8.2	9.5	11.9	49.6	5.9	1.143	5.9	0.797	0.798	0.819	0.830
10	5	0.953	0.953	1.016	0.948	11.6	11.6	12.0	10.2	12.5	100.0	12.5	1.179	17.4	0.507	0.507	0.514	0.537
10	10	1.045	1.045	1.121	1.062	11.0	10.9	11.5	9.6	18.6	100.0	18.6	1.275	25.8	0.370	0.371	0.378	0.398
10	15	0.935	0.935	1.008	0.975	13.1	13.0	11.4	10.5	25.8	100.0	25.8	1.400	36.8	0.318	0.318	0.327	0.344
10	25	1.002	1.002	1.072	1.061	11.6	11.5	11.4	8.6	40.5	100.0	40.5	1.639	51.8	0.269	0.270	0.279	0.292
10	50	0.995	0.995	1.029	1.030	14.1	13.3	12.8	10.3	67.1	100.0	67.1	2.231	75.2	0.231	0.232	0.237	0.243
25	2	0.978	0.978	1.066	0.997	10.3	10.3	9.9	9.7	12.7	5.5	0.7	1.092	0.7	0.478	0.478	0.490	0.491
25	5	0.968	0.968	1.022	1.017	10.9	11.0	10.7	10.2	17.2	46.5	8.0	1.107	8.0	0.308	0.308	0.314	0.323
25	10	0.955	0.955	1.008	1.034	12.4	12.5	12.1	10.1	29.0	66.2	19.2	1.202	19.2	0.227	0.226	0.232	0.242
25	15	0.949	0.949	1.009	1.038	12.5	12.3	11.0	9.6	42.3	71.6	30.3	1.309	30.3	0.193	0.192	0.200	0.206
25	25	1.022	1.022	1.063	1.087	11.3	11.6	10.8	9.7	66.0	86.7	57.2	1.580	57.2	0.167	0.166	0.172	0.175
25	50	0.942	0.943	0.953	0.954	11.3	10.9	10.9	10.7	92.7	97.2	90.1	2.214	90.1	0.143	0.143	0.145	0.145

The proportions of samples where $H_0 : \sigma_b^2 = 0$ is rejected are also shown, as well as the proportions of samples where $H_0 : \sigma_b^2 = 0$ is rejected and at the same time $\widehat{deff} \geq d$, $p\left[\widehat{deff} \geq d\text{\& Rej } H_0\right]$, and the proportion of samples where $\widehat{deff} \geq d$ given that $H_0 : \sigma_b^2 = 0$ is rejected, $p\left[\widehat{deff} \geq d|\text{Rej H}_0\right]$, are also shown.

Tables 5.1 and 5.2 showed the simulation results for the cutoff d=1.05 with both values of ρ. They showed that $p\left[\widehat{deff} \geq d|\text{Rej H}_0\right]$ was 100% for all designs. They showed that the variance ratios, non-coverage rates and average lengths of 90% confidence intervals were perfectly identical to the variance ratios, non-coverage rates and average lengths of 90% confidence intervals in Chapter 3.

Tables 5.3 and 5.4 showed the simulation results for the cutoff d=1.5 with both values of ρ. They showed that $p\left[\widehat{deff} \geq d|\text{Rej H}_0\right]$ was 100% for designs with $c \leq 10$, except in designs with c=10 with m=2. In these designs they showed that $p\left[\widehat{deff} \geq d|\text{Rej H}_0\right]$ was 100% for all designs. They showed that the variance ratios, non-coverage rates and average lengths of 90% confidence intervals were perfectly identical to the variance ratios, non-coverage rates and average lengths of 90% confidence intervals in Chapter 3. In designs with c=25, $p\left[\widehat{deff} \geq d|\text{Rej H}_0\right]$ was less than 100% for both values of ρ. Therefore, the variance ratios, non-coverage rates and average lengths of 90% confidence intervals were different from the variance ratios, non-coverage rates and average lengths of 90% confidence intervals in Chapter 3. There was no relevant differences in these designs.

$P[\widehat{deff} > d|rejH_0]$ was 100% or close to it almost all the time because H_0 is only rejected when $\hat{\rho}$ is reasonably large. Also, H_0 tends to be rejected when there is sufficient data on ρ, which occurs when neither c nor m are too small. As a result, the cases when H_0 is rejected are also the cases when $\widehat{deff}$ is large, so that $P[\widehat{deff} > d|rejH_0] \approx 1$. This means that applying the cutoff to the design effect has no effect, so that the results are very close or identical to those in Chapter 3.

5.2.2 Simulation Study Using Unbalanced Data case

A simulation study was conducted based on unequal PSU sizes to see the effect of using the estimated design effect on the adaptive strategies of estimating the variances of $\hat{\beta}$. Data were generated from model (2.3), with different PSU sizes, m_i. The values of ρ and c were varied. 1000 samples were generated in each case. The values of $\bar{m}$ were varied to be 3, 10 and 25. For this purpose three cases were used. In case 1, the number of observations was generated to be between 2 and 4 with average equal to 3 observations per PSU. In case 2, this number was varied from 5 to 15, with average equal to 10. Finally, in case 3, the average was 25, therefore the number of observations was varied from 15 to 35. Several cutoff values were used to define

Table 5.5: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.05 using RLRT in the unbalanced data case with ρ=0

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.05| Rej $H_0]$	$p[\widehat{deff}$ > 1.05 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.380	1.380	1.493	1.099	8.3	8.2	7.6	10.8	19.1	100.0	19.1	2.254	2.916	2.402	4.385
2	10	1.391	1.391	1.460	1.018	8.9	8.9	8.1	10.2	15.9	100.0	15.9	0.938	1.512	0.966	2.332
2	25	1.310	1.310	1.401	0.953	8.2	8.1	6.9	9.9	13.5	100.0	13.5	0.542	0.873	0.562	1.433
5	3	1.240	1.240	1.253	1.037	8.2	7.7	7.9	10.9	27.1	100.0	27.1	1.021	1.069	1.027	1.066
5	10	1.213	1.213	1.217	0.972	8.2	7.7	8.2	8.9	24.8	100.0	24.8	0.526	0.567	0.527	0.568
5	25	1.184	1.184	1.189	0.973	8.1	7.6	8.1	9.3	23.0	100.0	23.0	0.326	0.353	0.327	0.365
10	3	1.164	1.164	1.165	1.038	8.7	8.4	8.7	10.2	25.5	100.0	25.5	0.648	0.660	0.649	0.648
10	10	1.129	1.129	1.129	0.973	8.4	8.0	8.4	10.1	21.3	100.0	21.3	0.354	0.363	0.354	0.354
10	25	1.206	1.206	1.206	1.053	8.1	8.1	8.1	10.1	23.2	100.0	23.2	0.223	0.230	0.223	0.226
25	3	1.050	1.050	1.050	1.001	10.5	10.4	10.5	10.0	15.4	100.0	15.4	0.395	0.397	0.395	0.394
25	10	1.126	1.126	1.126	1.058	8.6	8.5	8.6	8.5	13.0	100.0	13.0	0.215	0.216	0.215	0.215
25	25	1.125	1.125	1.125	1.050	8.5	8.5	8.5	10.3	11.6	100.0	11.6	0.136	0.136	0.136	0.135
50	3	0.952	0.952	0.952	0.964	11.3	11.3	11.3	10.6	11.9	100.0	11.9	0.279	0.279	0.279	0.283
50	10	0.936	0.936	0.936	0.991	11.4	11.3	11.4	9.8	35.8	100.0	35.8	0.162	0.163	0.162	0.169
50	25	1.008	1.008	1.008	1.018	10.2	9.8	10.2	9.5	88.6	100.0	88.6	0.119	0.121	0.119	0.122

Table 5.6: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.05 using RLRT in the unbalanced data case with ρ=0.025

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.05| Rej $H_0]$	$p[\widehat{deff}$ > 1.05 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.343	1.343	1.454	1.109	9.3	9.2	8.2	10.1	19.7	100.0	19.7	2.339	3.044	2.502	4.625
2	10	1.291	1.291	1.365	1.031	11.0	10.8	9.6	9.8	19.8	100.0	19.8	1.017	1.761	1.054	2.602
2	25	1.141	1.141	1.195	0.982	15.1	14.9	13.9	8.9	27.0	100.0	27.0	0.637	1.347	0.657	1.890
5	3	1.220	1.220	1.233	1.046	7.8	7.4	7.7	9.1	30.9	100.0	30.9	1.048	1.103	1.055	1.109
5	10	1.096	1.097	1.102	0.950	10.3	8.8	10.1	9.2	34.8	100.0	34.8	0.560	0.619	0.561	0.631
5	25	1.008	1.009	1.011	0.939	13.7	11.6	13.6	10.7	49.1	100.0	49.1	0.383	0.446	0.384	0.460
10	3	1.137	1.137	1.138	1.037	10.1	9.8	10.1	11.5	30.1	100.0	30.1	0.665	0.678	0.665	0.671
10	10	1.069	1.070	1.069	1.007	10.1	9.5	10.1	10.0	38.4	100.0	38.4	0.381	0.397	0.381	0.401
10	25	1.028	1.028	1.028	1.007	12.0	10.0	12.0	10.6	61.0	100.0	61.0	0.268	0.287	0.268	0.291
25	3	1.011	1.011	1.011	0.991	10.2	10.2	10.2	10.3	20.4	100.0	20.4	0.404	0.406	0.404	0.409
25	10	1.057	1.057	1.057	1.075	10.1	9.8	10.1	9.4	38.9	100.0	38.9	0.233	0.237	0.233	0.243
25	25	0.973	0.973	0.973	0.992	12.1	11.6	12.1	10.5	71.9	100.0	71.9	0.166	0.171	0.166	0.174
50	3	0.952	0.952	0.952	0.964	11.3	11.3	11.3	10.6	11.9	100.0	11.9	0.279	0.279	0.279	0.283
50	10	0.936	0.936	0.936	0.991	11.4	11.3	11.4	9.8	35.8	100.0	35.8	0.162	0.163	0.162	0.169
50	25	1.008	1.008	1.008	1.018	10.2	9.8	10.2	9.5	88.6	100.0	88.6	0.119	0.121	0.119	0.122

Table 5.7: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.5 using RLRT in the unbalanced data case with ρ=0

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.5|	$p[\widehat{deff}$ > 1.5 &	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub		Rej H_0]	Rej H_0]	ADM	ADH	LMM	Hub
2	3	1.380	1.380	1.493	1.099	8.3	8.2	7.6	10.8	19.1	100.0	19.1	2.254	2.916	2.402	4.385
2	10	1.391	1.391	1.460	1.018	8.9	8.9	8.1	10.2	15.9	100.0	15.9	0.938	1.512	0.966	2.332
2	25	1.310	1.310	1.401	0.953	8.2	8.1	6.9	9.9	13.5	100.0	13.5	0.542	0.873	0.562	1.433
5	3	1.205	1.205	1.253	1.037	8.4	8.1	7.9	10.9	27.1	70.8	19.2	1.003	1.036	1.027	1.066
5	10	1.196	1.196	1.217	0.972	8.4	8.1	8.2	8.9	24.8	82.7	20.5	0.522	0.557	0.527	0.568
5	25	1.166	1.166	1.189	0.973	8.3	7.8	8.1	9.3	23.0	81.7	18.8	0.324	0.346	0.327	0.365
10	3	1.103	1.103	1.165	1.038	9.8	9.8	8.7	10.2	25.5	36.1	9.2	0.630	0.634	0.649	0.648
10	10	1.094	1.094	1.129	0.973	8.7	8.7	8.4	10.1	21.3	60.1	12.8	0.348	0.354	0.354	0.354
10	25	1.175	1.175	1.206	1.053	8.3	8.3	8.1	10.1	23.2	68.1	15.8	0.220	0.225	0.223	0.226
25	3	1.006	1.006	1.050	1.001	10.8	10.8	10.5	10.0	15.4	18.2	2.8	0.387	0.388	0.395	0.394
25	10	1.097	1.097	1.126	1.058	9.1	9.0	8.6	8.5	13.0	45.4	5.9	0.213	0.213	0.215	0.215
25	25	1.099	1.099	1.125	1.050	8.6	8.6	8.5	10.3	11.6	47.4	5.5	0.134	0.134	0.136	0.135
50	3	0.969	0.969	0.992	0.970	10.6	10.6	10.2	10.3	0.0	0.0	0.0	0.270	0.270	0.273	0.273
50	10	1.022	1.022	1.042	1.015	9.5	9.5	9.1	9.9	6.6	22.7	1.5	0.148	0.148	0.149	0.150
50	25	0.996	0.996	1.027	0.990	9.8	9.8	9.5	9.7	10.9	13.8	1.5	0.093	0.093	0.095	0.094

Table 5.8: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.5 using RLRT in the unbalanced data case with ρ=0.025

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.5|	$p[\widehat{deff}$ > 1.5 &	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub		Rej $H_0]$	Rej $H_0]$	ADM	ADH	LMM	Hub
2	3	1.343	1.343	1.454	1.109	9.3	9.2	8.2	10.1	19.7	100.0	19.7	2.339	3.044	2.502	4.625
2	10	1.291	1.291	1.365	1.031	11.0	10.8	9.6	9.8	19.8	100.0	19.8	1.017	1.761	1.054	2.602
2	25	1.141	1.141	1.195	0.982	15.1	14.9	13.9	8.9	27.0	100.0	27.0	0.637	1.347	0.657	1.890
5	3	1.185	1.185	1.233	1.046	7.9	7.8	7.7	9.1	30.9	73.5	22.7	1.029	1.069	1.055	1.109
5	10	1.080	1.080	1.102	0.950	10.6	9.5	10.1	9.2	34.8	86.8	30.2	0.555	0.607	0.561	0.631
5	25	0.997	0.997	1.011	0.939	14.3	12.6	13.6	10.7	49.1	90.4	44.4	0.381	0.438	0.384	0.460
10	3	1.074	1.074	1.138	1.037	10.9	10.9	10.1	11.5	30.1	43.2	13.0	0.645	0.651	0.665	0.671
10	10	1.031	1.031	1.069	1.007	11.5	11.0	10.1	10.0	38.4	69.3	26.6	0.373	0.385	0.381	0.401
10	25	1.006	1.006	1.028	1.007	12.6	10.9	12.0	10.6	61.0	85.4	52.1	0.264	0.281	0.268	0.291
25	3	0.957	0.957	1.011	0.991	11.1	11.1	10.2	10.3	20.4	20.6	4.2	0.394	0.394	0.404	0.409
25	10	0.998	0.998	1.057	1.075	10.9	10.7	10.1	9.4	38.9	55.8	21.7	0.227	0.229	0.233	0.243
25	25	0.937	0.937	0.973	0.992	13.1	12.8	12.1	10.5	71.9	80.0	57.5	0.162	0.166	0.166	0.174
50	3	0.917	0.917	0.952	0.964	12.0	12.0	11.3	10.6	11.9	5.0	0.6	0.274	0.274	0.279	0.283
50	10	0.877	0.877	0.936	0.991	13.2	13.2	11.4	9.8	35.8	46.9	16.8	0.157	0.157	0.162	0.169
50	25	0.953	0.953	1.008	1.018	11.4	11.0	10.2	9.5	88.6	74.7	66.2	0.115	0.117	0.119	0.122

the adaptive strategies as well as in the balanced data case, where d takes the values of 1.05, 1.1, 1.2 and 1.5. The results related to values of d of 1.1 and 1.2 are shown in Appendix D.

Tables 5.5 - 5.8 show the results of the simulation study for the unbalanced data case with two cutoff values, d=1.05 and 1.5. They show the ratio of the mean estimated variance of $\hat{\beta}$ using the four strategies of estimation (ADM, ADH, LMM and Huber) with a range of values of ρ of 0 and 0.025 for both cutoff values. Similar to what was done in Chapter 3, in all tables we used $\beta = 0$ and significance level $\alpha = 0.1$ for testing $\sigma_b^2 = 0$ and comparing the estimated design effect $\widehat{deff}$ to a cutoff value d. The tables show the non-coverage rates of 90% confidence intervals for β as well as the average lengths of these confidence intervals. The proportion of samples where $H_0 : \sigma_b^2 = 0$ is rejected, $H_0 : \sigma_b^2 = 0$ is rejected and at the same time $\widehat{deff} \geq d$ and the proportion of samples where $\widehat{deff} \geq d$ given that $H_0 : \sigma_b^2 = 0$ is rejected are also shown.

Tables 5.5 and 5.6 show the results for the cutoff value d=1.05 with both values of ρ. They show that $p\big[\widehat{deff} \geq d|\text{Rej H}_0\big]$ was 100% for all designs. They showed that the variance ratios, non-coverage rates and average lengths of 90% confidence intervals were perfectly identical to the variance ratios, non-coverage rates and average lengths of 90% confidence intervals in Chapter 3.

Tables 5.7 and 5.8 show the results for the cutoff value d=1.5 with both values of ρ. They showed that all simulation results were identical to simulation results in Chapter 3 in designs with c=2 with all values of $\bar{m}$ and for both values of ρ. When $c \geq 5$, $p\left[\widehat{deff} \geq d|\text{Rej H}_0\right]$ was less than 100%. Hence the adaptive variance estimators were smaller than the adaptive variance estimators in Chapter 3. They were less biased in designs with c=5 and 10 with $\bar{m} \leq 10$ and 25, respectively. The adaptive non-coverage rates were close to the nominal rate as in Chapter 3 when ρ=0. The ADM non-coverage rates were close to the nominal rate when ρ=0.025, except in designs with c=5, 10 and 25 with $\bar{m}$ = 25 and designs with c=50 with $\bar{m} \leq 10$. The ADH non-coverage rates were close to the nominal rate when ρ=0.025, except in designs with c=5 and 25 with $\bar{m}$ = 25 and in designs with c=50 with $\bar{m} \leq 10$. In these designs there was no relevant difference in the average lengths of the 90% adaptive confidence intervals from the average lengths of the 90% adaptive confidence intervals in Chapter 3.

5.3 Conclusions

The variance ratios, non-coverage rates and average lengths of 90% confidence intervals were perfectly identical to the variance ratios, non-coverage rates and average lengths of 90% confidence intervals in Chapter 3 in all designs when the cutoff value d=1.05, because $p\left[\widehat{deff} \geq d|\text{Rej H}_0\right]$=1 in all of these designs. When the cutoff value d=1.5, the variance ratios, non-coverage rates and average lengths of 90% confidence intervals were perfectly identical in balance designs with $c \leq 10$ and unbalanced designs with c=2.

In the unbalanced designs, the adaptive variance estimators were less biased than the adaptive variance estimators in Chapter 3 in designs with c=5 and $\bar{m} \leq 10$ and designs with c=10 and $\bar{m}$=25. For ρ=0.025, the adaptive non-coverage rates were closer to the nominal rate than in Chapter 3 when c=5, 10 and 25 with $\bar{m} = 25$ and when c=50 with $\bar{m} \leq 10$. Lengths of adaptive confidence intervals were relatively similar to lengths of adaptive confidence intervals in Chapter 3.

For balanced sampling, including a cutoff of 1.05 or 1.5 for the estimated $deff$ has no effect on adaptive strategies. Larger values for the cutoff may be worth evaluating in future research, but it seems unlikely that the approach will give useful benefits. When there are unequal sample sizes, there is a small benefit in including a cutoff of 1.5 for the estimated $deff$, perhaps because the restricted likelihood ratio test has a high type-I error rate for unbalanced data.

Chapter 6

Adaptive Design Using a Pilot Survey

6.1 Introduction

A pilot survey is a small survey conducted prior to a survey, in order to trial the operations, instrument design and possibly sample design for the main survey (Stopher and Metcalf, 1996, Chapter 4).

Pilot surveys are an important step in running a successful survey (Teijlingen and Hundley, 2002). They can save time and money by giving advance warning about the points where the main survey could fail (Teijlingen and Hundley, 2002). They should provide enough data for the researcher or survey manager to decide whether to continue with the main survey. They reduce the number of unexpected problems because there is an opportunity to redesign the main survey to be conducted according to the results revealed by the pilot survey (Skinner et al., 2007).

The number of units to select from each PSU is an important decision that has to be made in developing the design of a two-stage survey. A common approach is to assume a simple cost model such as (1.3). For two-stage sampling designs, assuming equal sample sizes from each PSU and simple random sampling at both stages, the optimal choice of number of observations per PSU is $m_{opt} = \sqrt{\frac{C_1}{C_2}\frac{1-\rho}{\rho}}$ (Hansen et al., 1953, p.286) where ρ is the intraclass correlation and C_1 and C_2 are the parameters of the cost model (1.3).

To develop the design, a value of ρ has to be assumed or estimated. One way to do this is to conduct a pilot survey. However, estimates of ρ are often quite small, for example 0.01 or 0.02 in human studies (Killip et al., 2004). When ρ is small even small changes to the assumed value can affect m_{opt}. The intraclass correlation is often quite small. It is 1 when there is perfect homogeneity within PSU. It can be negative when there is extreme heterogeneity within PSUs with smallest possible value of ρ equal to $-1/(M-1)$ (Hansen et al., 1953, p.260). When PSUs are geographic areas and final units are households in these areas, it is generally less than 0.1 (Verma et al., 1980). It is typically between 0 and 0.2, when PSUs are households and final units are people in households (Clark and Steel, 2002). Small values will lead to a large within PSU sample size (Steel and Clark, 2006). When the true ρ is small, the estimated $\hat{\rho}$ in multilevel analysis is often equal to 0 (Muthén and Satorra, 1995). Estimates of $\hat{\rho}$ calculated from a pilot survey would often be highly variable, given the small sample usually selected for pilot surveys.

In this chapter, it is assumed that the intraclass correlation is estimated from pilot survey data. It is then used to estimate the optimal sample PSU size based on minimizing the variance of the sample mean subject for fixed total cost.

Figure 6.1: Histograms of $\hat{\rho}$ from 1000 simulations when ρ=0.025

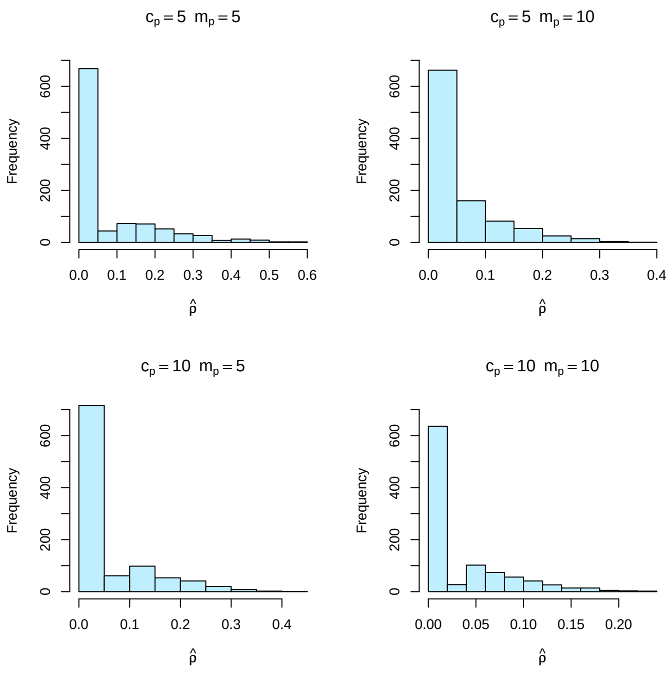


Figure 6.1 shows how variable $\hat{\rho}$ can be for typical pilot sample sizes. The distributions of $\hat{\rho}$ are shown for different numbers of sample PSUs, c_p, and

units per sample PSU, m_p, based on 1000 simulated data sets from model (2.3) with no covariates and ρ=0.025. The Figure shows that $\hat{\rho}$ is zero more than 70% of the time and even when nonzero is often much smaller than 0.025.

When $\hat{\rho}$ equals to zero, Equation (1.4) for m_{opt} cannot be applied. In this case, the optimal design involves setting m to the largest possible value, i.e. the PSU population size, M. The resulted number is truncated to be at least 2 to be able to estimate the intraclass correlation, and used as the number of PSU observations to design the main survey. In this case we can obtain the intraclass correlation. Even when $\hat{\rho}$ is positive, it may be very small, leading to large values of m. To avoid very large values of m in the main survey, truncation based on a maximum cutoff value A will be evaluated. The PSU sample size for the main study will therefore be

$$m_{main} = min(max\left(\sqrt{\frac{C_1}{C_2}\frac{1-\rho}{\rho}}, 2\right), A). \tag{6.1}$$

It will be assumed that the objective of the main survey is to estimate a regression coefficient β. Simulations to evaluate the procedures will be based on an intercept-only model, (2.3). Figure 6.2 shows the procedure of the pilot survey performed in this chapter.

Example

The following example shows the effect of small value of $\hat{\rho}$ on the optimal

PSU sample size (m_{opt}). It also shows the effect of the estimated intraclass correlation and the PSU sample size on the design effect.

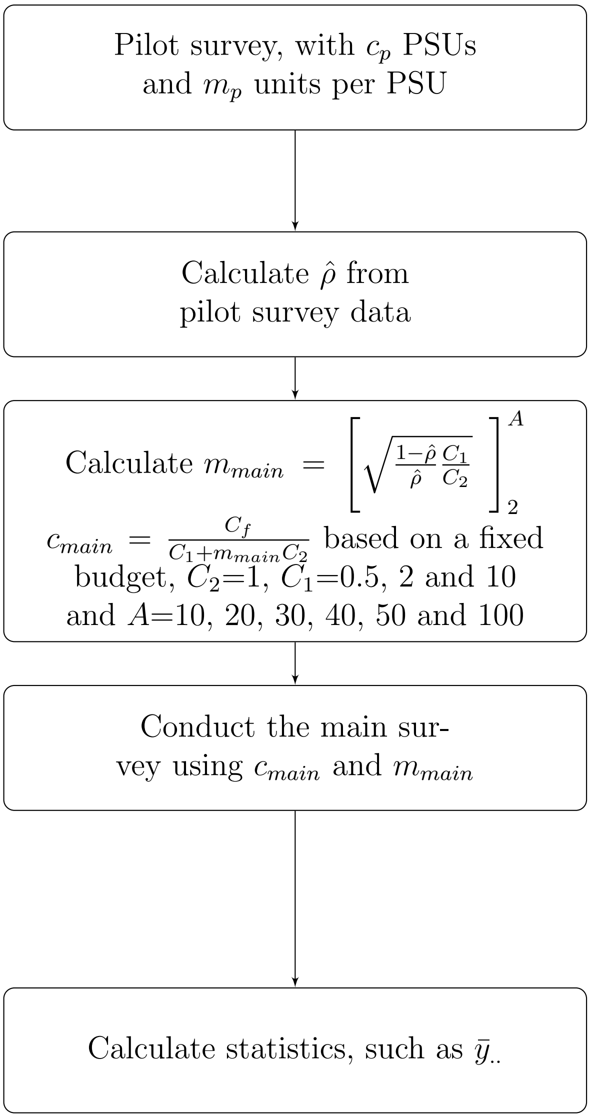


Figure 6.2: Flowchart explaining the adaptive procedures based on a pilot survey

Assume that the total cost, C_f is 5000, the cost of including an extra element in the sample, C_2, is 1 and the average cost of including an extra

PSU in the sample, C_1, is 0.5, 2 or 10.

Table 6.1: The effect of estimated intraclass correlation, $\hat{\rho}$ on the optimal number of PSUs, optimal number of observations per PSU and on the design effect, where $deff = 1 + \hat{\rho}(m-1)$, based on $\rho = 0.025$, a budget of C_f=5000 and different values of the cost of including a new PSU in the sample, C_1, of 0.5, 2 and 10.

Est ICC	$C_1 = 0.5$			$C_1 = 2$			$C_1 = 10$		
$\hat{\rho}$	m_{opt}	c_{opt}	$deff$	m_{opt}	c_{opt}	$deff$	m_{opt}	c_{opt}	$deff$
0.045	3	1429	1.05	7	667	1.15	15	323	1.35
0.025	4	1111	1.08	9	526	1.20	20	244	1.48
0.01	7	667	1.15	14	345	1.33	31	159	1.75
0.005	10	476	1.23	20	244	1.48	45	110	2.10
0.001	22	222	1.53	45	110	2.10	100	50	3.48
0.0005	32	154	1.78	63	79	2.55	141	35	4.50
0.0001	71	70	2.75	141	35	4.50	316	16	8.88
0.00005	100	50	3.48	200	25	5.98	447	11	12.15
0.00001	224	22	6.58	447	11	12.15	1000	5	25.98
0.000005	316	16	8.88	632	8	16.78	1414	4	36.33
0.000001	707	7	18.65	1414	4	36.33	3162	2	80.03

Table 6.1 shows that as $\hat{\rho}$ approaches zero, m_{opt} becomes very large, whereas the number of PSUs c_{opt} decreases. The value of m_{opt} is also larger when the cost of including a new PSU increases. The design effect, calculated from Equation (1.1), with ρ of 0.025 is also very large as $\hat{\rho}$ approaches zero. This demonstrates how small values of $\hat{\rho}$, which can easily occur when ρ=0.025, can lead to a very inefficient design.

This chapter is divided into 4 sections. In Section 6.2 a review of Brooks (1955) is given. Section 6.3 describes a simulation study conducted to evaluate the adaptive design based on a pilot, and to evaluate different settings for A, c_p and m_p. The parameters ρ, C_1 and C_2 were also varied. It discusses the best choice for A given ρ, C_1 and C_2 in order to minimize $var(\hat{\beta})$. In practice, however, ρ would not be known. Also, $var(\hat{\beta})$ is not the ideal measure for choosing m_p and c_p, because it does not reflect the cost of increasing m_p and c_p. Section 6.4 introduces the "cost-adjusted design effect" to compare the adaptive strategy where a pilot is conducted and used to design the main survey, to the strategy of conducting a simple random sampling (SRS) with no pilot, with same total cost.

6.2 Review of Brooks (1955)

Brooks (1955) described a very similar problem to the one covered by this chapter. He used the model

$$y_{ij} = \bar{Y}_{..} + b_i + e_{ij}, \quad i = 1, \ldots, c, \ j = 1, \ldots, m, \tag{6.2}$$

where $\bar{Y}_{..}$ is the population mean.

Fixing the two-stage sample cost model (1.3) and minimizing the variance of the sample estimate, he derived the optimal PSU sample size to be

$$m_{opt} = \sqrt{\frac{C_1}{C_2}} \frac{\sigma_e}{\sigma_b}, \tag{6.3}$$

m_{opt} could be estimated using a pilot sample to be

$$\hat{m}_{opt} = \sqrt{\frac{C_1}{C_2}}\frac{\hat{\sigma}_e}{\hat{\sigma}_b}. \tag{6.4}$$

But this estimate does not yield a fundamental value of $\hat{m}_{opt}$, therefore an integer k can be used such that

$$k(k-1) \leq \hat{m}_{opt}^2 \leq k(k+1), \tag{6.5}$$

$k = \infty$ is indicated whenever the variance ratio $\hat{\sigma}_b^2/\hat{\sigma}_e^2 \leq 1$, this means that every sampled PSU will have all of its elements enumerated.

Brooks (1955) assumed the same cost ratios for the pilot and the main samples. He varied the cost ratio C_1/C_2 and the ratio of the within- and between-PSU variance components, σ_e^2/σ_b^2, over ranges of values. Table 6.2 shows part of Brooks' table I which was based on the pilot sample designs corresponding to the value of $M = \infty$. He used different cost ratios of 0.01, 2 and 8 and variance components ratios of 0.25 , 1, 2, 8, 16, 32 and 64.

In this chapter we used the procedure of truncation the value of m if it is greater than a cutoff value, A. It also will be truncated below to be greater than or equal to 2. Whereas Brooks (1955) considered a use of $\hat{m}_{opt}$=1 in some cases, but this case is not considered in our work in this chapter as in this case we cannot estimate the intraclass correlation.

Brooks (1955) obtained approximate results, using an approximation which ignored the possibility of $\hat{\sigma}_b^2/\hat{\sigma}_e^2 \leq 1$. This was necessary given the

Table 6.2: Pilot sampling designs using cost ratio C_1/C_2 and variance components ratio σ_e^2/σ_b^2

C_1/C_2	0.01		2		8	
σ_e^2/σ_b^2	c_p	m_p	c_p	m_p	c_p	m_p
0.25	5	3	5	3	4	4
1	7	3	6	4	5	6
2	8	5	7	7	6	9
4	9	9	8	11	7	14
8	10	14	10	15	9	18
16	10	25	10	27	10	28
32	10	46	10	47	10	49
64	10	92	10	93	10	100

computing technology available in 1955, but it means that Brooks' results could be substantially in error. In contrast, we obtained results by simulation, so no such approximation was necessary in our case.

6.3 Simulation Study

A simulation study was conducted based on model (2.3). Different numbers of pilot PSUs (c_p) with equal within-PSU sample sizes (m_p) will be used. The cost of including a new PSU in the sample (C_1) was varied. The average cost of including an extra element in the sample (C_2) was fixed at 1.

The variance of $\hat{\beta}$ from the main survey was evaluated by calculating the variance over all 1000 estimated values of β.

The number of pilot sample PSUs (c_p) was varied over a range of values of 2, 5, 10 and 25. The number of units per PSU (m_p) was varied over a range of values of 2, 5, 10, 15, 25 and 50. A range of values of the cutoff A of 10, 20, 30, 40, 50 and 100 was evaluated.

The cost of including a new PSU in the sample (C_1) was varied over a range of values of 0.5, 2 and 10.

The value of ρ was estimated using Equation (2.7), using the estimated PSU-level variance components extracted from the random effects variances matrix (REmat) appeared in the summary of the *lmer*() function in the *lme*4 package in R (R Development Core Team, 2007).

Table 6.3 shows the simulation results for ρ=0 and ρ=0.05 with C_1=10 and various numbers of pilot PSUs, c_p, and numbers of observations per PSU, m_p. The true variance of $\hat{\beta}$ ($\times 10^3$) is calculated over the 1000 simulations.

Choice of A

For ρ=0, the minimum variance of $\hat{\beta}$ occurred at A=100 for almost all the values of c_p and m_p with a few exceptions. The first exception appeared when c_p was small (10 or less) with m_p=2, in this case the variance was minimized at A=40. The other exception appeared at $m_p = c_p$=25, where in this case the minimum variance occurred at A=50 because true $m_{opt} = \infty$.

When ρ=0.05, the best A was much lower at either 10 or 20 when there

Table 6.3: Variance of $\hat{\beta}$, ($\times 10^3$), calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_f=5000, C_1=10 and C_2=1. ρ=0 and 0.05)

Pilot		Variance of ($\hat{\beta}$), ($\times 10^3$), for ρ=0						Variance of ($\hat{\beta}$), ($\times 10^3$), for ρ=0.05					
PSUs	Obs	Cutoff for Within-PSU Sample Size (A)						Cutoff for Within-PSU Sample Size (A)					
c_p	m_p	10	20	30	40	50	100	10	20	30	40	50	100
2	2	0.594	0.582	0.528	0.512	0.520	0.536	0.829	0.751	0.881	0.887	0.864	1.245
2	5	0.485	0.395	0.351	0.348	0.367	0.330	0.687	0.659	0.740	0.797	0.885	1.320
2	10	0.436	0.385	0.330	0.327	0.342	0.307	0.619	0.688	0.726	0.713	0.785	1.188
2	15	0.422	0.335	0.297	0.294	0.297	0.293	0.623	0.629	0.667	0.706	0.787	1.191
2	25	0.406	0.345	0.309	0.300	0.273	0.259	0.628	0.659	0.672	0.731	0.785	1.013
2	50	0.402	0.323	0.276	0.267	0.260	0.249	0.637	0.648	0.667	0.728	0.800	0.990
5	2	0.514	0.486	0.516	0.441	0.476	0.430	0.710	0.794	0.783	0.820	0.879	1.006
5	5	0.438	0.378	0.325	0.341	0.310	0.302	0.661	0.642	0.723	0.791	0.857	1.206
5	10	0.427	0.337	0.303	0.298	0.286	0.274	0.648	0.648	0.670	0.716	0.782	1.053
5	15	0.440	0.294	0.315	0.286	0.274	0.253	0.627	0.621	0.676	0.642	0.731	0.951
5	25	0.408	0.314	0.293	0.283	0.264	0.262	0.582	0.617	0.662	0.690	0.715	0.891
5	50	0.410	0.302	0.269	0.269	0.265	0.248	0.602	0.605	0.660	0.696	0.665	0.784
10	2	0.501	0.434	0.474	0.396	0.402	0.418	0.731	0.720	0.753	0.801	0.813	1.000
10	5	0.431	0.340	0.337	0.313	0.286	0.285	0.658	0.621	0.679	0.778	0.819	1.096
10	10	0.414	0.308	0.303	0.281	0.258	0.248	0.628	0.612	0.667	0.717	0.743	0.951
10	15	0.416	0.301	0.291	0.278	0.277	0.260	0.648	0.635	0.623	0.647	0.708	0.884
10	25	0.421	0.299	0.281	0.270	0.260	0.237	0.644	0.591	0.618	0.661	0.661	0.868
10	50	0.408	0.317	0.265	0.253	0.246	0.234	0.664	0.633	0.607	0.636	0.614	0.690
25	2	0.463	0.382	0.375	0.365	0.349	0.345	0.662	0.681	0.724	0.686	0.838	1.045
25	5	0.413	0.333	0.295	0.277	0.265	0.259	0.613	0.605	0.609	0.719	0.757	1.074
25	10	0.393	0.324	0.283	0.275	0.263	0.246	0.616	0.624	0.586	0.677	0.698	0.950
25	15	0.393	0.311	0.290	0.254	0.254	0.245	0.635	0.635	0.612	0.688	0.677	0.706
25	25	0.402	0.287	0.281	0.273	0.246	0.253	0.651	0.594	0.608	0.632	0.610	0.633
25	50	0.422	0.334	0.280	0.265	0.252	0.242	0.626	0.594	0.594	0.570	0.570	0.570

Figure 6.3: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=10 and C_2=1, ρ=0)

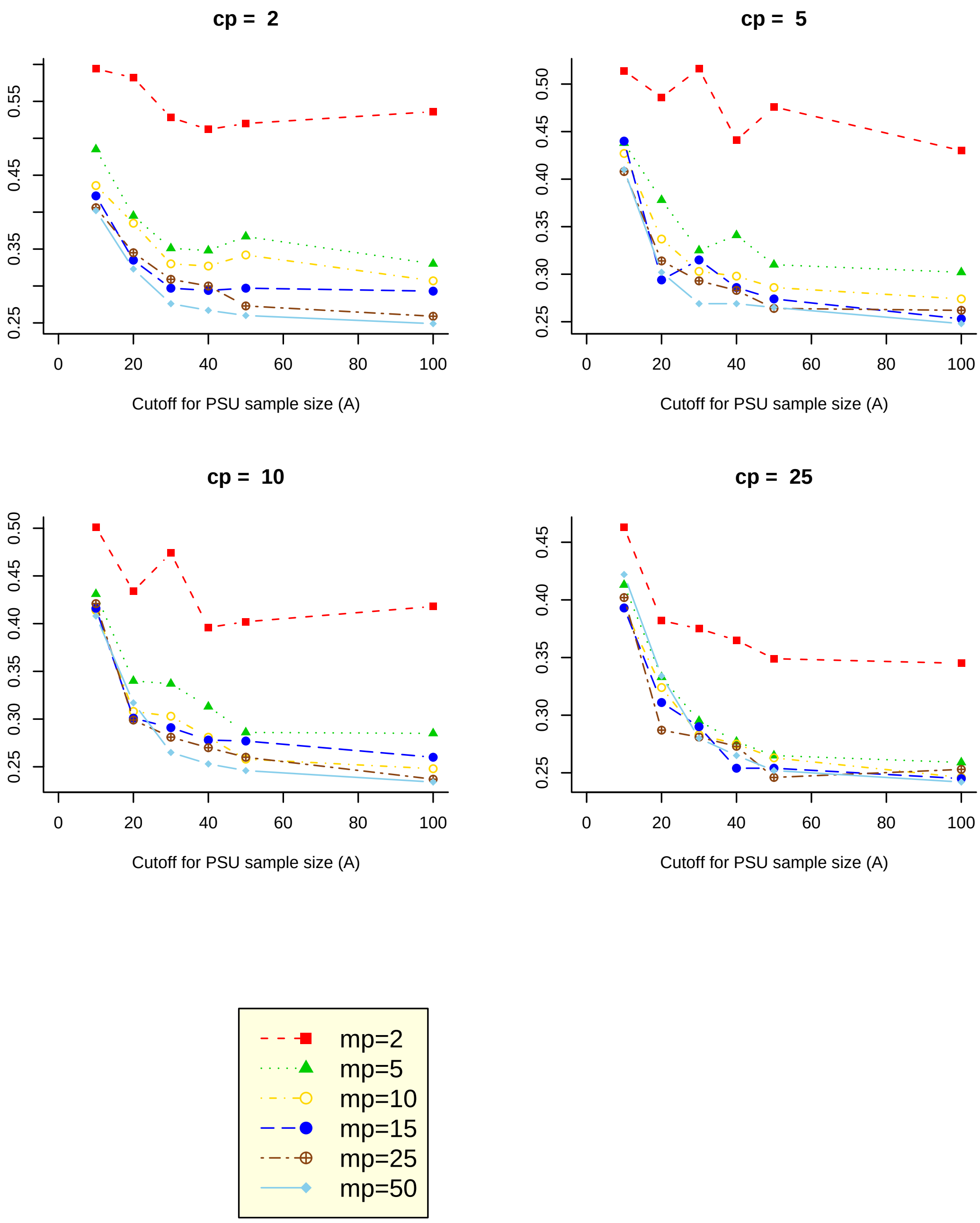
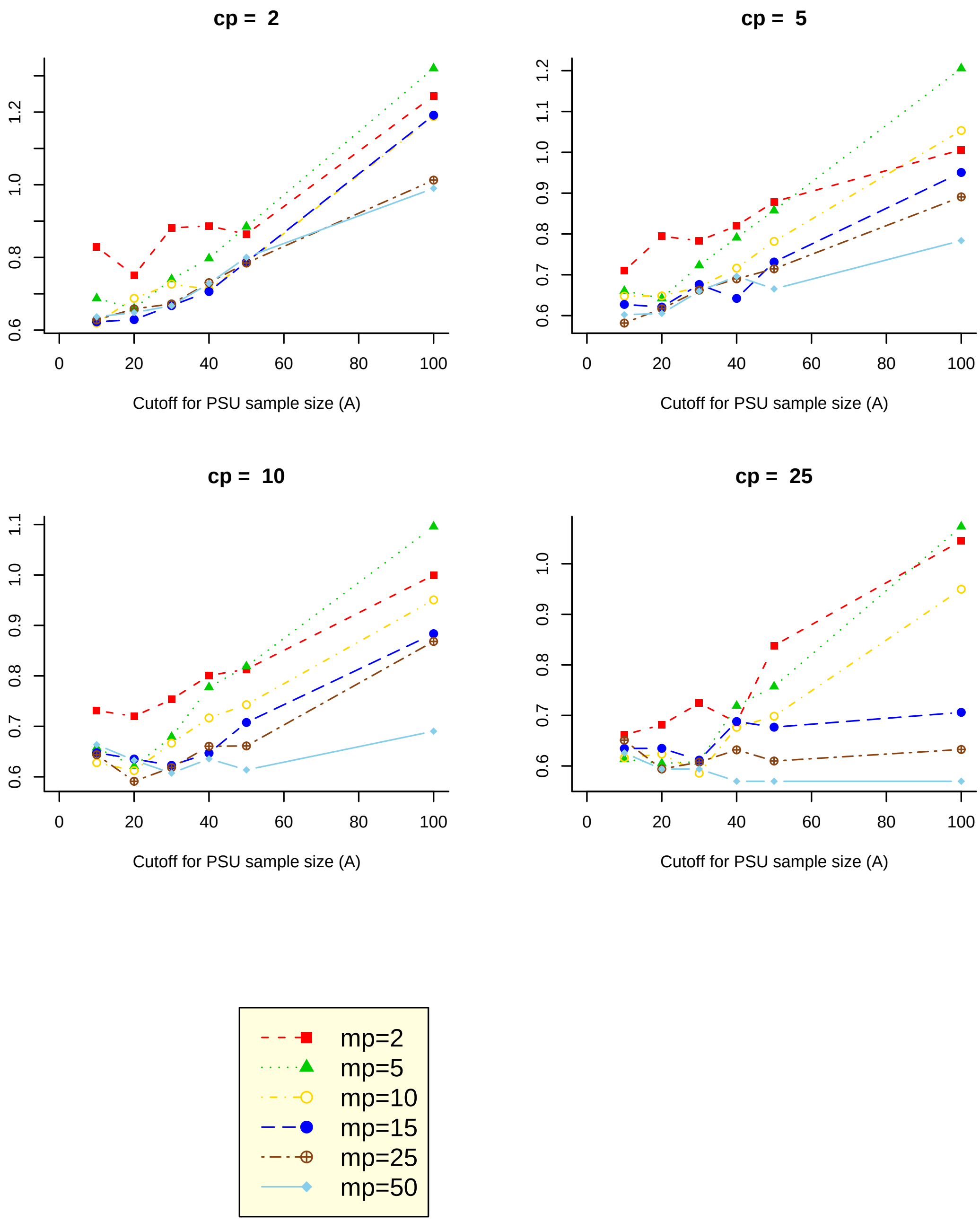


Figure 6.4: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=10 and C_2=1, ρ=0.05)



were small number of sample PSUs (5 or less) and 20 or 30 when there were large number of sample PSUs (10 or more). Table 6.4 shows the optimal A for each m_p and c_p.

Figures 6.3 and 6.4 show the plots of the variance of $\hat{\beta}$ versus the cutoff A, for all values of c_p and m_p with C_2=1, C_1=10 and ρ=0 and 0.05.

Choosing m_p and c_p

For ρ=0 and A=10, the minimum variance of $\hat{\beta}$ occurred when m_p=50 for $c_p \leq 10$ and when m_p=25 for c_p=25. For A=20 the minimum variance of $\hat{\beta}$ occurred at m_p=50 for $c_p \leq 5$ and at m_p=25 for $c_p \geq 10$. For A=30 and 100, the minimum variance of $\hat{\beta}$ occurred at m_p=50 for all values of c_p. For A=40, the minimum variance of $\hat{\beta}$ occurred at m_p=50 for $c_p \leq 10$ and at m_p=15 for c_p=25. For A=50, the minimum variance of $\hat{\beta}$ occurred when m_p=25 for c_p=5 and when m_p=50 for other values of c_p.

For ρ=0.05 and A=10, the minimum variance of $\hat{\beta}$ occurred at m_p=10 when c_p=2 and 10, at m_p=25 when c_p=5 and at m_p=5 when c_p=25. For A=20, the minimum variance of $\hat{\beta}$ occurred at m_p=15 when c_p=2, at m_p=50 when c_p=5, at m_p=25 when c_p=10 and at $m_p \geq$25 when c_p=25. A=30 gives minimum variance of $\hat{\beta}$ at m_p=15 and 50 when c_p=2, at m_p=50 when c_p=5 and 10 and at m_p=15 when c_p=25. A=40 gives minimum variance of $\hat{\beta}$ occurred at m_p=10 when $c_p \leq 5$ and at m_p=50 when $c_p \geq 10$. For A=40, the minimum variance of $\hat{\beta}$ occurred at m_p=10 and 25 when c_p=2 and at m_p=50

Table 6.4: Optimal A when ρ=0.05 for various m_p and c_p

		m_p					
		2	5	10	15	25	50
c_p	2	20	20	10	10	10	10
	5	10	20	10	20	10	10
	10	20	20	20	30	20	30
	25	10	20	30	30	20	40

for other values of c_p. For A=100, the minimum variance of $\hat{\beta}$ occurred at m_p=50 for all values of c_p.

6.4 Analysis of Simulation Results Using a Cost-Adjusted Design Effect

The discussion in the previous section was not enough to guide choice of pilot sample size, because the costs attached to a bigger pilot sample were not considered. In this section we will look at the total cost of the pilot and the main survey, and the variance of $\hat{\beta}$ from the main survey. We are comparing our strategy where a pilot is conducted and used to design the main survey, to the strategy of conducting a simple random sampling (SRS) with no pilot, with same total cost. For this purpose we defined the "cost-adjusted design effect" [1], $cdeff$, to be the ratio of the variance of an estimator under a complex design, to the variance of an estimator under simple random

[1]This contribution suggested by my supervisor Robert Clark

sampling with the same cost (or expected cost), according to a cost model. That is

$$cdeff = \frac{V}{V_{srs}}. \tag{6.6}$$

The difference between the $cdeff$ and the usual design effect is that: in the usual design effect the denominator is the variance from a SRS with the same sample size, whereas in the $cdeff$ the denominator is the variance from a SRS with the same cost. The $cdeff$ is useful for comparing the efficiency of designs with different costs.

Under the linear mixed model, the variance of the sample mean for a balanced two-stage design is given by

$$V = \frac{\sigma_b^2}{c} + \frac{\sigma_e^2}{n}. \tag{6.7}$$

The variance of the sample mean under a simple random sample is given by

$$V_{srs} = \frac{\sigma_b^2}{n_{srs}} + \frac{\sigma_e^2}{n_{srs}}, \tag{6.8}$$

because under simple random sampling, the number of PSUs (c) approximately equals the sample size (n), because provided the sampling fraction is small, 1 unit will be selected from each selected PSU in almost all cases.

Now suppose the cost under simple random sampling, $C_{srs} = n_{srs}C_1 + n_{srs}C_2$, to be equal to the cost of the two-stage design, including the ... test

$$C_{tot} \quad = \quad (c_{main} + c_p)C_1 + (n_{main} + n_p)C_2.$$

Therefore, the simple random sample size can be calculated to be

$$n_{srs} = \frac{C_{tot}}{C_1 + C_2}. \tag{6.9}$$

Therefore, $cdeff$ becomes

$$\begin{aligned} cdeff &= \frac{V}{\frac{\sigma_b^2}{n_{srs}} + \frac{\sigma_e^2}{n_{srs}}} = n_{srs}\left(\frac{V}{\sigma_b^2 + \sigma_e^2}\right) \\ &= \left(\frac{C_{tot}}{C_1 + C_2}\right)\left(\frac{V}{\sigma_b^2 + \sigma_e^2}\right) \\ &= \left(\frac{1}{C_1 + C_2}\right)\left(\frac{1}{\sigma_b^2 + \sigma_e^2}\right) C_{tot} V, \end{aligned} \tag{6.10}$$

where $V = var(\hat{\beta})$ from the main study, designed using a pilot. In the simulation study described in Section 6.3, the values of σ_b^2 and σ_e^2 were set to $\frac{\rho}{1-\rho}$ and 1, respectively, to ensure that the intraclass correlation was ρ. The value of C_2 was assumed to be 1. Therefore, $\sigma_b^2 + \sigma_e^2 = \frac{\rho}{1-\rho} + 1 = \frac{1}{1-\rho}$. Hence, Equation (6.10) reduces to

$$cdeff = \frac{1-\rho}{\rho}\, C_{tot}\, V.$$

We will now find the best choice of A, m_p and c_p by minimizing $cdeff$ from the simulation study.

Table 6.5 shows the best choice of A, m_p and c_p, based on the cost-adjusted design effect. For all values of C_1, the optimal A was generally small, A=10, with some exceptions. The first exception was when C_1=0.5 with ρ=0.05 and 0.1 the optimal A was 50 and 20 respectively. The second exception was when C_1=2 with ρ=0.1 as the optimal A was 40. Finally when C_1=10 with ρ=0 and 0.01 as the optimal A was 30 and 20 respectively. The table shows that the "cost-adjusted design effect" values became smaller for larger average cost of including an extra element in the sample.

Table 6.5: The best designs based on the cost-adjusted design effect based on perfect knowledge of ρ

		Optimal Design Setting			Cost Adjusted
PSU Cost	ICC	PSUs	Observations	Cutoff	Design Effect
C_1	ρ	c_p	m_p	A	$cdeff$
0.5	0	10	25	10	1.535
	0.01	25	15	10	1.731
	0.025	5	50	10	1.951
	0.05	25	50	50	1.965
	0.1	10	50	20	1.965
2	0	10	50	30	0.790
	0.01	10	25	10	0.931
	0.025	10	10	10	1.039
	0.05	10	10	10	1.196
	0.1	25	25	40	1.442
10	0	10	50	50	0.241
	0.01	5	25	30	0.329
	0.025	10	10	20	0.417
	0.05	5	25	10	0.509
	0.1	10	5	10	0.669

Figures 6.5 - 6.7 show the plots of the variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey for all costs of including a new PSU in the sample, C_1, where C_1=0.5, 2 and 10 and a fixed average cost of including an extra element in the sample, C_2=1 when

Table 6.6: The best designs based on the cost-adjusted design effect based on perfect knowledge of A, C_1=10, ρ=0.05

	Optimal Design Setting			Cost Adjusted
Cost	PSUs	Observations	Cutoff	Design Effect
C_f	c_p	m_p	A	$cdeff$
500	5	50	10	1.49
1000	10	10	10	1.03
2000	2	50	10	0.741
5000	5	25	10	0.509

ρ varies over a range of values of 0, 0.01, 0.025, 0.05 and 0.1.

Table 6.6 shows the optimal A based on different values of the total cost C_f of 500, 1000, 2000 and 5000 when the true ρ=0.05. It shows that the optimal A was 10 for all values of C_f. The table shows that the "cost-adjusted design effect" values became smaller for larger values of total cost. Values of c_p and m_p changed by varying C_f. For C_f=500, c_p=5 and m_p=50. For C_f=1000, c_p=m_p=10. For C_f=2000, c_p=2 and m_p=50. For C_f=5000, c_p=5 and m_p=25.

6.5 Conclusions

When ρ=0, a large value of A (generally 100) was most efficient, not surprisingly. When ρ=0.05, A=20 gave the best results in most cases. This suggests in practice, PSU sample sizes should be forced to be 20 or less unless a very large pilot is conducted to estimate ρ.

Figure 6.5: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey for different values of ρ (C_1=0.5 and C_2=1)

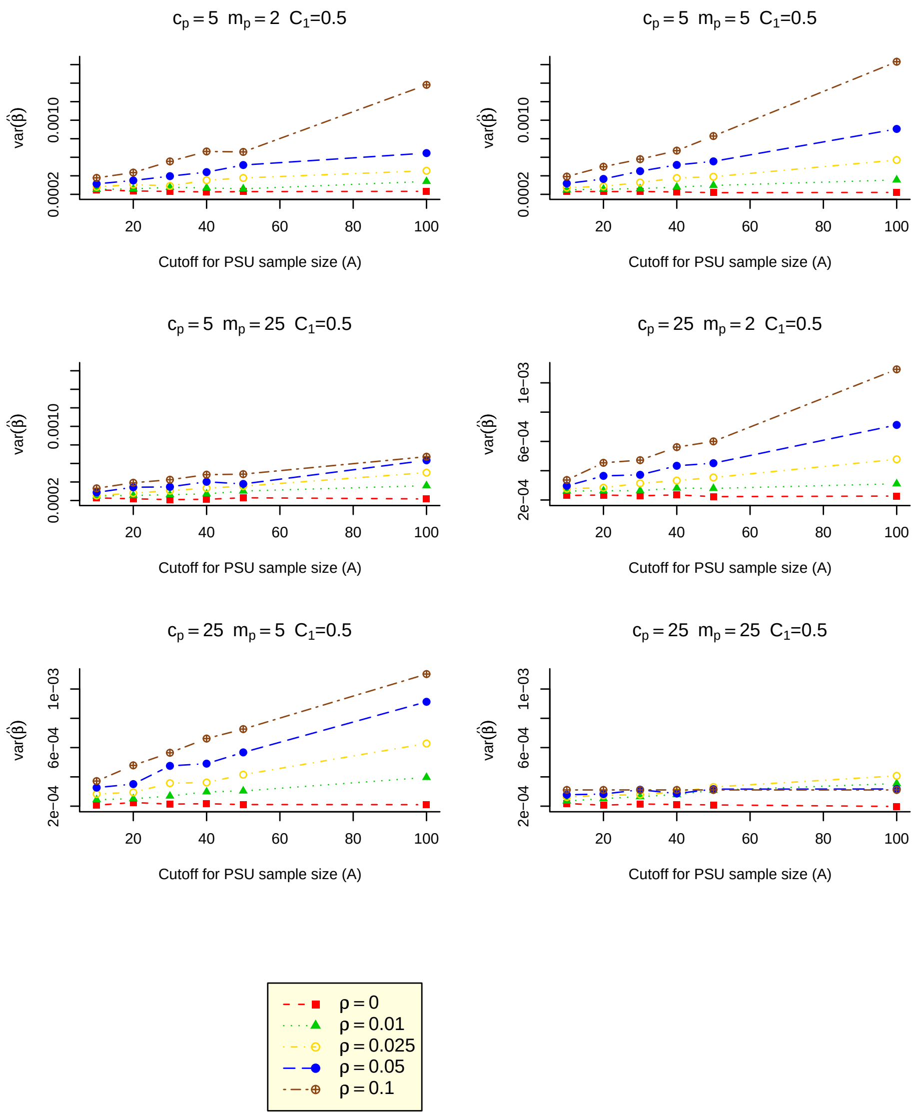


Figure 6.6: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey for different values of ρ (C_1=2 and C_2=1)

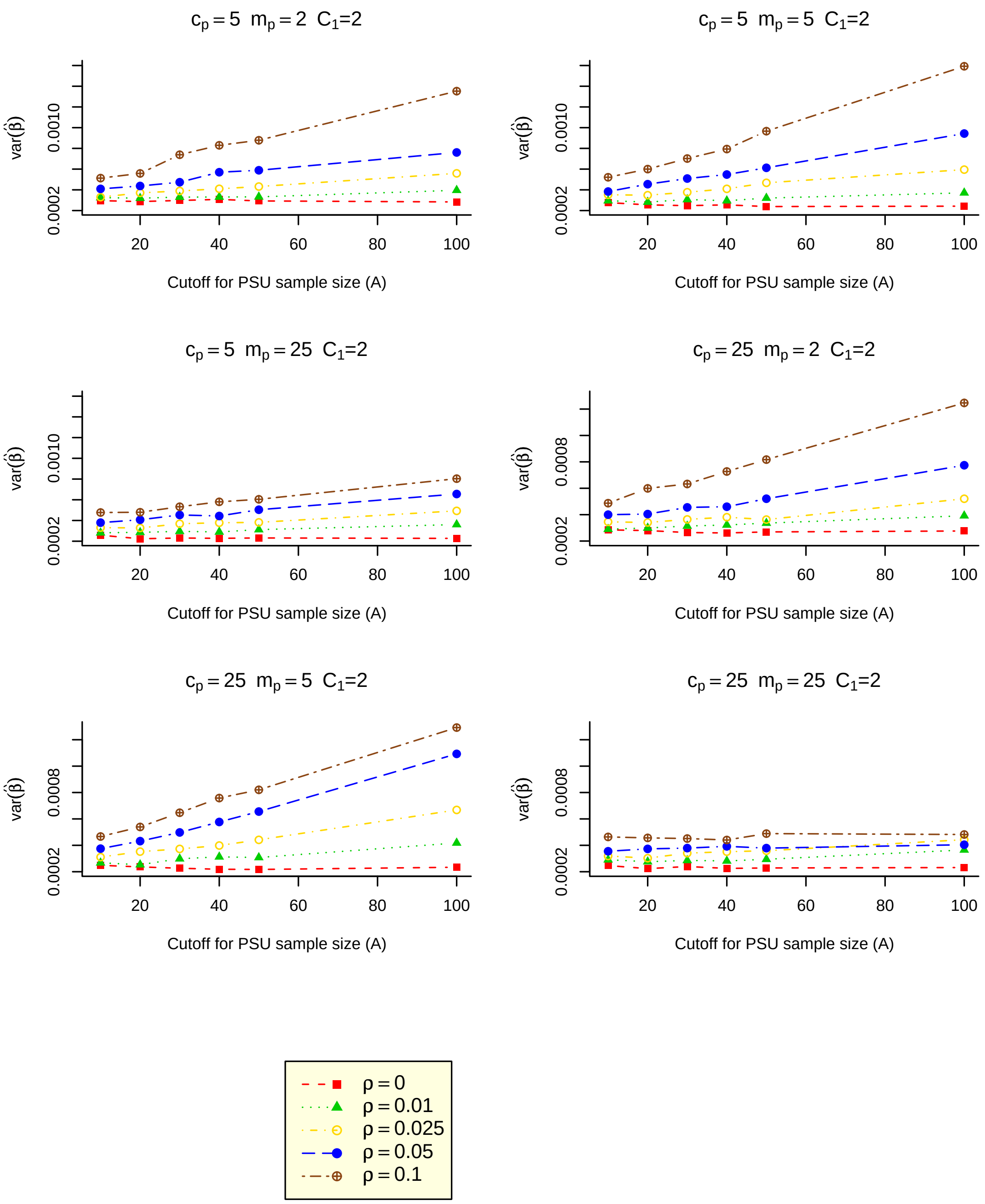
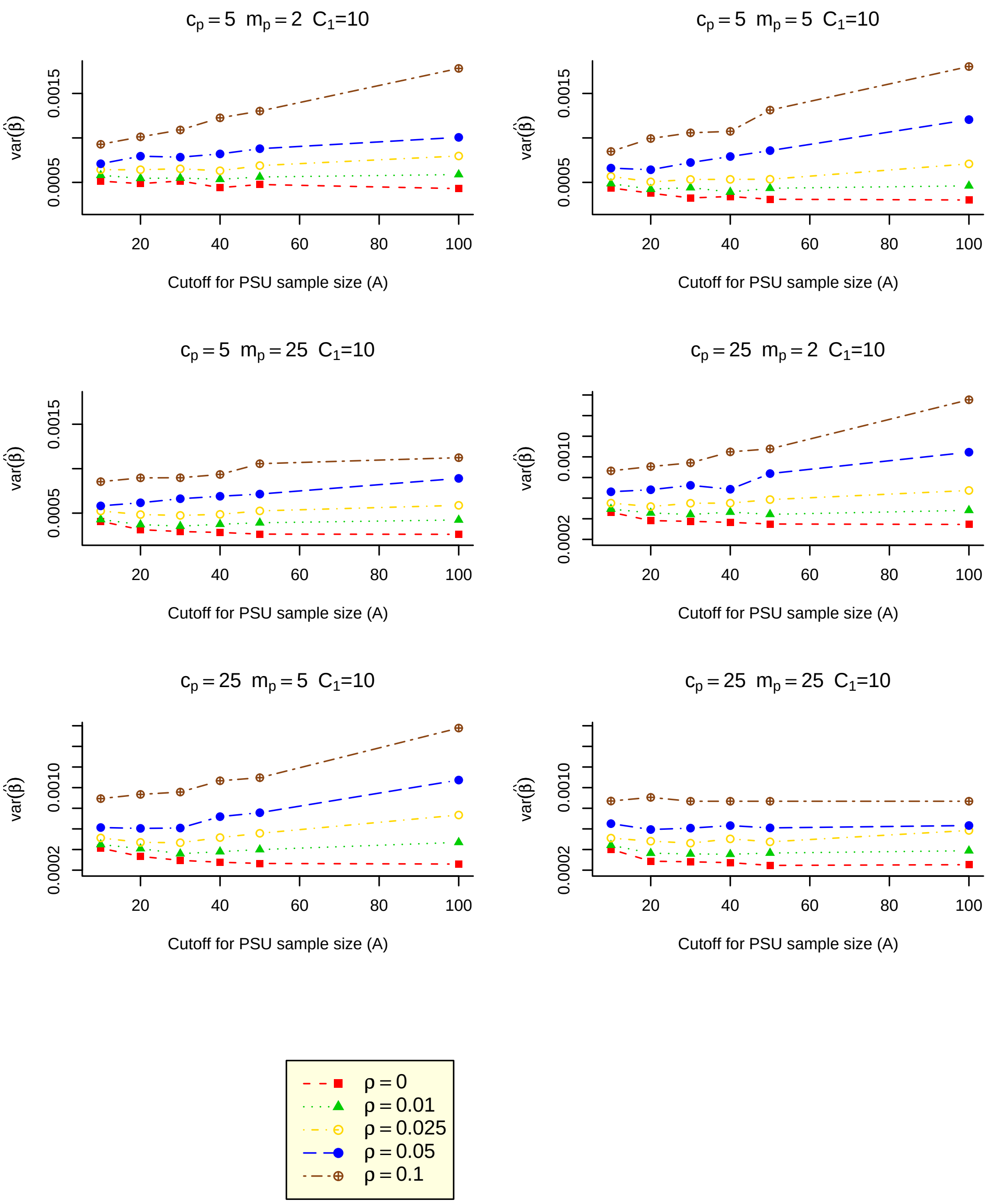


Figure 6.7: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey for different values of ρ (C_1=10 and C_2=1)



For a fixed total cost of 5000, and based on the variance of $\hat{\beta}$, when ρ=0, a small number of pilot PSUs (10 or less, in general) should be chosen with large number of observations per PSU (generally 50), for all values of A, in general. When ρ=0.05, the number of pilot PSUs should be 10 or less with 25 or less observations per PSU, in general, for $A \leq 30$. For A=40, the number of pilot PSUs should be 5 or less with 10 observations per PSU. When A=50, 2 pilot PSUs with 10 or 25 observations per PSU should be chosen. For A=100, a large number of observations per pilot PSU with any number of pilot PSUs should be chosen.

Based on the cost-adjusted design effect, when C_1=0.5 and C_f=5000, a large number of pilot PSUs (10 or more, in general) should be chosen with large number of observations per PSU (25 for ρ=0, 15 for ρ=0.01 and 50 for $\rho \geq 0.025$. When C_1=2, the number of pilot PSUs should be 10 in most cases with 10 or 25 observations per PSU. While when C_1=10, a small number of pilot PSUs (5 or 10) should be selected with 25 or more observations per PSU when ρ=0, 0.01 and 0.05. For other values of ρ, number of pilot PSUs should be 10 with 5 or 10 observations per PSU.

For a fixed total cost of 5000 and C_1=0.5, the best choice of A was 10 when $\rho \leq 0.025$. It was 50 when ρ=0.05 and 20 when ρ=0.1. For C_1=2, the best choice of A was 10 when ρ=0.01, 0.025 and 0.05, while it was 30 and 40, when ρ=0 and 0.1, respectively. For C_1=10, A=10 was the best choice when $\rho \geq 0.05$, and 20 or more otherwise.

For a range of values of C_f for fixed ρ of 0.05 and C_1=10, the optimal A was 10. The $cdeff$ decreased by increasing the C_f value. The best choice of c_p was 5 with 50 and 25 observations per PSU, when C_f is fixed at 500 and 5000, respectively. When C_f=1000, the best number of pilot PSUs was 10 with 10 observations each. Finally, when C_f=2000, the best number of pilot PSUs was 2 with 50 observations each.

Chapter 7

Conclusions

7.1 Summary and Conclusions

Regression coefficients and the variances of their estimates can be estimated using different methods when the intraclass correlation is believed to be small. The linear mixed model (LMM) is one alternative. Another alternative, when observations are assumed to be independent, is the linear model (LM). LMM variance estimators can be larger than LM variance estimators when the PSU sample size are large, and this leads to wider confidence intervals for β.

A third alternative is to use an adaptive strategy. The strategy developed in Chapter 3 is to test the null hypothesis that the PSU-level variance component, σ_b^2, is zero. The LM variance estimator is used if the null hypothesis is not rejected. Otherwise, the LMM or alternatively the Huber-White variance estimator is used.

Chapter 3 found that the adaptive confidence intervals in extreme designs with a small number of sample PSUs and a large number of observations per PSU. In these designs, the variance of the mean will be significantly boosted even when the intraclass correlation is small, however even with high intraclass correlation, the PSU-level variance component is unlikely to be statistically significant. Accordingly, for $c \leq 5$ with $\bar{m} \geq 25$, adaptive non-coverage rates were 15-20% higher than the nominal rate when $\rho \neq 0$, where c is the number of sample PSUs and $\bar{m}$ is the average number of observations per PSU. Therefore, even if clustering is not statistically significant for these extreme designs, it has to be allowed for in variances estimates.

The ADM, adaptive based on LMM as an alternative, confidence intervals were shorter than the LMM confidence intervals in designs with 2 sample PSUs with all average numbers of observations per PSU for all values of intraclass correlation, ρ. In the balanced designs, the ADM confidence intervals were a bit shorter for designs with 5 sample PSUs with $m \geq 25$ when $\rho=0$ and designs with $c=5$ for all numbers of observations per PSU, m, approximately, when $\rho \neq 0$. They were shorter in designs with number of sample PSUs, $c=10$ and $m \geq 10$ and $m=5$ and 10 when $\rho=0.025$ and 0.1, respectively. Otherwise, ADM and LMM confidence intervals performed similarly.

The ADH, adaptive based on Huber-White as an alternative, confidence intervals were much shorter than the Huber-White confidence intervals in designs with 2 and 5 sample PSUs with, approximately all average numbers of observations per PSU for all values ρ. In the balanced designs, the ADH confidence intervals were shorter for designs with 10 sample PSUs with $m \geq 10$, with $m \geq 15$ and $m \leq 15$ for ρ=0, 0.025 and 0.1, respectively and for designs with c=25 and m=10, 15 and 25 when ρ=0.025. There were no relevant differences, otherwise.

The same adaptive strategies were applied in Chapter 4 for log-normal data with two skewness levels, $\sigma = \frac{1}{3}$ and $\sigma = \frac{2}{3}$. Biases of adaptive variance estimators were similar to biases of adaptive variance estimators in Chapter 3. ADM variance estimators were less biased than the LMM variance estimators for designs with c=2 and c=5 with $m \leq 5$. In the unbalanced designs, ADM variance estimators were less biased than the LMM variance estimators for designs with $c \leq 5$ when $\sigma = \frac{1}{3}$ and in designs with c=2 when $\sigma = \frac{2}{3}$. ADH variance estimators were more biased than the Huber-White variance estimators in designs with $c \leq 5$ when $\rho = 0$ and in designs with c=2 when $\rho = 0.025$. There were no relevant differences otherwise.

ADH non-coverage rates were larger than Huber-White non-coverage rates except in designs with c=5, 25 with m=2 and c=10 with m=5. ADM non-coverage rates were larger than LMM non-coverage rates except in designs with c=2 and 10 with m=2 and 5, respectively; and designs with c=5 with $m \leq 10$ and designs with c=25.

ADM confidence intervals were shorter than LMM confidence intervals in designs with c=2, whereas ADH confidence intervals were shorter than Huber-White confidence intervals in designs with $c \leq 5$ and designs with c=10 with $m \geq 5$. In the unbalanced designs, the adaptive confidence intervals were shorter than the non-adaptive confidence intervals in designs with c=2 with all $\bar{m}$ similar to what was in Chapter 3 and unlike what was in Chapter 3 in designs with c=2.

Rejecting $H_0 : \sigma_b^2 = 0$ is possible even if the estimated intraclass correlation and the estimated design effect are relatively small. It may be desirable to use the linear model rather than the linear mixed model in these cases. To assess this possibility a new adaptive strategy was used in Chapter 5. We used the LMM or alternatively the Huber-White variance estimators were used if H_0 is rejected and $\widehat{deff} \geq d$, where d is a cutoff value. Otherwise, the LM variance estimators were used.

A simulation study showed that for balanced designs, cutoffs of d=1.05 and 1.5 had no effect - results were identical to the adaptive strategy described in Chapter 3. For unbalanced designs, a cutoff of d=1.5 slightly improved adaptive confidence intervals and variance estimates.

In Chapter 6 we considered a pilot survey to estimate the intraclass correlation assuming the intercept-only model. This estimator was used to estimate the optimal within-PSU sample size for the main survey, for fixed cost based on a simple cost model. The estimated value of ρ could be zero or close to zero and this might lead to a very large PSU sample size being calculated, which could lead to very high variances from the main survey. To deal with this problem, m was truncated above at a cutoff, A. The value of m was also truncated below to be greater than or equal to 2. A range of values of the cutoff A were evaluated by simulation. A range of values of the pilot sample sizes of PSUs (c_p) and units per PSU (m_p) were also evaluated.

Based on the variance of $\hat{\beta}$ when C_1=10, the best choice of A (out of possible values) occurred at:

- A=100 when ρ=0 for all values of c_p and m_p except for the extreme case $c_p = m_p$=2;
- A between 10 and 40 depending on the value of m_p and c_p.

Based on the variance of $\hat{\beta}$, when C_1=10 and ρ=0, the best choice of m_p was 50. When ρ=0.025, the best choice was at m_p=10 if A is 10 or 50, at m_p=15 if A is 20, 30 or 40 and at m_p=50 if A is 100.

Designs were also evaluated in terms of their cost-adjusted design effect ($cdeff$), a measure of efficiency reflecting both cost and variance. Based on the cost-adjusted design effect, when C_1=10, the optimal A was

- 50 when ρ =0 when c_p=10 and m_p=50;
- 30 when ρ =0.01 when c_p=5 and m_p=25;

- 20 when ρ =0.025 when c_p=m_p=10;
- 10 when ρ =0.05 when c_p=5 and m_p=25;
- 40 when ρ =0.1 when c_p=25 and m_p=5.

Chapter 6 also gives results for other values of C_1 and C_f.

7.2 Further Research

Chapter 3 found that adaptive confidence intervals perform poorly in designs with small numbers of PSUs and large numbers of observations per PSU. ADM and LMM non-coverage rates are high for these extreme designs. A possible reason is that there is not much power to detect the PSU-level variance component in the adaptive approach, even when it is substantial. One way to do this was the adaptive approaches developed in this thesis. Another possible approach is model averaging of the LMM and LM models. This would be more computationally intensive but would perhaps give better results than adopting either the LMM or LM.

Another possible reason is that the LMM confidence intervals are not exact and do not do well for small sample sizes. Confidence intervals rely on the degrees of freedom and we do not have exact degrees of freedom in the LMM case. We tried the approach suggested by Faes(2009). Other approaches such as Kenward and Roger (1997) or Satterthwaite (1941) would be worth trying and might result in confidence intervals with better coverage properties when the number of clusters is small.

Chapter 6 developed optimal design strategies for using a pilot study to guide the sample design of a main study. Optimal choices of m_p, c_p and a cutoff A for the within-PSU sample size for the main study, were obtained by simulation, for given values of ρ and other parameters. In practice, however, ρ would be unknown, and the pilot/main design strategy would need to be developed in ignorance of ρ. Future research could focus on choices of m_p, c_p and A that perform well across a range of possibilities for ρ.

Appendix A

Proofs for Chapter 2

A.1 Unbalanced Data Case

A.1.1 The Maximum Likelihood Estimators

Under H_A

The likelihood function for the sample observations $y_{ij}s$ from model 2.3 is given by

$$L = \prod_{i=1}^{c} f(\mathbf{y}_i), \tag{A.1}$$

where

$$f(\mathbf{y}_i) = \frac{1}{(2\pi)^{m_i/2}|\mathbf{V}_i|^{1/2}} exp\Big\{ -\frac{1}{2}\sum_{i=1}^{c}(\mathbf{y}_i - \beta)'\mathbf{V}_i^{-1}(\mathbf{y}_i - \beta)\Big\}.$$

Therefore, the likelihood function (A.1) is given by

$$L = \frac{1}{(2\pi)^{m_i/2}\prod_{i=1}^{c}|\mathbf{V}_i|^{1/2}} exp\Big\{ -\frac{1}{2}\sum_{i=1}^{c}(\mathbf{y}_i - \beta)'\mathbf{V}_i^{-1}(\mathbf{y}_i - \beta)\Big\}. \tag{A.2}$$

But

$$\begin{aligned} |\mathbf{V}_i| &= \eta_i(\sigma_e^2)^{m_i-1}; \\ \mathbf{V}_i^{-1} &= \frac{1}{\sigma_e^2}\mathbf{I}_{m_i} - \frac{\sigma_b^2}{\eta_i\ \sigma_e^2}\mathbf{J}_{m_i}. \end{aligned}$$

where $\eta_i = \sigma_e^2 + m_i\sigma_b^2$. Substituting for $\beta = \beta\mathbf{1}_{m_i}$, $|\mathbf{V}_i|$ and $\mathbf{V}_i^{-1}$ in (A.2), we obtain

$$L = \frac{exp\Big[-\frac{1}{2}\Big\{\frac{n-c}{\sigma_e^2}MSE + \sum_{i=1}^{c}\frac{m_i(\bar{y}_{i.}-\beta)^2}{\eta_i}\Big\}\Big]}{(2\pi)^{n/2}(\sigma_e^2)^{(n-c)/2}\prod_{i=1}^{c}(\eta_i)^{1/2}}. \quad (A.3)$$

The natural logarithm of the likelihood function is determined by taking the logarithm for both sides of (A.3), which is simplified to

$$\begin{aligned} \ell &= -\frac{n}{2}ln(2\pi) - \frac{n-c}{2}ln(\sigma_e^2) - \frac{1}{2}ln(\eta_i) \\ &\quad -\frac{(n-c)MSE}{2\sigma_e^2} - \frac{1}{2}\sum_{i=1}^{c}\frac{m_i(\bar{y}_{i.}-\beta)^2}{\eta_i}. \end{aligned} \quad (A.4)$$

The partial derivatives of (A.4) with respect to β, σ_e^2 and η_i are obtained as

$$\left.\begin{aligned} \frac{\partial\ell}{\partial\beta} &= \sum_{i=1}^{c}\frac{m_i(\bar{y}_{i.}-\beta)}{\eta_i}; \\ \frac{\partial\ell}{\partial\sigma_e^2} &= \frac{n-c}{2\sigma_e^2} + \frac{(n-c)MSE}{2(\sigma_e^2)^2}; \\ \frac{\partial\ell}{\partial\eta_i} &= -\frac{1}{2}\Big[\sum_{i=1}^{c}\frac{1}{\eta_i} + \frac{m_i(\bar{y}_{i.}-\beta)^2}{\eta_i^2}\Big]. \end{aligned}\right\} \quad (A.5)$$

Equating to zero the partial derivatives in (A.5) and solving with respect to β, η_i and σ_e^2 and denoting the solutions by $\hat{\beta}$, $\hat{\eta}_i$ and $\hat{\sigma}_e^2$ and after some simplifications we obtain

$$\left.\begin{array}{lcl}\hat{\beta} & = & \frac{\sum_{i=1}^{c}\left(m_i\bar{y}_{i.}/\hat{\eta}_i\right)}{\sum_{i=1}^{c}\left(m_i/\hat{\eta}_i\right)} \\ & = & \bar{y}_w; \\ \hat{\sigma}_e^2 & = & MSE; \\ \sum_{i=1}^{c}\frac{1}{\hat{\eta}_i} & = & \sum_{i=1}^{c}\frac{m_i(\bar{y}_{i.}-\bar{y}_w)^2}{\hat{\eta}_i^2}.\end{array}\right\} \tag{A.6}$$

It is obvious that the system of equations (A.6) has no explicit solutions for $\hat{\beta}$ and $\hat{\eta}_i$, therefore there is no explicit solution for $\hat{\sigma}_b^2$.

Under H_0

Under H_0 we have $\sigma_b^2 = 0$ therefore, the log-likelihood function (A.4) reduces to

$$\begin{aligned}\ell_0 &= -\frac{n}{2}ln(2\pi) - \frac{n-c-1}{2}ln(\sigma_e^2) \\ &\quad -\frac{1}{2\sigma_e^2}\Big[(n-c)MSE + \sum_{i=1}^{c} m_i(\bar{y}_{i.}-\beta)^2\Big].\end{aligned} \tag{A.7}$$

Differentiating (A.7) partially with respect to σ_e^2, we obtain

$$\frac{\partial \ell_0}{\sigma_e^2} = -\frac{n-c-1}{2\sigma_e^2} + \frac{1}{2(\sigma_e^2)^2}\Big[(n-c)MSE + \sum_{i=1}^{c} m_i(\bar{y}_{i.}-\beta)^2\Big]. \tag{A.8}$$

Equating (A.8) to zero and solving with respect to σ_e^2 and denoting the solution by $\hat{\sigma}_e^2$, we find

$$\begin{aligned}\hat{\sigma}_e^2 &= \frac{n-c}{n-c-1}MSE + \frac{1}{n-c-1}\sum_{i=1}^{c} m_i(\bar{y}_{i.}-\beta)^2 \\ &= \frac{1}{n-c-1}\sum_{i=1}^{c}\sum_{j=1}^{m_i} m_i(\bar{y}_{ij}-\bar{y}_{..})^2.\end{aligned} \tag{A.9}$$

A.1.2 Derivation of Equation (2.31)

A.1.3 The Restricted Maximum Likelihood Estimators (RLRT)

Proceeding from the general case considered in Subsection 2.3.2, the restricted log-likelihood function for the sample observations y_{ij}s, from the model defined in (2.3) is given as

$$\begin{aligned} \ell_R &= -\frac{1}{2}\Big[(n-c)ln(\sigma^2_{eA}) + \sum_{i=1}^{c} ln(\eta_i) - ln\Big(\sum_{i=1}^{c}\frac{m_i}{\eta_i}\Big) \\ &\quad + \frac{(n-c)MSE_A}{\sigma^2_{eA}} + \sum_{i=1}^{c}\frac{m_i(\bar{y}_{i.}-\bar{y}_w)^2}{\eta_i}\Big]. \end{aligned} \tag{A.10}$$

The partial derivatives of (A.10) with respect to σ^2_e and η_i are given by

$$\left.\begin{aligned} \frac{\partial \ell_R}{\partial \sigma^2_{eA}} &= -\frac{1}{2}\Big[\frac{n-c}{\sigma^2_{eA}} - \frac{(n-c)MSE_A}{(\sigma^2_{eA})^2}\Big]; \\ \frac{\partial \ell_R}{\partial \eta_i} &= -\frac{1}{2}\Big[\sum_{i=1}^{c}\frac{1}{\eta_i} + \frac{\sum_{i=1}^{c}\frac{m_i}{\eta_i^2}}{\sum_{i=1}^{c}\frac{m_i}{\eta_i}} - \frac{m_i(\bar{y}_{i.}-\bar{y}_w)^2}{\eta_i^2}\Big]. \end{aligned}\right\} \tag{A.11}$$

Equating to zero the partial derivatives in (A.11) and solving with respect to σ^2_e and η_i and representing the solutions by $\hat{\sigma}^2_{eA}$ and $\hat{\eta}_i$, we get

$$\left.\begin{aligned} \hat{\sigma}^2_{eA} &= MSE_A; \\ \sum_{i=1}^{c}\frac{m_i(\bar{y}_{i.}-\bar{y}_w)^2}{\hat{\eta}_i^2} &= \sum_{i=1}^{c}\frac{1}{\hat{\eta}_i} - \frac{\sum_{i=1}^{c}\frac{m_i}{\hat{\eta}_i^2}}{\sum_{i=1}^{c}\frac{m_i}{\hat{\eta}_i}}. \end{aligned}\right\} \tag{A.12}$$

Therefore, there is no explicit form for $\hat{\eta}_i$. Hence, $\hat{\sigma}^2_b$ has no explicit form.

The restricted maximum likelihood under H_A is given by

$$\begin{aligned} -2 \overset{MAX}{H_A}\ \ell_R &= (n-c)ln(MSE_A) + \sum_{i=1}^{c} ln(\hat{\eta}_i) + ln\Big(\sum_{i=1}^{c}(\hat{\lambda}_i)\Big) \\ &\quad + n - c + \sum_{i=1}^{c}\hat{\lambda}_i(\bar{y}_{i.}-\hat{\beta})^2. \end{aligned} \tag{A.13}$$

Under H_0

Under H_0, we have $\sigma_b^2 = 0$ therefore, the log-likelihood function (A.10) reduces to

$$\begin{aligned}\ell_{R_0} &= -\frac{1}{2}\Big[(n-1)ln(\sigma_{e0}^2) + ln(n) + \frac{1}{\sigma_{e0}^2}\Big((n-c)MSE_0 \\ &\quad + \sum_{i=1}^{c} m_i(\bar{y}_{i.} - \bar{y}_w)^2\Big)\Big]. \end{aligned} \tag{A.14}$$

But, under H_0, $\bar{y}_w$ reduces to $\bar{y}_{..}$, because

$$\begin{aligned}\bar{y}_w &= \frac{\sum_{i=1}^{c}(m_i\bar{y}_{i.}/\sigma_{e0}^2)}{\sum_{i=1}^{c}(m_i/\sigma_{e0}^2)} \\ &= \frac{\sum_{i=1}^{c} m_i\bar{y}_{i.}}{\sum_{i=1}^{c} m_i} \\ &= \frac{\sum_{i=1}^{c}\sum_{j=1}^{m_i} y_{ij}}{n} \\ &= \bar{y}_{..}. \end{aligned} \tag{A.15}$$

Therefore, Equation (A.14) reduces to

$$\begin{aligned}\ell_{R_0} &= -\frac{1}{2}\Big[(n-1)ln(\sigma_{e0}^2) + ln(n) + \frac{1}{\sigma_{e0}^2}\Big((n-c)MSE_0 \\ &\quad + \sum_{i=1}^{c} m_i(\bar{y}_{i.} - \bar{y}_{..})^2\Big)\Big]. \end{aligned} \tag{A.16}$$

Differentiating (A.16) partially with respect to σ_{e0}^2, we obtain

$$\begin{aligned}\frac{\partial\ell_{R_0}}{\sigma_{e0}^2} &= -\frac{1}{2}\Big[\frac{n-1}{\sigma_{e0}^2} - \frac{1}{\sigma_{e0}^4}\Big((n-c)MSE_0 \\ &\quad + \sum_{i=1}^{c} m_i(\bar{y}_{i.} - \bar{y}_{..})^2\Big)\Big]. \end{aligned} \tag{A.17}$$

Equating (A.17) to zero and solving with respect to σ_{e0}^2 and denoting the solution by $\hat{\sigma}_{e0}^2$, we find

$$\begin{aligned}
\hat{\sigma}_{e0}^2 &= \frac{1}{n-1}\Big[\sum_{i=1}^{c}(\bar{y}_{ij}-\bar{y}_{i.})^2+\sum_{i=1}^{c}\sum_{j=1}^{m_i} m_i(\bar{y}_{i.}-\bar{y}_{..})^2\Big] \\
&= \frac{1}{n-1}\sum_{i=1}^{c}\sum_{j=1}^{m_i}(y_{ij}-\bar{y}_{..})^2 \\
&= MSE_0. \qquad \text{(A.18)}
\end{aligned}$$

The restricted maximum likelihood under H_0 is given by

$$\begin{aligned}
-2 \overset{MAX}{H_0}\ \ell_R &= (n-c)ln(MSE_0)+\sum_{i=1}^{c} ln(MSE_0)+ln\Big(\sum_{i=1}^{c}\frac{m_i}{MSE_0}\Big) \\
&\quad +n-c+\frac{\sum_{i=1}^{c} m_i(\bar{y}_{i.}-\bar{y}_{..})^2}{MSE_0} \\
&= (n-c)ln(MSE_0)+c\ ln(MSE_0)+ln\Big(\frac{n}{MSE_0}\Big) \\
&\quad +n-c+\frac{\sum_{i=1}^{c} m_i(\bar{y}_{i.}-\bar{y}_{..})^2}{MSE_0} \\
&= n-c+ln(n)+(n-1)ln(MSE_0) \\
&\quad +\frac{\sum_{i=1}^{c} m_i(\bar{y}_{i.}-\bar{y}_{..})^2}{MSE_0}. \qquad \text{(A.19)}
\end{aligned}$$

One way to define the restricted likelihood ratio test is to subtract Equation (A.19) from (A.13). Therefore, the restricted likelihood ratio test can be given as

$$\begin{aligned}
\Lambda &= -2\Big(\overset{MAX}{H_0}\ \ell_R - \overset{MAX}{H_A}\ \ell_R\Big) \\
&= ln(n) + (n-1)\ ln(MSE_0) + \frac{\sum_{i=1}^{c} m_i(\bar{y}_{i.} - \bar{y}_{..})^2}{MSE_0} \\
&\quad -(n-c)ln(MSE_A) - \sum_{i=1}^{c} ln(\hat{\eta}_i) - ln\Big(\sum_{i=1}^{c}(\hat{\lambda}_i)\Big) \\
&\quad -\sum_{i=1}^{c}\hat{\lambda}_i(\bar{y}_{i.} - \hat{\beta})^2. \qquad (A.20)
\end{aligned}$$

Substituting $m_i = m$ into the unbalanced case, we get

$$\begin{aligned}
-2\ell_{RA} &= (n-c)ln(MSE_A) + \sum_{i=1}^{c} ln(\hat{\eta}) + ln\Big(\sum_{i=1}^{c}(\hat{\lambda})\Big) \\
&\quad +n - c + \sum_{i=1}^{c}\hat{\lambda}(\bar{y}_{i.} - \hat{\beta})^2 \\
&= n - c + (n-c)ln(MSE) + c\ ln(\hat{\eta}) + ln\Big(\frac{n}{\hat{\eta}}\Big) + \sum_{i=1}^{c}\frac{m}{\hat{\eta}}(\bar{y}_{i.} - \bar{y}_{..})^2 \\
&= n - c + (n-c)ln(MSE) + c\ ln(MSA) + ln(n) - ln(MSA) + \frac{1}{MSA}.(c-1)MSA \\
&= n - 1 + (n-c)ln(MSE) + (c-1)ln(MSA) + ln(n). \qquad (A.21)
\end{aligned}$$

$$\begin{aligned}
-2\ell_{R0} &= (n-1)ln\Big(\frac{SSE+SSA}{n-1}\Big) + (n-c)ln(MSE) + (c-1)ln(MSA) + ln(n) \\
&= (n-1)ln\Big(\frac{SSE+SSA}{n-1}\Big) + ln(n) + \frac{(c-1)MSA + (n-c)MSE}{\frac{SSE+SSA}{n-1}} \\
&= (n-1)ln\Big(\frac{SSE+SSA}{n-1}\Big) + ln(n) + \frac{(c-1)MSA + (n-c)MSE}{\frac{(n-c)MSE+(c-1)MSA}{n-1}} \\
&= n - 1 + ln(n) + (n-1)ln\Big(\frac{(n-c)MSE + (c-1)MSA}{n-1}\Big). \qquad (A.22)
\end{aligned}$$

Subtracting (A.21) from (A.22), we obtain

$$\begin{aligned}
\Lambda &= n - 1 + ln(n) + (n-1)ln\Big(\frac{(n-c)MSE + (c-1)MSA}{n-1}\Big) \\
&\quad -n + 1 - (n-c)ln(MSE) - (c-1)ln(MSA) - ln(n) \\
&= -(n-1)ln(MSE) + (c-1)ln(MSE) - (c-1)ln(MSA) \\
&= +(n-1)ln\Big(\frac{(n-c)MSE + (c-1)MSA}{n-1}\Big) \\
&= (n-1)ln\Big(\frac{n-c}{n-1} + \frac{c-1}{n-1}F\Big) - (c-1)ln(F).
\end{aligned} \tag{A.23}$$

where $\hat{\eta}_i = \hat{\sigma}_e^2 + m_i \hat{\sigma}_b^2$, $\hat{\lambda}_i = \frac{m_i}{\hat{\eta}_i}$.

A.2 Restricted Maximum Likelihood Method Likelihood Ratio Test (RLRT) for Testing $H_0 : \sigma_b^2 = 0$

Under model 2.1, the likelihood function using restricted maximum likelihood is given by

$$\begin{aligned}
\ell_R &= -\frac{1}{2}(n-1)log(2\pi) - \frac{1}{2}log(n) - \frac{1}{2}(n-c)log(\sigma_e^2) \\
&\quad -\frac{1}{2}(c-1)log(\eta) - \frac{SSE}{2\sigma_e^2} - \frac{SSA}{2\eta}.
\end{aligned} \tag{A.24}$$

Differentiating this likelihood Equation with respect to the parameters η and σ_e^2, we get

$$\begin{aligned}
\frac{\partial \ell_R}{\partial \sigma_e^2} &= -\frac{n-c}{2\sigma_e^2} + \frac{SSE}{2(\sigma_e^2)^2} \\
\frac{\partial \ell_R}{\partial \eta} &= -\frac{c-1}{2\eta} + \frac{SSA}{2\eta^2}.
\end{aligned} \tag{A.25}$$

Equating the partial derivatives in (A.25) to zero and referring to the solutions as $\hat{\eta}$ and $\hat{\sigma}_e^2$ we obtain

$$\begin{aligned} \hat{\sigma}_e^2 &= \frac{SSE}{(n-c)} = MSE; \\ \hat{\eta} &= \frac{SSA}{c-1} = MSA. \end{aligned} \tag{A.26}$$

Therefore

$$\hat{\sigma_b^2} = \frac{1}{m}(MSA - MSE).$$

Hence, multiplying by 2 the restricted maximum likelihood Equation under the full model

$$\begin{aligned} 2 \overset{MAX}{H_A} \ell_R &= (1-n)log(2\pi e) - log(n) \\ &\quad -(n-c)log(MSE) - (c-1)log(MSA). \end{aligned} \tag{A.27}$$

A.2.1 Under the null hypothesis H_0

We know that under H_0, $\sigma_b^2 = 0$, so in this case η reduces to σ_e^2. Therefore, if we substitute this quantity in (A.24), we obtain

$$\begin{aligned} \ell_R &= -\frac{1}{2}(n-1)log(2\pi) - \frac{1}{2}log(n) - \frac{1}{2}(n-1)log(\sigma_e^2) \\ &\quad -\frac{SSE+SSA}{2\sigma_e^2}. \end{aligned} \tag{A.28}$$

Differentiating Equation (A.28) with respect to σ_e^2, we get

$$\frac{\partial \ell_R}{\partial \sigma_e^2} = -\frac{n-1}{2\sigma_e^2} + \frac{SSE+SSA}{2(\sigma_e^2)^2}. \tag{A.29}$$

Equating to zero the partial derivative in (A.29) with respect to σ_e^2, and representing the solution by $\hat{\sigma}_e^2$, we obtain

$$\begin{aligned}\hat{\sigma}_e^2 &= \frac{SSE+SSA}{n-1} \\ &= \frac{n-c}{n-1}MSE+\frac{c-1}{n-1}MSA. \end{aligned} \tag{A.30}$$

Therefore, -2 multiplied by the restricted maximum likelihood becomes

$$\begin{aligned}-2 \overset{MAX}{H_0}\ \ell_R &= (n-1)log(2\pi e)+log(n) \\ &\quad +(n-1)log\Big(\frac{SSE+SSA}{n-1}.\Big) \end{aligned} \tag{A.31}$$

Adding equations (A.27) and (A.31), we obtain the restricted likelihood ratio test as

$$\Lambda_R = \begin{cases} (n-1)\,log\Big(\frac{n-c}{n-1}+\frac{c-1}{n-1}\,F\Big)-(c-1)\,log(F) & if \quad F>1, \\ 0 & if \quad F\leq 1. \end{cases} \tag{A.32}$$

where $F=\frac{MSA}{MSE}$.

A.3 Proof of 2.11

$$\begin{aligned} var(\hat{\beta}) &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}var(\mathbf{Y})\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{V}\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \\ &= \mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \end{aligned}$$

A.4 Proof of 2.23

$$\begin{aligned}
\sum_{i=1}^{c}\mathbf{x}_i'\hat{V}_i^{-1}\mathbf{x}_i &= \sum_{i=1}^{c}\mathbf{1}_{m_i}'\Big[\frac{1}{\hat{\sigma}_e^2}(\mathbf{I}_{m_i}) - \frac{\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}\mathbf{J}_{m_i}\Big]\mathbf{1}_{m_i} \\
&= \sum_{i=1}^{c}\mathbf{1}_{m_i}'\Big[\frac{1}{\hat{\sigma}_e^2}(\mathbf{I}_{m_i}) - \frac{\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}(\mathbf{1}_{m_i}\mathbf{1}_{m_i}')\Big]\mathbf{1}_{m_i} \\
&= \sum_{i=1}^{c}\Big[\frac{1}{\hat{\sigma}_e^2}(\mathbf{1}_{m_i}'\mathbf{1}_{m_i}) - \frac{\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}\mathbf{1}_{m_i}'\mathbf{1}_{m_i}\mathbf{1}_{m_i}'\mathbf{1}_{m_i}\Big] \\
\sum_{i=1}^{c}\mathbf{x}_i'\hat{V}_i^{-1}\mathbf{x}_i &= \sum_{i=1}^{c}\Big[\frac{m_i}{\hat{\sigma}_e^2} - \frac{m_i^2\,\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}\Big] \\
&= \sum_{i=1}^{c}\frac{1}{\hat{\sigma}_e^2}\Big[\frac{m_i\hat{\sigma}_e^2+m_i^2\hat{\sigma}_b^2-m_i^2\hat{\sigma}_b^2}{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}\Big] \\
&= \sum_{i=1}^{c}\frac{1}{\hat{\sigma}_e^2}\Big[\frac{m_i\hat{\sigma}_e^2}{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}\Big] \\
&= \sum_{i=1}^{c}\Big[\frac{m_i}{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}\Big] \\
&= \sum_{i=1}^{c}\hat{\lambda}_i \qquad\qquad (A.33)
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^{c}\mathbf{x}_i'\hat{V}_i^{-1}\hat{\mathbf{e}}_i\hat{\mathbf{e}}_i'\hat{V}_i^{-1}\mathbf{x}_i &= \sum_{i=1}^{c}\mathbf{1}_{m_i}'\Big\{\frac{1}{\hat{\sigma}_e^2}(\mathbf{I}_{m_i}) - \frac{\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}(\mathbf{J}_{m_i})\Big\}\hat{\mathbf{e}}_i\hat{\mathbf{e}}_i' \\
&\quad\times\Big\{\frac{1}{\hat{\sigma}_e^2}(\mathbf{I}_{m_i}) - \frac{\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}(\mathbf{J}_{m_i})\Big\}\mathbf{1}_{m_i} \\
&= \sum_{i=1}^{c}\mathbf{1}_{m_i}'\Big\{\frac{1}{\hat{\sigma}_e^2}(\mathbf{I}_{m_i}) - \frac{\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}(\mathbf{1}_{m_i}\mathbf{1}_{m_i}')\Big\}\hat{\mathbf{e}}_i\hat{\mathbf{e}}_i' \\
&\quad\times\Big\{\frac{1}{\hat{\sigma}_e^2}(\mathbf{I}_{m_i}) - \frac{\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}(\mathbf{1}_{m_i}\mathbf{1}_{m_i}')\Big\}\mathbf{1}_{m_i} \\
&= \sum_{i=1}^{c}\Big\{\frac{1}{\hat{\sigma}_e^2}(\mathbf{1}_{m_i}'\mathbf{I}_{m_i}) - \frac{\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}(\mathbf{1}_{m_i}'\mathbf{1}_{m_i}\mathbf{1}_{m_i}')\Big\}\hat{\mathbf{e}}_i\hat{\mathbf{e}}_i' \\
&\quad\times\Big\{\frac{1}{\hat{\sigma}_e^2}(\mathbf{I}_{m_i}\mathbf{1}_{m_i}) - \frac{\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}(\mathbf{1}_{m_i}\mathbf{1}_{m_i}'\mathbf{1}_{m_i})\Big\} \\
&= \sum_{i=1}^{c}\Big\{\frac{1}{\hat{\sigma}_e^2}(\mathbf{1}_{m_i}') - \frac{m_i\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}(\mathbf{1}_{m_i}')\Big\}\hat{\mathbf{e}}_i\hat{\mathbf{e}}_i'
\end{aligned}$$

$$
\begin{aligned}
& \times\Big\{\frac{1}{\hat{\sigma}_e^2}(\mathbf{1}_{m_i}) - \frac{m_i\hat{\sigma}_b^2}{\hat{\sigma}_e^2(\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2)}(\mathbf{1}_{m_i})\Big\} \\
= & \Big(\frac{1}{\hat{\sigma}_e^2}\Big)^2\Big[\sum_{i=1}^{c}\Big\{1-\frac{m_i\hat{\sigma}_b^2}{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}\Big\}\mathbf{1}'_{m_i}\hat{\mathbf{e}}_i\hat{\mathbf{e}}'_i\mathbf{1}_{m_i}\Big\{1-\frac{m_i\hat{\sigma}_b^2}{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}\Big\}\Big]. \\
= & \Big(\frac{1}{\hat{\sigma}_e^2}\Big)^2\Big[\sum_{i=1}^{c}\Big\{1-\frac{m_i\hat{\sigma}_b^2}{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}\Big\}^2\mathbf{1}'_{m_i}\hat{\mathbf{e}}_i\hat{\mathbf{e}}'_i\mathbf{1}_{m_i}\Big] \\
= & \Big(\frac{1}{\hat{\sigma}_e^2}\Big)^2\Big[\sum_{i=1}^{c}\Big\{\frac{\hat{\sigma}_e^2}{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}\Big\}^2\mathbf{1}'_{m_i}\hat{\mathbf{e}}_i\hat{\mathbf{e}}'_i\mathbf{1}_{m_i}\Big] \\
= & \sum_{i=1}^{c}\Big\{\frac{\hat{\sigma}_e^2}{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}\Big\}^2(\mathbf{1}'_{m_i}\hat{\mathbf{e}}_i)^2 \\
= & \sum_{i=1}^{c}\Big\{\frac{\hat{\sigma}_e^2}{\hat{\sigma}_e^2+m_i\hat{\sigma}_b^2}\Big\}^2(m_i(\bar{y}_{i.}-\hat{\beta}))^2 \\
= & \sum_{i=1}^{c}\hat{\lambda}^2(m_i(\bar{y}_{i.}-\hat{\beta}))^2 \qquad \text{(A.34)}
\end{aligned}
$$

Substituting (A.34) and (A.34) into (2.23) gives

$$
\widehat{var}(\hat{\beta}) = \frac{\sum_{i=1}^{c}\hat{\lambda}_i^2(\bar{y}_{i.}-\hat{\beta})^2}{\Big(\sum_{i=1}^{c}\hat{\lambda}_i\Big)^2}. \qquad \text{(A.35)}
$$

A.5 Proof of 2.24

In this case $\lambda_i = \lambda$ for all i, therefore

$$
\begin{aligned}
\widehat{var}(\hat{\beta}) &= \widehat{var}(\bar{y}_{..}) \qquad \text{(A.36)} \\
&= \frac{1}{c(c-1)}\sum_{i=1}^{c}(\bar{y}_{i.}-\bar{y}_{..})^2.
\end{aligned}
$$

Appendix B

Proofs and Additional Tables for Chapter 3

B.1 Derivation of the Multiplication Factor Used to Correct the Huber-White Variance estimator in the Unbalanced Data case, Equation (3.3)

$$
\begin{aligned}
E\big(\widehat{var(\hat{\beta})}\big) &= \frac{1}{\Big(\sum_{i=1}^{c}\lambda_i\Big)^2} E\Big[\sum_{i=1}^{c}\lambda_i^2(\bar{y}_{i.}-\hat{\beta})^2\Big] \\
&= \frac{1}{\Big(\sum_{i=1}^{c}\lambda_i\Big)^2} \Big[\sum_{i=1}^{c}\lambda_i^2[(\bar{y}_{i.}-\beta)-(\hat{\beta}-\beta)]^2\Big] \\
&= \frac{1}{\Big(\sum_{i=1}^{c}\lambda_i\Big)^2} \Big[\sum_{i=1}^{c}\lambda_i^2\Big\{E(\bar{y}_{i.}-\beta)^2+E(\hat{\beta}-\beta)^2 \\
&\quad -2E(\bar{y}_{i.}-\hat{\beta})(\hat{\beta}-\beta)\Big\}\Big] \\
&= \frac{1}{\Big(\sum_{i=1}^{c}\lambda_i\Big)^2} \Big[\sum_{i=1}^{c}\lambda_i^2 var(\bar{y}_{i.})+\sum_{i=1}^{c}\lambda_i^2 var(\hat{\beta}) \\
&\quad -2\sum_{i=1}^{c}\lambda_i^2 cov(\bar{y}_{i.},\hat{\beta})\Big].
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{\left(\sum_{i=1}^{c}\lambda_i\right)^2}\Big[\sum_{i=1}^{c}\lambda_i^2 var(\bar{y}_{i.}) + \sum_{i=1}^{c}\lambda_i^2 var(\hat{\beta}) \\
&\quad -2\sum_{i=1}^{c}\lambda_i^2 cov\Big(\bar{y}_{i.}, \frac{\sum_{j=1}^{c}\lambda_j\bar{y}_{j.}}{\sum_{j=1}^{c}\lambda_j}\Big)\Big] \\
&= \frac{1}{\left(\sum_{i=1}^{c}\lambda_i\right)^2}\Big[\sum_{i=1}^{c}\Big(\lambda_i^2\frac{1}{\lambda_i}\Big) + (var(\hat{\beta}))\sum_{i=1}^{c}(\lambda_i^2) \\
&\quad -2\sum_{i=1}^{c}\Big(\lambda_i^2 var(\bar{y}_{i.})\frac{\lambda_i}{\sum_{j=1}^{c}\lambda_j}\Big)\Big] \\
&= \frac{1}{\left(\sum_{i=1}^{c}\lambda_i\right)^2}\Big[\sum_{i=1}^{c}\Big(\lambda_i^2\frac{1}{\lambda_i}\Big) + (var(\hat{\beta}))\sum_{i=1}^{c}(\lambda_i^2) \\
&\quad -2\sum_{i=1}^{c}(\lambda_i^2)\frac{1}{\lambda_i}\Big(\frac{\lambda_i}{\sum_{j=1}^{c}\lambda_j}\Big)\Big] \\
&= \frac{1}{\left(\sum_{i=1}^{c}\lambda_i\right)^2}\Big[\sum_{i=1}^{c}\Big(\lambda_i^2\frac{1}{\lambda_i}\Big) + (var(\hat{\beta}))\sum_{i=1}^{c}(\lambda_i^2) - 2\frac{\sum_{i=1}^{c}(\lambda_i^2)}{\sum_{j=1}^{c}\lambda_j}\Big] \\
&= \frac{1}{\left(\sum_{i=1}^{c}\lambda_i\right)^2}\Big[\sum_{i=1}^{c}\Big(\lambda_i^2\frac{1}{\lambda_i}\Big) + (var(\hat{\beta}))\sum_{i=1}^{c}(\lambda_i^2) \\
&\quad -2(var(\hat{\beta}))\sum_{i=1}^{c}(\lambda_i^2)\Big] \\
&= var(\hat{\beta}) - (var(\hat{\beta}))\frac{\sum_{i=1}^{c}(\lambda_i^2)}{(\sum_{i=1}^{c}\lambda_i)^2} \\
&= var(\hat{\beta})\Big[1 - \frac{\sum_{i=1}^{c}(\lambda_i^2)}{(\sum_{i=1}^{c}\lambda_i)^2}\Big] \\
&= \frac{(\sum_{i=1}^{c}\lambda_i)^2 - \sum_{i=1}^{c}(\lambda_i^2)}{(\sum_{i=1}^{c}\lambda_i)^2} var(\hat{\beta})
\end{aligned}
$$

$$
\therefore \quad \frac{E\big(\widehat{var}(\hat{\beta})\big)}{var(\hat{\beta})} = \frac{(\sum_{i=1}^{c}\lambda_i)^2 - \sum_{i=1}^{c}(\lambda_i^2)}{(\sum_{i=1}^{c}\lambda_i)^2}.
$$

B.2 Extra Tables and Plots

Figure B.1: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.01

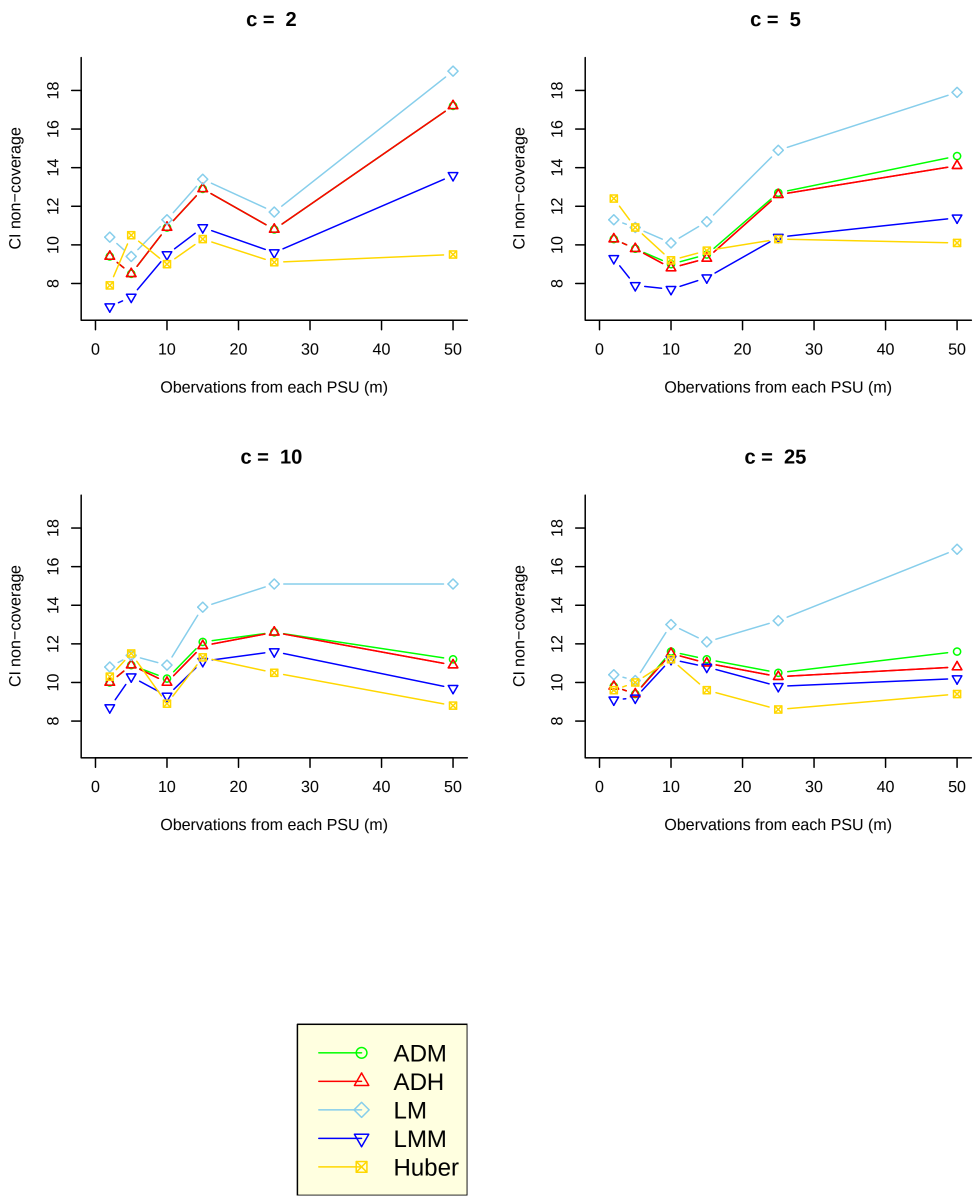


Table B.1: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the balanced data case with ρ=0, using parametric bootstrap to estimate $\widehat{var}(\widehat{var}(\hat{\beta}))$

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(reject H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	Lrt	ADM	ADH	LMM	Hub
2	2	1.236	1.236	1.460	1.028	8.0	8.0	3.6	9.3	10.1	15.519	2.962	17.119	5.098
2	5	1.208	1.208	1.457	1.001	9.1	9.1	0.9	9.3	5.7	1.157	1.524	1.395	3.207
2	10	1.154	1.154	1.374	0.935	9.7	9.7	0.6	11.3	5.1	0.807	1.050	0.973	2.221
2	15	1.184	1.184	1.444	0.983	10.7	10.7	0.4	10.3	4.5	0.649	0.812	0.790	1.830
2	25	1.162	1.162	1.418	0.941	8.6	8.6	0.2	10.6	4.7	0.502	0.637	0.610	1.399
2	50	1.272	1.272	1.546	1.017	9.2	9.2	0.3	10.5	3.4	0.346	0.419	0.417	0.969
5	2	1.156	1.156	1.292	1.092	8.5	8.5	2.3	8.3	10.5	1.237	1.213	1.276	1.301
5	5	1.141	1.141	1.266	1.044	8.7	8.5	4.1	9.2	7.2	0.711	0.728	0.750	0.812
5	10	1.101	1.101	1.225	0.985	9.6	9.6	5.2	10.1	6.4	0.492	0.504	0.526	0.565
5	15	1.131	1.131	1.256	1.010	9.4	9.4	5.2	9.8	6.2	0.400	0.410	0.429	0.461
5	25	1.140	1.140	1.249	1.004	10.0	9.9	5.3	9.3	6.6	0.310	0.318	0.331	0.357
5	50	1.125	1.125	1.254	1.011	9.8	9.7	5.6	9.6	5.8	0.217	0.222	0.234	0.253
10	2	1.057	1.057	1.136	1.006	9.0	8.9	5.3	10.9	10.0	0.786	0.789	0.802	0.800
10	5	1.065	1.065	1.118	0.977	9.2	9.1	7.6	11.6	9.5	0.489	0.495	0.500	0.502
10	10	1.021	1.021	1.089	0.965	9.4	9.4	7.8	10.4	7.5	0.341	0.344	0.354	0.358
10	15	1.071	1.071	1.130	0.994	9.6	9.4	7.8	10.6	8.7	0.279	0.282	0.288	0.292
10	25	1.077	1.077	1.142	0.991	8.5	8.4	7.2	9.8	7.9	0.216	0.218	0.224	0.225
10	50	1.016	1.016	1.077	0.954	10.2	10.2	8.4	10.4	8.2	0.153	0.154	0.159	0.162
25	2	1.038	1.038	1.080	0.995	9.8	9.8	8.2	9.4	10.7	0.483	0.483	0.487	0.480
25	5	0.980	0.980	0.990	0.946	11.3	11.3	11.1	11.7	8.9	0.303	0.304	0.304	0.305
25	10	1.044	1.044	1.058	1.007	10.3	10.2	9.8	10.7	10.1	0.214	0.215	0.216	0.216
25	15	1.042	1.042	1.055	0.999	9.2	9.2	8.8	10.0	11.3	0.175	0.176	0.176	0.176
25	25	1.164	1.164	1.177	1.108	7.4	7.3	7.0	8.3	8.5	0.134	0.135	0.136	0.135
25	50	1.117	1.117	1.151	1.073	8.1	8.1	7.5	8.3	9.1	0.095	0.096	0.097	0.096

Table B.2: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the balanced data case with ρ=0.05, using parametric bootstrap to estimate $\widehat{var}(\widehat{var}(\hat{\beta}))$

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(reject H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	Lrt	ADM	ADH	LMM	Hub
2	2	1.196	1.196	1.414	1.035	8.7	8.7	3.5	8.3	10.8	12.404	3.117	15.705	5.371
2	5	1.128	1.128	1.368	1.031	11.3	11.3	1.1	10.7	8.6	1.199	1.800	1.484	3.557
2	10	0.985	0.985	1.204	0.949	17.0	17.0	1.1	11.0	9.7	0.866	1.371	1.099	2.725
2	15	0.911	0.911	1.099	0.909	20.4	20.4	2.2	11.4	12.8	0.734	1.283	0.949	2.421
2	25	0.836	0.836	0.985	0.859	24.2	24.1	3.1	9.9	16.8	0.616	1.198	0.804	2.116
2	50	0.902	0.902	1.001	0.926	29.5	29.1	4.6	10.9	26.2	0.531	1.251	0.675	1.838
5	2	1.147	1.147	1.272	1.109	9.6	9.5	3.1	9.5	13.4	1.283	1.264	1.319	1.368
5	5	1.097	1.097	1.235	1.103	10.1	10.0	5.4	7.4	14.1	0.754	0.794	0.806	0.917
5	10	0.968	0.968	1.090	1.006	14.0	13.7	8.3	9.6	21.7	0.556	0.601	0.606	0.700
5	15	0.919	0.919	1.030	0.973	16.0	15.6	9.8	9.7	27.2	0.477	0.525	0.523	0.613
5	25	0.938	0.938	1.017	0.994	19.5	18.6	12.9	11.4	41.2	0.418	0.479	0.454	0.546
5	50	0.980	0.981	1.016	1.008	17.3	14.6	11.3	10.1	65.5	0.381	0.455	0.400	0.485
10	2	1.069	1.069	1.156	1.051	9.5	9.5	6.6	9.6	13.7	0.815	0.820	0.832	0.844
10	5	0.940	0.940	1.004	0.952	12.7	12.6	10.3	11.5	17.6	0.518	0.529	0.537	0.564
10	10	0.936	0.936	1.008	0.990	13.9	13.2	10.1	9.4	35.6	0.397	0.413	0.416	0.447
10	15	0.951	0.951	1.012	1.008	14.4	13.9	11.0	9.5	46.4	0.345	0.364	0.362	0.392
10	25	1.017	1.017	1.048	1.048	14.5	13.2	11.4	11.3	67.2	0.309	0.331	0.318	0.346
10	50	1.057	1.057	1.066	1.067	13.1	10.7	9.9	9.5	89.6	0.279	0.302	0.281	0.306
25	2	1.012	1.012	1.061	1.000	11.1	11.1	9.6	10.9	15.0	0.498	0.500	0.505	0.504
25	5	0.921	0.921	0.936	0.964	12.4	12.2	11.5	10.7	28.5	0.326	0.329	0.329	0.344
25	10	0.920	0.920	0.933	0.957	14.3	14.1	13.1	12.0	54.7	0.252	0.257	0.254	0.266
25	15	0.938	0.938	0.948	0.960	11.6	10.8	10.7	10.2	77.1	0.225	0.230	0.227	0.235
25	25	0.981	0.981	0.983	0.987	11.3	10.4	10.4	10.1	93.1	0.199	0.205	0.200	0.206
25	50	1.059	1.059	1.059	1.059	10.0	9.3	9.5	9.3	99.4	0.176	0.181	0.176	0.181

Table B.3: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT with ρ=0.01.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.056	1.056	1.245	0.919	9.8	9.8	13.1	10.1	9.5	4.376	2.925	4.847	5.142
2	5	1.183	1.183	1.429	1.027	11.2	10.9	11.1	11.3	6.6	1.313	1.615	1.430	3.298
2	10	1.091	1.091	1.312	0.925	10.4	10.4	11.4	8.2	6.1	0.894	1.112	0.983	2.369
2	15	1.230	1.231	1.515	1.139	10.6	10.6	10.1	10.3	7.1	0.753	0.964	0.853	2.065
2	25	1.059	1.059	1.310	0.983	13.1	13.1	13.9	11.1	6.4	0.580	0.729	0.660	1.600
2	50	0.994	0.994	1.231	0.975	17.0	17.0	15.6	10.8	10.3	0.443	0.610	0.507	1.236
5	2	0.999	0.999	1.106	0.938	10.6	10.3	10.8	10.8	11.2	1.190	1.202	1.219	1.291
5	5	0.983	0.983	1.083	0.886	11.8	11.8	11.6	11.9	8.0	0.723	0.728	0.737	0.805
5	10	1.038	1.038	1.161	0.980	10.6	10.6	10.5	9.8	10.2	0.519	0.525	0.541	0.597
5	15	1.128	1.128	1.267	1.085	9.2	9.2	9.6	9.0	9.9	0.423	0.427	0.440	0.494
5	25	0.969	0.969	1.088	0.944	12.0	12.0	11.8	10.4	12.1	0.334	0.340	0.349	0.395
5	50	1.077	1.077	1.211	1.119	12.5	12.4	11.8	9.3	22.2	0.263	0.268	0.279	0.315
10	2	1.074	1.074	1.148	1.012	9.9	9.9	10.7	10.2	11.1	0.795	0.797	0.802	0.804
10	5	0.982	0.982	1.039	0.929	10.6	10.5	10.9	11.3	9.5	0.495	0.497	0.498	0.512
10	10	1.113	1.113	1.184	1.066	9.8	9.5	10.9	9.5	11.3	0.354	0.354	0.357	0.370
10	15	0.984	0.984	1.053	0.970	10.4	10.4	10.3	9.9	11.9	0.289	0.289	0.293	0.308
10	25	0.985	0.985	1.056	0.981	11.9	11.8	12.2	11.8	20.1	0.236	0.236	0.241	0.250
10	50	0.899	0.899	0.970	0.957	13.2	13.0	12.5	10.9	33.4	0.181	0.182	0.188	0.199
25	2	1.086	1.086	1.136	1.049	8.0	8.0	9.1	8.7	9.4	0.480	0.480	0.481	0.478
25	5	0.985	0.985	0.995	0.965	10.7	10.7	11.0	10.8	11.9	0.306	0.306	0.304	0.310
25	10	0.983	0.983	0.999	0.988	10.2	10.3	11.0	9.3	15.5	0.220	0.219	0.218	0.225
25	15	0.928	0.928	0.940	0.943	11.3	11.5	11.8	10.7	20.0	0.182	0.182	0.180	0.188
25	25	0.985	0.985	1.003	1.017	11.5	11.4	11.4	9.7	30.5	0.147	0.146	0.146	0.152
25	50	1.001	1.001	1.033	1.042	11.3	11.3	10.9	9.3	54.0	0.113	0.113	0.115	0.117

Figure B.2: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.05

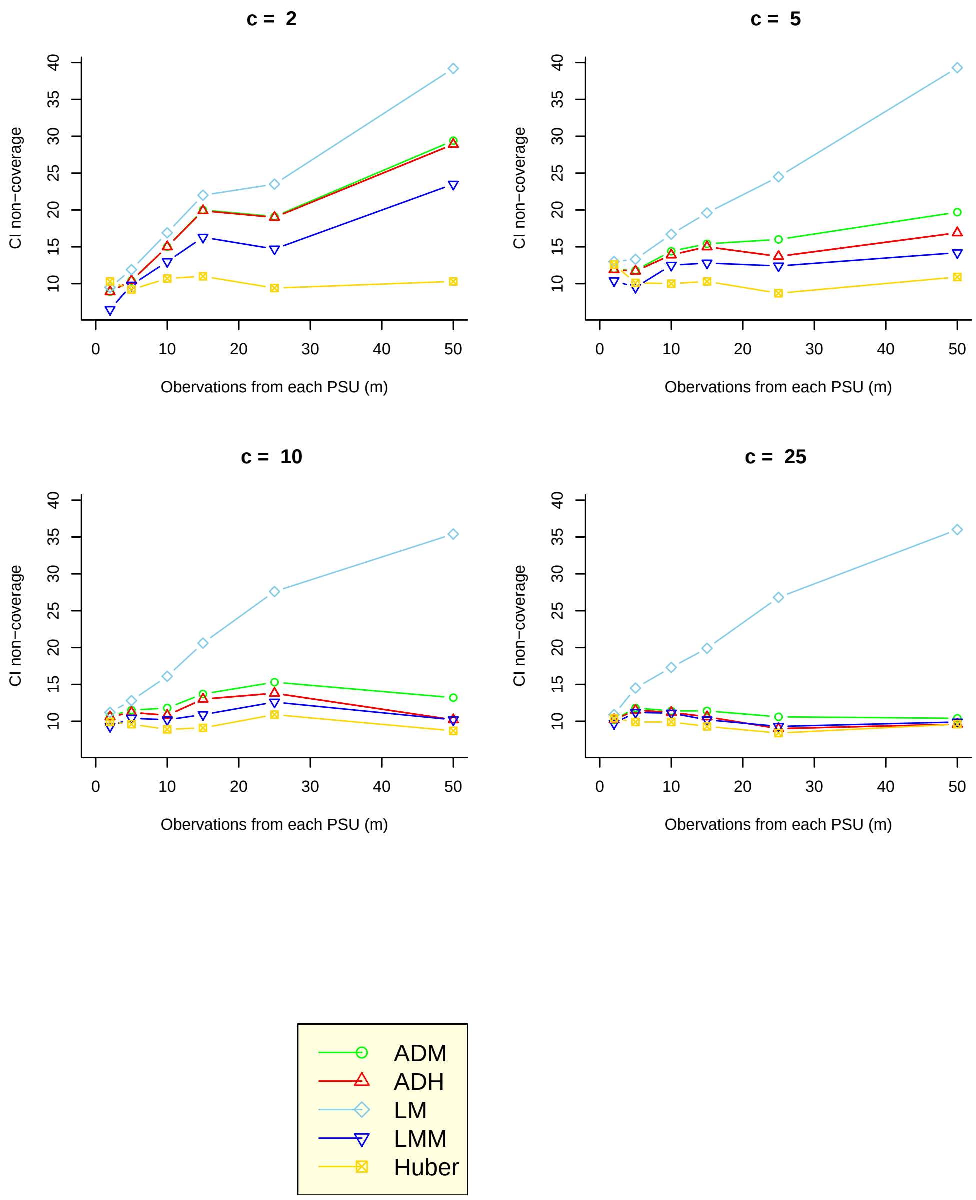


Figure B.3: Confidence interval lengths using different variance estimation methods and for various values of m and c, ρ=0.01

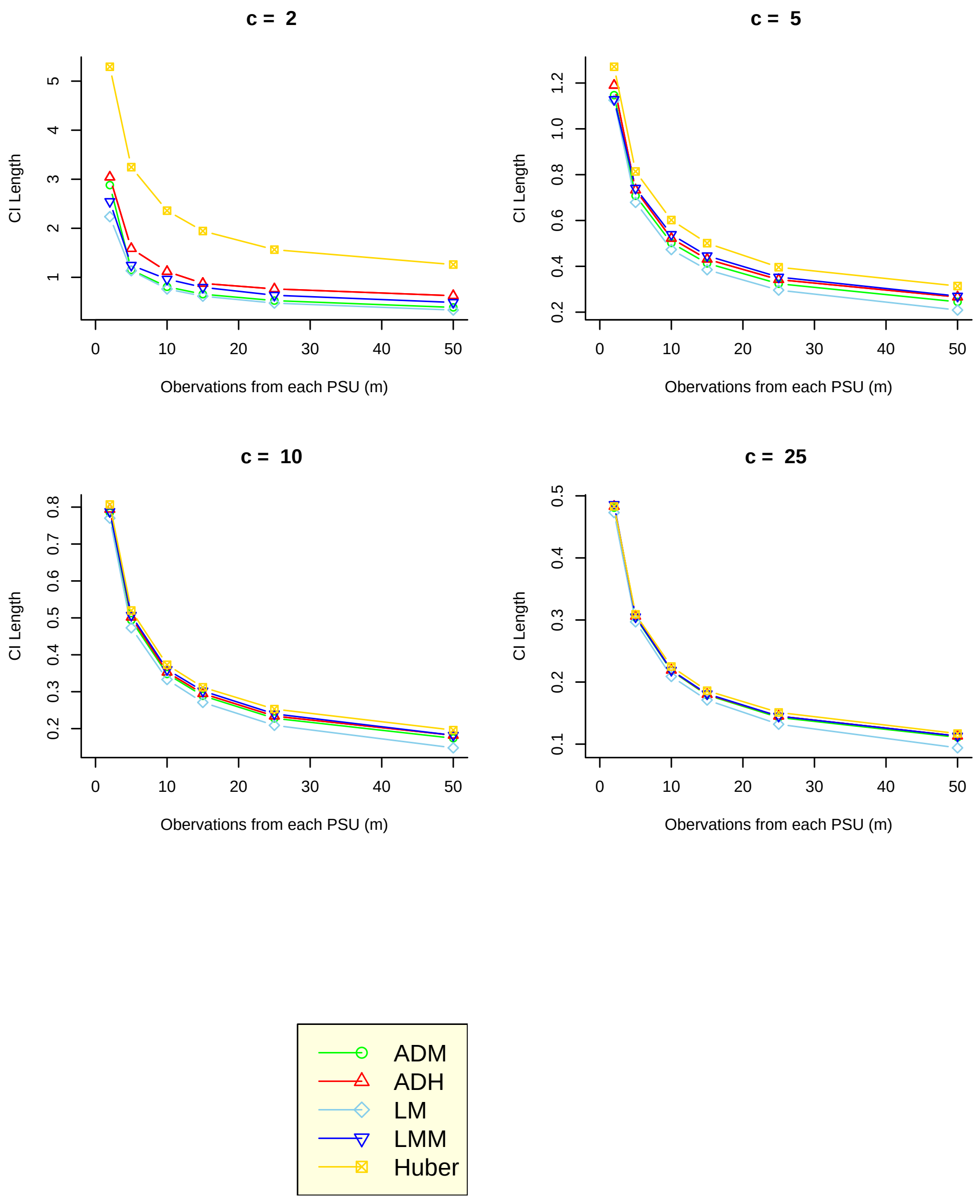


Figure B.4: Confidence interval lengths using different variance estimation methods and for various values of m and c, ρ=0.05

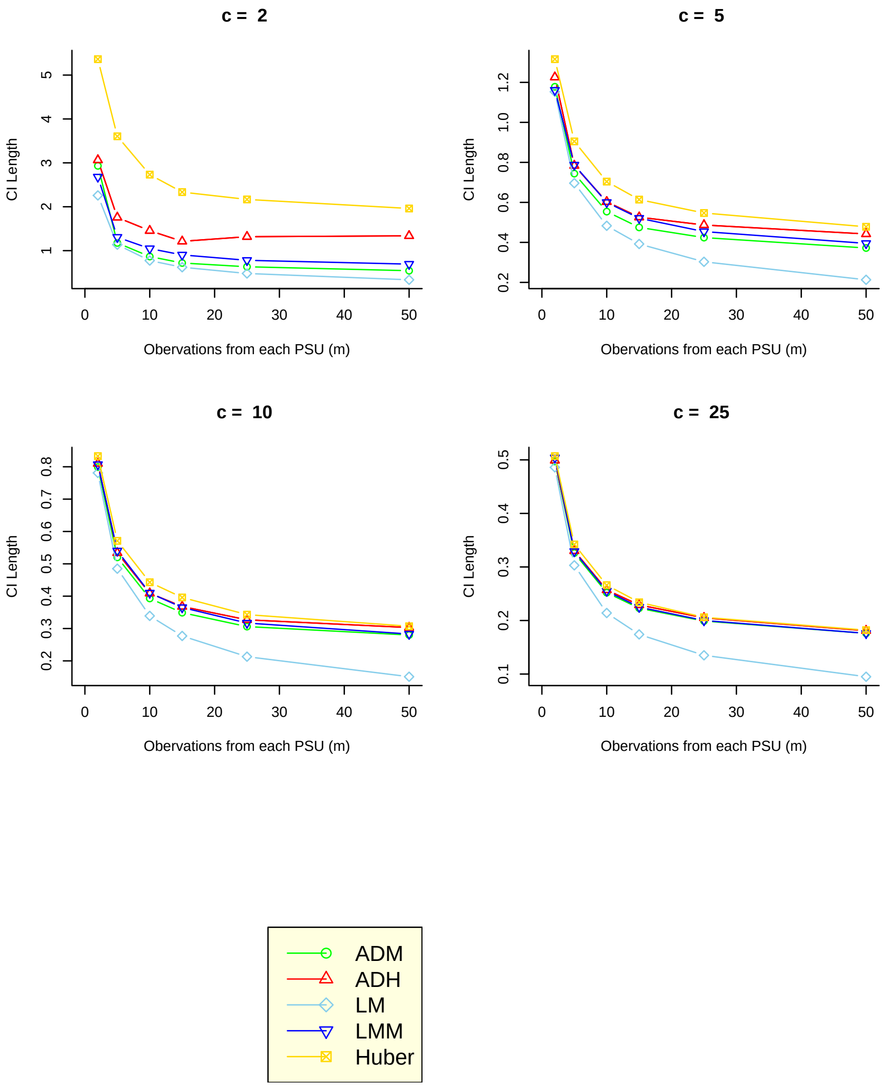


Table B.4: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT with ρ=0.05.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.218	1.218	1.447	1.083	7.7	7.7	12.4	8.8	10.5	5.365	3.066	5.864	5.350
2	5	1.042	1.042	1.243	0.930	13.3	13.1	13.3	11.2	7.6	1.355	1.743	1.469	3.489
2	10	1.145	1.145	1.364	1.097	14.8	14.8	14.8	10.2	11.6	1.022	1.516	1.147	2.825
2	15	1.057	1.057	1.275	1.054	16.7	16.7	15.0	10.2	14.7	0.900	1.349	1.018	2.458
2	25	1.045	1.045	1.226	1.096	21.5	21.5	16.9	8.2	20.6	0.849	1.366	0.965	2.277
2	50	1.044	1.044	1.159	1.089	27.7	27.7	22.1	9.3	29.0	0.838	1.360	0.919	1.981
5	2	1.129	1.129	1.255	1.070	8.9	8.8	9.0	9.8	11.9	1.215	1.221	1.249	1.307
5	5	1.086	1.086	1.218	1.082	10.5	10.5	10.1	9.2	15.1	0.784	0.795	0.822	0.908
5	10	1.032	1.032	1.155	1.067	12.8	12.7	11.4	9.9	21.4	0.596	0.605	0.631	0.704
5	15	1.020	1.020	1.140	1.076	13.4	13.3	11.0	9.0	27.1	0.520	0.528	0.554	0.615
5	25	0.919	0.919	1.001	0.977	16.1	15.5	13.8	9.3	42.3	0.468	0.480	0.497	0.546
5	50	0.999	0.999	1.031	1.023	14.5	14.1	12.4	9.2	66.5	0.450	0.457	0.463	0.485
10	2	0.996	0.996	1.076	0.974	10.3	10.3	9.8	10.0	12.9	0.817	0.817	0.831	0.840
10	5	1.033	1.034	1.110	1.058	9.7	9.7	9.8	8.7	18.6	0.527	0.529	0.539	0.566
10	10	0.986	0.986	1.060	1.039	10.2	10.0	10.0	8.8	32.8	0.406	0.407	0.420	0.441
10	15	0.987	0.987	1.052	1.045	13.2	12.7	11.8	9.1	46.8	0.363	0.364	0.378	0.392
10	25	0.887	0.888	0.925	0.924	15.2	15.1	13.8	11.4	61.7	0.318	0.319	0.326	0.336
10	50	1.063	1.063	1.072	1.072	10.3	9.7	9.3	8.2	88.6	0.300	0.298	0.302	0.302
25	2	0.958	0.958	1.007	0.955	10.6	10.6	11.0	11.0	15.1	0.497	0.498	0.500	0.504
25	5	1.045	1.045	1.059	1.074	8.8	8.7	9.5	8.3	29.6	0.331	0.329	0.330	0.340
25	10	0.950	0.950	0.970	0.995	11.4	11.7	11.5	9.9	53.5	0.256	0.254	0.259	0.265
25	15	0.943	0.944	0.954	0.972	11.0	11.1	10.7	9.7	72.5	0.230	0.227	0.231	0.233
25	25	0.964	0.964	0.967	0.971	12.2	12.3	12.1	12.0	93.1	0.208	0.206	0.208	0.207
25	50	0.966	0.967	0.967	0.967	8.9	9.3	8.9	9.3	99.7	0.183	0.182	0.183	0.182

Table B.5: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0.01.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(reject H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	Lrt	ADM	ADH	LMM	Hub
2	3	1.367	1.367	1.480	1.108	8.8	8.8	7.5	9.7	19.2	2.286	2.965	2.447	4.507
2	10	1.349	1.349	1.424	1.029	9.4	9.2	8.8	10.0	17.1	0.970	1.600	1.003	2.436
2	25	1.209	1.209	1.286	0.975	12.2	12.0	11.1	9.0	19.4	0.579	1.065	0.601	1.645
5	3	1.236	1.236	1.249	1.044	7.7	7.3	7.7	9.8	28.8	1.032	1.083	1.039	1.084
5	10	1.152	1.152	1.157	0.959	9.3	8.6	8.8	9.0	28.3	0.537	0.585	0.539	0.593
5	25	1.086	1.086	1.089	0.954	10.9	9.5	10.6	10.0	36.4	0.349	0.393	0.350	0.404
10	3	1.154	1.154	1.156	1.039	9.9	9.7	9.9	11.1	26.9	0.654	0.666	0.655	0.657
10	10	1.105	1.105	1.105	0.995	9.7	9.2	9.7	10.2	28.2	0.364	0.376	0.364	0.374
10	25	1.088	1.088	1.088	1.028	10.6	10.0	10.6	10.4	40.0	0.240	0.252	0.240	0.255
25	3	1.027	1.027	1.027	0.993	10.4	10.3	10.4	10.3	16.4	0.398	0.400	0.398	0.400
25	10	1.087	1.087	1.087	1.068	8.8	8.8	8.8	9.1	22.5	0.222	0.224	0.222	0.226
25	25	0.988	0.988	0.988	1.002	11.5	11.2	11.5	10.4	36.3	0.145	0.148	0.145	0.152
50	3	0.972	0.972	0.972	0.965	11.0	11.0	11.0	11.0	8.6	0.275	0.275	0.275	0.277
50	10	0.973	0.973	0.973	0.998	10.2	10.1	10.2	9.9	16.5	0.153	0.154	0.153	0.158
50	25	0.986	0.986	0.986	1.012	10.4	9.8	10.4	8.6	50.6	0.103	0.104	0.103	0.106

Table B.6: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0.05.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(reject H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	Lrt	ADM	ADH	LMM	Hub
2	3	1.311	1.311	1.418	1.107	10.0	9.9	8.3	9.7	21.0	2.435	3.208	2.603	4.801
2	10	1.224	1.224	1.285	1.031	13.0	12.2	11.8	9.7	24.3	1.101	2.040	1.138	2.871
2	25	1.072	1.072	1.114	0.985	18.8	17.9	17.1	8.9	34.4	0.727	1.711	0.749	2.241
5	3	1.191	1.191	1.205	1.045	8.7	8.5	8.6	9.1	33.8	1.075	1.136	1.082	1.150
5	10	1.033	1.033	1.036	0.941	11.6	9.9	11.6	10.1	43.2	0.598	0.674	0.599	0.693
5	25	0.959	0.959	0.960	0.931	16.7	12.7	16.7	11.0	64.9	0.442	0.531	0.442	0.544
10	3	1.111	1.111	1.112	1.034	10.1	9.6	10.0	10.9	34.6	0.683	0.699	0.683	0.697
10	10	1.046	1.046	1.046	1.024	11.9	10.5	11.9	9.7	53.4	0.413	0.436	0.413	0.445
10	25	0.989	0.989	0.989	0.988	14.1	11.3	14.1	10.8	80.1	0.313	0.342	0.313	0.345
25	3	0.984	0.984	0.984	0.985	10.6	10.5	10.6	10.0	26.7	0.415	0.417	0.415	0.424
25	10	1.059	1.059	1.059	1.080	11.0	10.3	11.0	9.2	64.6	0.257	0.263	0.257	0.269
25	25	0.987	0.987	0.987	0.992	11.1	10.0	11.1	9.9	94.0	0.199	0.206	0.199	0.207
50	3	0.930	0.930	0.930	0.962	11.7	11.7	11.7	11.2	19.4	0.286	0.287	0.286	0.294
50	10	0.956	0.956	0.956	0.988	12.0	11.7	12.0	10.4	71.7	0.181	0.183	0.181	0.187
50	25	1.016	1.016	1.016	1.017	9.6	9.2	9.6	9.2	99.7	0.143	0.145	0.143	0.145

Figure B.5: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.01

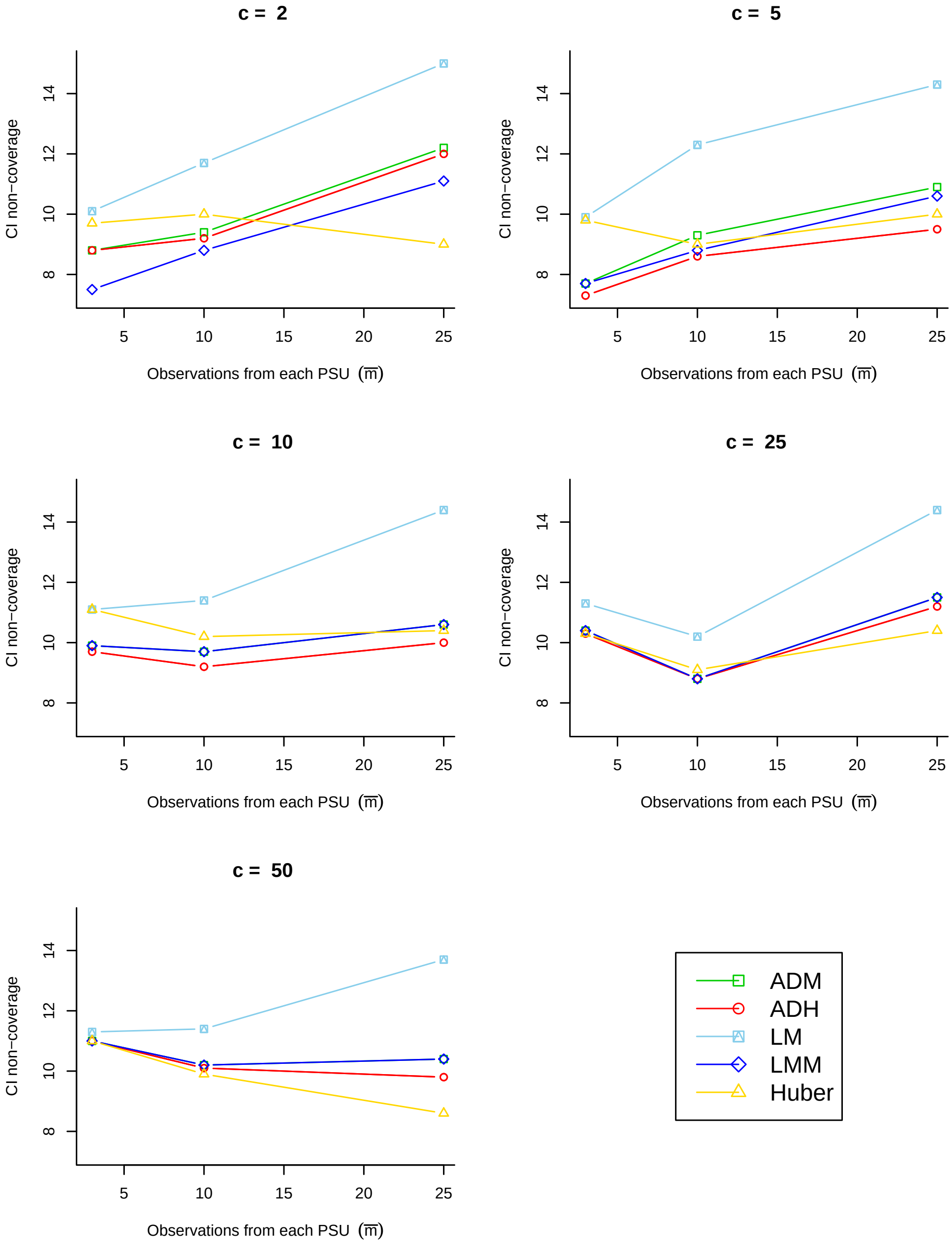


Figure B.6: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.05

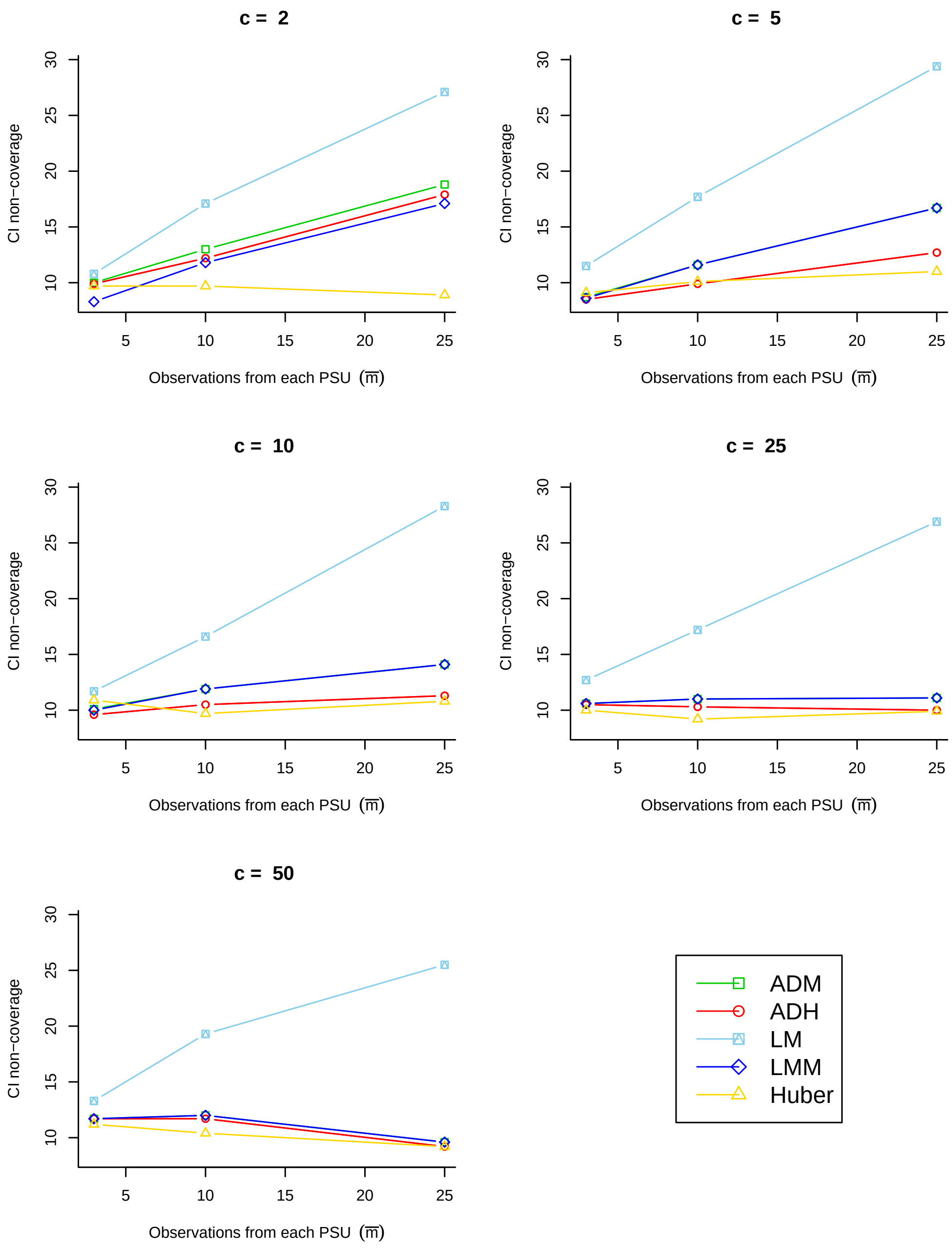


Figure B.7: Confidence interval lengths using different variance estimation methods and for various values of m and c, ρ=0.01

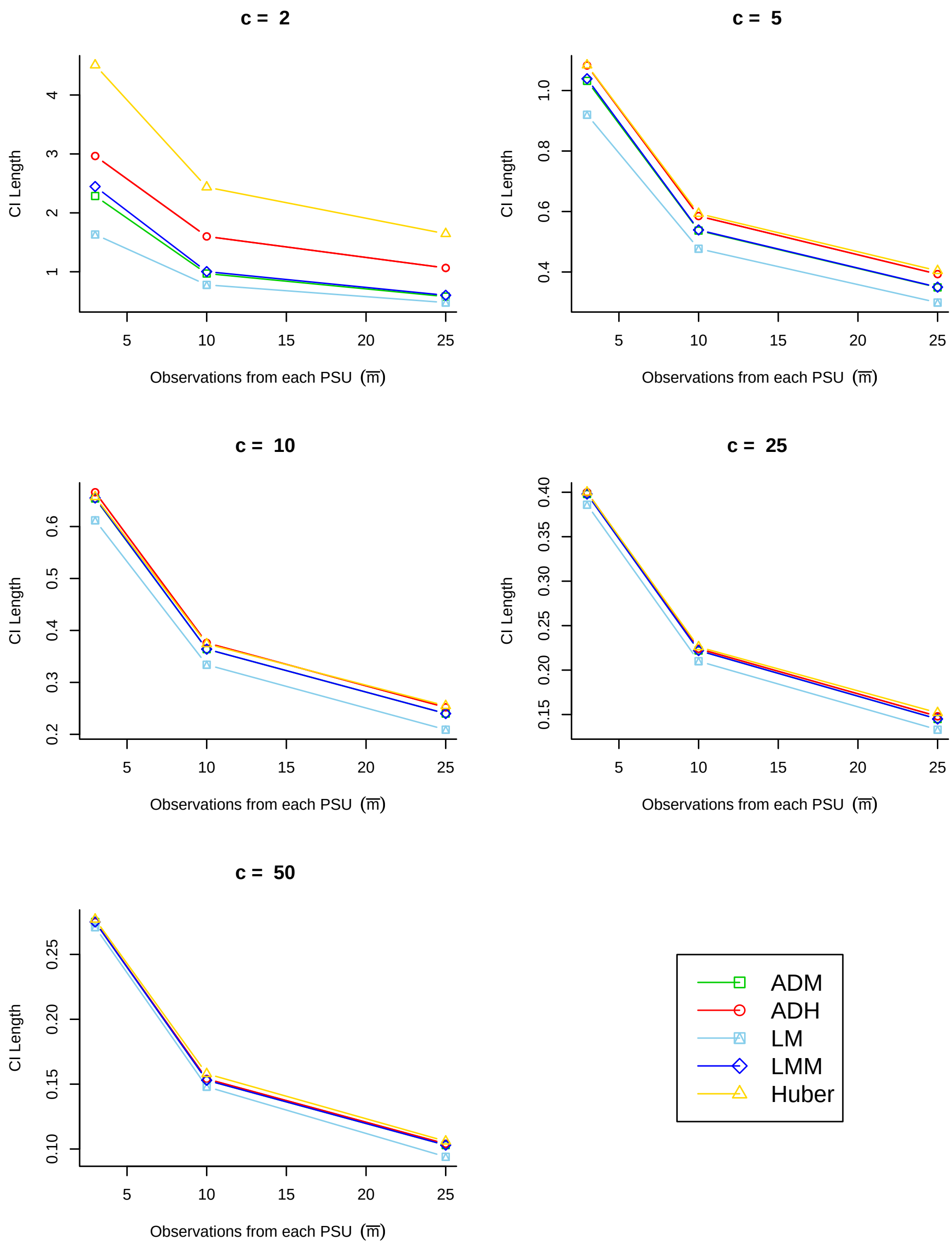


Figure B.8: Confidence interval lengths using different variance estimation methods and for various values of m and c, ρ=0.05

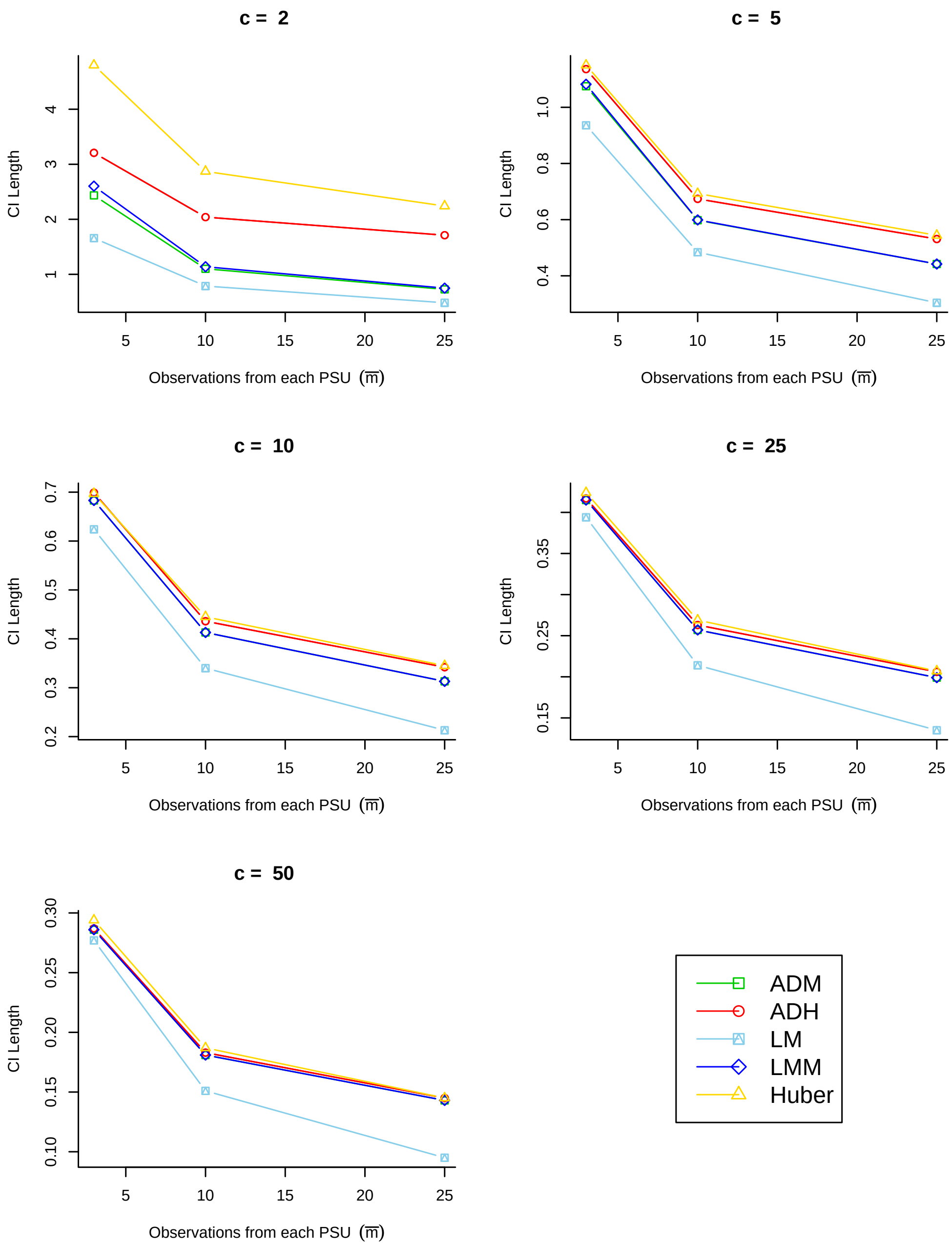


Figure B.9: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0

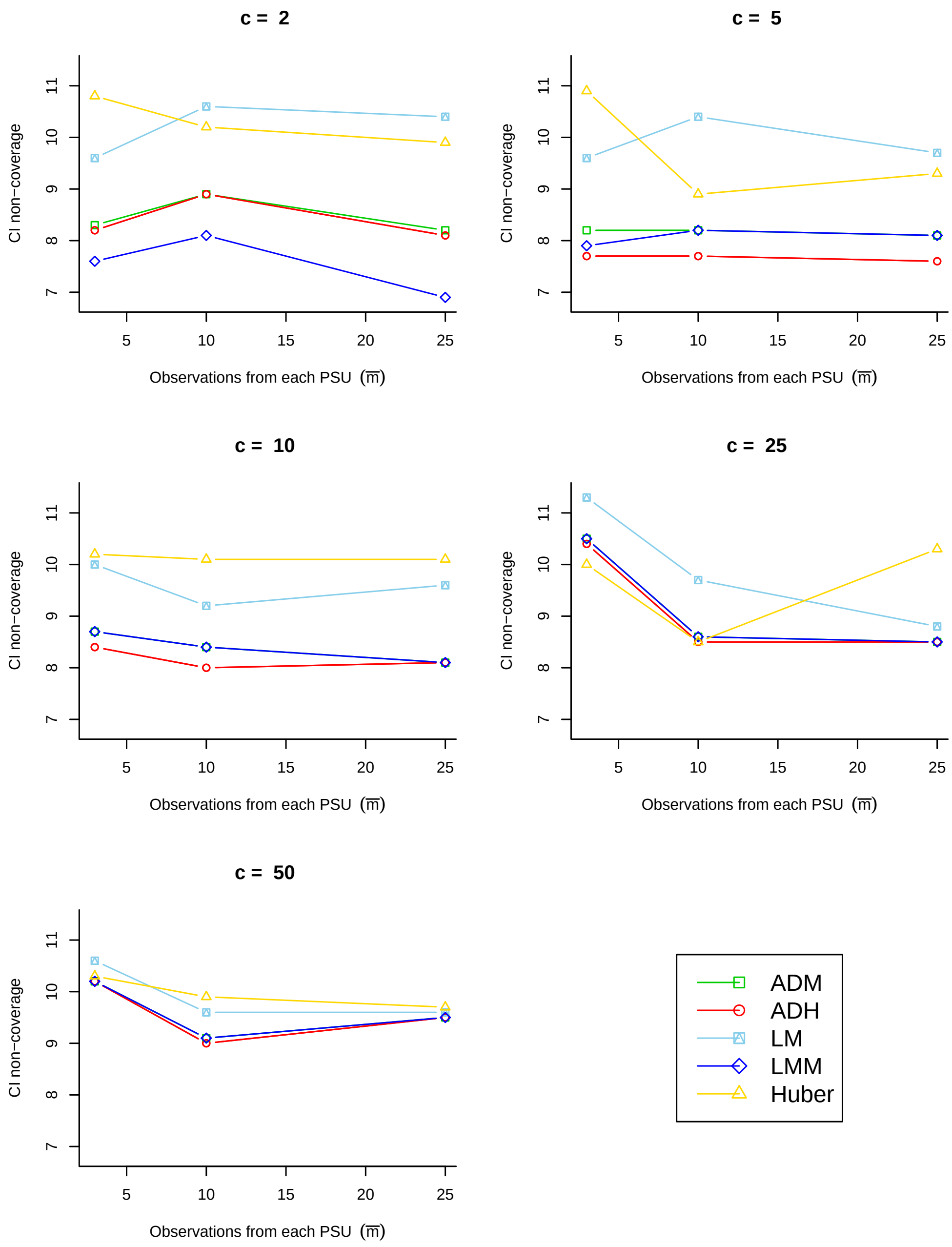


Figure B.10: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.025

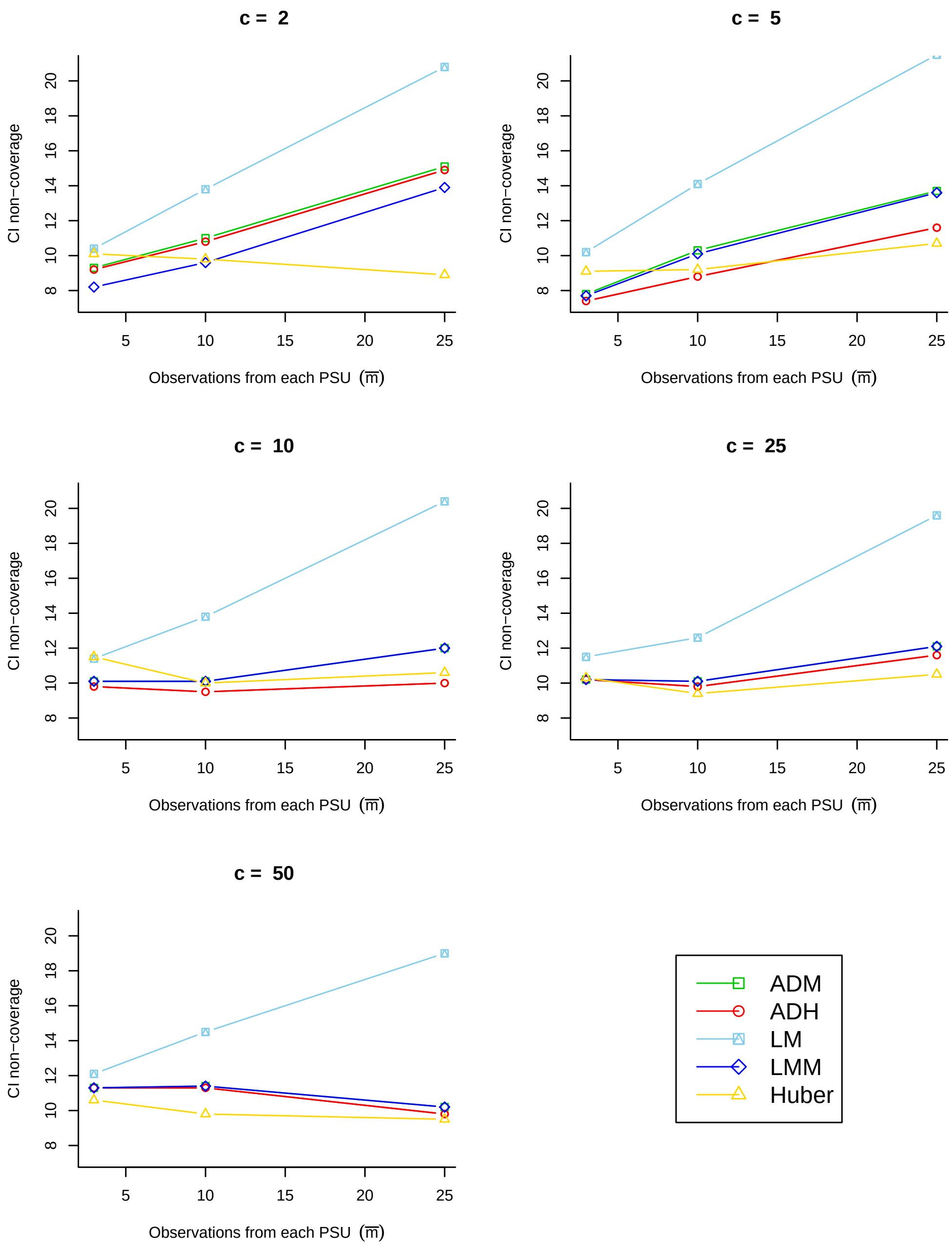


Figure B.11: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.1

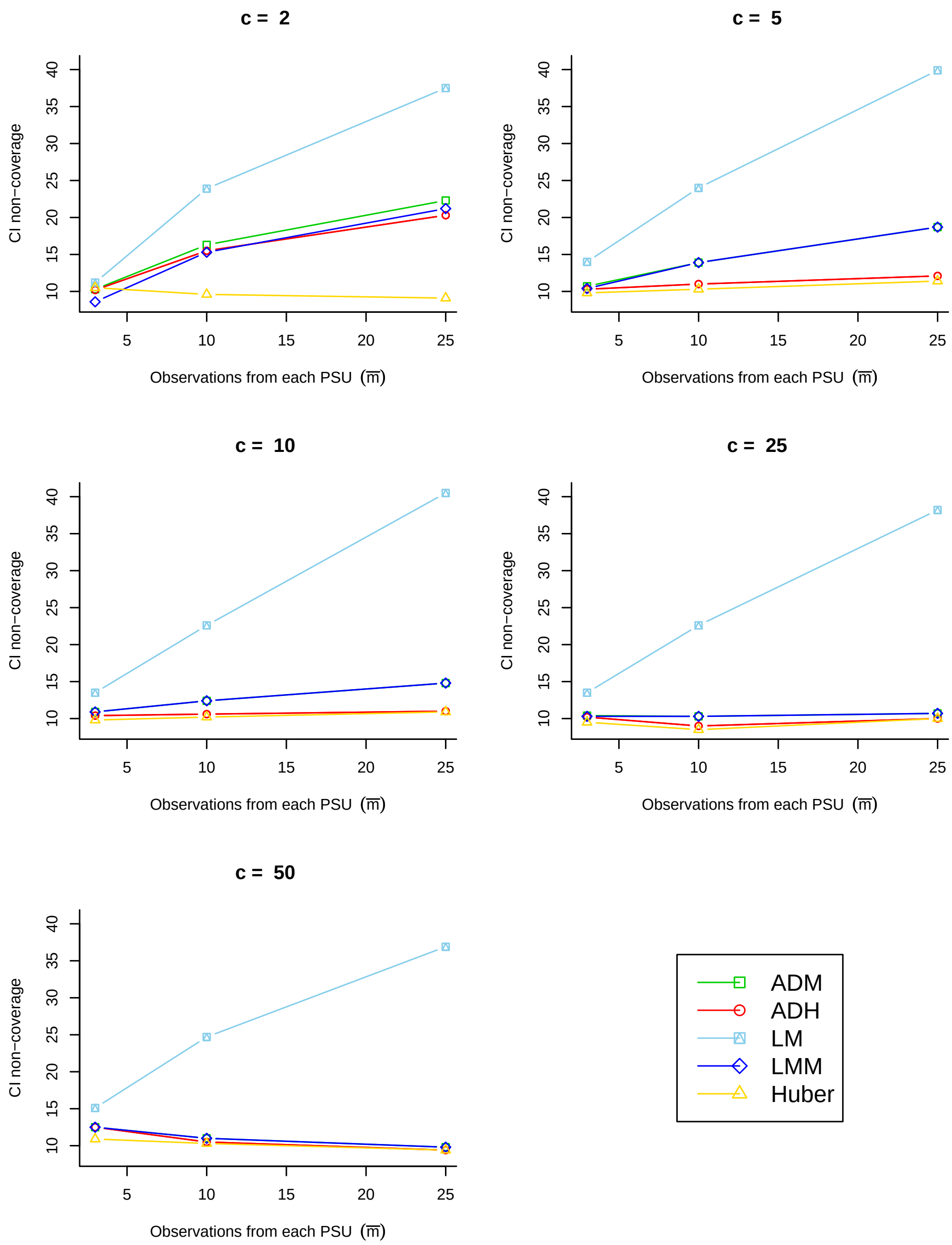


Figure B.12: Confidence interval lengths using different variance estimation methods and for various values of m and c, ρ=0

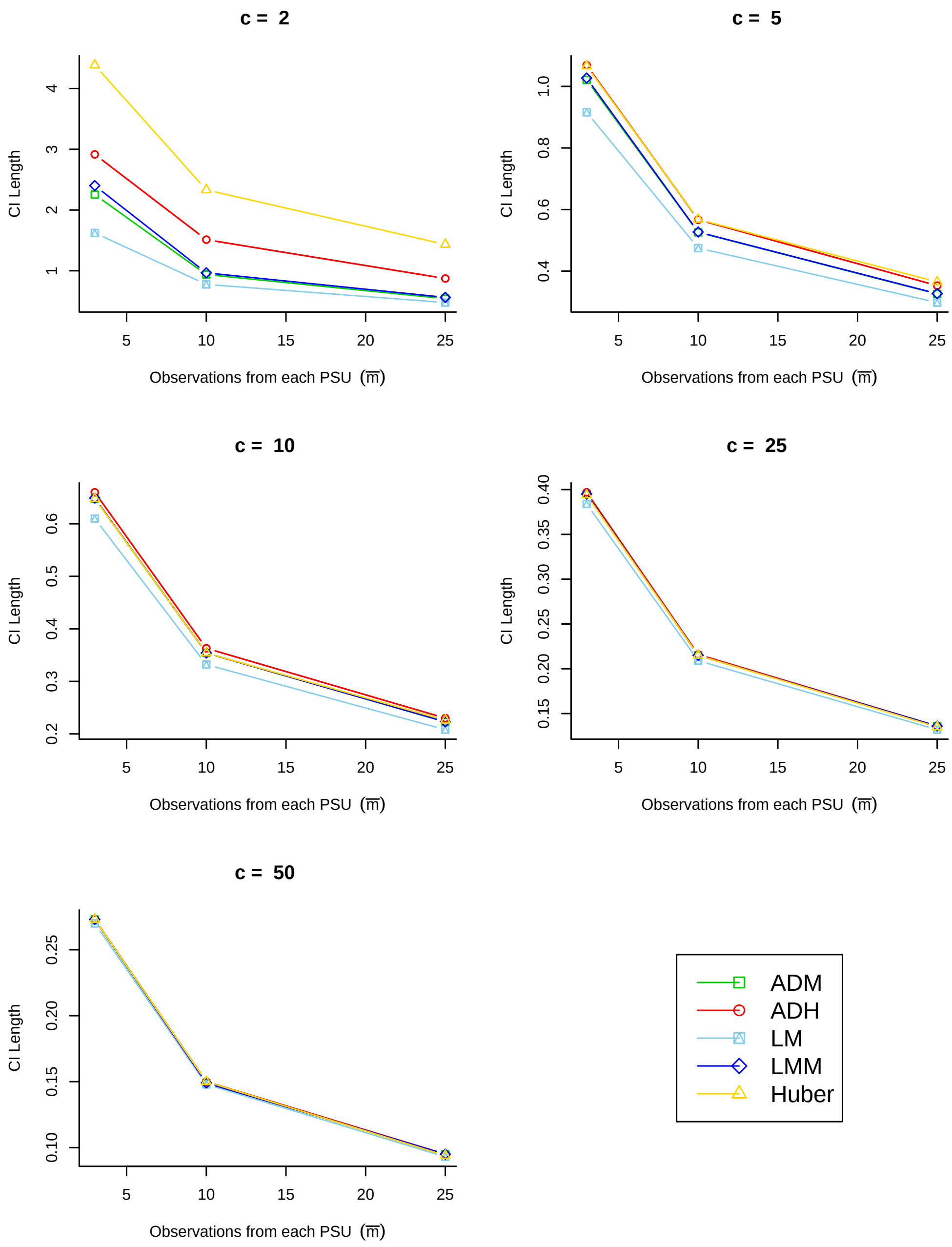


Figure B.13: Confidence interval lengths using different variance estimation methods and for various values of m and c, ρ=0.025

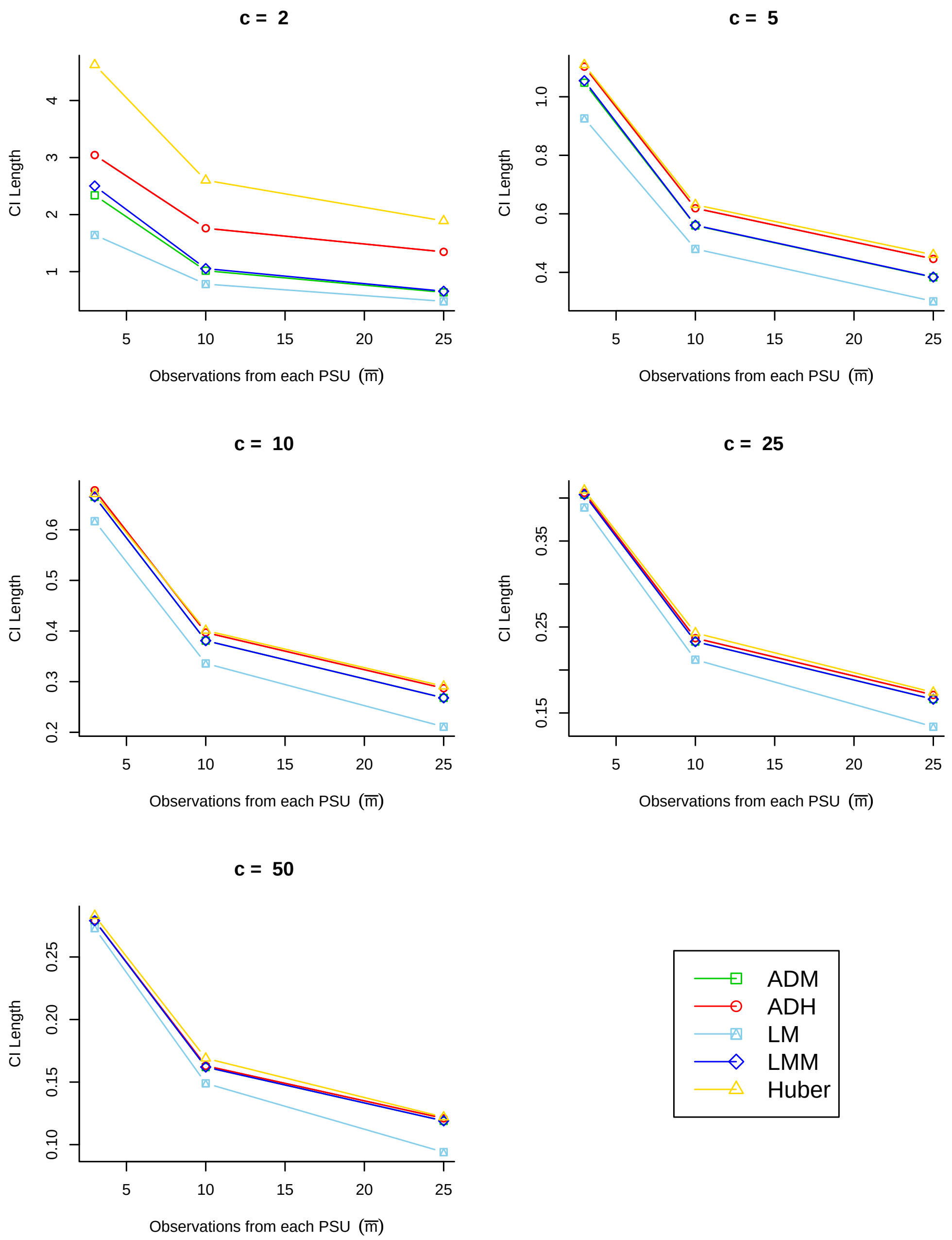
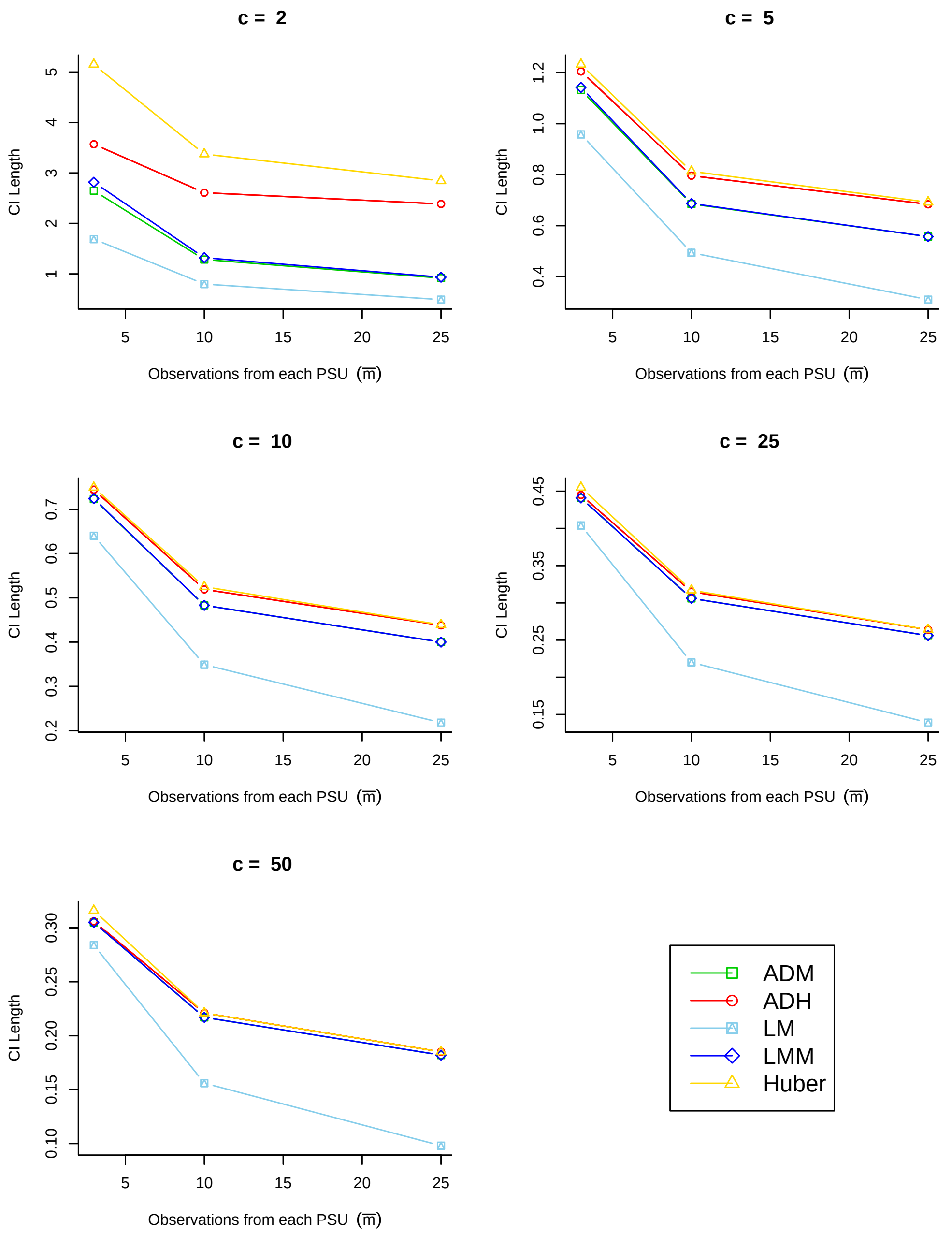


Figure B.14: Confidence interval lengths using different variance estimation methods and for various values of m and c, ρ=0.1



Appendix C

Extra Tables and Plots for Chapter 4

Table C.1: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0.05, $\sigma = \frac{1}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.133	1.133	1.314	0.987	11.4	11.4	14.8	11.3	10.1	1.801	1.026	1.934	1.790
2	5	1.048	1.048	1.282	0.990	14.7	14.6	15.0	11.6	9.3	0.490	0.643	0.536	1.268
2	10	1.047	1.047	1.276	1.017	14.7	14.7	13.5	10.2	10.9	0.350	0.497	0.395	0.977
2	15	1.122	1.122	1.325	1.126	17.5	17.5	16.6	8.8	16.2	0.331	0.515	0.373	0.909
2	25	0.977	0.977	1.141	1.004	23.3	23.3	19.2	10.4	18.9	0.294	0.459	0.331	0.766
2	50	1.033	1.033	1.160	1.088	25.4	25.4	20.4	8.4	28.2	0.269	0.447	0.298	0.671
5	2	1.052	1.052	1.166	1.008	11.4	11.4	11.8	11.2	10.6	0.426	0.429	0.436	0.467
5	5	1.017	1.017	1.145	1.016	12.3	12.3	12.2	10.9	12.5	0.267	0.270	0.279	0.312
5	10	0.969	0.969	1.081	1.005	15.0	15.0	14.0	11.5	21.7	0.208	0.211	0.217	0.246
5	15	0.954	0.954	1.070	1.008	14.6	14.6	13.4	10.2	27.0	0.179	0.183	0.191	0.214
5	25	0.926	0.926	1.017	0.986	15.7	15.3	12.5	9.4	39.4	0.160	0.164	0.170	0.187
5	50	0.912	0.912	0.953	0.947	18.5	17.1	16.3	11.5	62.8	0.151	0.154	0.157	0.167
10	2	1.113	1.113	1.186	1.074	10.1	10.1	9.9	10.3	13.5	0.286	0.287	0.289	0.293
10	5	0.923	0.923	0.989	0.941	12.8	12.6	13.4	10.9	17.8	0.185	0.186	0.188	0.198
10	10	0.956	0.956	1.033	1.013	12.8	12.6	11.7	10.1	30.1	0.142	0.142	0.148	0.155
10	15	0.864	0.864	0.932	0.923	16.2	16.3	14.9	12.9	38.0	0.122	0.122	0.128	0.133
10	25	0.942	0.942	0.977	0.974	13.7	13.3	12.4	10.6	63.6	0.113	0.113	0.116	0.118
10	50	1.007	1.007	1.017	1.018	11.1	10.7	10.6	9.0	87.5	0.105	0.104	0.106	0.106
25	2	1.001	1.001	1.049	0.994	10.3	10.3	10.3	10.4	13.4	0.174	0.174	0.174	0.176
25	5	0.943	0.943	0.959	0.972	13.0	13.0	12.6	12.0	26.5	0.115	0.115	0.116	0.119
25	10	0.940	0.940	0.956	0.981	11.4	11.7	11.7	10.7	51.9	0.090	0.090	0.090	0.093
25	15	1.027	1.027	1.037	1.056	10.5	10.5	10.8	9.1	70.9	0.080	0.079	0.080	0.081
25	25	0.923	0.923	0.928	0.931	10.5	11.1	10.4	10.7	90.6	0.072	0.071	0.072	0.072
25	50	1.070	1.070	1.070	1.071	9.0	8.7	9.2	8.7	99.3	0.063	0.063	0.063	0.063

Table C.2: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0.1, $\sigma = \frac{1}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.145	1.145	1.373	1.062	10.8	10.8	14.2	10.7	10.2	1.868	1.067	2.064	1.890
2	5	1.100	1.100	1.354	1.099	14.7	14.1	15.8	10.2	10.4	0.493	0.677	0.548	1.338
2	10	0.896	0.896	1.077	0.918	20.2	20.2	19.0	10.7	15.7	0.382	0.599	0.433	1.084
2	15	1.000	1.000	1.143	1.010	22.6	22.5	19.6	10.0	20.6	0.370	0.608	0.412	0.979
2	25	0.986	0.986	1.081	1.011	28.4	28.4	23.2	10.9	29.8	0.391	0.659	0.425	0.934
2	50	0.921	0.921	0.982	0.949	33.4	33.2	25.6	12.5	39.2	0.388	0.641	0.414	0.824
5	2	1.024	1.024	1.142	1.028	11.3	11.1	11.0	10.0	14.0	0.433	0.437	0.446	0.483
5	5	1.042	1.042	1.175	1.106	13.5	13.1	13.4	9.9	22.1	0.290	0.295	0.308	0.345
5	10	0.875	0.875	0.965	0.934	18.1	17.9	15.4	10.4	34.9	0.236	0.242	0.251	0.279
5	15	0.947	0.947	1.030	1.009	17.3	16.5	13.4	9.4	45.7	0.220	0.225	0.234	0.254
5	25	0.982	0.982	1.022	1.014	14.8	14.4	12.5	9.7	64.0	0.214	0.217	0.221	0.233
5	50	1.048	1.048	1.058	1.056	12.9	12.2	11.6	9.0	84.2	0.217	0.215	0.220	0.220
10	2	0.876	0.876	0.949	0.885	14.0	14.1	13.6	13.7	17.0	0.288	0.289	0.294	0.301
10	5	0.950	0.950	1.029	1.019	12.6	12.4	12.2	10.2	26.9	0.193	0.193	0.199	0.211
10	10	0.977	0.977	1.027	1.028	13.6	13.5	11.0	9.6	53.4	0.163	0.164	0.168	0.175
10	15	0.878	0.878	0.904	0.905	15.0	14.6	13.7	12.1	72.0	0.154	0.154	0.157	0.160
10	25	0.991	0.991	1.001	1.002	12.7	12.1	12.1	11.0	88.7	0.147	0.147	0.149	0.149
10	50	0.981	0.981	0.982	0.982	10.4	10.7	10.3	10.6	97.9	0.139	0.137	0.139	0.137
25	2	0.978	0.978	1.033	0.997	10.6	10.6	10.5	10.4	19.6	0.176	0.176	0.178	0.180
25	5	0.940	0.940	0.957	0.983	12.4	12.5	11.9	11.4	48.9	0.124	0.124	0.125	0.128
25	10	0.930	0.930	0.936	0.947	11.6	12.1	11.3	10.8	83.4	0.104	0.103	0.104	0.104
25	15	1.078	1.079	1.080	1.082	8.4	8.1	8.5	7.9	96.1	0.098	0.096	0.098	0.097
25	25	0.975	0.975	0.975	0.976	9.5	9.5	9.5	9.5	99.7	0.089	0.089	0.089	0.089
25	50	1.063	1.064	1.063	1.064	8.7	8.1	8.7	8.1	100.0	0.082	0.082	0.082	0.082

Table C.3: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0, $\sigma = \frac{1}{2}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.126	1.126	1.289	0.960	11.2	11.2	14.6	10.2	10.3	2.370	1.603	2.639	2.872
2	5	1.110	1.110	1.349	0.938	13.0	12.9	12.8	11.8	4.4	0.710	0.819	0.766	1.818
2	10	1.191	1.191	1.499	1.065	10.5	10.5	10.5	10.0	3.5	0.488	0.558	0.555	1.369
2	15	1.173	1.173	1.434	1.010	11.6	11.5	12.2	9.4	4.8	0.407	0.490	0.454	1.109
2	25	1.228	1.228	1.487	0.999	8.5	8.5	10.5	9.3	4.7	0.316	0.377	0.347	0.844
2	50	1.204	1.204	1.495	1.047	9.8	9.8	10.0	9.5	5.1	0.228	0.275	0.257	0.626
5	2	1.087	1.087	1.172	1.002	13.6	13.7	12.9	13.6	10.7	0.689	0.691	0.690	0.736
5	5	1.126	1.126	1.232	1.016	10.6	10.4	11.3	10.4	6.3	0.420	0.424	0.426	0.472
5	10	1.061	1.061	1.173	0.957	10.7	10.8	11.9	10.8	5.9	0.294	0.295	0.298	0.333
5	15	1.165	1.165	1.302	1.068	9.7	9.7	10.6	10.5	5.9	0.244	0.246	0.251	0.281
5	25	1.092	1.093	1.211	0.969	9.3	9.2	9.8	11.6	7.9	0.190	0.192	0.196	0.215
5	50	1.150	1.150	1.271	1.014	8.8	8.8	8.9	10.6	6.8	0.135	0.136	0.138	0.152
10	2	1.148	1.148	1.221	1.091	10.3	10.2	10.2	11.3	10.9	0.460	0.461	0.462	0.469
10	5	1.088	1.088	1.150	1.035	10.3	10.3	11.4	10.8	8.5	0.295	0.296	0.295	0.307
10	10	1.059	1.059	1.117	0.979	10.8	10.9	11.8	10.7	5.1	0.204	0.205	0.203	0.211
10	15	1.042	1.042	1.101	0.959	10.8	10.8	10.6	11.4	9.7	0.170	0.170	0.170	0.174
10	25	1.047	1.047	1.105	0.954	9.3	9.5	10.0	9.6	7.5	0.130	0.131	0.130	0.134
10	50	0.978	0.978	1.043	0.909	12.1	12.1	12.3	13.1	8.7	0.093	0.094	0.094	0.097
25	2	1.085	1.085	1.130	1.048	11.2	11.1	11.2	11.1	10.1	0.287	0.287	0.285	0.286
25	5	1.056	1.056	1.067	1.008	8.7	8.7	10.4	9.1	9.1	0.183	0.182	0.180	0.182
25	10	0.987	0.987	0.996	0.935	11.0	11.0	11.8	11.1	7.7	0.128	0.128	0.126	0.128
25	15	1.049	1.049	1.067	0.999	10.2	10.2	10.6	10.0	8.5	0.106	0.105	0.105	0.105
25	25	1.103	1.103	1.118	1.060	8.6	8.6	10.1	9.6	9.9	0.082	0.082	0.081	0.082
25	50	1.020	1.020	1.052	0.972	10.5	10.4	10.8	10.7	7.4	0.058	0.058	0.057	0.058

Table C.4: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0.025, $\sigma = \frac{1}{2}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.170	1.170	1.403	1.074	13.5	13.5	14.9	9.5	10.5	2.562	1.624	2.945	3.000
2	5	1.117	1.117	1.397	1.066	13.5	13.1	13.9	9.5	7.6	0.763	0.961	0.849	2.047
2	10	1.049	1.049	1.315	0.984	13.6	13.6	13.5	9.7	6.2	0.531	0.669	0.606	1.492
2	15	0.956	0.956	1.185	0.904	16.8	16.8	16.7	8.6	7.1	0.436	0.560	0.493	1.240
2	25	1.013	1.013	1.232	0.991	17.7	17.7	16.6	9.5	11.5	0.388	0.546	0.439	1.061
2	50	0.958	0.958	1.119	0.973	23.3	23.3	19.2	9.8	17.0	0.328	0.505	0.368	0.875
5	2	1.099	1.099	1.199	1.033	12.2	12.2	13.1	12.9	10.9	0.694	0.700	0.707	0.754
5	5	1.164	1.164	1.297	1.133	11.4	11.4	10.9	11.4	11.3	0.437	0.441	0.451	0.503
5	10	1.019	1.019	1.160	1.046	13.3	13.2	12.9	9.9	12.0	0.314	0.317	0.330	0.377
5	15	1.029	1.029	1.157	1.037	13.3	13.1	12.0	11.5	16.7	0.268	0.273	0.282	0.318
5	25	1.013	1.013	1.139	1.056	16.0	16.0	13.3	10.6	22.7	0.224	0.229	0.239	0.267
5	50	0.964	0.964	1.060	1.024	17.4	16.6	14.0	11.0	36.1	0.185	0.190	0.197	0.219
10	2	1.156	1.156	1.240	1.121	9.7	9.7	10.0	9.9	10.2	0.461	0.463	0.465	0.474
10	5	0.990	0.990	1.055	0.971	13.0	13.0	13.1	11.7	11.2	0.295	0.296	0.297	0.311
10	10	1.016	1.016	1.088	1.029	12.4	12.4	11.5	10.5	16.6	0.216	0.217	0.220	0.233
10	15	0.939	0.939	1.023	0.982	12.8	12.8	12.2	10.7	23.3	0.185	0.186	0.193	0.202
10	25	0.957	0.957	1.032	1.019	13.4	13.5	11.5	9.7	35.3	0.154	0.155	0.160	0.169
10	50	1.006	1.007	1.048	1.046	11.8	12.3	10.9	8.8	60.2	0.133	0.132	0.137	0.140
25	2	1.007	1.007	1.052	0.990	11.3	11.2	11.7	11.4	11.1	0.287	0.288	0.286	0.289
25	5	0.931	0.931	0.946	0.943	11.9	12.1	12.3	11.3	14.9	0.184	0.184	0.183	0.189
25	10	0.962	0.962	0.979	0.990	10.4	10.5	10.5	9.9	26.0	0.137	0.136	0.137	0.141
25	15	0.862	0.862	0.880	0.902	12.9	13.1	12.7	11.2	39.1	0.117	0.116	0.118	0.121
25	25	1.023	1.023	1.038	1.059	9.8	10.0	10.4	9.1	59.0	0.099	0.098	0.099	0.101
25	50	1.101	1.101	1.109	1.112	8.2	7.4	8.2	7.2	89.7	0.083	0.083	0.084	0.084

Table C.5: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0.05, $\sigma = \frac{1}{2}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.133	1.133	1.307	0.991	13.3	13.3	15.8	10.9	10.5	2.819	1.650	3.058	2.925
2	5	1.026	1.026	1.284	0.980	15.4	15.0	15.4	12.4	7.0	0.755	0.943	0.835	2.031
2	10	1.029	1.029	1.280	1.008	15.3	15.3	13.9	10.5	9.7	0.566	0.774	0.645	1.606
2	15	0.998	0.998	1.197	0.988	18.2	18.2	17.5	9.1	12.2	0.497	0.721	0.565	1.392
2	25	1.026	1.026	1.202	1.048	21.0	21.0	18.8	10.4	18.5	0.471	0.737	0.529	1.232
2	50	1.012	1.013	1.147	1.068	25.5	25.5	19.2	9.6	27.0	0.435	0.714	0.485	1.088
5	2	1.064	1.064	1.170	1.025	12.6	12.5	12.0	11.2	12.1	0.693	0.698	0.705	0.763
5	5	0.965	0.965	1.093	0.974	13.7	13.6	13.6	11.3	12.2	0.438	0.444	0.460	0.516
5	10	0.987	0.987	1.118	1.039	14.4	14.4	13.0	10.3	17.9	0.332	0.337	0.350	0.401
5	15	1.021	1.021	1.140	1.067	14.3	14.3	13.8	11.4	26.0	0.295	0.302	0.313	0.350
5	25	0.878	0.878	0.960	0.929	18.2	18.0	15.2	10.3	38.1	0.260	0.267	0.275	0.305
5	50	0.899	0.899	0.940	0.929	18.1	17.6	14.8	11.3	62.5	0.248	0.253	0.258	0.272
10	2	1.018	1.018	1.094	0.998	13.7	13.6	13.5	12.8	9.9	0.465	0.465	0.468	0.480
10	5	0.906	0.906	0.973	0.938	13.6	13.7	13.6	11.3	16.8	0.306	0.306	0.311	0.331
10	10	0.950	0.950	1.024	1.004	14.9	14.9	13.6	12.0	28.2	0.233	0.234	0.240	0.254
10	15	0.938	0.938	0.999	0.989	13.5	13.5	13.2	10.7	41.1	0.206	0.206	0.213	0.222
10	25	0.992	0.993	1.033	1.033	13.7	13.3	12.4	10.3	61.9	0.185	0.185	0.190	0.195
10	50	1.009	1.009	1.022	1.024	11.4	11.0	10.7	9.0	84.8	0.169	0.168	0.171	0.171
25	2	0.955	0.955	1.000	0.949	12.0	12.1	12.4	12.1	13.5	0.284	0.285	0.284	0.288
25	5	0.880	0.880	0.894	0.909	13.9	14.0	14.3	12.4	25.8	0.191	0.190	0.190	0.197
25	10	0.918	0.918	0.934	0.962	12.5	12.2	12.4	10.6	47.2	0.147	0.146	0.148	0.153
25	15	0.957	0.958	0.968	0.989	12.5	13.0	13.1	12.1	67.5	0.131	0.130	0.132	0.134
25	25	0.940	0.940	0.945	0.953	11.5	11.3	11.1	10.1	87.7	0.116	0.115	0.117	0.116
25	50	1.078	1.078	1.079	1.079	8.9	8.8	8.9	8.7	99.0	0.102	0.102	0.102	0.102

Table C.6: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0.1, $\sigma = \frac{1}{2}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.066	1.066	1.279	1.004	13.0	13.0	15.9	11.2	10.2	2.914	1.678	3.245	3.049
2	5	0.980	0.980	1.251	1.020	15.4	14.6	17.0	10.2	9.3	0.743	1.029	0.845	2.186
2	10	0.850	0.850	1.055	0.896	22.0	21.8	19.5	10.3	13.7	0.594	0.899	0.688	1.739
2	15	1.065	1.065	1.205	1.077	24.1	24.1	20.2	11.7	21.4	0.624	1.048	0.691	1.650
2	25	0.980	0.980	1.095	1.016	27.8	27.8	23.5	10.5	27.1	0.587	0.991	0.645	1.463
2	50	0.868	0.868	0.933	0.899	31.7	31.4	24.3	11.9	38.9	0.605	1.020	0.650	1.331
5	2	1.104	1.104	1.227	1.091	11.4	11.4	11.0	10.5	11.2	0.695	0.697	0.714	0.771
5	5	0.937	0.937	1.069	0.993	14.3	14.2	14.4	11.7	17.0	0.450	0.457	0.477	0.538
5	10	0.956	0.956	1.057	1.011	17.1	16.8	13.8	10.2	33.5	0.383	0.391	0.406	0.449
5	15	0.968	0.969	1.045	1.030	16.3	15.7	13.5	10.9	47.8	0.365	0.375	0.386	0.422
5	25	1.008	1.008	1.049	1.040	14.7	14.4	11.8	9.8	63.7	0.351	0.357	0.363	0.383
5	50	1.060	1.060	1.070	1.069	12.5	11.3	10.6	8.6	84.4	0.349	0.346	0.352	0.355
10	2	0.984	0.984	1.054	0.982	13.4	13.4	13.5	13.0	15.9	0.469	0.470	0.473	0.487
10	5	0.949	0.949	1.028	1.019	14.5	13.9	14.2	10.8	27.4	0.320	0.321	0.331	0.351
10	10	0.961	0.961	1.012	1.007	13.9	13.9	12.1	10.1	51.7	0.269	0.269	0.276	0.287
10	15	0.867	0.867	0.898	0.899	14.8	14.3	13.5	11.5	67.4	0.248	0.248	0.254	0.260
10	25	0.916	0.916	0.926	0.927	15.2	14.5	13.8	12.5	86.6	0.237	0.235	0.239	0.239
10	50	0.976	0.976	0.977	0.977	10.5	11.2	10.2	10.9	97.4	0.225	0.221	0.225	0.222
25	2	0.960	0.960	1.008	0.965	12.3	12.0	12.2	11.6	17.5	0.288	0.288	0.290	0.293
25	5	0.943	0.943	0.960	0.989	12.7	12.6	13.0	11.1	46.5	0.205	0.204	0.206	0.212
25	10	1.036	1.036	1.042	1.053	10.3	10.5	10.3	10.1	83.0	0.173	0.171	0.173	0.174
25	15	0.931	0.931	0.933	0.935	10.8	11.2	10.3	10.8	94.4	0.159	0.157	0.159	0.157
25	25	0.976	0.976	0.976	0.976	9.7	10.0	9.7	10.0	99.5	0.145	0.144	0.145	0.144
25	50	1.065	1.066	1.065	1.066	8.7	7.7	8.7	7.7	100.0	0.133	0.132	0.133	0.132

Table C.7: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0.05, $\sigma = \frac{2}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.168	1.168	1.363	1.087	15.2	15.2	18.1	11.6	10.7	3.968	2.347	4.366	4.390
2	5	1.057	1.057	1.309	1.022	16.8	16.6	16.2	11.9	7.2	1.128	1.393	1.242	3.078
2	10	1.041	1.042	1.305	1.063	18.1	18.0	16.8	9.8	10.0	0.849	1.182	0.973	2.472
2	15	0.875	0.875	1.068	0.871	19.8	19.7	18.7	12.4	10.7	0.715	0.995	0.811	2.019
2	25	0.986	0.986	1.201	1.034	21.2	21.2	20.0	10.9	15.4	0.657	0.989	0.756	1.796
2	50	1.007	1.007	1.148	1.072	27.0	27.0	21.5	9.4	26.4	0.688	1.123	0.772	1.715
5	2	0.994	0.994	1.086	0.953	15.5	15.3	15.7	14.9	9.7	0.992	0.998	1.004	1.088
5	5	1.020	1.020	1.165	1.052	14.3	14.2	13.1	11.2	10.6	0.649	0.656	0.682	0.776
5	10	0.965	0.965	1.098	1.002	15.0	15.0	13.8	11.8	16.0	0.488	0.496	0.517	0.590
5	15	0.964	0.964	1.113	1.051	16.0	16.0	15.1	11.4	22.5	0.439	0.449	0.474	0.539
5	25	0.861	0.861	0.952	0.919	17.5	17.2	15.1	9.7	34.4	0.386	0.396	0.411	0.460
5	50	0.968	0.969	1.019	1.005	16.6	16.2	14.0	11.2	60.2	0.368	0.378	0.384	0.409
10	2	0.994	0.994	1.063	0.978	14.7	14.6	14.4	14.0	10.2	0.707	0.709	0.712	0.734
10	5	0.998	0.998	1.070	1.020	13.6	13.9	13.6	11.2	15.3	0.465	0.465	0.471	0.499
10	10	0.932	0.932	1.001	0.981	14.4	14.3	13.9	11.3	26.6	0.353	0.354	0.363	0.385
10	15	0.942	0.942	1.011	0.999	14.0	14.0	13.2	11.0	37.3	0.308	0.308	0.320	0.334
10	25	1.005	1.005	1.052	1.051	15.7	15.3	14.0	12.0	57.7	0.279	0.278	0.287	0.295
10	50	1.001	1.001	1.016	1.018	11.8	10.7	10.2	7.9	83.2	0.251	0.249	0.254	0.255
25	2	0.985	0.985	1.027	0.974	12.6	12.6	12.9	12.2	13.4	0.436	0.437	0.435	0.441
25	5	0.916	0.916	0.930	0.947	12.7	12.5	13.0	11.6	23.6	0.293	0.292	0.293	0.303
25	10	0.969	0.969	0.987	1.018	11.2	11.4	11.0	10.1	41.6	0.223	0.220	0.223	0.230
25	15	0.906	0.906	0.921	0.940	11.7	12.2	11.3	10.7	61.9	0.200	0.198	0.201	0.204
25	25	0.962	0.962	0.968	0.977	10.8	11.1	10.5	10.0	84.5	0.176	0.174	0.177	0.177
25	50	1.084	1.084	1.085	1.085	9.0	8.7	8.9	8.6	98.6	0.153	0.152	0.153	0.152

Table C.8: Variance ratios, average length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced Log-Normal data case with ρ=0.05, $\sigma = \frac{2}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	2	1.092	1.092	1.309	1.061	14.9	14.9	17.5	12.0	10.5	4.093	2.438	4.477	4.573
2	5	0.911	0.911	1.190	0.977	18.3	17.1	17.6	10.1	8.5	1.112	1.486	1.282	3.346
2	10	0.927	0.927	1.140	0.978	21.7	21.5	21.0	10.6	13.8	0.904	1.369	1.035	2.636
2	15	0.956	0.956	1.139	1.010	25.6	25.6	22.4	10.6	18.6	0.860	1.374	0.975	2.362
2	25	1.037	1.037	1.177	1.095	27.6	27.5	23.3	10.3	26.6	0.890	1.497	0.989	2.253
2	50	0.847	0.847	0.918	0.884	32.5	31.9	24.9	11.5	37.9	0.873	1.488	0.946	1.981
5	2	1.091	1.091	1.204	1.077	13.8	13.5	14.1	12.5	11.2	1.030	1.037	1.052	1.150
5	5	0.901	0.901	1.018	0.955	16.0	15.6	16.2	12.4	17.1	0.692	0.705	0.728	0.832
5	10	0.906	0.906	1.016	0.976	18.2	17.7	16.5	11.2	30.2	0.565	0.579	0.605	0.678
5	15	0.995	0.995	1.085	1.066	16.4	16.1	13.5	9.8	43.8	0.547	0.560	0.581	0.638
5	25	0.912	0.912	0.964	0.952	19.0	18.9	15.4	12.6	57.0	0.501	0.510	0.525	0.557
5	50	1.036	1.036	1.049	1.048	12.7	12.2	10.6	9.4	81.6	0.509	0.508	0.516	0.524
10	2	0.972	0.972	1.041	0.978	13.4	13.2	13.5	13.0	15.7	0.721	0.723	0.727	0.754
10	5	0.931	0.931	1.006	0.996	15.1	14.9	14.4	12.0	23.7	0.479	0.481	0.492	0.526
10	10	0.963	0.963	1.020	1.018	15.5	15.3	14.0	11.0	46.9	0.399	0.401	0.412	0.431
10	15	0.977	0.977	1.015	1.016	15.2	14.9	13.3	10.8	64.8	0.372	0.372	0.381	0.391
10	25	0.948	0.948	0.964	0.965	14.3	14.4	13.1	12.7	82.7	0.349	0.347	0.354	0.356
10	50	0.993	0.993	0.995	0.995	10.8	11.5	10.3	10.8	96.5	0.334	0.328	0.335	0.329
25	2	0.966	0.966	1.012	0.975	12.1	12.1	13.0	12.2	17.5	0.445	0.446	0.447	0.454
25	5	0.970	0.970	0.986	1.016	11.3	11.7	11.8	10.4	44.6	0.315	0.312	0.316	0.325
25	10	0.999	1.000	1.010	1.026	12.5	13.0	11.8	12.0	75.3	0.257	0.255	0.258	0.261
25	15	0.932	0.933	0.936	0.940	12.1	11.7	12.0	11.1	91.5	0.239	0.235	0.239	0.237
25	25	0.977	0.977	0.977	0.977	10.2	10.1	10.1	10.1	99.1	0.217	0.216	0.217	0.216
25	50	1.065	1.065	1.065	1.065	8.5	8.3	8.5	8.3	100.0	0.197	0.197	0.197	0.197

Table C.9: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0.1, $\sigma = \frac{1}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	3	1.199	1.199	1.314	1.005	11.9	11.8	10.2	9.7	19.8	0.788	1.029	0.845	1.551
2	10	0.903	0.903	0.958	0.812	21.1	20.7	19.2	10.1	27.3	0.400	0.776	0.417	1.100
2	25	1.123	1.123	1.156	1.072	22.6	20.9	21.0	10.7	41.5	0.296	0.752	0.304	0.913
5	3	1.090	1.090	1.108	1.006	11.0	10.5	11.0	11.1	37.8	0.374	0.398	0.378	0.409
5	10	1.051	1.051	1.053	1.014	15.1	12.0	15.1	10.8	60.6	0.235	0.274	0.235	0.282
5	25	1.002	1.002	1.002	0.997	14.3	9.8	14.3	9.0	81.9	0.189	0.232	0.189	0.236
10	3	1.042	1.042	1.044	1.002	9.7	9.5	9.7	9.2	40.7	0.246	0.253	0.246	0.255
10	10	1.006	1.006	1.006	1.000	13.1	11.1	13.1	10.2	72.8	0.163	0.174	0.163	0.177
10	25	1.016	1.017	1.016	1.017	12.8	10.6	12.8	10.4	94.1	0.134	0.147	0.134	0.147
25	3	0.999	1.000	1.000	1.018	10.3	10.2	10.3	9.6	36.8	0.150	0.151	0.150	0.155
25	10	0.926	0.927	0.926	0.936	12.3	11.3	12.3	10.4	87.3	0.102	0.105	0.102	0.106
25	25	1.039	1.039	1.039	1.039	10.3	9.4	10.3	9.4	99.4	0.086	0.089	0.086	0.089
50	3	0.932	0.932	0.932	0.981	12.5	12.4	12.5	11.2	36.1	0.103	0.104	0.103	0.107
50	10	0.976	0.976	0.976	0.979	10.9	10.5	10.9	10.5	97.2	0.074	0.075	0.074	0.075
50	25	0.996	0.996	0.996	0.996	9.8	9.3	9.8	9.3	100.0	0.061	0.062	0.061	0.062

Table C.10: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0, $\sigma = \frac{1}{2}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	3	1.317	1.317	1.440	1.012	11.4	11.2	10.0	11.3	15.7	1.155	1.438	1.239	2.327
2	10	1.287	1.287	1.370	0.946	10.0	9.6	9.2	8.6	13.5	0.535	0.809	0.555	1.371
2	25	1.464	1.464	1.544	1.085	8.5	8.4	7.9	9.1	15.1	0.330	0.556	0.341	0.886
5	3	1.228	1.228	1.249	1.080	10.3	9.9	10.2	10.2	24.1	0.583	0.608	0.588	0.628
5	10	1.290	1.290	1.295	1.045	9.2	8.3	9.1	9.7	23.8	0.311	0.334	0.311	0.339
5	25	1.284	1.284	1.287	1.044	9.1	8.5	9.0	10.1	23.0	0.196	0.212	0.196	0.217
10	3	1.142	1.142	1.143	1.010	8.7	8.3	8.7	10.1	22.4	0.386	0.392	0.386	0.384
10	10	1.094	1.094	1.094	0.951	10.1	9.3	10.1	12.0	21.6	0.212	0.217	0.212	0.214
10	25	1.077	1.077	1.077	0.940	9.1	9.0	9.1	9.6	19.6	0.133	0.136	0.133	0.135
25	3	1.194	1.194	1.194	1.127	7.6	7.6	7.6	8.2	12.1	0.237	0.238	0.237	0.235
25	10	1.003	1.003	1.003	0.940	9.7	9.6	9.7	10.2	13.5	0.130	0.131	0.130	0.129
25	25	1.097	1.097	1.097	1.013	8.3	8.2	8.3	9.1	11.7	0.082	0.082	0.082	0.081
50	3	1.038	1.038	1.038	1.018	10.9	10.8	10.9	11.0	6.3	0.162	0.163	0.162	0.162
50	10	0.998	0.998	0.998	0.962	9.6	9.6	9.6	10.1	5.7	0.090	0.090	0.090	0.089
50	25	1.064	1.064	1.064	1.022	8.8	8.8	8.8	9.7	9.2	0.057	0.057	0.057	0.057

Table C.11: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0.025, $\sigma = \frac{1}{2}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	3	1.276	1.276	1.405	1.010	11.7	11.4	10.5	11.2	15.9	1.171	1.464	1.261	2.360
2	10	1.082	1.082	1.166	0.861	14.2	13.9	12.8	10.2	17.7	0.555	0.917	0.581	1.474
2	25	1.279	1.279	1.347	1.082	14.2	13.9	13.1	11.3	23.3	0.364	0.733	0.376	1.055
5	3	1.180	1.180	1.198	1.051	10.6	10.1	10.4	11.1	28.2	0.589	0.618	0.595	0.638
5	10	1.160	1.160	1.165	1.022	12.0	11.3	12.0	11.0	34.3	0.327	0.361	0.328	0.373
5	25	1.070	1.071	1.073	0.994	13.0	10.9	12.9	9.8	49.3	0.222	0.259	0.223	0.267
10	3	1.102	1.102	1.104	1.006	9.6	9.1	9.6	9.7	25.2	0.389	0.396	0.389	0.393
10	10	1.043	1.044	1.043	0.975	10.9	10.1	10.9	10.1	37.0	0.224	0.233	0.224	0.235
10	25	0.993	0.993	0.993	0.975	12.4	10.4	12.4	10.5	54.7	0.153	0.164	0.153	0.167
25	3	1.116	1.116	1.116	1.082	8.8	8.5	8.8	8.8	16.7	0.239	0.240	0.239	0.240
25	10	0.918	0.918	0.918	0.926	11.2	11.1	11.2	10.1	33.1	0.137	0.138	0.137	0.142
25	25	1.012	1.012	1.012	1.035	11.3	11.0	11.3	9.8	64.7	0.097	0.099	0.097	0.101
50	3	0.976	0.977	0.977	0.988	11.8	11.8	11.8	11.8	10.4	0.163	0.164	0.163	0.166
50	10	0.916	0.916	0.916	0.967	11.5	11.3	11.5	9.5	31.3	0.095	0.096	0.095	0.099
50	25	1.003	1.003	1.003	1.016	10.6	10.0	10.6	9.5	82.0	0.069	0.070	0.069	0.070

Table C.12: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0.1, $\sigma = \frac{1}{2}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	3	1.163	1.163	1.290	0.994	13.0	12.6	11.7	11.4	19.1	1.243	1.608	1.339	2.507
2	10	0.916	0.916	0.968	0.821	21.7	21.1	19.9	9.6	27.2	0.647	1.260	0.673	1.791
2	25	1.119	1.119	1.151	1.064	23.1	21.3	21.1	11.4	40.9	0.477	1.210	0.489	1.476
5	3	1.094	1.094	1.112	1.016	12.1	11.6	12.1	12.5	37.5	0.610	0.648	0.615	0.670
5	10	1.049	1.049	1.052	1.012	14.8	11.8	14.8	11.1	58.3	0.383	0.445	0.383	0.459
5	25	0.996	0.996	0.997	0.990	15.4	10.5	15.4	9.9	80.9	0.306	0.376	0.306	0.382
10	3	1.047	1.047	1.049	1.009	10.3	9.9	10.3	10.3	38.7	0.403	0.413	0.404	0.419
10	10	1.011	1.011	1.011	1.004	13.6	11.8	13.6	11.1	71.6	0.266	0.285	0.266	0.289
10	25	1.015	1.015	1.015	1.016	13.3	10.4	13.3	10.4	93.6	0.217	0.238	0.217	0.239
25	3	1.011	1.011	1.011	1.028	10.0	9.9	10.0	9.5	34.9	0.248	0.250	0.248	0.256
25	10	0.936	0.936	0.936	0.946	12.5	11.8	12.5	10.9	85.7	0.168	0.172	0.168	0.174
25	25	1.037	1.037	1.037	1.038	10.6	9.4	10.6	9.4	99.3	0.141	0.146	0.141	0.146
50	3	0.928	0.928	0.928	0.975	12.9	12.9	12.9	11.2	33.7	0.170	0.171	0.170	0.177
50	10	0.977	0.977	0.977	0.981	11.4	10.9	11.4	10.7	95.9	0.121	0.123	0.121	0.123
50	25	1.006	1.006	1.006	1.006	9.8	9.1	9.8	9.1	100.0	0.099	0.101	0.099	0.101

Table C.13: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ using RLRT in the unbalanced data case with ρ=0.1, $\sigma = \frac{2}{3}$.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β (%)				Pr(Rej H_0) (%)	Confidence Interval Length			
c	$\bar{m}$	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub	RLRT	ADM	ADH	LMM	Hub
2	3	1.129	1.129	1.263	0.981	16.5	15.8	15.0	13.2	18.3	1.785	2.286	1.929	3.703
2	10	0.926	0.926	0.979	0.833	22.1	20.9	20.4	11.8	26.7	0.954	1.847	0.993	2.673
2	25	1.112	1.113	1.147	1.056	23.2	21.5	21.5	12.1	39.6	0.703	1.769	0.721	2.186
5	3	1.099	1.099	1.116	1.028	14.3	13.9	14.2	13.2	36.3	0.910	0.966	0.918	1.005
5	10	1.046	1.047	1.052	1.007	14.8	12.5	14.8	12.6	56.3	0.573	0.664	0.575	0.687
5	25	0.993	0.993	0.994	0.985	16.2	10.7	16.2	9.8	79.1	0.455	0.558	0.455	0.568
10	3	1.048	1.048	1.050	1.011	11.4	11.1	11.4	10.4	36.5	0.608	0.622	0.608	0.632
10	10	1.011	1.011	1.011	1.004	13.9	12.3	13.9	11.7	68.7	0.401	0.429	0.401	0.436
10	25	1.014	1.014	1.014	1.015	13.9	11.6	13.9	11.4	91.8	0.323	0.354	0.323	0.356
25	3	1.023	1.023	1.023	1.038	10.8	10.7	10.8	10.1	32.3	0.378	0.381	0.378	0.390
25	10	0.945	0.945	0.945	0.957	13.0	12.2	13.0	11.6	82.2	0.253	0.260	0.253	0.263
25	25	1.037	1.037	1.037	1.037	10.3	9.4	10.3	9.4	99.1	0.210	0.218	0.210	0.218
50	3	0.927	0.927	0.927	0.974	12.9	12.9	12.9	11.8	30.2	0.259	0.260	0.259	0.269
50	10	0.977	0.977	0.977	0.985	11.8	11.1	11.8	10.9	93.0	0.182	0.185	0.182	0.186
50	25	1.018	1.018	1.018	1.018	9.9	9.4	9.9	9.4	100.0	0.149	0.151	0.149	0.151

Appendix D

Extra Tables and Plots for Chapter 5

Table D.1: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.05$ with ρ=0.05, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.05| Rej $H_0]$	$p[\widehat{deff}$ > 1.05 & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.218	1.218	1.447	1.083	7.7	7.7	12.4	8.8	10.5	10.5	100.0	1.223	41.3	5.365	3.066	5.864	5.350
2	5	1.042	1.042	1.243	0.930	13.3	13.1	13.3	11.2	7.6	7.6	100.0	1.380	30.2	1.355	1.743	1.469	3.489
2	10	1.145	1.145	1.364	1.097	14.8	14.8	14.8	10.2	11.6	11.6	100.0	1.682	35.0	1.022	1.516	1.147	2.825
2	15	1.057	1.057	1.275	1.054	16.7	16.7	15.0	10.2	14.7	14.7	100.0	1.867	38.0	0.900	1.349	1.018	2.458
2	25	1.045	1.045	1.226	1.096	21.5	21.5	16.9	8.2	20.6	20.6	100.0	2.414	47.9	0.849	1.366	0.965	2.277
2	50	1.044	1.044	1.159	1.089	27.7	27.7	22.1	9.3	29.0	29.0	100.0	3.530	55.8	0.838	1.360	0.919	1.981
5	2	1.129	1.129	1.255	1.070	8.9	8.8	9.0	9.8	11.9	11.9	100.0	1.189	48.5	1.215	1.221	1.249	1.307
5	5	1.086	1.086	1.218	1.082	10.5	10.5	10.1	9.2	15.1	15.1	100.0	1.337	38.8	0.784	0.795	0.822	0.908
5	10	1.032	1.032	1.155	1.067	12.8	12.7	11.4	9.9	21.4	21.4	100.0	1.545	46.6	0.596	0.605	0.631	0.704
5	15	1.020	1.020	1.140	1.076	13.4	13.3	11.0	9.0	27.1	27.1	100.0	1.733	54.3	0.520	0.528	0.554	0.615
5	25	0.919	0.919	1.001	0.977	16.1	15.5	13.8	9.3	42.3	42.3	100.0	2.204	68.0	0.468	0.480	0.497	0.546
5	50	0.999	0.999	1.031	1.023	14.5	14.1	12.4	9.2	66.5	66.5	100.0	3.442	81.1	0.450	0.457	0.463	0.485
10	2	0.996	0.996	1.076	0.974	10.3	10.3	9.8	10.0	12.9	12.9	100.0	1.147	49.8	0.817	0.817	0.831	0.840
10	5	1.033	1.034	1.110	1.058	9.7	9.7	9.8	8.7	18.6	18.6	100.0	1.253	37.8	0.527	0.529	0.539	0.566
10	10	0.986	0.986	1.060	1.039	10.2	10.0	10.0	8.8	32.8	32.8	100.0	1.470	53.5	0.406	0.407	0.420	0.441
10	15	0.987	0.987	1.052	1.045	13.2	12.7	11.8	9.1	46.8	46.8	100.0	1.719	68.6	0.363	0.364	0.378	0.392
10	25	0.887	0.888	0.925	0.924	15.2	15.1	13.8	11.4	61.7	61.7	100.0	2.099	78.2	0.318	0.319	0.326	0.336
10	50	1.063	1.063	1.072	1.072	10.3	9.7	9.3	8.2	88.6	88.6	100.0	3.384	94.1	0.300	0.298	0.302	0.302
25	2	0.958	0.958	1.007	0.955	10.6	10.6	11.0	11.0	15.1	15.1	100.0	1.103	50.3	0.497	0.498	0.500	0.504
25	5	1.045	1.045	1.059	1.074	8.8	8.7	9.5	8.3	29.6	29.6	100.0	1.182	34.5	0.331	0.329	0.330	0.340
25	10	0.950	0.950	0.970	0.995	11.4	11.7	11.5	9.9	53.5	53.5	100.0	1.405	61.6	0.256	0.254	0.259	0.265
25	15	0.943	0.944	0.954	0.972	11.0	11.1	10.7	9.7	72.5	72.5	100.0	1.660	77.6	0.230	0.227	0.231	0.233
25	25	0.964	0.964	0.967	0.971	12.2	12.3	12.1	12.0	93.1	93.1	100.0	2.197	94.9	0.208	0.206	0.208	0.207
25	50	0.966	0.967	0.967	0.967	8.9	9.3	8.9	9.3	99.7	99.7	100.0	3.437	99.9	0.183	0.182	0.183	0.182

Table D.2: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.05$ with ρ=0.1, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.05| Rej $H_0]$	$p[\widehat{deff}$ > 1.05 & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.078	1.078	1.316	1.044	8.8	8.8	12.4	9.8	12.3	12.3	100.0	1.256	44.9	5.750	3.243	6.586	5.700
2	5	1.252	1.252	1.470	1.217	12.7	12.5	12.7	9.1	14.8	14.8	100.0	1.574	39.5	1.699	2.440	1.849	4.249
2	10	0.935	0.935	1.092	0.934	21.7	21.7	18.6	8.7	16.4	16.4	100.0	1.893	41.1	1.136	1.849	1.276	3.198
2	15	0.990	0.990	1.125	1.014	23.3	23.3	20.0	10.8	23.1	23.1	100.0	2.350	48.5	1.158	1.940	1.286	3.034
2	25	0.987	0.988	1.094	1.029	26.8	26.8	21.7	8.7	29.8	29.8	100.0	3.234	56.1	1.159	1.984	1.275	2.839
2	50	1.004	1.004	1.055	1.029	31.5	31.1	24.8	8.3	44.1	44.1	100.0	5.435	66.5	1.268	2.140	1.338	2.659
5	2	0.872	0.872	0.968	0.850	11.7	11.7	11.4	11.7	14.2	14.2	100.0	1.204	50.5	1.263	1.275	1.305	1.388
5	5	1.025	1.025	1.159	1.078	13.8	13.7	12.5	8.7	19.7	19.7	100.0	1.434	48.1	0.844	0.858	0.897	1.003
5	10	0.890	0.890	0.978	0.947	16.9	16.6	14.4	10.9	36.3	36.3	100.0	1.897	62.7	0.708	0.724	0.752	0.830
5	15	0.935	0.935	1.000	0.981	16.1	15.6	13.5	9.5	50.4	50.4	100.0	2.340	72.5	0.662	0.679	0.695	0.752
5	25	0.972	0.972	1.007	1.004	13.6	13.0	11.4	7.9	68.3	68.3	100.0	3.352	86.5	0.653	0.664	0.674	0.709
5	50	0.973	0.973	0.983	0.982	13.1	12.7	12.3	10.6	83.6	83.6	100.0	5.691	92.4	0.642	0.634	0.650	0.650
10	2	0.992	0.992	1.071	0.994	9.8	9.7	9.2	9.3	17.8	17.8	100.0	1.174	55.6	0.850	0.850	0.865	0.882
10	5	1.055	1.055	1.120	1.108	10.8	10.7	11.1	8.1	33.9	33.9	100.0	1.408	52.1	0.590	0.589	0.603	0.634
10	10	0.980	0.980	1.029	1.030	13.1	12.5	12.1	9.4	58.2	58.2	100.0	1.869	76.7	0.489	0.490	0.505	0.521
10	15	1.020	1.020	1.045	1.046	11.2	10.8	10.3	9.1	75.6	75.6	100.0	2.323	86.8	0.460	0.458	0.468	0.473
10	25	0.950	0.950	0.958	0.960	12.2	11.6	11.3	10.2	89.3	89.3	100.0	3.283	94.1	0.430	0.428	0.433	0.434
10	50	0.973	0.973	0.974	0.974	10.7	10.7	10.5	10.5	98.0	98.0	100.0	5.724	99.4	0.411	0.406	0.411	0.407
25	2	1.076	1.076	1.134	1.097	9.1	9.2	8.9	9.0	23.2	23.2	100.0	1.138	59.6	0.517	0.517	0.524	0.529
25	5	0.928	0.928	0.943	0.970	11.6	11.9	11.2	11.0	51.6	51.6	100.0	1.343	57.9	0.366	0.363	0.367	0.377
25	10	1.025	1.025	1.034	1.042	10.0	9.8	9.8	8.9	84.6	84.6	100.0	1.873	89.0	0.309	0.307	0.311	0.311
25	15	0.981	0.981	0.982	0.983	10.6	10.6	10.5	10.0	97.0	97.0	100.0	2.364	97.9	0.286	0.285	0.287	0.285
25	25	0.924	0.924	0.924	0.924	11.6	11.7	11.6	11.7	99.9	99.9	100.0	3.366	99.9	0.264	0.263	0.264	0.263
25	50	0.968	0.968	0.968	0.968	10.1	10.2	10.1	10.2	100.0	100.0	100.0	5.853	100.0	0.245	0.245	0.245	0.245

Figure D.1: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.025, using adaptive using RLRT and deffle1.05.

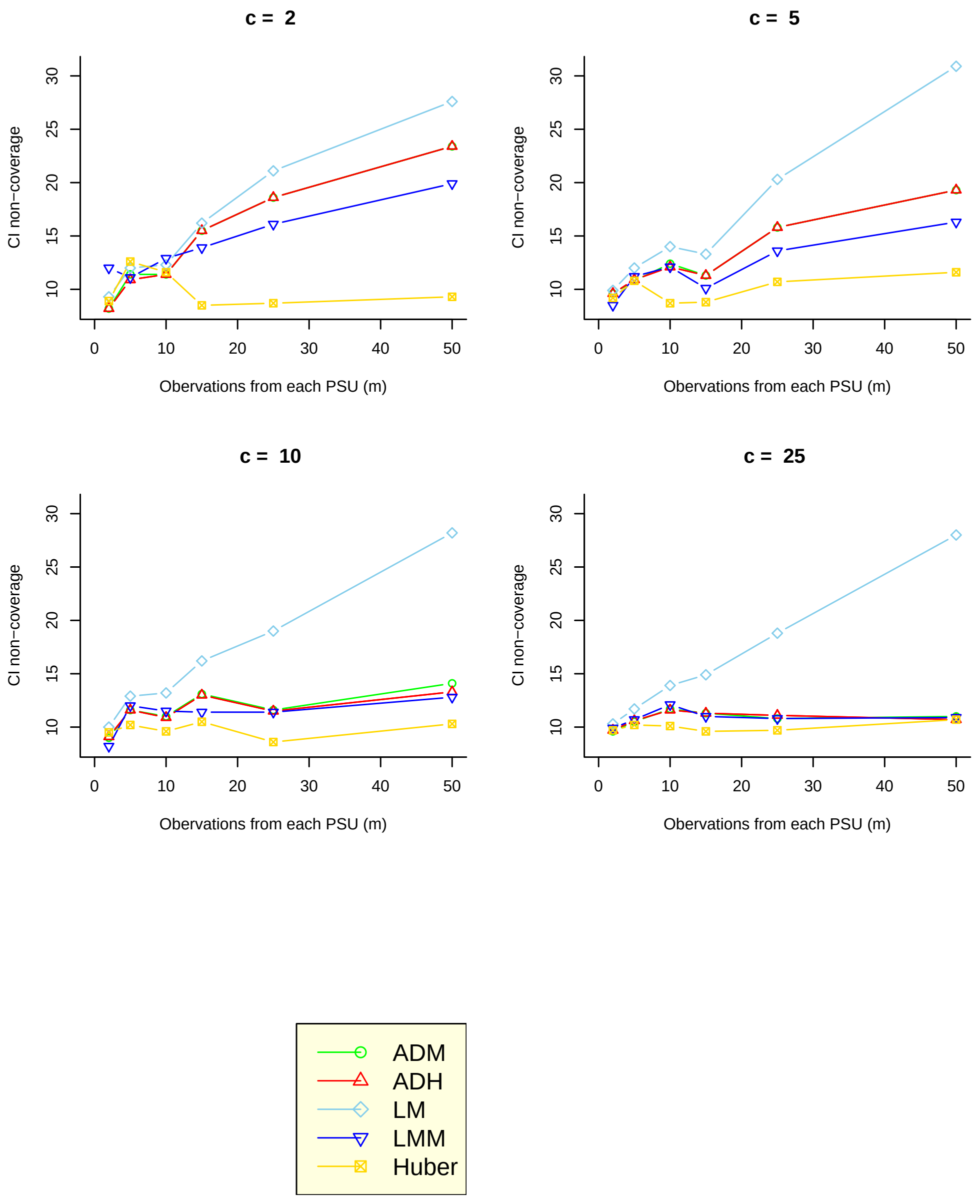


Figure D.2: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.025, using adaptive using RLRT and deff $\geq$ 1.05.

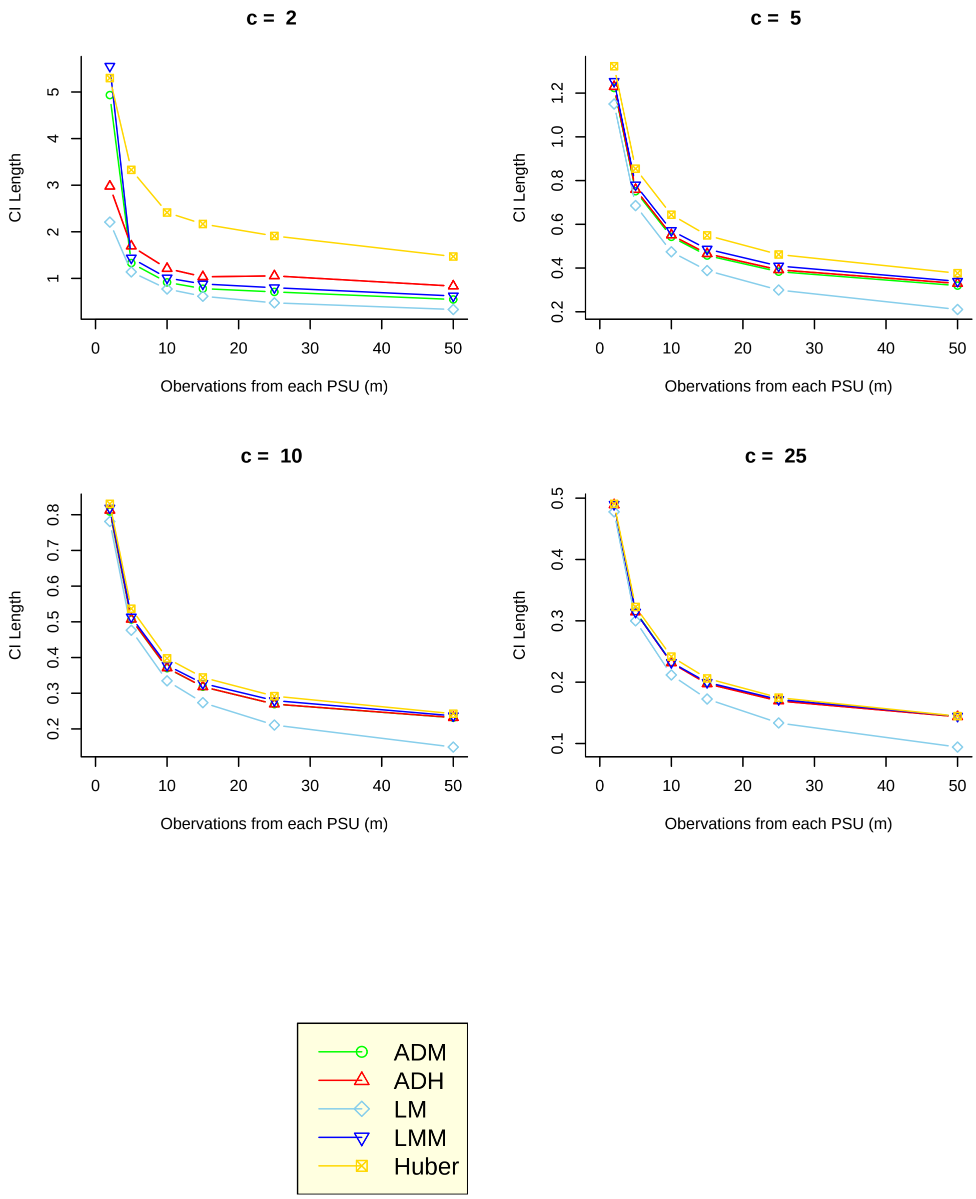


Table D.3: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.1$ with ρ=0, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}>1.1 \mid \text{Rej } H_0]$	$p[\widehat{deff}>1.1 \,\&\, \text{Rej } H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.290	1.290	1.553	1.183	8.4	8.4	11.6	7.8	11.2	11.2	100.0	1.241	41.3	5.676	3.059	6.330	5.429
2	5	1.283	1.283	1.517	1.042	9.2	9.0	9.1	10.3	6.3	6.3	100.0	1.309	25.1	1.243	1.564	1.327	3.142
2	10	1.259	1.259	1.523	1.055	8.9	8.9	9.2	10.7	5.1	5.1	100.0	1.386	26.4	0.862	1.053	0.952	2.270
2	15	1.179	1.179	1.412	0.927	9.8	9.8	10.1	10.4	3.8	3.8	100.0	1.365	24.8	0.683	0.794	0.749	1.772
2	25	1.165	1.165	1.419	0.976	10.8	10.8	11.3	8.9	4.2	4.2	100.0	1.407	24.8	0.528	0.621	0.584	1.442
2	50	1.318	1.318	1.581	1.087	7.9	7.9	9.4	9.5	5.5	5.5	100.0	1.493	24.5	0.389	0.477	0.426	1.015
5	2	1.074	1.074	1.183	0.986	9.4	9.2	10.2	9.4	9.9	9.9	100.0	1.165	40.4	1.173	1.181	1.190	1.255
5	5	1.163	1.163	1.288	1.057	9.3	9.3	10.0	8.5	7.6	7.6	100.0	1.207	27.3	0.716	0.721	0.732	0.801
5	10	1.152	1.152	1.282	1.044	8.0	8.0	8.8	9.5	6.7	6.7	100.0	1.223	26.2	0.500	0.505	0.513	0.569
5	15	1.133	1.133	1.259	1.017	9.2	9.2	10.1	10.1	7.9	7.9	100.0	1.242	26.4	0.412	0.417	0.423	0.465
5	25	1.124	1.124	1.234	0.999	9.4	9.4	10.0	10.4	7.9	7.9	100.0	1.242	25.2	0.317	0.321	0.324	0.360
5	50	1.157	1.157	1.294	1.059	8.0	7.9	8.5	8.7	7.0	7.0	100.0	1.257	26.8	0.224	0.226	0.232	0.258
10	2	1.036	1.036	1.103	0.976	10.9	10.8	10.9	11.0	11.4	11.4	100.0	1.124	36.4	0.788	0.787	0.793	0.794
10	5	1.148	1.148	1.221	1.071	8.1	8.1	9.1	9.3	8.0	8.0	100.0	1.133	23.2	0.489	0.490	0.492	0.503
10	10	1.033	1.033	1.095	0.972	10.8	11.0	10.9	10.6	8.5	8.5	100.0	1.145	21.7	0.347	0.348	0.348	0.361
10	15	1.205	1.205	1.282	1.116	7.2	7.1	8.9	9.1	8.6	8.6	100.0	1.151	23.2	0.282	0.283	0.285	0.292
10	25	1.203	1.203	1.268	1.103	7.3	7.3	7.9	9.4	8.3	8.3	100.0	1.147	20.0	0.219	0.219	0.219	0.225
10	50	1.137	1.137	1.209	1.045	9.7	9.7	10.0	10.7	8.0	8.0	100.0	1.149	21.7	0.154	0.154	0.155	0.159
25	2	0.950	0.950	0.994	0.920	11.4	11.4	11.7	12.3	10.1	10.1	100.0	1.082	32.9	0.483	0.483	0.483	0.482
25	5	0.956	0.956	0.963	0.913	10.6	10.7	11.9	11.1	8.1	8.1	100.0	1.050	10.6	0.303	0.302	0.298	0.302
25	10	1.038	1.038	1.051	0.984	10.6	10.6	11.1	11.5	8.7	8.7	100.0	1.058	12.2	0.214	0.214	0.212	0.214
25	15	1.013	1.013	1.025	0.961	9.5	9.4	10.8	10.9	6.8	6.8	100.0	1.049	10.2	0.174	0.173	0.171	0.173
25	25	1.123	1.123	1.137	1.069	8.4	8.4	8.8	9.0	9.1	9.1	100.0	1.068	12.8	0.135	0.135	0.134	0.136
25	50	1.045	1.046	1.068	0.971	9.4	9.4	10.3	10.3	8.8	8.8	100.0	1.076	16.0	0.096	0.096	0.095	0.095

Table D.4: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.1$ with ρ=0.025, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.1| Rej $H_0]$	$p[\widehat{deff}$ > 1.1 & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.189	1.189	1.440	1.074	8.2	8.2	12.0	8.9	9.9	9.9	100.0	1.234	40.2	4.934	2.977	5.552	5.299
2	5	1.182	1.182	1.415	1.012	11.4	10.9	11.1	12.6	7.5	7.5	100.0	1.381	29.5	1.330	1.693	1.437	3.330
2	10	1.139	1.139	1.356	0.985	11.4	11.4	12.9	11.6	7.5	7.5	100.0	1.475	27.9	0.915	1.211	1.009	2.413
2	15	1.005	1.005	1.227	0.947	15.5	15.5	13.9	8.5	8.6	8.6	100.0	1.630	33.2	0.780	1.034	0.880	2.167
2	25	1.074	1.074	1.273	1.058	18.6	18.6	16.1	8.7	14.4	14.4	100.0	1.985	38.6	0.710	1.054	0.800	1.911
2	50	0.886	0.886	1.056	0.911	23.4	23.4	19.9	9.3	16.9	16.9	100.0	2.291	43.7	0.548	0.835	0.623	1.469
5	2	1.123	1.123	1.245	1.061	9.6	9.6	8.5	9.2	11.9	11.9	100.0	1.190	44.5	1.224	1.230	1.254	1.323
5	5	0.986	0.986	1.102	0.941	10.8	10.9	11.2	10.8	11.2	11.2	100.0	1.276	33.2	0.751	0.759	0.780	0.854
5	10	0.967	0.967	1.093	0.975	12.4	12.1	12.1	8.7	14.7	14.7	100.0	1.398	39.9	0.543	0.551	0.572	0.644
5	15	1.080	1.080	1.233	1.118	11.3	11.3	10.1	8.8	17.1	17.1	100.0	1.487	44.2	0.457	0.465	0.487	0.549
5	25	0.895	0.895	1.007	0.945	15.8	15.8	13.6	10.7	24.7	24.7	100.0	1.691	51.6	0.383	0.392	0.409	0.462
5	50	0.812	0.812	0.888	0.861	19.3	19.3	16.3	11.6	40.2	40.2	100.0	2.185	66.1	0.320	0.330	0.339	0.376
10	2	1.044	1.044	1.125	1.008	9.0	9.1	8.2	9.5	11.9	11.9	100.0	1.143	42.9	0.809	0.812	0.819	0.830
10	5	0.953	0.953	1.016	0.948	11.6	11.6	12.0	10.2	12.5	12.5	100.0	1.179	28.4	0.507	0.507	0.514	0.537
10	10	1.045	1.045	1.121	1.062	11.0	10.9	11.5	9.6	18.6	18.6	100.0	1.275	36.4	0.370	0.371	0.378	0.398
10	15	0.935	0.935	1.008	0.975	13.1	13.0	11.4	10.5	25.8	25.8	100.0	1.400	45.2	0.318	0.318	0.327	0.344
10	25	1.002	1.002	1.072	1.061	11.6	11.5	11.4	8.6	40.5	40.5	100.0	1.639	61.2	0.269	0.270	0.279	0.292
10	50	0.995	0.995	1.029	1.030	14.1	13.3	12.8	10.3	67.1	67.1	100.0	2.231	81.8	0.231	0.232	0.237	0.243
25	2	1.018	1.018	1.066	0.997	9.6	9.7	9.9	9.7	12.7	12.7	100.0	1.092	34.5	0.489	0.489	0.490	0.491
25	5	1.007	1.007	1.022	1.017	10.5	10.6	10.7	10.2	17.2	17.2	100.0	1.107	22.3	0.316	0.314	0.314	0.323
25	10	0.991	0.991	1.008	1.034	11.7	11.6	12.1	10.1	29.0	29.0	100.0	1.202	34.7	0.232	0.231	0.232	0.242
25	15	0.990	0.990	1.009	1.038	11.3	11.3	11.0	9.6	42.3	42.3	100.0	1.309	49.2	0.199	0.197	0.200	0.206
25	25	1.049	1.049	1.063	1.087	10.8	11.1	10.8	9.7	66.0	66.0	100.0	1.580	71.8	0.171	0.169	0.172	0.175
25	50	0.947	0.948	0.953	0.954	11.0	10.7	10.9	10.7	92.7	92.7	100.0	2.214	96.2	0.144	0.144	0.145	0.145

Table D.5: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.1$ with ρ=0.05, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.1| Rej $H_0]$	$p[\widehat{deff}$ > 1.1 & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.218	1.218	1.447	1.083	7.7	7.7	12.4	8.8	10.5	10.5	100.0	1.223	38.1	5.365	3.066	5.864	5.350
2	5	1.042	1.042	1.243	0.930	13.3	13.1	13.3	11.2	7.6	7.6	100.0	1.380	30.2	1.355	1.743	1.469	3.489
2	10	1.145	1.145	1.364	1.097	14.8	14.8	14.8	10.2	11.6	11.6	100.0	1.682	35.0	1.022	1.516	1.147	2.825
2	15	1.057	1.057	1.275	1.054	16.7	16.7	15.0	10.2	14.7	14.7	100.0	1.867	38.0	0.900	1.349	1.018	2.458
2	25	1.045	1.045	1.226	1.096	21.5	21.5	16.9	8.2	20.6	20.6	100.0	2.414	47.9	0.849	1.366	0.965	2.277
2	50	1.044	1.044	1.159	1.089	27.7	27.7	22.1	9.3	29.0	29.0	100.0	3.530	55.8	0.838	1.360	0.919	1.981
5	2	1.129	1.129	1.255	1.070	8.9	8.8	9.0	9.8	11.9	11.9	100.0	1.189	45.3	1.215	1.221	1.249	1.307
5	5	1.086	1.086	1.218	1.082	10.5	10.5	10.1	9.2	15.1	15.1	100.0	1.337	38.8	0.784	0.795	0.822	0.908
5	10	1.032	1.032	1.155	1.067	12.8	12.7	11.4	9.9	21.4	21.4	100.0	1.545	46.6	0.596	0.605	0.631	0.704
5	15	1.020	1.020	1.140	1.076	13.4	13.3	11.0	9.0	27.1	27.1	100.0	1.733	54.3	0.520	0.528	0.554	0.615
5	25	0.919	0.919	1.001	0.977	16.1	15.5	13.8	9.3	42.3	42.3	100.0	2.204	68.0	0.468	0.480	0.497	0.546
5	50	0.999	0.999	1.031	1.023	14.5	14.1	12.4	9.2	66.5	66.5	100.0	3.442	81.1	0.450	0.457	0.463	0.485
10	2	0.996	0.996	1.076	0.974	10.3	10.3	9.8	10.0	12.9	12.9	100.0	1.147	44.1	0.817	0.817	0.831	0.840
10	5	1.033	1.034	1.110	1.058	9.7	9.7	9.8	8.7	18.6	18.6	100.0	1.253	37.8	0.527	0.529	0.539	0.566
10	10	0.986	0.986	1.060	1.039	10.2	10.0	10.0	8.8	32.8	32.8	100.0	1.470	53.5	0.406	0.407	0.420	0.441
10	15	0.987	0.987	1.052	1.045	13.2	12.7	11.8	9.1	46.8	46.8	100.0	1.719	68.6	0.363	0.364	0.378	0.392
10	25	0.887	0.888	0.925	0.924	15.2	15.1	13.8	11.4	61.7	61.7	100.0	2.099	78.2	0.318	0.319	0.326	0.336
10	50	1.063	1.063	1.072	1.072	10.3	9.7	9.3	8.2	88.6	88.6	100.0	3.384	94.1	0.300	0.298	0.302	0.302
25	2	0.958	0.958	1.007	0.955	10.6	10.6	11.0	11.0	15.1	15.1	100.0	1.103	39.6	0.497	0.498	0.500	0.504
25	5	1.045	1.045	1.059	1.074	8.8	8.7	9.5	8.3	29.6	29.6	100.0	1.182	34.5	0.331	0.329	0.330	0.340
25	10	0.950	0.950	0.970	0.995	11.4	11.7	11.5	9.9	53.5	53.5	100.0	1.405	61.6	0.256	0.254	0.259	0.265
25	15	0.943	0.944	0.954	0.972	11.0	11.1	10.7	9.7	72.5	72.5	100.0	1.660	77.6	0.230	0.227	0.231	0.233
25	25	0.964	0.964	0.967	0.971	12.2	12.3	12.1	12.0	93.1	93.1	100.0	2.197	94.9	0.208	0.206	0.208	0.207
25	50	0.966	0.967	0.967	0.967	8.9	9.3	8.9	9.3	99.7	99.7	100.0	3.437	99.9	0.183	0.182	0.183	0.182

Table D.6: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.1$ with ρ=0.1, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat\beta))/var(\hat\beta)$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.1| Rej $H_0]$	$p[\widehat{deff}$ > 1.1 & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.078	1.078	1.316	1.044	8.8	8.8	12.4	9.8	12.3	12.3	100.0	1.256	42.7	5.750	3.243	6.586	5.700
2	5	1.252	1.252	1.470	1.217	12.7	12.5	12.7	9.1	14.8	14.8	100.0	1.574	39.5	1.699	2.440	1.849	4.249
2	10	0.935	0.935	1.092	0.934	21.7	21.7	18.6	8.7	16.4	16.4	100.0	1.893	41.1	1.136	1.849	1.276	3.198
2	15	0.990	0.990	1.125	1.014	23.3	23.3	20.0	10.8	23.1	23.1	100.0	2.350	48.5	1.158	1.940	1.286	3.034
2	25	0.987	0.988	1.094	1.029	26.8	26.8	21.7	8.7	29.8	29.8	100.0	3.234	56.1	1.159	1.984	1.275	2.839
2	50	1.004	1.004	1.055	1.029	31.5	31.1	24.8	8.3	44.1	44.1	100.0	5.435	66.5	1.268	2.140	1.338	2.659
5	2	0.872	0.872	0.968	0.850	11.7	11.7	11.4	11.7	14.2	14.2	100.0	1.204	47.1	1.263	1.275	1.305	1.388
5	5	1.025	1.025	1.159	1.078	13.8	13.7	12.5	8.7	19.7	19.7	100.0	1.434	48.1	0.844	0.858	0.897	1.003
5	10	0.890	0.890	0.978	0.947	16.9	16.6	14.4	10.9	36.3	36.3	100.0	1.897	62.7	0.708	0.724	0.752	0.830
5	15	0.935	0.935	1.000	0.981	16.1	15.6	13.5	9.5	50.4	50.4	100.0	2.340	72.5	0.662	0.679	0.695	0.752
5	25	0.972	0.972	1.007	1.004	13.6	13.0	11.4	7.9	68.3	68.3	100.0	3.352	86.5	0.653	0.664	0.674	0.709
5	50	0.973	0.973	0.983	0.982	13.1	12.7	12.3	10.6	83.6	83.6	100.0	5.691	92.4	0.642	0.634	0.650	0.650
10	2	0.992	0.992	1.071	0.994	9.8	9.7	9.2	9.3	17.8	17.8	100.0	1.174	49.8	0.850	0.850	0.865	0.882
10	5	1.055	1.055	1.120	1.108	10.8	10.7	11.1	8.1	33.9	33.9	100.0	1.408	52.1	0.590	0.589	0.603	0.634
10	10	0.980	0.980	1.029	1.030	13.1	12.5	12.1	9.4	58.2	58.2	100.0	1.869	76.7	0.489	0.490	0.505	0.521
10	15	1.020	1.020	1.045	1.046	11.2	10.8	10.3	9.1	75.6	75.6	100.0	2.323	86.8	0.460	0.458	0.468	0.473
10	25	0.950	0.950	0.958	0.960	12.2	11.6	11.3	10.2	89.3	89.3	100.0	3.283	94.1	0.430	0.428	0.433	0.434
10	50	0.973	0.973	0.974	0.974	10.7	10.7	10.5	10.5	98.0	98.0	100.0	5.724	99.4	0.411	0.406	0.411	0.407
25	2	1.076	1.076	1.134	1.097	9.1	9.2	8.9	9.0	23.2	23.2	100.0	1.138	49.6	0.517	0.517	0.524	0.529
25	5	0.928	0.928	0.943	0.970	11.6	11.9	11.2	11.0	51.6	51.6	100.0	1.343	57.9	0.366	0.363	0.367	0.377
25	10	1.025	1.025	1.034	1.042	10.0	9.8	9.8	8.9	84.6	84.6	100.0	1.873	89.0	0.309	0.307	0.311	0.311
25	15	0.981	0.981	0.982	0.983	10.6	10.6	10.5	10.0	97.0	97.0	100.0	2.364	97.9	0.286	0.285	0.287	0.285
25	25	0.924	0.924	0.924	0.924	11.6	11.7	11.6	11.7	99.9	99.9	100.0	3.366	99.9	0.264	0.263	0.264	0.263
25	50	0.968	0.968	0.968	0.968	10.1	10.2	10.1	10.2	100.0	100.0	100.0	5.853	100.0	0.245	0.245	0.245	0.245

Figure D.3: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.025, using adaptive using RLRT and deff $\geq$ 1.1.

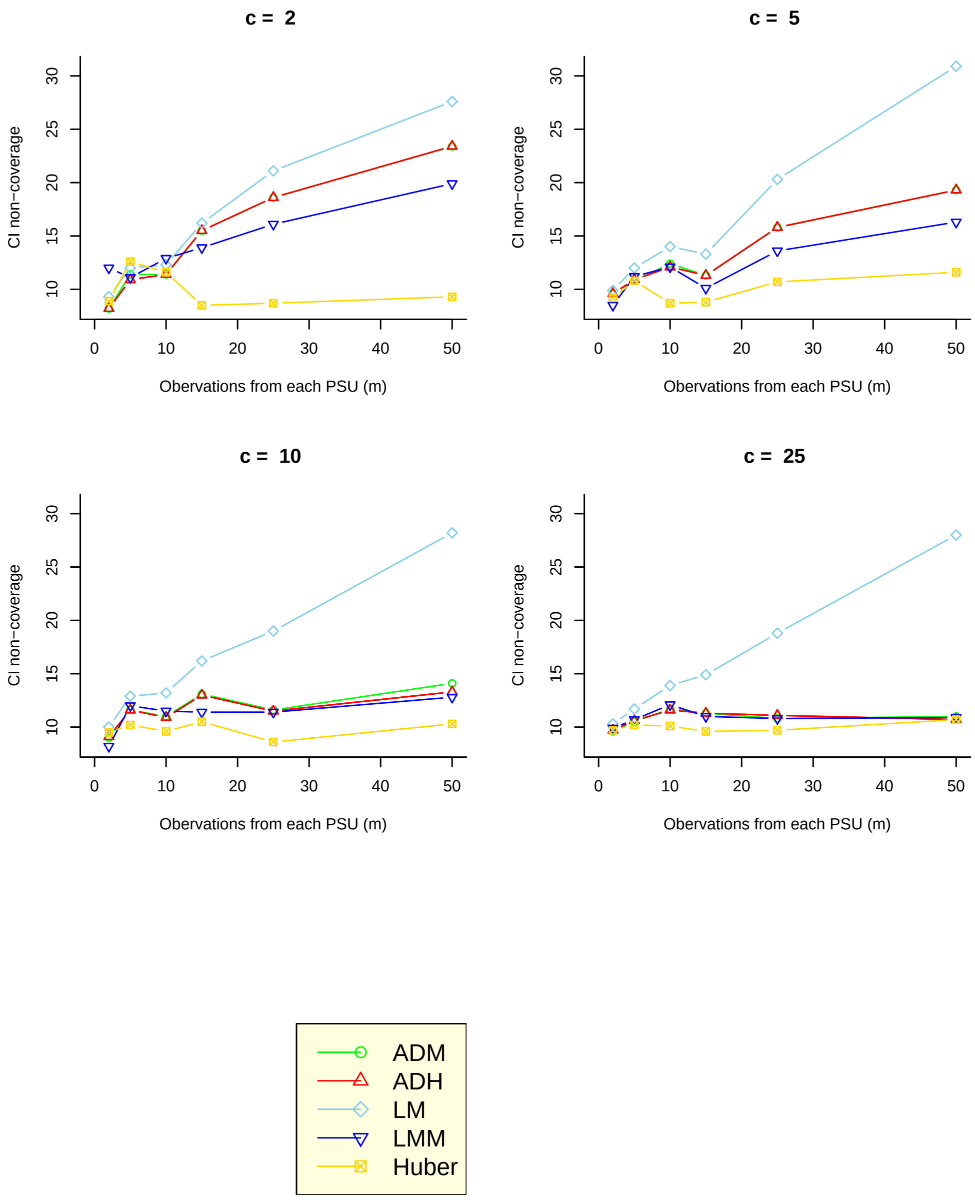


Figure D.4: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.025, using adaptive using RLRT and deff $\geq$ 1.1.

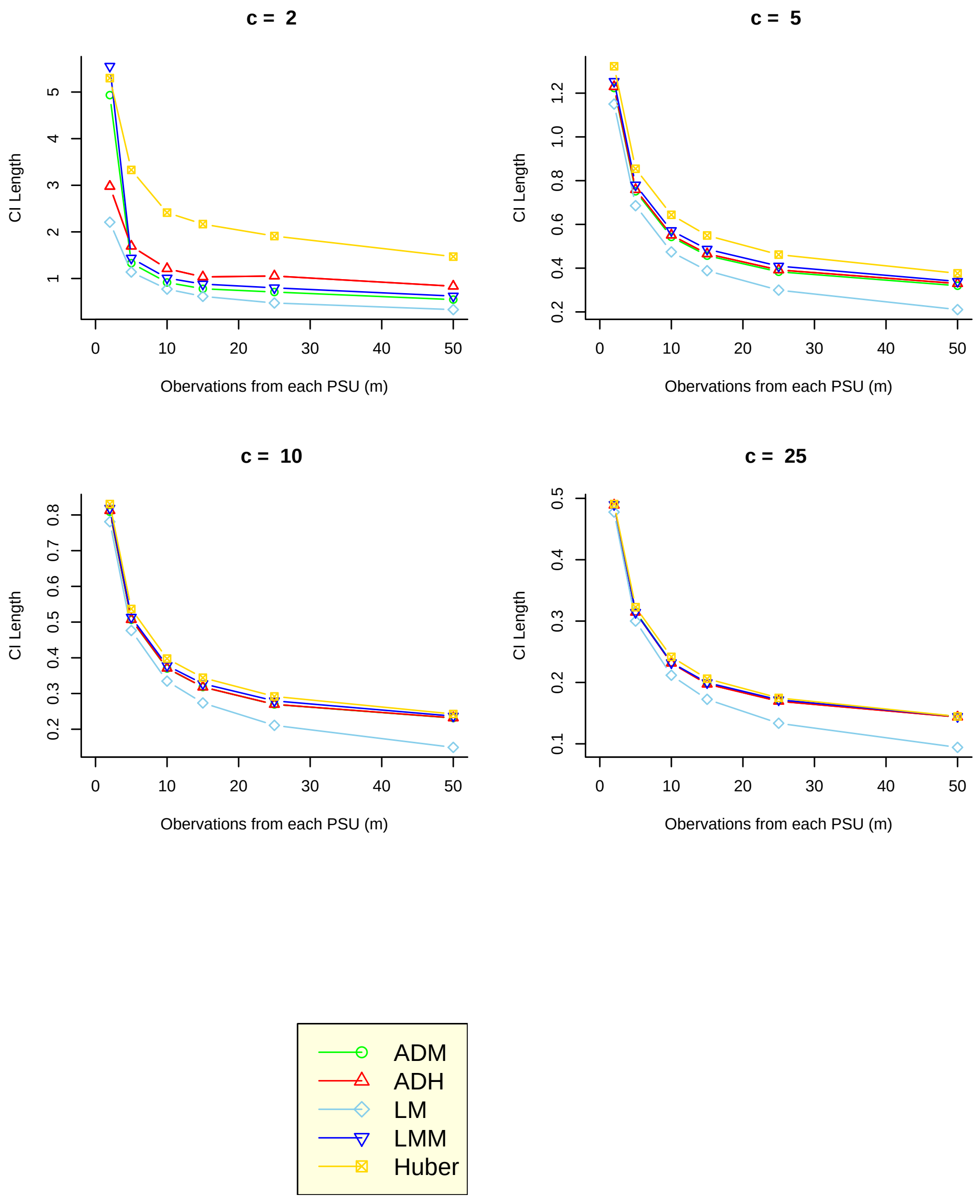


Table D.7: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.2$ with ρ=0, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}>1.2\|$ Rej $H_0]$	$p[\widehat{deff}>1.2$ & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.290	1.290	1.553	1.183	8.4	8.4	11.6	7.8	11.2	11.2	100.0	1.241	37.0	5.676	3.059	6.330	5.429
2	5	1.283	1.283	1.517	1.042	9.2	9.0	9.1	10.3	6.3	6.3	100.0	1.309	25.1	1.243	1.564	1.327	3.142
2	10	1.259	1.259	1.523	1.055	8.9	8.9	9.2	10.7	5.1	5.1	100.0	1.386	26.4	0.862	1.053	0.952	2.270
2	15	1.179	1.179	1.412	0.927	9.8	9.8	10.1	10.4	3.8	3.8	100.0	1.365	24.8	0.683	0.794	0.749	1.772
2	25	1.165	1.165	1.419	0.976	10.8	10.8	11.3	8.9	4.2	4.2	100.0	1.407	24.8	0.528	0.621	0.584	1.442
2	50	1.318	1.318	1.581	1.087	7.9	7.9	9.4	9.5	5.5	5.5	100.0	1.493	24.5	0.389	0.477	0.426	1.015
5	2	1.074	1.074	1.183	0.986	9.4	9.2	10.2	9.4	9.9	9.9	100.0	1.165	31.9	1.173	1.181	1.190	1.255
5	5	1.163	1.163	1.288	1.057	9.3	9.3	10.0	8.5	7.6	7.6	100.0	1.207	27.3	0.716	0.721	0.732	0.801
5	10	1.152	1.152	1.282	1.044	8.0	8.0	8.8	9.5	6.7	6.7	100.0	1.223	26.2	0.500	0.505	0.513	0.569
5	15	1.133	1.133	1.259	1.017	9.2	9.2	10.1	10.1	7.9	7.9	100.0	1.242	26.4	0.412	0.417	0.423	0.465
5	25	1.124	1.124	1.234	0.999	9.4	9.4	10.0	10.4	7.9	7.9	100.0	1.242	25.2	0.317	0.321	0.324	0.360
5	50	1.157	1.157	1.294	1.059	8.0	7.9	8.5	8.7	7.0	7.0	100.0	1.257	26.8	0.224	0.226	0.232	0.258
10	2	1.036	1.036	1.103	0.976	10.9	10.8	10.9	11.0	11.4	11.4	100.0	1.124	27.3	0.788	0.787	0.793	0.794
10	5	1.148	1.148	1.221	1.071	8.1	8.1	9.1	9.3	8.0	8.0	100.0	1.133	23.2	0.489	0.490	0.492	0.503
10	10	1.033	1.033	1.095	0.972	10.8	11.0	10.9	10.6	8.5	8.5	100.0	1.145	21.7	0.347	0.348	0.348	0.361
10	15	1.205	1.205	1.282	1.116	7.2	7.1	8.9	9.1	8.6	8.6	100.0	1.151	23.2	0.282	0.283	0.285	0.292
10	25	1.203	1.203	1.268	1.103	7.3	7.3	7.9	9.4	8.3	8.3	100.0	1.147	20.0	0.219	0.219	0.219	0.225
10	50	1.137	1.137	1.209	1.045	9.7	9.7	10.0	10.7	8.0	8.0	100.0	1.149	21.7	0.154	0.154	0.155	0.159
25	2	0.950	0.950	0.994	0.920	11.4	11.4	11.7	12.3	10.1	10.1	100.0	1.082	16.2	0.483	0.483	0.483	0.482
25	5	0.956	0.956	0.963	0.913	10.6	10.7	11.9	11.1	8.1	8.1	100.0	1.050	10.6	0.303	0.302	0.298	0.302
25	10	1.038	1.038	1.051	0.984	10.6	10.6	11.1	11.5	8.7	8.7	100.0	1.058	12.2	0.214	0.214	0.212	0.214
25	15	1.013	1.013	1.025	0.961	9.5	9.4	10.8	10.9	6.8	6.8	100.0	1.049	10.2	0.174	0.173	0.171	0.173
25	25	1.123	1.123	1.137	1.069	8.4	8.4	8.8	9.0	9.1	9.1	100.0	1.068	12.8	0.135	0.135	0.134	0.136
25	50	1.045	1.046	1.068	0.971	9.4	9.4	10.3	10.3	8.8	8.8	100.0	1.076	16.0	0.096	0.096	0.095	0.095

Table D.8: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.2$ with ρ=0.025, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.2| Rej $H_0]$	$p[\widehat{deff}$ > 1.2 & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.189	1.189	1.440	1.074	8.2	8.2	12.0	8.9	9.9	9.9	100.0	1.234	36.2	4.934	2.977	5.552	5.299
2	5	1.182	1.182	1.415	1.012	11.4	10.9	11.1	12.6	7.5	7.5	100.0	1.381	29.5	1.330	1.693	1.437	3.330
2	10	1.139	1.139	1.356	0.985	11.4	11.4	12.9	11.6	7.5	7.5	100.0	1.475	27.9	0.915	1.211	1.009	2.413
2	15	1.005	1.005	1.227	0.947	15.5	15.5	13.9	8.5	8.6	8.6	100.0	1.630	33.2	0.780	1.034	0.880	2.167
2	25	1.074	1.074	1.273	1.058	18.6	18.6	16.1	8.7	14.4	14.4	100.0	1.985	38.6	0.710	1.054	0.800	1.911
2	50	0.886	0.886	1.056	0.911	23.4	23.4	19.9	9.3	16.9	16.9	100.0	2.291	43.7	0.548	0.835	0.623	1.469
5	2	1.123	1.123	1.245	1.061	9.6	9.6	8.5	9.2	11.9	11.9	100.0	1.190	36.9	1.224	1.230	1.254	1.323
5	5	0.986	0.986	1.102	0.941	10.8	10.9	11.2	10.8	11.2	11.2	100.0	1.276	33.2	0.751	0.759	0.780	0.854
5	10	0.967	0.967	1.093	0.975	12.4	12.1	12.1	8.7	14.7	14.7	100.0	1.398	39.9	0.543	0.551	0.572	0.644
5	15	1.080	1.080	1.233	1.118	11.3	11.3	10.1	8.8	17.1	17.1	100.0	1.487	44.2	0.457	0.465	0.487	0.549
5	25	0.895	0.895	1.007	0.945	15.8	15.8	13.6	10.7	24.7	24.7	100.0	1.691	51.6	0.383	0.392	0.409	0.462
5	50	0.812	0.812	0.888	0.861	19.3	19.3	16.3	11.6	40.2	40.2	100.0	2.185	66.1	0.320	0.330	0.339	0.376
10	2	1.044	1.044	1.125	1.008	9.0	9.1	8.2	9.5	11.9	11.9	100.0	1.143	30.8	0.809	0.812	0.819	0.830
10	5	0.953	0.953	1.016	0.948	11.6	11.6	12.0	10.2	12.5	12.5	100.0	1.179	28.4	0.507	0.507	0.514	0.537
10	10	1.045	1.045	1.121	1.062	11.0	10.9	11.5	9.6	18.6	18.6	100.0	1.275	36.4	0.370	0.371	0.378	0.398
10	15	0.935	0.935	1.008	0.975	13.1	13.0	11.4	10.5	25.8	25.8	100.0	1.400	45.2	0.318	0.318	0.327	0.344
10	25	1.002	1.002	1.072	1.061	11.6	11.5	11.4	8.6	40.5	40.5	100.0	1.639	61.2	0.269	0.270	0.279	0.292
10	50	0.995	0.995	1.029	1.030	14.1	13.3	12.8	10.3	67.1	67.1	100.0	2.231	81.8	0.231	0.232	0.237	0.243
25	2	1.018	1.018	1.066	0.997	9.6	9.7	9.9	9.7	12.7	12.7	100.0	1.092	18.4	0.489	0.489	0.490	0.491
25	5	1.007	1.007	1.022	1.017	10.5	10.6	10.7	10.2	17.2	17.2	100.0	1.107	22.3	0.316	0.314	0.314	0.323
25	10	0.991	0.991	1.008	1.034	11.7	11.6	12.1	10.1	29.0	29.0	100.0	1.202	34.7	0.232	0.231	0.232	0.242
25	15	0.990	0.990	1.009	1.038	11.3	11.3	11.0	9.6	42.3	42.3	100.0	1.309	49.2	0.199	0.197	0.200	0.206
25	25	1.049	1.049	1.063	1.087	10.8	11.1	10.8	9.7	66.0	66.0	100.0	1.580	71.8	0.171	0.169	0.172	0.175
25	50	0.947	0.948	0.953	0.954	11.0	10.7	10.9	10.7	92.7	92.7	100.0	2.214	96.2	0.144	0.144	0.145	0.145

Table D.9: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.2$ with ρ=0.05, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.2| Rej $H_0]$	$p[\widehat{deff}$ > 1.2 & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.218	1.218	1.447	1.083	7.7	7.7	12.4	8.8	10.5	10.5	100.0	1.223	34.6	5.365	3.066	5.864	5.350
2	5	1.042	1.042	1.243	0.930	13.3	13.1	13.3	11.2	7.6	7.6	100.0	1.380	30.2	1.355	1.743	1.469	3.489
2	10	1.145	1.145	1.364	1.097	14.8	14.8	14.8	10.2	11.6	11.6	100.0	1.682	35.0	1.022	1.516	1.147	2.825
2	15	1.057	1.057	1.275	1.054	16.7	16.7	15.0	10.2	14.7	14.7	100.0	1.867	38.0	0.900	1.349	1.018	2.458
2	25	1.045	1.045	1.226	1.096	21.5	21.5	16.9	8.2	20.6	20.6	100.0	2.414	47.9	0.849	1.366	0.965	2.277
2	50	1.044	1.044	1.159	1.089	27.7	27.7	22.1	9.3	29.0	29.0	100.0	3.530	55.8	0.838	1.360	0.919	1.981
5	2	1.129	1.129	1.255	1.070	8.9	8.8	9.0	9.8	11.9	11.9	100.0	1.189	36.7	1.215	1.221	1.249	1.307
5	5	1.086	1.086	1.218	1.082	10.5	10.5	10.1	9.2	15.1	15.1	100.0	1.337	38.8	0.784	0.795	0.822	0.908
5	10	1.032	1.032	1.155	1.067	12.8	12.7	11.4	9.9	21.4	21.4	100.0	1.545	46.6	0.596	0.605	0.631	0.704
5	15	1.020	1.020	1.140	1.076	13.4	13.3	11.0	9.0	27.1	27.1	100.0	1.733	54.3	0.520	0.528	0.554	0.615
5	25	0.919	0.919	1.001	0.977	16.1	15.5	13.8	9.3	42.3	42.3	100.0	2.204	68.0	0.468	0.480	0.497	0.546
5	50	0.999	0.999	1.031	1.023	14.5	14.1	12.4	9.2	66.5	66.5	100.0	3.442	81.1	0.450	0.457	0.463	0.485
10	2	0.996	0.996	1.076	0.974	10.3	10.3	9.8	10.0	12.9	12.9	100.0	1.147	32.0	0.817	0.817	0.831	0.840
10	5	1.033	1.034	1.110	1.058	9.7	9.7	9.8	8.7	18.6	18.6	100.0	1.253	37.8	0.527	0.529	0.539	0.566
10	10	0.986	0.986	1.060	1.039	10.2	10.0	10.0	8.8	32.8	32.8	100.0	1.470	53.5	0.406	0.407	0.420	0.441
10	15	0.987	0.987	1.052	1.045	13.2	12.7	11.8	9.1	46.8	46.8	100.0	1.719	68.6	0.363	0.364	0.378	0.392
10	25	0.887	0.888	0.925	0.924	15.2	15.1	13.8	11.4	61.7	61.7	100.0	2.099	78.2	0.318	0.319	0.326	0.336
10	50	1.063	1.063	1.072	1.072	10.3	9.7	9.3	8.2	88.6	88.6	100.0	3.384	94.1	0.300	0.298	0.302	0.302
25	2	0.958	0.958	1.007	0.955	10.6	10.6	11.0	11.0	15.1	15.1	100.0	1.103	22.2	0.497	0.498	0.500	0.504
25	5	1.045	1.045	1.059	1.074	8.8	8.7	9.5	8.3	29.6	29.6	100.0	1.182	34.5	0.331	0.329	0.330	0.340
25	10	0.950	0.950	0.970	0.995	11.4	11.7	11.5	9.9	53.5	53.5	100.0	1.405	61.6	0.256	0.254	0.259	0.265
25	15	0.943	0.944	0.954	0.972	11.0	11.1	10.7	9.7	72.5	72.5	100.0	1.660	77.6	0.230	0.227	0.231	0.233
25	25	0.964	0.964	0.967	0.971	12.2	12.3	12.1	12.0	93.1	93.1	100.0	2.197	94.9	0.208	0.206	0.208	0.207
25	50	0.966	0.967	0.967	0.967	8.9	9.3	8.9	9.3	99.7	99.7	100.0	3.437	99.9	0.183	0.182	0.183	0.182

Table D.10: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $deff \geq 1.2$ with ρ=0.1, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat\beta))/var(\hat\beta)$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.2| Rej $H_0]$	$p[\widehat{deff}$ > 1.2 & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.078	1.078	1.316	1.044	8.8	8.8	12.4	9.8	12.3	12.3	100.0	1.256	38.8	5.750	3.243	6.586	5.700
2	5	1.252	1.252	1.470	1.217	12.7	12.5	12.7	9.1	14.8	14.8	100.0	1.574	39.5	1.699	2.440	1.849	4.249
2	10	0.935	0.935	1.092	0.934	21.7	21.7	18.6	8.7	16.4	16.4	100.0	1.893	41.1	1.136	1.849	1.276	3.198
2	15	0.990	0.990	1.125	1.014	23.3	23.3	20.0	10.8	23.1	23.1	100.0	2.350	48.5	1.158	1.940	1.286	3.034
2	25	0.987	0.988	1.094	1.029	26.8	26.8	21.7	8.7	29.8	29.8	100.0	3.234	56.1	1.159	1.984	1.275	2.839
2	50	1.004	1.004	1.055	1.029	31.5	31.1	24.8	8.3	44.1	44.1	100.0	5.435	66.5	1.268	2.140	1.338	2.659
5	2	0.872	0.872	0.968	0.850	11.7	11.7	11.4	11.7	14.2	14.2	100.0	1.204	38.5	1.263	1.275	1.305	1.388
5	5	1.025	1.025	1.159	1.078	13.8	13.7	12.5	8.7	19.7	19.7	100.0	1.434	48.1	0.844	0.858	0.897	1.003
5	10	0.890	0.890	0.978	0.947	16.9	16.6	14.4	10.9	36.3	36.3	100.0	1.897	62.7	0.708	0.724	0.752	0.830
5	15	0.935	0.935	1.000	0.981	16.1	15.6	13.5	9.5	50.4	50.4	100.0	2.340	72.5	0.662	0.679	0.695	0.752
5	25	0.972	0.972	1.007	1.004	13.6	13.0	11.4	7.9	68.3	68.3	100.0	3.352	86.5	0.653	0.664	0.674	0.709
5	50	0.973	0.973	0.983	0.982	13.1	12.7	12.3	10.6	83.6	83.6	100.0	5.691	92.4	0.642	0.634	0.650	0.650
10	2	0.992	0.992	1.071	0.994	9.8	9.7	9.2	9.3	17.8	17.8	100.0	1.174	37.7	0.850	0.850	0.865	0.882
10	5	1.055	1.055	1.120	1.108	10.8	10.7	11.1	8.1	33.9	33.9	100.0	1.408	52.1	0.590	0.589	0.603	0.634
10	10	0.980	0.980	1.029	1.030	13.1	12.5	12.1	9.4	58.2	58.2	100.0	1.869	76.7	0.489	0.490	0.505	0.521
10	15	1.020	1.020	1.045	1.046	11.2	10.8	10.3	9.1	75.6	75.6	100.0	2.323	86.8	0.460	0.458	0.468	0.473
10	25	0.950	0.950	0.958	0.960	12.2	11.6	11.3	10.2	89.3	89.3	100.0	3.283	94.1	0.430	0.428	0.433	0.434
10	50	0.973	0.973	0.974	0.974	10.7	10.7	10.5	10.5	98.0	98.0	100.0	5.724	99.4	0.411	0.406	0.411	0.407
25	2	1.076	1.076	1.134	1.097	9.1	9.2	8.9	9.0	23.2	23.2	100.0	1.138	31.8	0.517	0.517	0.524	0.529
25	5	0.928	0.928	0.943	0.970	11.6	11.9	11.2	11.0	51.6	51.6	100.0	1.343	57.9	0.366	0.363	0.367	0.377
25	10	1.025	1.025	1.034	1.042	10.0	9.8	9.8	8.9	84.6	84.6	100.0	1.873	89.0	0.309	0.307	0.311	0.311
25	15	0.981	0.981	0.982	0.983	10.6	10.6	10.5	10.0	97.0	97.0	100.0	2.364	97.9	0.286	0.285	0.287	0.285
25	25	0.924	0.924	0.924	0.924	11.6	11.7	11.6	11.7	99.9	99.9	100.0	3.366	99.9	0.264	0.263	0.264	0.263
25	50	0.968	0.968	0.968	0.968	10.1	10.2	10.1	10.2	100.0	100.0	100.0	5.853	100.0	0.245	0.245	0.245	0.245

Figure D.5: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.025, using adaptive using RLRT and deff $\geq$ 1.2.

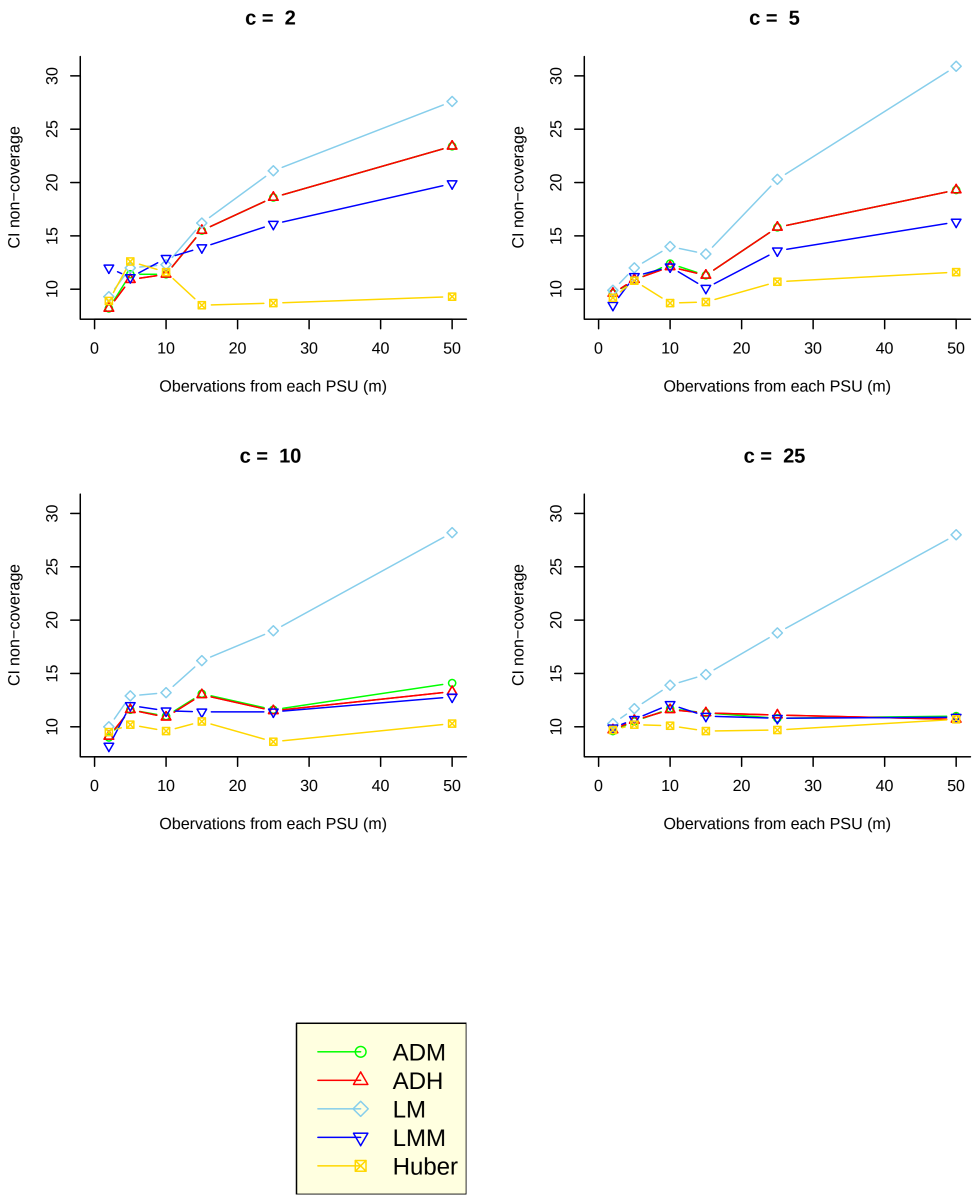


Figure D.6: Confidence interval non-coverage using different variance estimation methods and for various values of m and c, ρ=0.025, using adaptive using RLRT and deff $\geq$ 1.2.

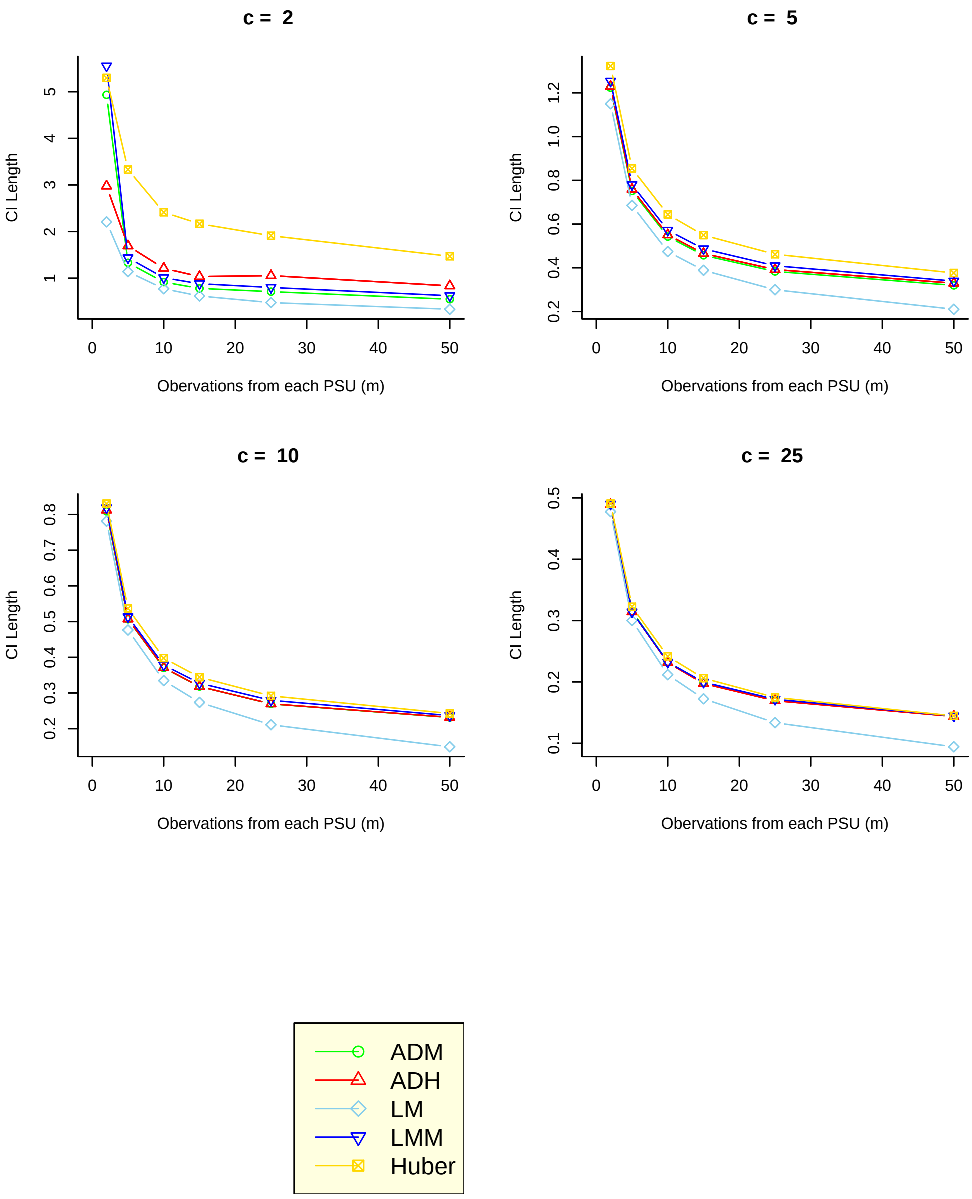


Figure D.7: Histograms for $\hat{\rho}$ and $\hat{\rho}$ when H_0 is rejected and accepted and $\widehat{deff}$ when H_0 is rejected (c=10, m=2, $\rho = 0.025$)

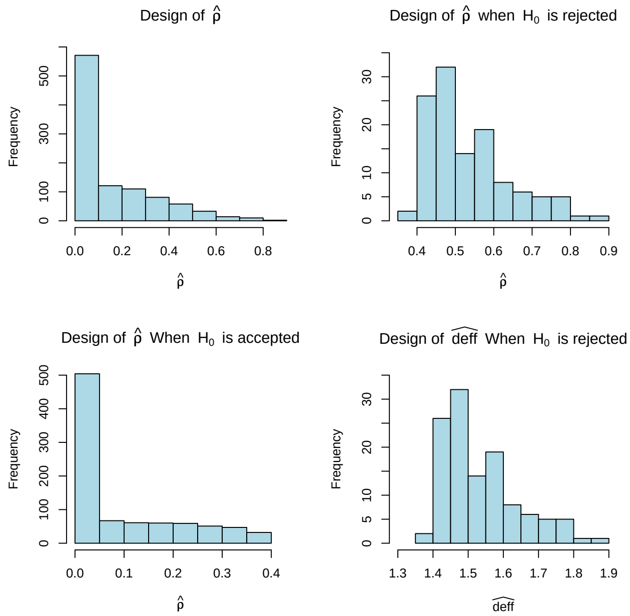


Figure D.8: Histograms for $\hat{\rho}$ and $\hat{\rho}$ when H_0 is rejected and accepted and $\widehat{deff}$ when H_0 is rejected (c=10, m=10, $\rho = 0.025$)

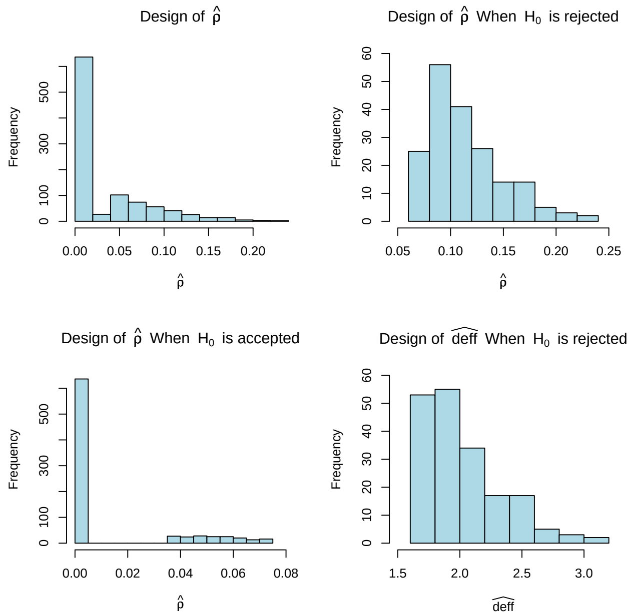


Table D.11: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $\widehat{deff} \geq 1.5$ with ρ=0.05, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}>1.5 \mid \text{Rej } H_0]$	$p[\widehat{deff} > 1.5 \text{ \& Rej } H_0]$	$E(\widehat{deff})$	$p[\widehat{deff}>1.5]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.218	1.218	1.447	1.083	7.7	7.7	12.4	8.8	10.5	10.5	100.0	1.223	22.0	5.365	3.066	5.864	5.350
2	5	1.042	1.042	1.243	0.930	13.3	13.1	13.3	11.2	7.6	7.6	100.0	1.380	25.3	1.355	1.743	1.469	3.489
2	10	1.145	1.145	1.364	1.097	14.8	14.8	14.8	10.2	11.6	11.6	100.0	1.682	31.8	1.022	1.516	1.147	2.825
2	15	1.057	1.057	1.275	1.054	16.7	16.7	15.0	10.2	14.7	14.7	100.0	1.867	34.9	0.900	1.349	1.018	2.458
2	25	1.045	1.045	1.226	1.096	21.5	21.5	16.9	8.2	20.6	20.6	100.0	2.414	44.5	0.849	1.366	0.965	2.277
2	50	1.044	1.044	1.159	1.089	27.7	27.7	22.1	9.3	29.0	29.0	100.0	3.530	53.6	0.838	1.360	0.919	1.981
5	2	1.129	1.129	1.255	1.070	8.9	8.8	9.0	9.8	11.9	11.9	100.0	1.189	13.6	1.215	1.221	1.249	1.307
5	5	1.086	1.086	1.218	1.082	10.5	10.5	10.1	9.2	15.1	15.1	100.0	1.337	30.7	0.784	0.795	0.822	0.908
5	10	1.032	1.032	1.155	1.067	12.8	12.7	11.4	9.9	21.4	21.4	100.0	1.545	38.8	0.596	0.605	0.631	0.704
5	15	1.020	1.020	1.140	1.076	13.4	13.3	11.0	9.0	27.1	27.1	100.0	1.733	47.0	0.520	0.528	0.554	0.615
5	25	0.919	0.919	1.001	0.977	16.1	15.5	13.8	9.3	42.3	42.3	100.0	2.204	61.3	0.468	0.480	0.497	0.546
5	50	0.999	0.999	1.031	1.023	14.5	14.1	12.4	9.2	66.5	66.5	100.0	3.442	78.2	0.450	0.457	0.463	0.485
10	2	0.970	0.970	1.076	0.974	10.9	10.9	9.8	10.0	12.9	6.6	51.2	1.147	6.6	0.804	0.804	0.831	0.840
10	5	1.033	1.034	1.110	1.058	9.7	9.7	9.8	8.7	18.6	18.6	100.0	1.253	23.9	0.527	0.529	0.539	0.566
10	10	0.986	0.986	1.060	1.039	10.2	10.0	10.0	8.8	32.8	32.8	100.0	1.470	42.7	0.406	0.407	0.420	0.441
10	15	0.987	0.987	1.052	1.045	13.2	12.7	11.8	9.1	46.8	46.8	100.0	1.719	57.2	0.363	0.364	0.378	0.392
10	25	0.887	0.888	0.925	0.924	15.2	15.1	13.8	11.4	61.7	61.7	100.0	2.099	71.4	0.318	0.319	0.326	0.336
10	50	1.063	1.063	1.072	1.072	10.3	9.7	9.3	8.2	88.6	88.6	100.0	3.384	92.3	0.300	0.298	0.302	0.302
25	2	0.913	0.913	1.007	0.955	11.7	11.7	11.0	11.0	15.1	0.6	4.0	1.103	0.6	0.485	0.485	0.500	0.504
25	5	0.989	0.989	1.059	1.074	9.5	9.7	9.5	8.3	29.6	15.2	51.4	1.182	15.2	0.320	0.319	0.330	0.340
25	10	0.913	0.913	0.970	0.995	12.2	12.4	11.5	9.9	53.5	41.4	77.4	1.405	41.4	0.249	0.248	0.259	0.265
25	15	0.920	0.920	0.954	0.972	12.0	12.2	10.7	9.7	72.5	63.7	87.9	1.660	63.7	0.226	0.224	0.231	0.233
25	25	0.955	0.955	0.967	0.971	12.7	13.0	12.1	12.0	93.1	88.5	95.1	2.197	88.5	0.206	0.204	0.208	0.207
25	50	0.966	0.966	0.967	0.967	8.9	9.3	8.9	9.3	99.7	99.6	99.9	3.437	99.6	0.183	0.182	0.183	0.182

Table D.12: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and $\widehat{deff} \geq 1.5$ with ρ=0.1, balanced data case.

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.5| Rej $H_0]$	$p[\widehat{deff}$ > 1.5 & Rej $H_0]$	$E(\widehat{deff})$		Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub						ADM	ADH	LMM	Hub
2	2	1.078	1.078	1.316	1.044	8.8	8.8	12.4	9.8	12.3	12.3	100.0	1.256	26.4	5.750	3.243	6.586	5.700
2	5	1.252	1.252	1.470	1.217	12.7	12.5	12.7	9.1	14.8	14.8	100.0	1.574	34.3	1.699	2.440	1.849	4.249
2	10	0.935	0.935	1.092	0.934	21.7	21.7	18.6	8.7	16.4	16.4	100.0	1.893	38.2	1.136	1.849	1.276	3.198
2	15	0.990	0.990	1.125	1.014	23.3	23.3	20.0	10.8	23.1	23.1	100.0	2.350	45.6	1.158	1.940	1.286	3.034
2	25	0.987	0.988	1.094	1.029	26.8	26.8	21.7	8.7	29.8	29.8	100.0	3.234	53.9	1.159	1.984	1.275	2.839
2	50	1.004	1.004	1.055	1.029	31.5	31.1	24.8	8.3	44.1	44.1	100.0	5.435	64.0	1.268	2.140	1.338	2.659
5	2	0.872	0.872	0.968	0.850	11.7	11.7	11.4	11.7	14.2	14.2	100.0	1.204	17.4	1.263	1.275	1.305	1.388
5	5	1.025	1.025	1.159	1.078	13.8	13.7	12.5	8.7	19.7	19.7	100.0	1.434	37.1	0.844	0.858	0.897	1.003
5	10	0.890	0.890	0.978	0.947	16.9	16.6	14.4	10.9	36.3	36.3	100.0	1.897	55.6	0.708	0.724	0.752	0.830
5	15	0.935	0.935	1.000	0.981	16.1	15.6	13.5	9.5	50.4	50.4	100.0	2.340	68.0	0.662	0.679	0.695	0.752
5	25	0.972	0.972	1.007	1.004	13.6	13.0	11.4	7.9	68.3	68.3	100.0	3.352	82.1	0.653	0.664	0.674	0.709
5	50	0.973	0.973	0.983	0.982	13.1	12.7	12.3	10.6	83.6	83.6	100.0	5.691	90.2	0.642	0.634	0.650	0.650
10	2	0.954	0.954	1.071	0.994	10.3	10.2	9.2	9.3	17.8	9.0	50.6	1.174	9.0	0.830	0.829	0.865	0.882
10	5	1.055	1.055	1.120	1.108	10.8	10.7	11.1	8.1	33.9	33.9	100.0	1.408	39.1	0.590	0.589	0.603	0.634
10	10	0.980	0.980	1.029	1.030	13.1	12.5	12.1	9.4	58.2	58.2	100.0	1.869	67.2	0.489	0.490	0.505	0.521
10	15	1.020	1.020	1.045	1.046	11.2	10.8	10.3	9.1	75.6	75.6	100.0	2.323	82.2	0.460	0.458	0.468	0.473
10	25	0.950	0.950	0.958	0.960	12.2	11.6	11.3	10.2	89.3	89.3	100.0	3.283	92.4	0.430	0.428	0.433	0.434
10	50	0.973	0.973	0.974	0.974	10.7	10.7	10.5	10.5	98.0	98.0	100.0	5.724	98.8	0.411	0.406	0.411	0.407
25	2	0.995	0.995	1.134	1.097	10.1	10.1	8.9	9.0	23.2	0.9	3.9	1.138	0.9	0.496	0.496	0.524	0.529
25	5	0.878	0.878	0.943	0.970	12.8	12.8	11.2	11.0	51.6	35.2	68.2	1.343	35.2	0.353	0.351	0.367	0.377
25	10	1.009	1.010	1.034	1.042	10.6	10.4	9.8	8.9	84.6	77.9	92.1	1.873	77.9	0.306	0.303	0.311	0.311
25	15	0.977	0.977	0.982	0.983	10.7	10.7	10.5	10.0	97.0	94.9	97.8	2.364	94.9	0.285	0.284	0.287	0.285
25	25	0.924	0.924	0.924	0.924	11.6	11.7	11.6	11.7	99.9	99.9	100.0	3.366	99.9	0.264	0.263	0.264	0.263
25	50	0.968	0.968	0.968	0.968	10.1	10.2	10.1	10.2	100.0	100.0	100.0	5.853	100.0	0.245	0.245	0.245	0.245

Table D.13: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.05 using RLRT in the unbalanced data case with ρ=0.1

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.05| Rej $H_0]$	$p[\widehat{deff}$ > 1.05 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.258	1.258	1.356	1.100	10.3	10.2	8.6	10.5	24.2	100.0	24.2	2.648	3.570	2.817	5.150
2	10	1.145	1.145	1.188	1.025	16.3	15.5	15.3	9.6	32.7	100.0	32.7	1.285	2.609	1.321	3.375
2	25	1.028	1.028	1.050	0.987	22.3	20.3	21.2	9.1	45.8	100.0	45.8	0.917	2.386	0.935	2.845
5	3	1.149	1.149	1.165	1.047	10.7	10.3	10.4	9.8	39.5	100.0	39.5	1.132	1.205	1.142	1.232
5	10	0.986	0.986	0.989	0.943	13.9	11.0	13.9	10.3	59.4	100.0	59.4	0.685	0.796	0.687	0.813
5	25	0.934	0.934	0.934	0.926	18.7	12.1	18.7	11.4	81.3	100.0	81.3	0.557	0.684	0.557	0.692
10	3	1.073	1.073	1.074	1.029	10.9	10.4	10.9	9.8	44.3	100.0	44.3	0.723	0.744	0.724	0.749
10	10	1.055	1.055	1.055	1.052	12.4	10.6	12.4	10.2	76.7	100.0	76.7	0.483	0.519	0.483	0.525
10	25	0.972	0.973	0.972	0.973	14.8	11.0	14.8	10.9	94.6	100.0	94.6	0.400	0.438	0.400	0.439
25	3	0.963	0.963	0.963	0.980	10.4	10.2	10.3	9.5	43.1	100.0	43.1	0.441	0.445	0.441	0.455
25	10	1.087	1.087	1.087	1.093	10.3	9.0	10.3	8.5	91.3	100.0	91.3	0.306	0.315	0.306	0.317
25	25	0.998	0.998	0.998	0.998	10.7	10.0	10.7	10.0	99.8	100.0	99.8	0.256	0.264	0.256	0.264
50	3	0.916	0.916	0.916	0.962	12.5	12.5	12.5	10.9	38.3	100.0	38.3	0.305	0.306	0.305	0.316
50	10	0.988	0.988	0.988	0.990	11.0	10.5	11.0	10.3	98.1	100.0	98.1	0.217	0.221	0.217	0.221
50	25	1.008	1.008	1.008	1.008	9.8	9.4	9.8	9.4	100.0	100.0	100.0	0.182	0.185	0.182	0.185

Figure D.9: Confidence interval non-coverage using different variance estimation methods and for various values of $\bar{m}$ and c, ρ=0.025, $deff = 1.05$

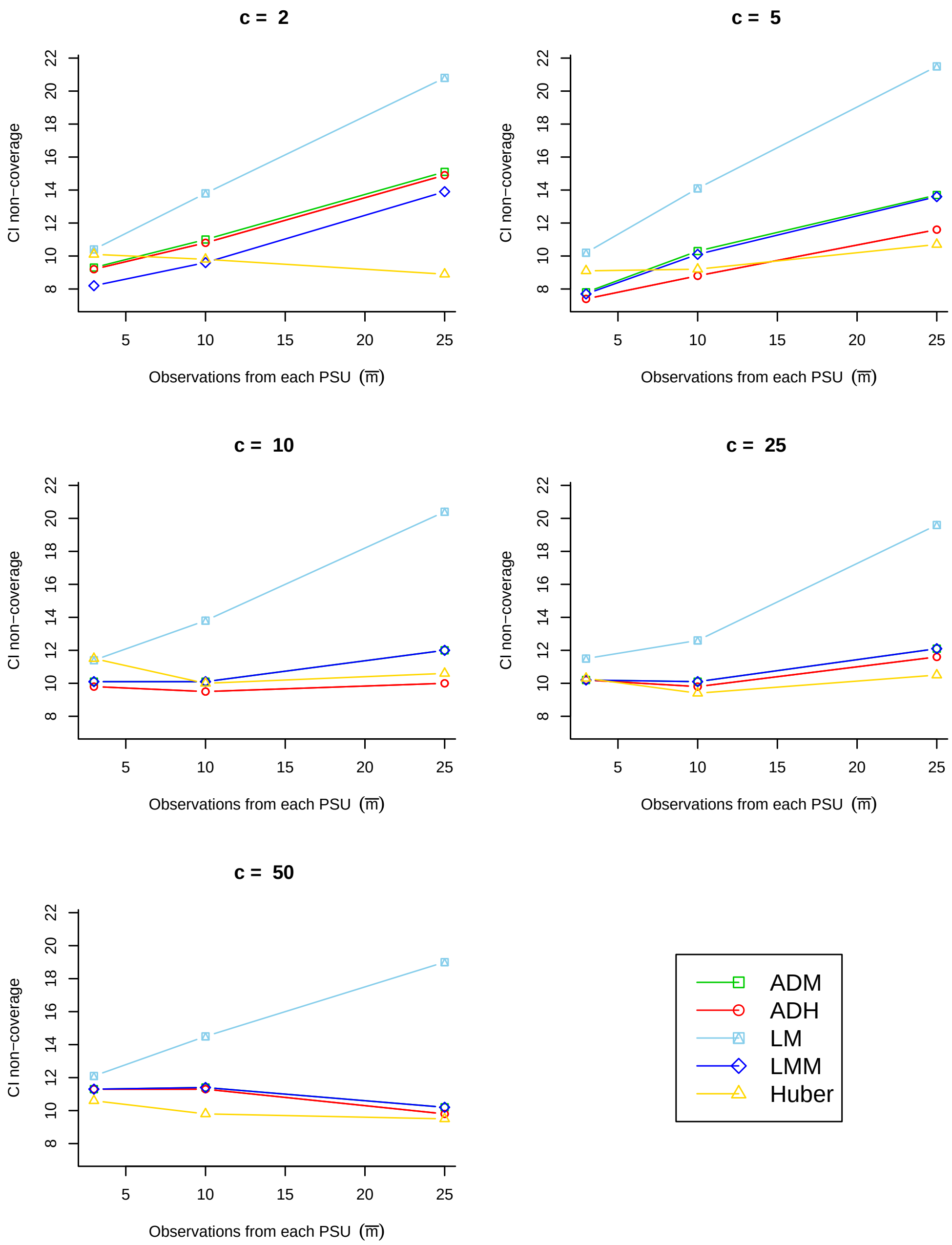


Figure D.10: Confidence interval lengths using different variance estimation methods and for various values of $\bar{m}$ and c, ρ=0.025, $deff = 1.05$

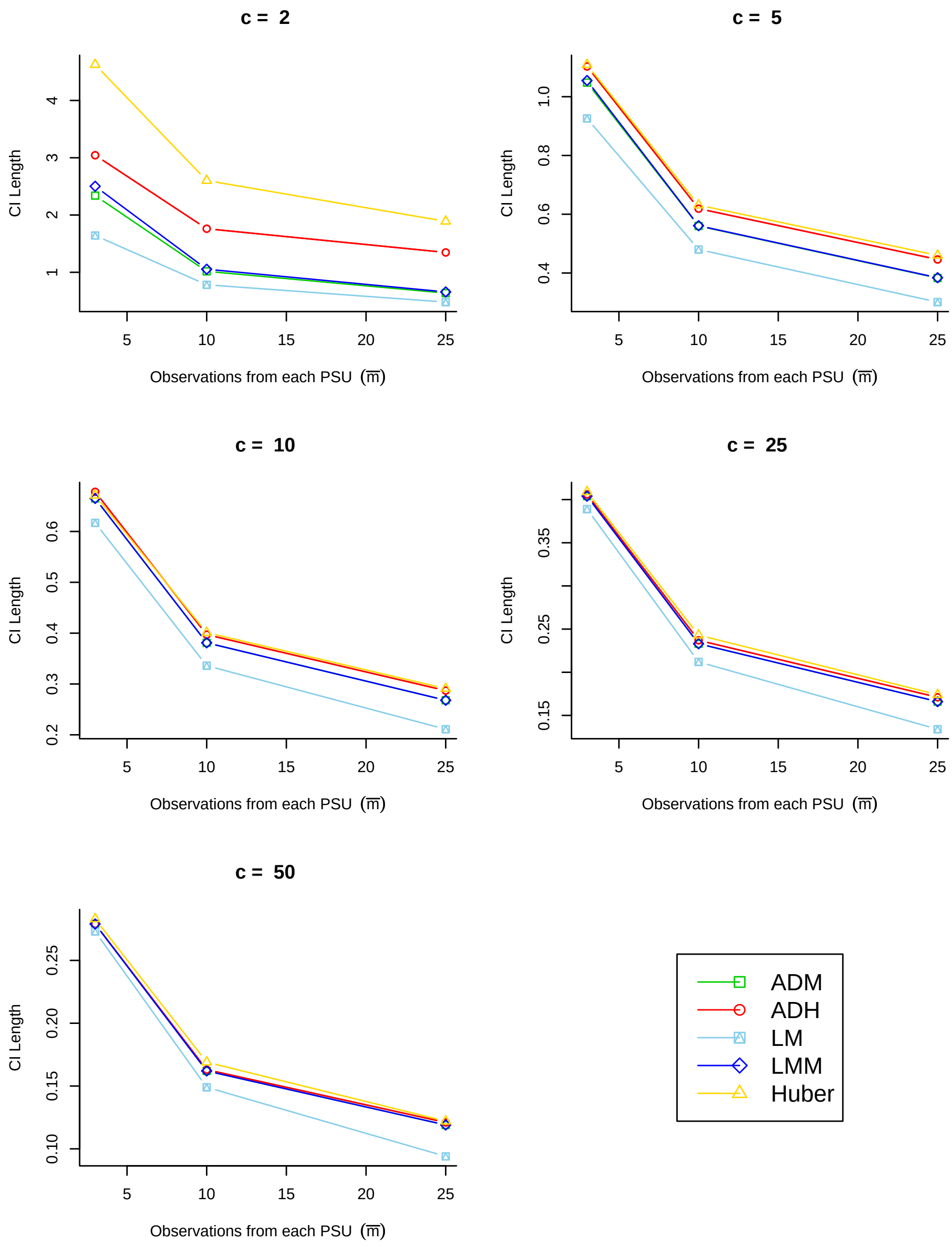


Table D.14: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.1 using RLRT in the unbalanced data case with ρ=0

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.1| Rej $H_0]$	$p[\widehat{deff}$ > 1.1 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.380	1.380	1.493	1.099	8.3	8.2	7.6	10.8	19.1	100.0	19.1	2.254	2.916	2.402	4.385
2	10	1.391	1.391	1.460	1.018	8.9	8.9	8.1	10.2	15.9	100.0	15.9	0.938	1.512	0.966	2.332
2	25	1.310	1.310	1.401	0.953	8.2	8.1	6.9	9.9	13.5	100.0	13.5	0.542	0.873	0.562	1.433
5	3	1.240	1.240	1.253	1.037	8.2	7.7	7.9	10.9	27.1	100.0	27.1	1.021	1.069	1.027	1.066
5	10	1.213	1.213	1.217	0.972	8.2	7.7	8.2	8.9	24.8	100.0	24.8	0.526	0.567	0.527	0.568
5	25	1.184	1.184	1.189	0.973	8.1	7.6	8.1	9.3	23.0	100.0	23.0	0.326	0.353	0.327	0.365
10	3	1.164	1.164	1.165	1.038	8.7	8.4	8.7	10.2	25.5	100.0	25.5	0.648	0.660	0.649	0.648
10	10	1.129	1.129	1.129	0.973	8.4	8.0	8.4	10.1	21.3	100.0	21.3	0.354	0.363	0.354	0.354
10	25	1.206	1.206	1.206	1.053	8.1	8.1	8.1	10.1	23.2	100.0	23.2	0.223	0.230	0.223	0.226
25	3	1.050	1.050	1.050	1.001	10.5	10.4	10.5	10.0	15.4	100.0	15.4	0.395	0.397	0.395	0.394
25	10	1.126	1.126	1.126	1.058	8.6	8.5	8.6	8.5	13.0	100.0	13.0	0.215	0.216	0.215	0.215
25	25	1.125	1.125	1.125	1.050	8.5	8.5	8.5	10.3	11.6	100.0	11.6	0.136	0.136	0.136	0.135
50	3	0.992	0.992	0.992	0.970	10.2	10.2	10.2	10.3	6.8	100.0	6.8	0.273	0.273	0.273	0.273
50	10	1.042	1.042	1.042	1.015	9.1	9.0	9.1	9.9	6.6	100.0	6.6	0.149	0.150	0.149	0.150
50	25	1.027	1.027	1.027	0.990	9.5	9.5	9.5	9.7	10.9	100.0	10.9	0.095	0.095	0.095	0.094

Table D.15: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.1 using RLRT in the unbalanced data case with ρ=0.025

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.1| Rej $H_0]$	$p[\widehat{deff}$ > 1.1 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.343	1.343	1.454	1.109	9.3	9.2	8.2	10.1	19.7	100.0	19.7	2.339	3.044	2.502	4.625
2	10	1.291	1.291	1.365	1.031	11.0	10.8	9.6	9.8	19.8	100.0	19.8	1.017	1.761	1.054	2.602
2	25	1.141	1.141	1.195	0.982	15.1	14.9	13.9	8.9	27.0	100.0	27.0	0.637	1.347	0.657	1.890
5	3	1.220	1.220	1.233	1.046	7.8	7.4	7.7	9.1	30.9	100.0	30.9	1.048	1.103	1.055	1.109
5	10	1.096	1.097	1.102	0.950	10.3	8.8	10.1	9.2	34.8	100.0	34.8	0.560	0.619	0.561	0.631
5	25	1.008	1.009	1.011	0.939	13.7	11.6	13.6	10.7	49.1	100.0	49.1	0.383	0.446	0.384	0.460
10	3	1.137	1.137	1.138	1.037	10.1	9.8	10.1	11.5	30.1	100.0	30.1	0.665	0.678	0.665	0.671
10	10	1.069	1.070	1.069	1.007	10.1	9.5	10.1	10.0	38.4	100.0	38.4	0.381	0.397	0.381	0.401
10	25	1.028	1.028	1.028	1.007	12.0	10.0	12.0	10.6	61.0	100.0	61.0	0.268	0.287	0.268	0.291
25	3	1.011	1.011	1.011	0.991	10.2	10.2	10.2	10.3	20.4	100.0	20.4	0.404	0.406	0.404	0.409
25	10	1.057	1.057	1.057	1.075	10.1	9.8	10.1	9.4	38.9	100.0	38.9	0.233	0.237	0.233	0.243
25	25	0.973	0.973	0.973	0.992	12.1	11.6	12.1	10.5	71.9	100.0	71.9	0.166	0.171	0.166	0.174
50	3	0.952	0.952	0.952	0.964	11.3	11.3	11.3	10.6	11.9	100.0	11.9	0.279	0.279	0.279	0.283
50	10	0.936	0.936	0.936	0.991	11.4	11.3	11.4	9.8	35.8	100.0	35.8	0.162	0.163	0.162	0.169
50	25	1.008	1.008	1.008	1.018	10.2	9.8	10.2	9.5	88.6	100.0	88.6	0.119	0.121	0.119	0.122

Table D.16: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.1 using RLRT in the unbalanced data case with ρ=0.1

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.1| Rej $H_0]$	$p[\widehat{deff}$ > 1.1 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.258	1.258	1.356	1.100	10.3	10.2	8.6	10.5	24.2	100.0	24.2	2.648	3.570	2.817	5.150
2	10	1.145	1.145	1.188	1.025	16.3	15.5	15.3	9.6	32.7	100.0	32.7	1.285	2.609	1.321	3.375
2	25	1.028	1.028	1.050	0.987	22.3	20.3	21.2	9.1	45.8	100.0	45.8	0.917	2.386	0.935	2.845
5	3	1.149	1.149	1.165	1.047	10.7	10.3	10.4	9.8	39.5	100.0	39.5	1.132	1.205	1.142	1.232
5	10	0.986	0.986	0.989	0.943	13.9	11.0	13.9	10.3	59.4	100.0	59.4	0.685	0.796	0.687	0.813
5	25	0.934	0.934	0.934	0.926	18.7	12.1	18.7	11.4	81.3	100.0	81.3	0.557	0.684	0.557	0.692
10	3	1.073	1.073	1.074	1.029	10.9	10.4	10.9	9.8	44.3	100.0	44.3	0.723	0.744	0.724	0.749
10	10	1.055	1.055	1.055	1.052	12.4	10.6	12.4	10.2	76.7	100.0	76.7	0.483	0.519	0.483	0.525
10	25	0.972	0.973	0.972	0.973	14.8	11.0	14.8	10.9	94.6	100.0	94.6	0.400	0.438	0.400	0.439
25	3	0.963	0.963	0.963	0.980	10.4	10.2	10.3	9.5	43.1	100.0	43.1	0.441	0.445	0.441	0.455
25	10	1.087	1.087	1.087	1.093	10.3	9.0	10.3	8.5	91.3	100.0	91.3	0.306	0.315	0.306	0.317
25	25	0.998	0.998	0.998	0.998	10.7	10.0	10.7	10.0	99.8	100.0	99.8	0.256	0.264	0.256	0.264
50	3	0.916	0.916	0.916	0.962	12.5	12.5	12.5	10.9	38.3	100.0	38.3	0.305	0.306	0.305	0.316
50	10	0.988	0.988	0.988	0.990	11.0	10.5	11.0	10.3	98.1	100.0	98.1	0.217	0.221	0.217	0.221
50	25	1.008	1.008	1.008	1.008	9.8	9.4	9.8	9.4	100.0	100.0	100.0	0.182	0.185	0.182	0.185

Figure D.11: Confidence interval non-coverage using different variance estimation methods and for various values of $\bar{m}$ and c, ρ=0.025, $deff = 1.1$

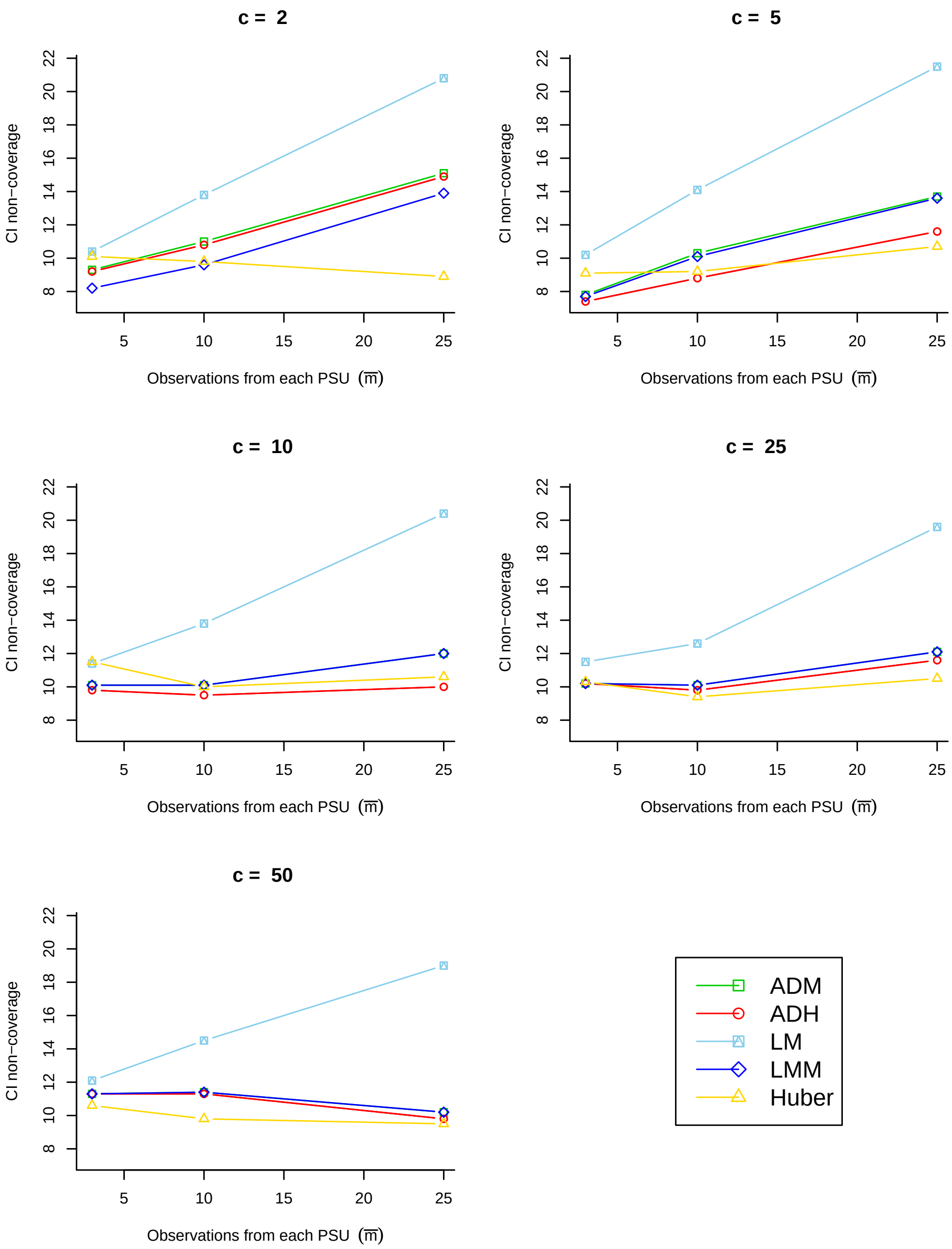


Figure D.12: Confidence interval lengths using different variance estimation methods and for various values of $\bar{m}$ and c, ρ=0.025, $deff = 1.1$

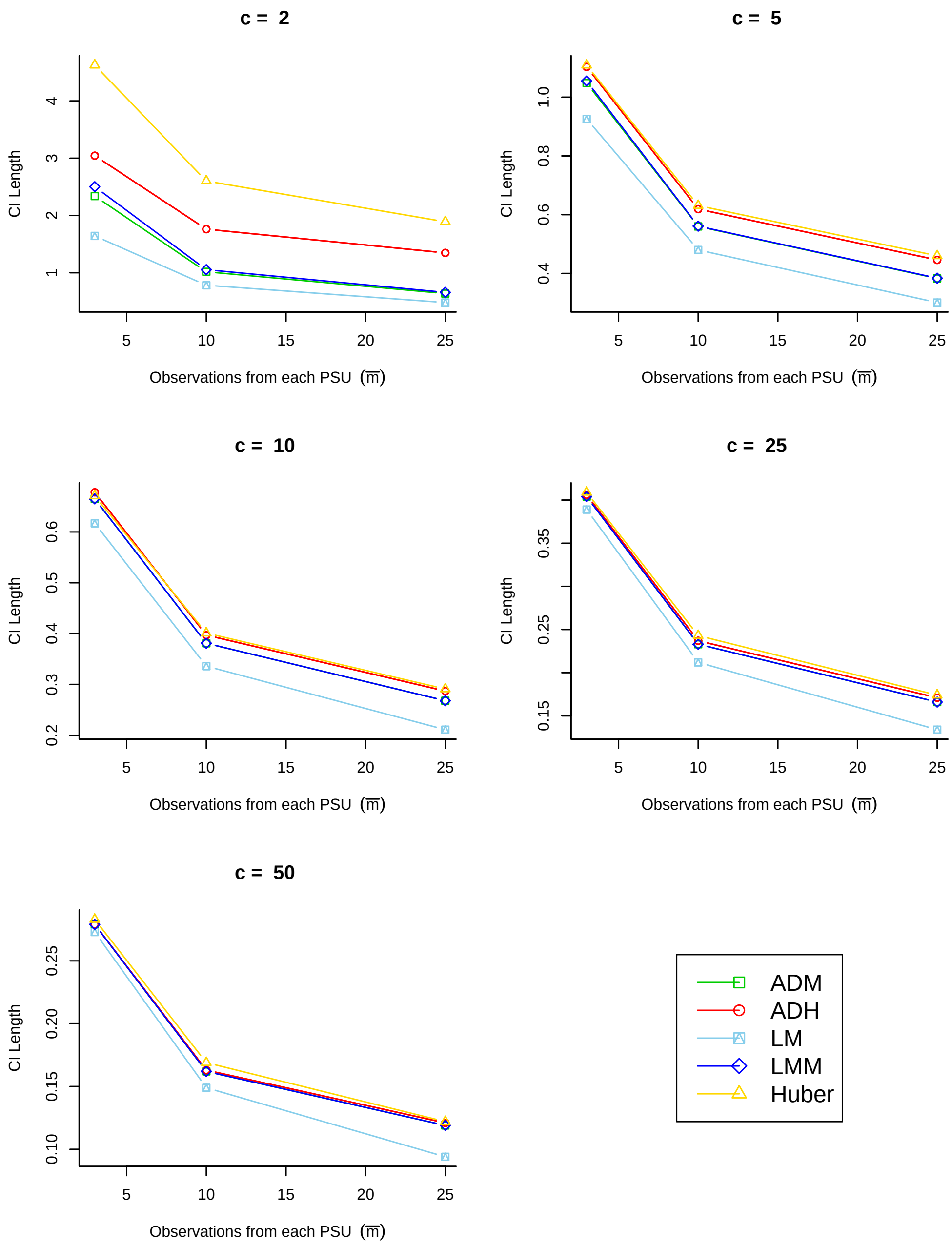


Table D.17: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.2 using RLRT in the unbalanced data case with ρ=0

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.2| Rej $H_0]$	$p[\widehat{deff}$ > 1.2 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.380	1.380	1.493	1.099	8.3	8.2	7.6	10.8	19.1	100.0	19.1	2.254	2.916	2.402	4.385
2	10	1.391	1.391	1.460	1.018	8.9	8.9	8.1	10.2	15.9	100.0	15.9	0.938	1.512	0.966	2.332
2	25	1.310	1.310	1.401	0.953	8.2	8.1	6.9	9.9	13.5	100.0	13.5	0.542	0.873	0.562	1.433
5	3	1.240	1.240	1.253	1.037	8.2	7.7	7.9	10.9	27.1	100.0	27.1	1.021	1.069	1.027	1.066
5	10	1.213	1.213	1.217	0.972	8.2	7.7	8.2	8.9	24.8	100.0	24.8	0.526	0.567	0.527	0.568
5	25	1.184	1.184	1.189	0.973	8.1	7.6	8.1	9.3	23.0	100.0	23.0	0.326	0.353	0.327	0.365
10	3	1.163	1.163	1.165	1.038	8.8	8.5	8.7	10.2	25.5	100.0	25.0	0.648	0.659	0.649	0.648
10	10	1.129	1.129	1.129	0.973	8.4	8.0	8.4	10.1	21.3	100.0	21.3	0.354	0.363	0.354	0.354
10	25	1.206	1.206	1.206	1.053	8.1	8.1	8.1	10.1	23.2	100.0	23.2	0.223	0.230	0.223	0.226
25	3	1.049	1.049	1.050	1.001	10.5	10.4	10.5	10.0	15.4	96.8	14.9	0.395	0.397	0.395	0.394
25	10	1.126	1.126	1.126	1.058	8.6	8.5	8.6	8.5	13.0	100.0	13.0	0.215	0.216	0.215	0.215
25	25	1.125	1.125	1.125	1.050	8.5	8.5	8.5	10.3	11.6	100.0	11.6	0.136	0.136	0.136	0.135
50	3	0.991	0.991	0.992	0.970	10.2	10.2	10.2	10.3	6.8	98.5	6.7	0.273	0.273	0.273	0.273
50	10	1.042	1.042	1.042	1.015	9.1	9.0	9.1	9.9	6.6	100.0	6.6	0.149	0.150	0.149	0.150
50	25	1.027	1.027	1.027	0.990	9.5	9.5	9.5	9.7	10.9	100.0	10.9	0.095	0.095	0.095	0.094

Table D.18: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.2 using RLRT in the unbalanced data case with ρ=0.025

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.2| Rej $H_0]$	$p[\widehat{deff}$ > 1.2 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.343	1.343	1.454	1.109	9.3	9.2	8.2	10.1	19.7	100.0	19.7	2.339	3.044	2.502	4.625
2	10	1.291	1.291	1.365	1.031	11.0	10.8	9.6	9.8	19.8	100.0	19.8	1.017	1.761	1.054	2.602
2	25	1.141	1.141	1.195	0.982	15.1	14.9	13.9	8.9	27.0	100.0	27.0	0.637	1.347	0.657	1.890
5	3	1.220	1.220	1.233	1.046	7.8	7.4	7.7	9.1	30.9	100.0	30.9	1.048	1.103	1.055	1.109
5	10	1.096	1.097	1.102	0.950	10.3	8.8	10.1	9.2	34.8	100.0	34.8	0.560	0.619	0.561	0.631
5	25	1.008	1.009	1.011	0.939	13.7	11.6	13.6	10.7	49.1	100.0	49.1	0.383	0.446	0.384	0.460
10	3	1.136	1.136	1.138	1.037	10.1	9.8	10.1	11.5	30.1	98.7	29.7	0.665	0.678	0.665	0.671
10	10	1.069	1.070	1.069	1.007	10.1	9.5	10.1	10.0	38.4	100.0	38.4	0.381	0.397	0.381	0.401
10	25	1.028	1.028	1.028	1.007	12.0	10.0	12.0	10.6	61.0	100.0	61.0	0.268	0.287	0.268	0.291
25	3	1.010	1.010	1.011	0.991	10.2	10.2	10.2	10.3	20.4	97.1	19.8	0.404	0.406	0.404	0.409
25	10	1.057	1.057	1.057	1.075	10.1	9.8	10.1	9.4	38.9	100.0	38.9	0.233	0.237	0.233	0.243
25	25	0.973	0.973	0.973	0.992	12.1	11.6	12.1	10.5	71.9	100.0	71.9	0.166	0.171	0.166	0.174
50	3	0.952	0.952	0.952	0.964	11.3	11.3	11.3	10.6	11.9	99.2	11.8	0.279	0.279	0.279	0.283
50	10	0.936	0.936	0.936	0.991	11.4	11.3	11.4	9.8	35.8	100.0	35.8	0.162	0.163	0.162	0.169
50	25	1.008	1.008	1.008	1.018	10.2	9.8	10.2	9.5	88.6	100.0	88.6	0.119	0.121	0.119	0.122

Table D.19: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.2 using RLRT in the unbalanced data case with ρ=0.1

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.2| Rej $H_0]$	$p[\widehat{deff}$ > 1.2 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.258	1.258	1.356	1.100	10.3	10.2	8.6	10.5	24.2	100.0	24.2	2.648	3.570	2.817	5.150
2	10	1.145	1.145	1.188	1.025	16.3	15.5	15.3	9.6	32.7	100.0	32.7	1.285	2.609	1.321	3.375
2	25	1.028	1.028	1.050	0.987	22.3	20.3	21.2	9.1	45.8	100.0	45.8	0.917	2.386	0.935	2.845
5	3	1.149	1.149	1.165	1.047	10.7	10.3	10.4	9.8	39.5	100.0	39.5	1.132	1.205	1.142	1.232
5	10	0.986	0.986	0.989	0.943	13.9	11.0	13.9	10.3	59.4	100.0	59.4	0.685	0.796	0.687	0.813
5	25	0.934	0.934	0.934	0.926	18.7	12.1	18.7	11.4	81.3	100.0	81.3	0.557	0.684	0.557	0.692
10	3	1.073	1.073	1.074	1.029	10.9	10.4	10.9	9.8	44.3	100.0	44.3	0.723	0.744	0.724	0.749
10	10	1.055	1.055	1.055	1.052	12.4	10.6	12.4	10.2	76.7	100.0	76.7	0.483	0.519	0.483	0.525
10	25	0.972	0.973	0.972	0.973	14.8	11.0	14.8	10.9	94.6	100.0	94.6	0.400	0.438	0.400	0.439
25	3	0.962	0.962	0.963	0.980	10.4	10.2	10.3	9.5	43.1	98.6	42.5	0.441	0.445	0.441	0.455
25	10	1.087	1.087	1.087	1.093	10.3	9.0	10.3	8.5	91.3	100.0	91.3	0.306	0.315	0.306	0.317
25	25	0.998	0.998	0.998	0.998	10.7	10.0	10.7	10.0	99.8	100.0	99.8	0.256	0.264	0.256	0.264
50	3	0.915	0.915	0.916	0.962	12.5	12.5	12.5	10.9	38.3	99.7	38.2	0.305	0.306	0.305	0.316
50	10	0.988	0.988	0.988	0.990	11.0	10.5	11.0	10.3	98.1	100.0	98.1	0.217	0.221	0.217	0.221
50	25	1.008	1.008	1.008	1.008	9.8	9.4	9.8	9.4	100.0	100.0	100.0	0.182	0.185	0.182	0.185

Figure D.13: Confidence interval non-coverage using different variance estimation methods and for various values of $\bar{m}$ and c, ρ=0.025, $deff = 1.2$

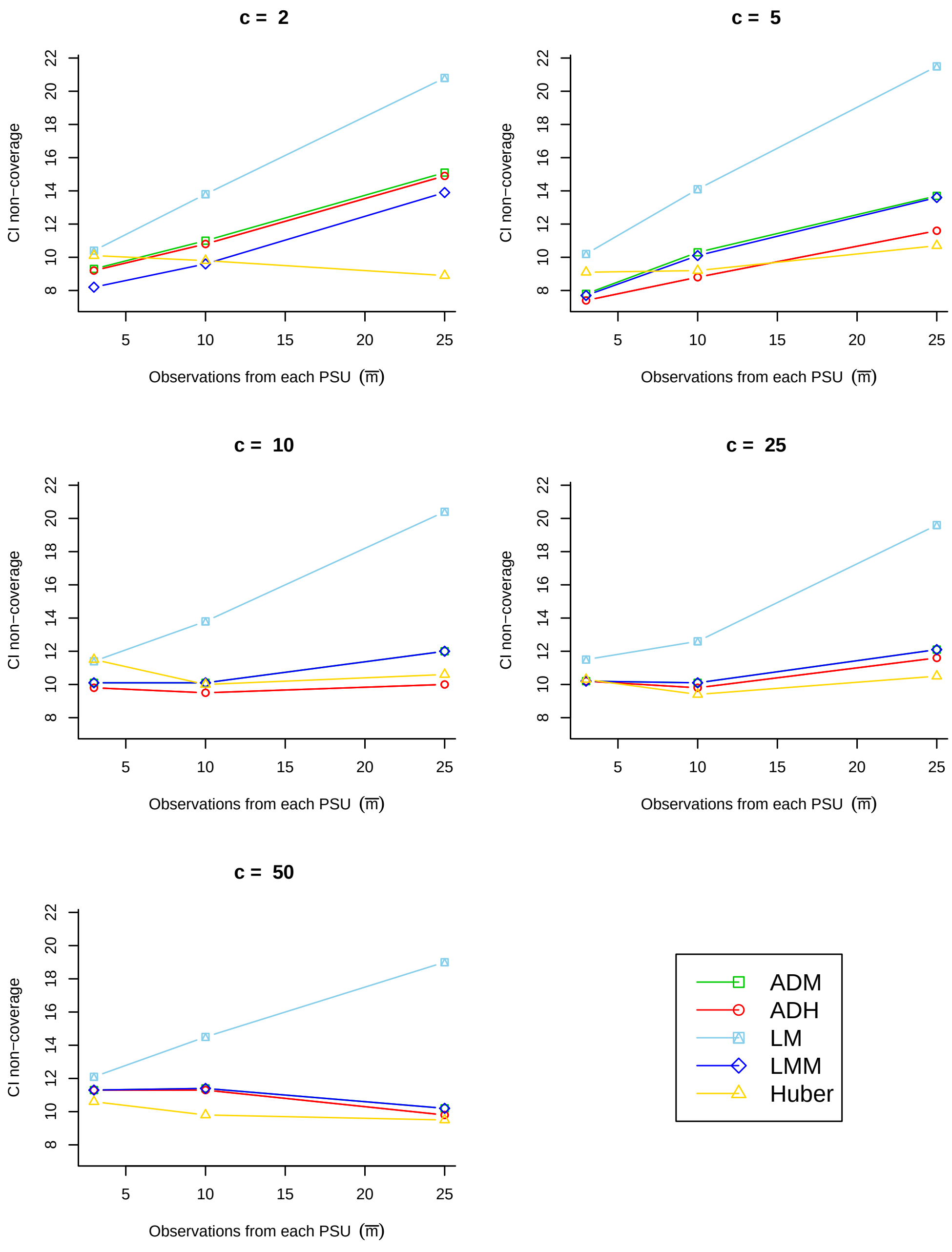


Figure D.14: Confidence interval lengths using different variance estimation methods and for various values of $\bar{m}$ and c, ρ=0.025, $deff = 1.2$

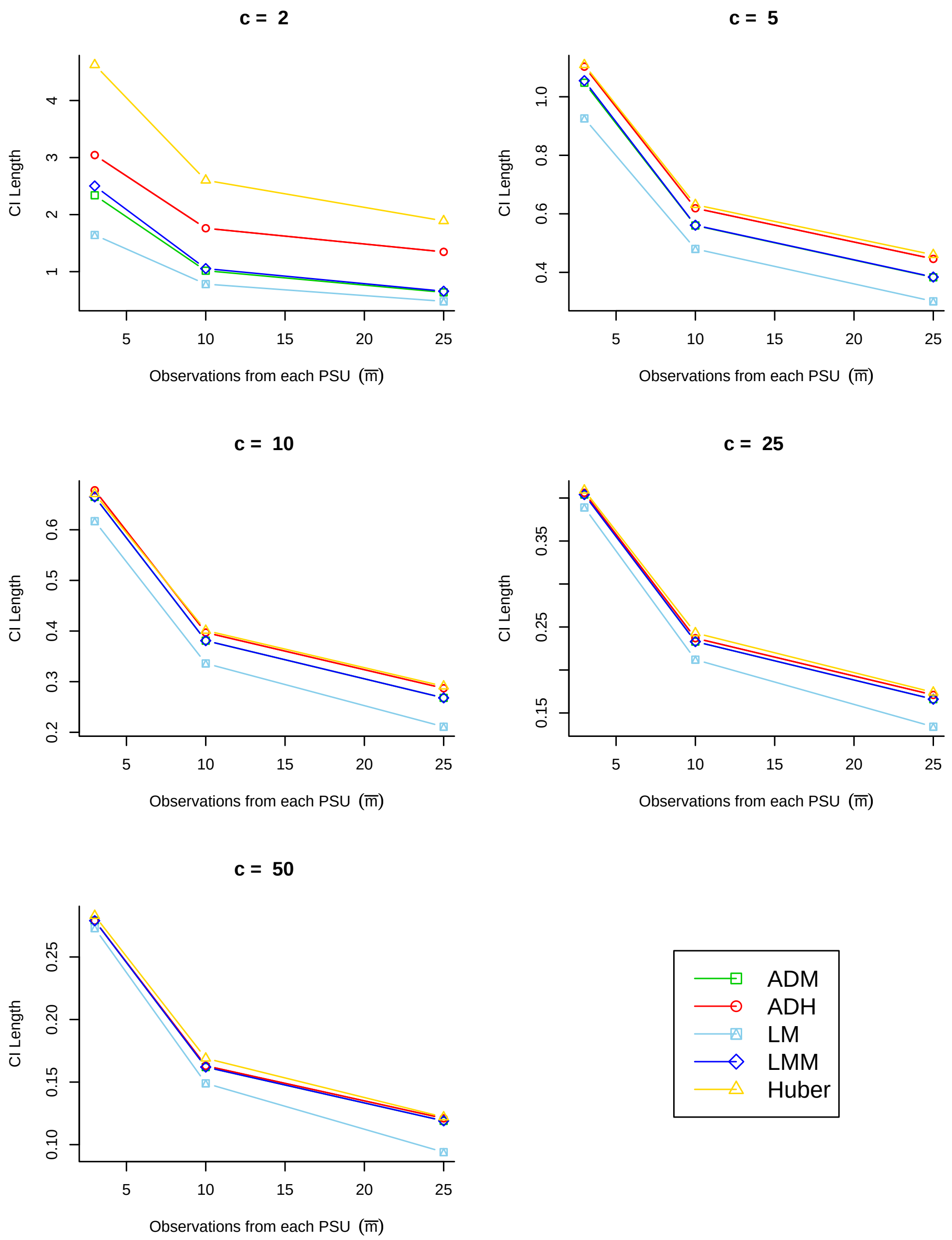


Table D.20: Variance ratios, length and non-coverage of the 90% confidence intervals for β, and power of testing $H_0 : \sigma_b^2 = 0$ and deff ≥ 1.5 using RLRT in the unbalanced data case with ρ=0.1

PSUs	Obs	$E(\widehat{var}(\hat{\beta}))/var(\hat{\beta})$				Non-Coverage of CI for β				RLRT	$p[\widehat{deff}$ >1.5| Rej $H_0]$	$p[\widehat{deff}$ > 1.5 & Rej $H_0]$	Confidence Interval Length			
c	m	ADM	ADH	LMM	Hub	ADM	ADH	LMM	Hub				ADM	ADH	LMM	Hub
2	3	1.258	1.258	1.356	1.100	10.3	10.2	8.6	10.5	24.2	100.0	24.2	2.648	3.570	2.817	5.150
2	10	1.145	1.145	1.188	1.025	16.3	15.5	15.3	9.6	32.7	100.0	32.7	1.285	2.609	1.321	3.375
2	25	1.028	1.028	1.050	0.987	22.3	20.3	21.2	9.1	45.8	100.0	45.8	0.917	2.386	0.935	2.845
5	3	1.112	1.112	1.165	1.047	11.0	10.8	10.4	9.8	39.5	75.4	29.8	1.109	1.163	1.142	1.232
5	10	0.977	0.978	0.989	0.943	14.1	11.5	13.9	10.3	59.4	93.3	55.4	0.681	0.786	0.687	0.813
5	25	0.932	0.932	0.934	0.926	18.8	12.5	18.7	11.4	81.3	97.8	79.5	0.556	0.681	0.557	0.692
10	3	1.004	1.004	1.074	1.029	12.0	12.0	10.9	9.8	44.3	53.5	23.7	0.698	0.708	0.724	0.749
10	10	1.034	1.034	1.055	1.052	13.4	11.7	12.4	10.2	76.7	87.7	67.3	0.477	0.509	0.483	0.525
10	25	0.970	0.970	0.972	0.973	15.2	11.5	14.8	10.9	94.6	97.6	92.3	0.399	0.436	0.400	0.439
25	3	0.873	0.873	0.963	0.980	12.0	12.0	10.3	9.5	43.1	30.1	13.1	0.420	0.421	0.441	0.455
25	10	1.066	1.067	1.087	1.093	10.9	9.6	10.3	8.5	91.3	90.5	82.6	0.302	0.311	0.306	0.317
25	25	0.998	0.998	0.998	0.998	10.7	10.0	10.7	10.0	99.8	99.9	99.7	0.255	0.264	0.256	0.264
50	3	0.820	0.820	0.916	0.962	13.9	13.9	12.5	10.9	38.3	15.1	5.8	0.289	0.289	0.305	0.316
50	10	0.974	0.974	0.988	0.990	11.6	11.1	11.0	10.3	98.1	93.0	91.2	0.215	0.218	0.217	0.221
50	25	1.008	1.008	1.008	1.008	9.8	9.4	9.8	9.4	100.0	100.0	100.0	0.182	0.185	0.182	0.185

Appendix E

Extra Tables for Chapter 6

Table E.1: Variance of $\hat{\beta}$, $(\times 10^3)$, calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=0.5 and C_2=1. ρ=0 and 0.01)

Pilot		True Variance of $(\hat{\beta})$ for ρ=0						True Variance of $(\hat{\beta})$ for ρ=0.01					
PSUs	Obs	Cutoff for Within-PSU Sample Size (A)						Cutoff for Within-PSU Sample Size (A)					
c_p	m_p	10	20	30	40	50	100	10	20	30	40	50	100
2	2	0.223	0.221	0.235	0.229	0.230	0.230	0.253	0.263	0.262	0.271	0.290	0.335
2	5	0.230	0.222	0.220	0.217	0.218	0.223	0.247	0.262	0.270	0.273	0.290	0.371
2	10	0.229	0.213	0.219	0.224	0.218	0.218	0.238	0.261	0.260	0.265	0.301	0.353
2	15	0.231	0.233	0.217	0.209	0.222	0.215	0.254	0.254	0.271	0.272	0.308	0.355
2	25	0.219	0.218	0.225	0.216	0.225	0.214	0.237	0.251	0.255	0.294	0.290	0.357
2	50	0.223	0.219	0.210	0.222	0.224	0.217	0.233	0.241	0.264	0.275	0.279	0.368
5	2	0.248	0.237	0.231	0.226	0.230	0.232	0.253	0.259	0.275	0.270	0.260	0.339
5	5	0.231	0.230	0.231	0.226	0.218	0.221	0.249	0.256	0.262	0.278	0.296	0.356
5	10	0.229	0.216	0.230	0.223	0.225	0.219	0.252	0.249	0.264	0.281	0.271	0.350
5	15	0.223	0.212	0.223	0.218	0.210	0.207	0.230	0.250	0.269	0.293	0.291	0.346
5	25	0.228	0.219	0.210	0.213	0.230	0.217	0.250	0.255	0.263	0.272	0.302	0.360
5	50	0.221	0.207	0.205	0.211	0.203	0.215	0.240	0.239	0.260	0.262	0.265	0.310
10	2	0.231	0.239	0.228	0.227	0.235	0.233	0.261	0.245	0.264	0.262	0.265	0.324
10	5	0.228	0.217	0.232	0.227	0.221	0.224	0.254	0.252	0.265	0.296	0.298	0.372
10	10	0.220	0.217	0.206	0.217	0.205	0.213	0.230	0.245	0.265	0.271	0.282	0.370
10	15	0.210	0.216	0.214	0.213	0.204	0.219	0.241	0.257	0.275	0.274	0.292	0.335
10	25	0.213	0.218	0.221	0.218	0.215	0.214	0.252	0.242	0.257	0.277	0.298	0.341
10	50	0.217	0.203	0.214	0.219	0.211	0.208	0.238	0.244	0.255	0.270	0.272	0.296
25	2	0.231	0.234	0.229	0.235	0.223	0.227	0.261	0.263	0.264	0.282	0.279	0.311
25	5	0.207	0.225	0.214	0.217	0.211	0.210	0.244	0.251	0.269	0.298	0.304	0.396
25	10	0.218	0.222	0.214	0.211	0.209	0.207	0.246	0.236	0.260	0.296	0.294	0.371
25	15	0.226	0.216	0.204	0.218	0.208	0.207	0.233	0.249	0.249	0.272	0.289	0.361
25	25	0.218	0.207	0.215	0.211	0.208	0.198	0.235	0.251	0.265	0.279	0.308	0.354
25	50	0.233	0.227	0.206	0.225	0.213	0.221	0.242	0.241	0.258	0.253	0.256	0.278

Figure E.1: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=0.5 and C_2=1, ρ=0)

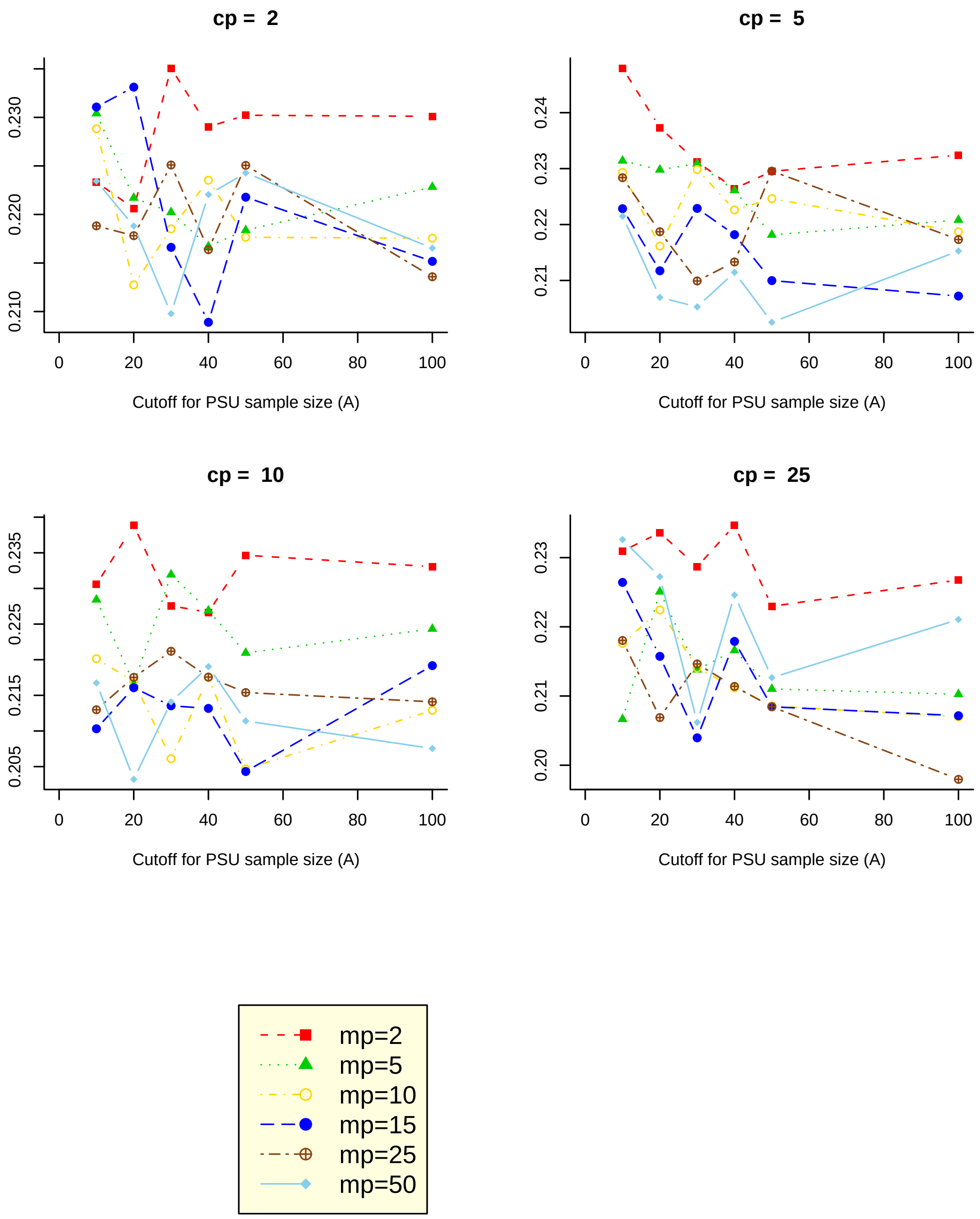


Figure E.2: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=0.5 and C_2=1, ρ=0.01)

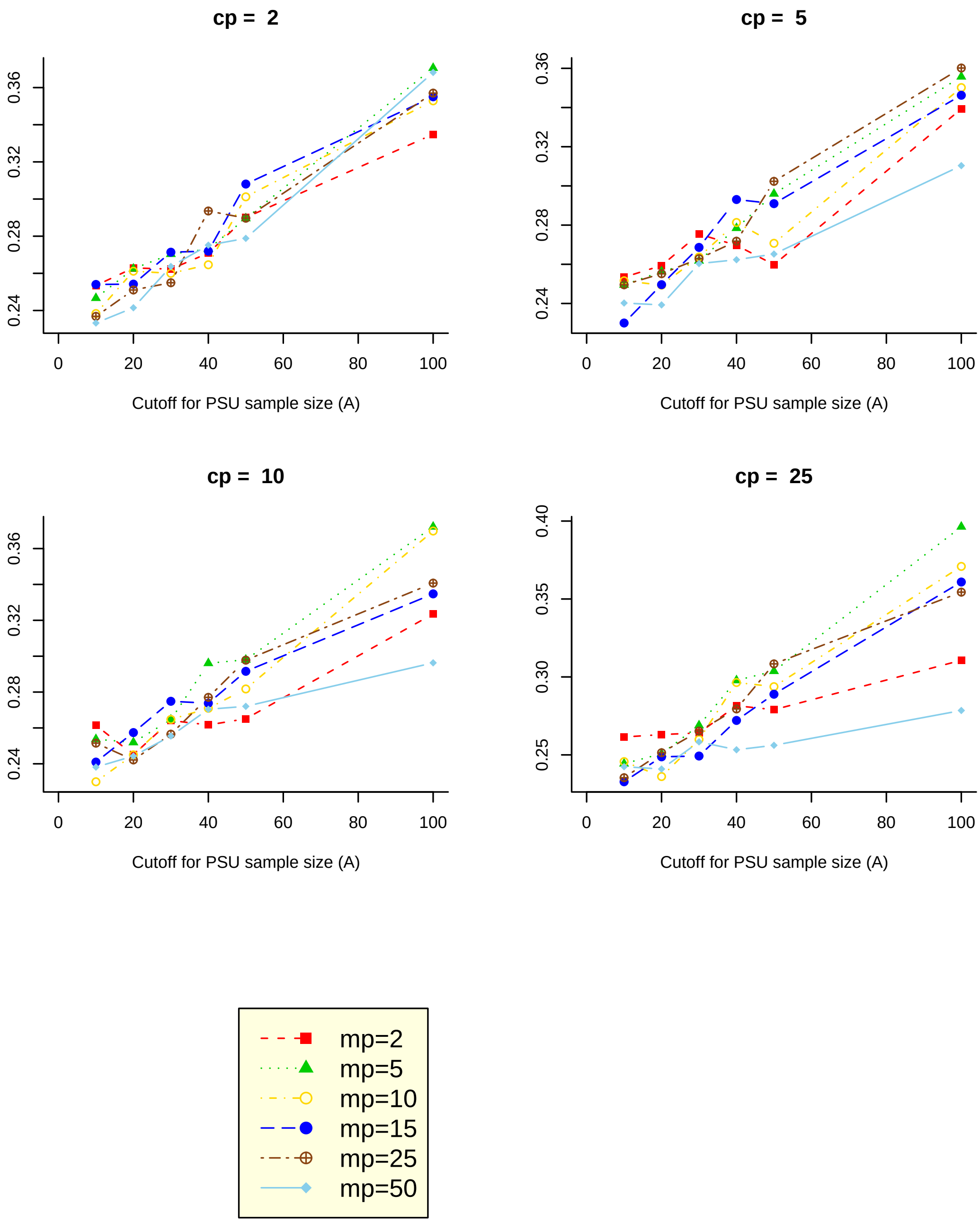


Table E.2: Variance of $\hat{\beta}$, ($\times 10^3$), calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=0.5 and C_2=1. ρ=0.025 and 0.05)

Pilot		True Variance of ($\hat{\beta}$) for ρ=0.025						True Variance of ($\hat{\beta}$) for ρ=0.05					
PSUs	Obs	Cutoff for Within-PSU Sample Size (A)						Cutoff for Within-PSU Sample Size (A)					
c_p	m_p	10	20	30	40	50	100	10	20	30	40	50	100
2	2	0.273	0.305	0.313	0.361	0.357	0.522	0.319	0.347	0.412	0.467	0.524	0.829
2	5	0.270	0.323	0.327	0.371	0.416	0.587	0.316	0.405	0.428	0.561	0.596	0.952
2	10	0.262	0.307	0.346	0.360	0.422	0.573	0.306	0.356	0.447	0.506	0.522	0.868
2	15	0.287	0.293	0.329	0.348	0.399	0.546	0.312	0.366	0.433	0.509	0.583	0.841
2	25	0.253	0.291	0.320	0.349	0.398	0.575	0.322	0.355	0.414	0.462	0.542	0.769
2	50	0.257	0.284	0.314	0.354	0.394	0.533	0.320	0.352	0.402	0.443	0.473	0.802
5	2	0.281	0.306	0.291	0.352	0.377	0.455	0.313	0.350	0.396	0.440	0.516	0.645
5	5	0.274	0.286	0.328	0.376	0.390	0.571	0.318	0.366	0.450	0.517	0.556	0.906
5	10	0.272	0.298	0.323	0.370	0.393	0.565	0.302	0.360	0.404	0.462	0.522	0.785
5	15	0.277	0.292	0.326	0.346	0.374	0.509	0.304	0.341	0.391	0.455	0.468	0.750
5	25	0.254	0.280	0.306	0.334	0.355	0.501	0.288	0.343	0.347	0.402	0.379	0.634
5	50	0.254	0.274	0.279	0.304	0.332	0.424	0.301	0.304	0.337	0.331	0.335	0.442
10	2	0.275	0.301	0.318	0.337	0.362	0.485	0.302	0.361	0.365	0.417	0.496	0.722
10	5	0.276	0.287	0.333	0.398	0.433	0.578	0.306	0.359	0.452	0.482	0.557	0.848
10	10	0.265	0.286	0.316	0.380	0.397	0.546	0.311	0.352	0.434	0.437	0.515	0.744
10	15	0.274	0.285	0.276	0.330	0.366	0.477	0.294	0.361	0.378	0.380	0.424	0.595
10	25	0.270	0.277	0.286	0.315	0.349	0.443	0.308	0.305	0.304	0.358	0.411	0.454
10	50	0.261	0.280	0.279	0.301	0.301	0.330	0.290	0.281	0.297	0.298	0.302	0.293
25	2	0.278	0.284	0.314	0.333	0.354	0.478	0.297	0.365	0.372	0.434	0.451	0.713
25	5	0.283	0.294	0.357	0.361	0.415	0.628	0.327	0.351	0.475	0.490	0.567	0.913
25	10	0.272	0.293	0.339	0.369	0.385	0.467	0.303	0.338	0.358	0.432	0.492	0.689
25	15	0.269	0.286	0.299	0.345	0.365	0.484	0.280	0.320	0.340	0.368	0.383	0.446
25	25	0.263	0.274	0.279	0.296	0.331	0.407	0.278	0.283	0.312	0.286	0.316	0.317
25	50	0.272	0.258	0.266	0.249	0.262	0.267	0.281	0.276	0.272	0.272	0.277	0.282

Figure E.3: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=0.5 and C_2=1, ρ=0.025)

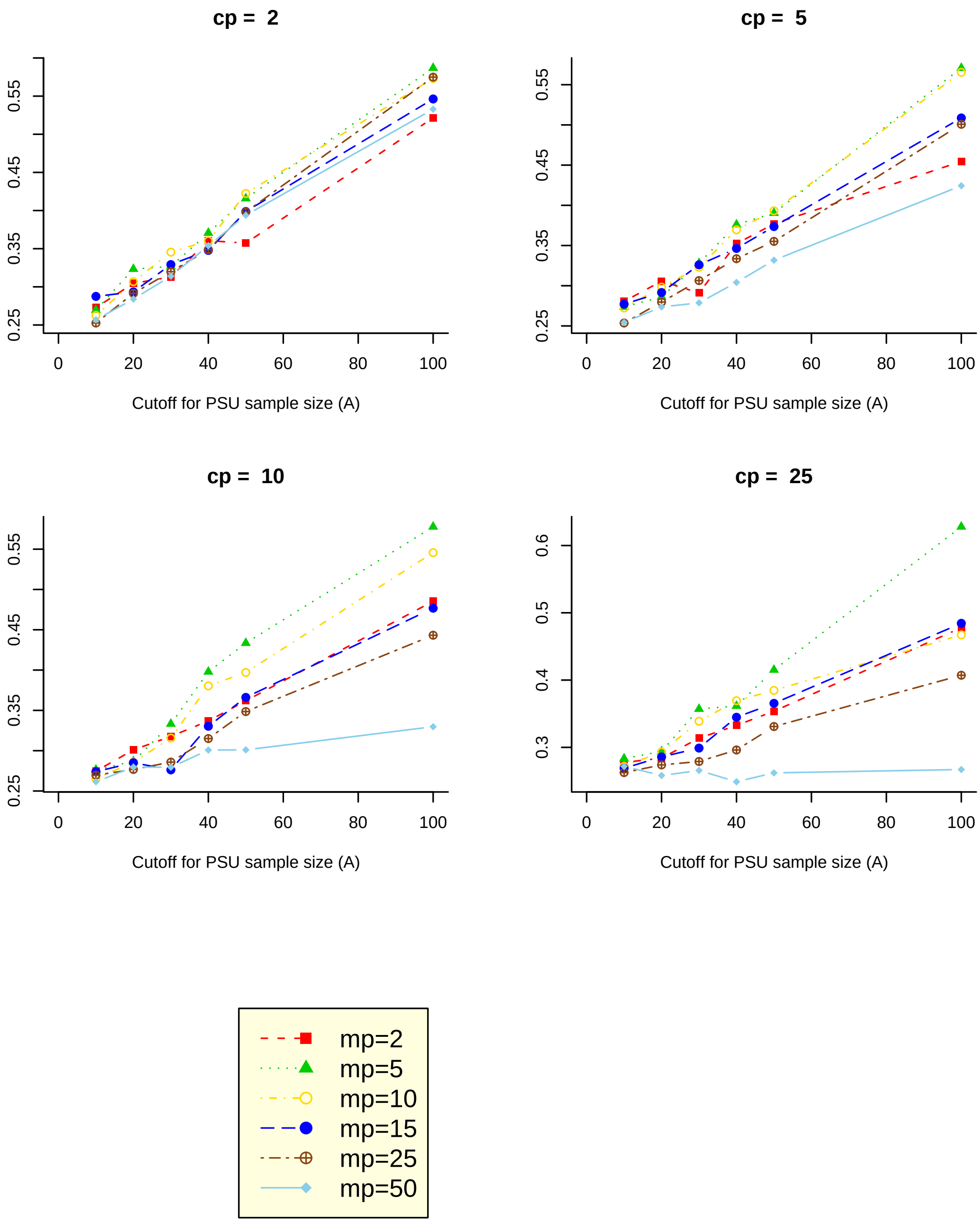


Figure E.4: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=0.5 and C_2=1, ρ=0.05)

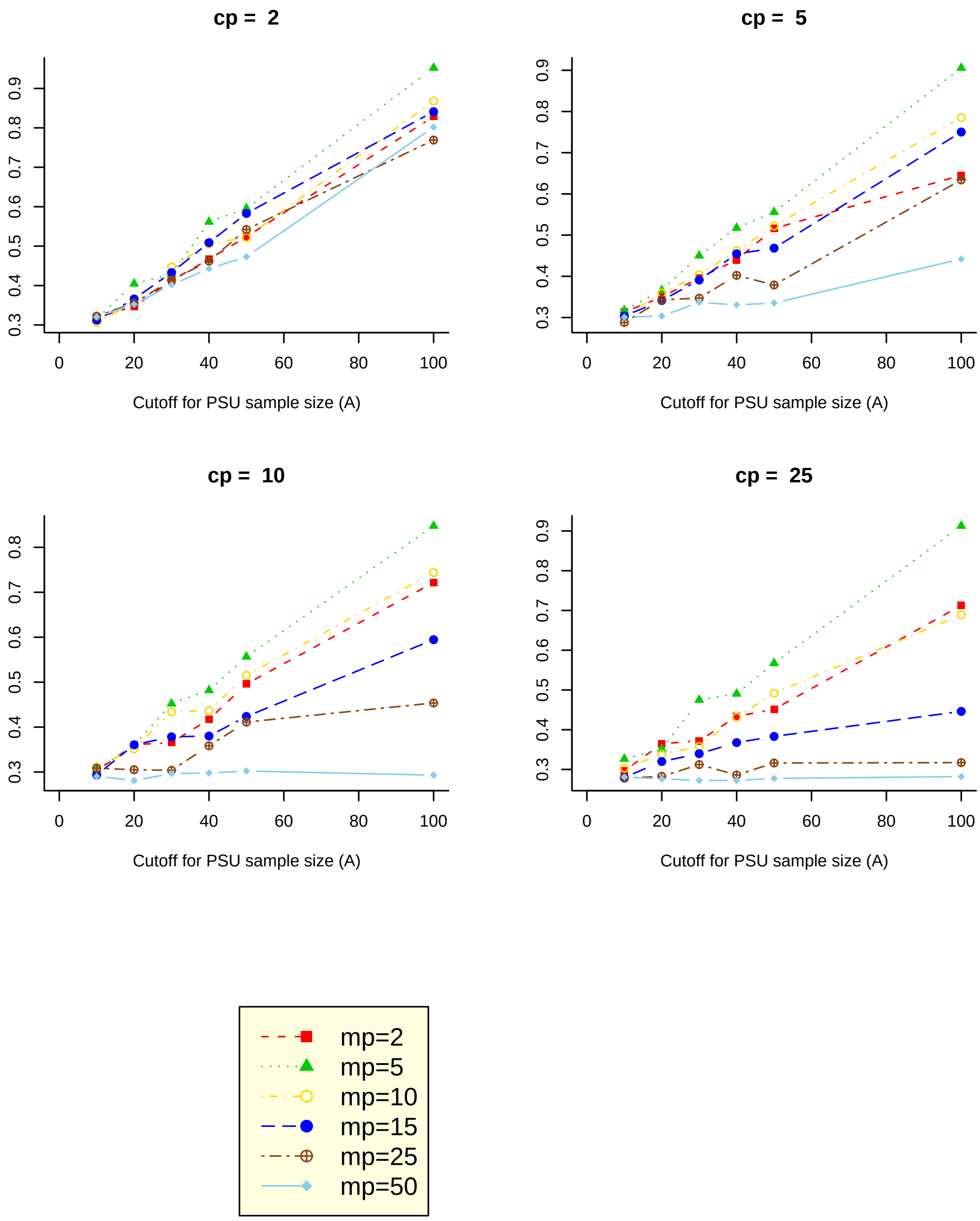


Table E.3: Variance of $\hat{\beta}$, $(\times 10^3)$, calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=0.5 and C_2=1. ρ=0.1)

Pilot		True Variance of $(\hat{\beta})$ for ρ=0.1					
PSUs	Obs	Cutoff for Within-PSU Sample Size (A)					
c_p	m_p	10	20	30	40	50	100
2	2	0.375	0.500	0.656	0.708	0.887	1.467
2	5	0.405	0.575	0.620	0.705	0.937	1.616
2	10	0.376	0.485	0.641	0.729	0.861	1.591
2	15	0.398	0.468	0.590	0.736	0.760	1.546
2	25	0.392	0.480	0.569	0.570	0.702	1.227
2	50	0.372	0.452	0.525	0.593	0.597	1.101
5	2	0.378	0.435	0.556	0.663	0.658	1.382
5	5	0.392	0.498	0.580	0.672	0.829	1.631
5	10	0.360	0.439	0.571	0.572	0.677	1.245
5	15	0.362	0.413	0.503	0.529	0.575	0.977
5	25	0.332	0.392	0.424	0.479	0.485	0.673
5	50	0.338	0.352	0.356	0.385	0.393	0.496
10	2	0.367	0.435	0.503	0.626	0.740	1.173
10	5	0.391	0.496	0.587	0.715	0.753	1.309
10	10	0.355	0.427	0.446	0.483	0.575	0.794
10	15	0.347	0.352	0.404	0.442	0.436	0.604
10	25	0.324	0.346	0.347	0.339	0.343	0.422
10	50	0.310	0.310	0.338	0.304	0.323	0.308
25	2	0.336	0.454	0.472	0.561	0.600	1.092
25	5	0.371	0.479	0.565	0.662	0.726	1.102
25	10	0.306	0.343	0.352	0.418	0.391	0.539
25	15	0.319	0.337	0.330	0.352	0.361	0.363
25	25	0.311	0.311	0.311	0.311	0.311	0.311
25	50	0.305	0.305	0.305	0.305	0.305	0.305

Figure E.5: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=0.5 and C_2=1, ρ=0.1)

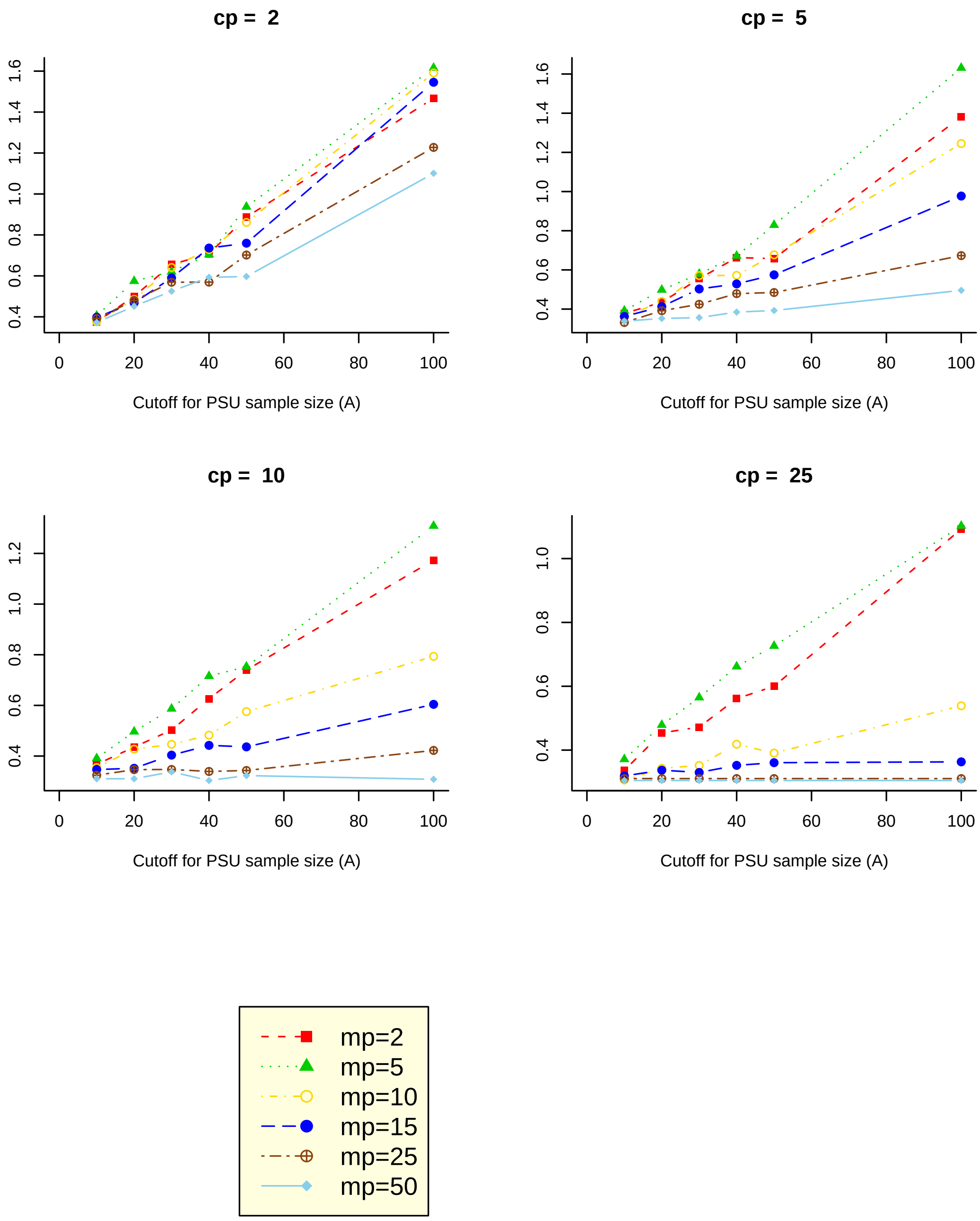


Table E.4: Variance of $\hat{\beta}$, ($\times 10^3$), calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=2 and C_2=1. ρ=0 and 0.01)

Pilot		True Variance of $(\hat{\beta})$ for ρ=0						True Variance of $(\hat{\beta})$ for ρ=0.01					
PSUs	Obs	Cutoff for Within-PSU Sample Size (A)						Cutoff for Within-PSU Sample Size (A)					
c_p	m_p	10	20	30	40	50	100	10	20	30	40	50	100
2	2	0.297	0.301	0.280	0.292	0.274	0.297	0.337	0.320	0.322	0.348	0.354	0.426
2	5	0.280	0.246	0.268	0.263	0.255	0.260	0.309	0.309	0.295	0.326	0.344	0.398
2	10	0.271	0.249	0.241	0.249	0.248	0.239	0.294	0.277	0.305	0.299	0.336	0.392
2	15	0.267	0.241	0.242	0.234	0.244	0.242	0.272	0.276	0.283	0.313	0.325	0.396
2	25	0.257	0.235	0.235	0.225	0.237	0.245	0.278	0.288	0.287	0.309	0.324	0.384
2	50	0.257	0.230	0.235	0.224	0.227	0.222	0.284	0.265	0.271	0.293	0.301	0.380
5	2	0.296	0.287	0.297	0.308	0.293	0.283	0.325	0.319	0.330	0.331	0.332	0.396
5	5	0.277	0.255	0.247	0.255	0.239	0.242	0.295	0.282	0.307	0.295	0.320	0.371
5	10	0.244	0.240	0.246	0.232	0.240	0.234	0.272	0.273	0.291	0.299	0.307	0.370
5	15	0.260	0.241	0.235	0.224	0.234	0.242	0.283	0.279	0.292	0.298	0.321	0.381
5	25	0.258	0.224	0.231	0.228	0.231	0.227	0.281	0.286	0.295	0.289	0.312	0.361
5	50	0.239	0.235	0.219	0.221	0.223	0.212	0.277	0.267	0.278	0.281	0.299	0.345
10	2	0.306	0.303	0.274	0.286	0.299	0.292	0.323	0.318	0.326	0.322	0.329	0.391
10	5	0.288	0.248	0.247	0.243	0.234	0.231	0.294	0.290	0.287	0.303	0.318	0.380
10	10	0.246	0.236	0.234	0.222	0.240	0.223	0.281	0.281	0.284	0.300	0.314	0.374
10	15	0.259	0.244	0.227	0.222	0.225	0.221	0.272	0.281	0.281	0.298	0.323	0.360
10	25	0.250	0.240	0.226	0.224	0.219	0.219	0.276	0.284	0.265	0.291	0.301	0.358
10	50	0.240	0.223	0.214	0.220	0.212	0.221	0.273	0.270	0.285	0.275	0.288	0.343
25	2	0.285	0.279	0.265	0.261	0.269	0.276	0.290	0.299	0.314	0.322	0.336	0.392
25	5	0.249	0.236	0.226	0.219	0.217	0.234	0.267	0.253	0.298	0.312	0.308	0.418
25	10	0.259	0.245	0.228	0.229	0.223	0.231	0.285	0.277	0.284	0.305	0.309	0.406
25	15	0.251	0.222	0.225	0.230	0.224	0.209	0.269	0.267	0.296	0.301	0.304	0.417
25	25	0.247	0.223	0.239	0.225	0.228	0.232	0.291	0.276	0.286	0.282	0.294	0.365
25	50	0.247	0.228	0.232	0.216	0.221	0.226	0.270	0.288	0.272	0.274	0.293	0.322

Figure E.6: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=2 and C_2=1, ρ=0)

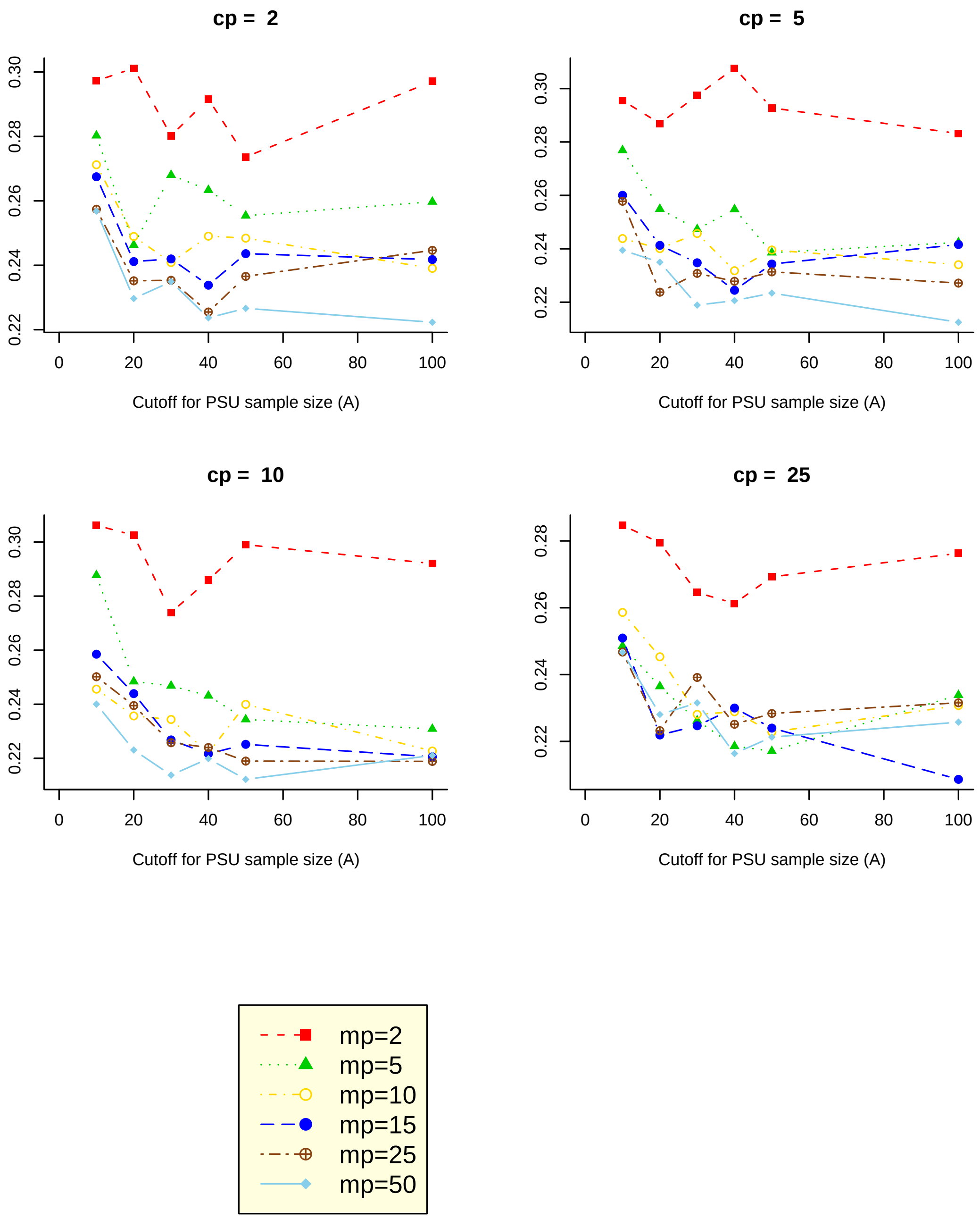


Figure E.7: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=2 and C_2=1, ρ=0.01)

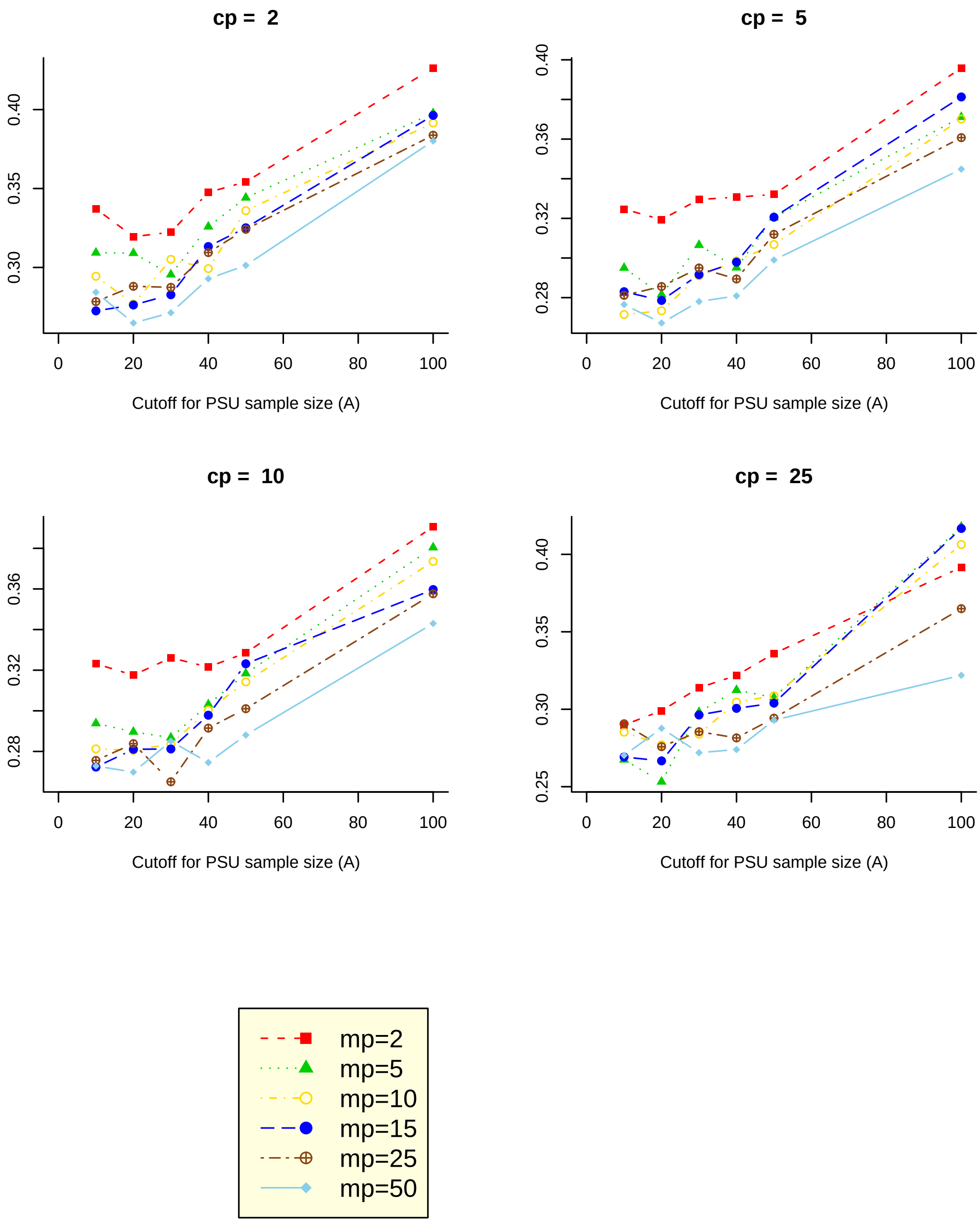


Table E.5: Variance of $\hat{\beta}$, $(\times 10^3)$, calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_f=5000, C_1=2 and C_2=1. ρ=0.025 and 0.05)

Pilot		True Variance of ($\hat{\beta}$) for ρ=0.025						True Variance of ($\hat{\beta}$) for ρ=0.05					
PSUs	Obs	Cutoff for Within-PSU Sample Size (A)						Cutoff for Within-PSU Sample Size (A)					
c_p	m_p	10	20	30	40	50	100	10	20	30	40	50	100
2	2	0.367	0.376	0.382	0.408	0.442	0.581	0.410	0.451	0.518	0.551	0.599	0.964
2	5	0.342	0.347	0.386	0.420	0.448	0.612	0.407	0.444	0.509	0.538	0.673	0.904
2	10	0.344	0.340	0.387	0.391	0.436	0.662	0.373	0.425	0.523	0.585	0.660	0.917
2	15	0.325	0.334	0.363	0.401	0.428	0.601	0.359	0.439	0.498	0.523	0.620	0.893
2	25	0.316	0.342	0.373	0.405	0.444	0.570	0.366	0.426	0.477	0.499	0.584	0.829
2	50	0.303	0.346	0.352	0.399	0.440	0.524	0.378	0.398	0.435	0.493	0.523	0.745
5	2	0.329	0.371	0.389	0.410	0.432	0.559	0.409	0.437	0.473	0.570	0.589	0.761
5	5	0.352	0.348	0.377	0.410	0.468	0.595	0.384	0.454	0.509	0.548	0.613	0.943
5	10	0.308	0.329	0.374	0.372	0.436	0.609	0.401	0.422	0.474	0.475	0.568	0.851
5	15	0.326	0.322	0.359	0.380	0.429	0.529	0.365	0.421	0.468	0.516	0.519	0.813
5	25	0.331	0.328	0.369	0.378	0.382	0.493	0.379	0.407	0.453	0.443	0.504	0.656
5	50	0.307	0.302	0.351	0.380	0.385	0.426	0.367	0.389	0.398	0.428	0.449	0.505
10	2	0.361	0.364	0.388	0.384	0.434	0.581	0.382	0.450	0.482	0.518	0.597	0.800
10	5	0.311	0.348	0.377	0.420	0.454	0.601	0.389	0.409	0.468	0.547	0.587	0.944
10	10	0.304	0.350	0.355	0.387	0.411	0.549	0.355	0.387	0.460	0.509	0.562	0.809
10	15	0.311	0.344	0.344	0.391	0.409	0.527	0.365	0.406	0.437	0.476	0.531	0.697
10	25	0.319	0.329	0.344	0.361	0.369	0.514	0.365	0.402	0.398	0.424	0.444	0.559
10	50	0.316	0.309	0.316	0.341	0.352	0.419	0.373	0.388	0.364	0.380	0.375	0.425
25	2	0.345	0.341	0.364	0.381	0.363	0.521	0.400	0.404	0.455	0.460	0.521	0.774
25	5	0.311	0.352	0.373	0.398	0.441	0.668	0.375	0.432	0.498	0.577	0.656	1.093
25	10	0.311	0.341	0.348	0.386	0.419	0.624	0.367	0.380	0.417	0.473	0.495	0.722
25	15	0.321	0.330	0.341	0.374	0.387	0.542	0.367	0.422	0.412	0.466	0.449	0.566
25	25	0.319	0.303	0.341	0.352	0.359	0.443	0.355	0.373	0.379	0.392	0.379	0.404
25	50	0.325	0.300	0.323	0.323	0.318	0.324	0.363	0.365	0.359	0.367	0.364	0.379

Figure E.8: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=2 and C_2=1, ρ=0.025)

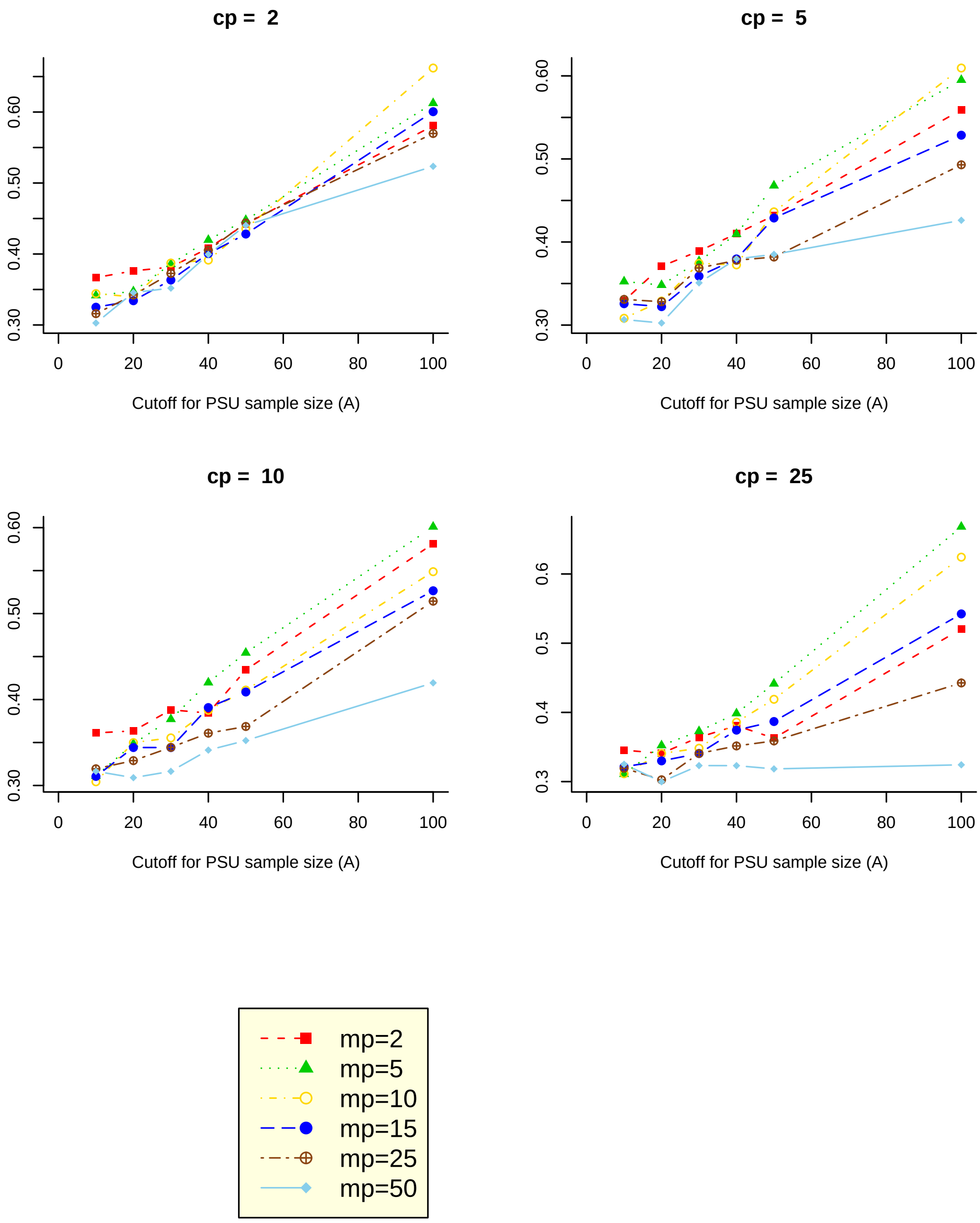


Figure E.9: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=2 and C_2=1, ρ=0.05)

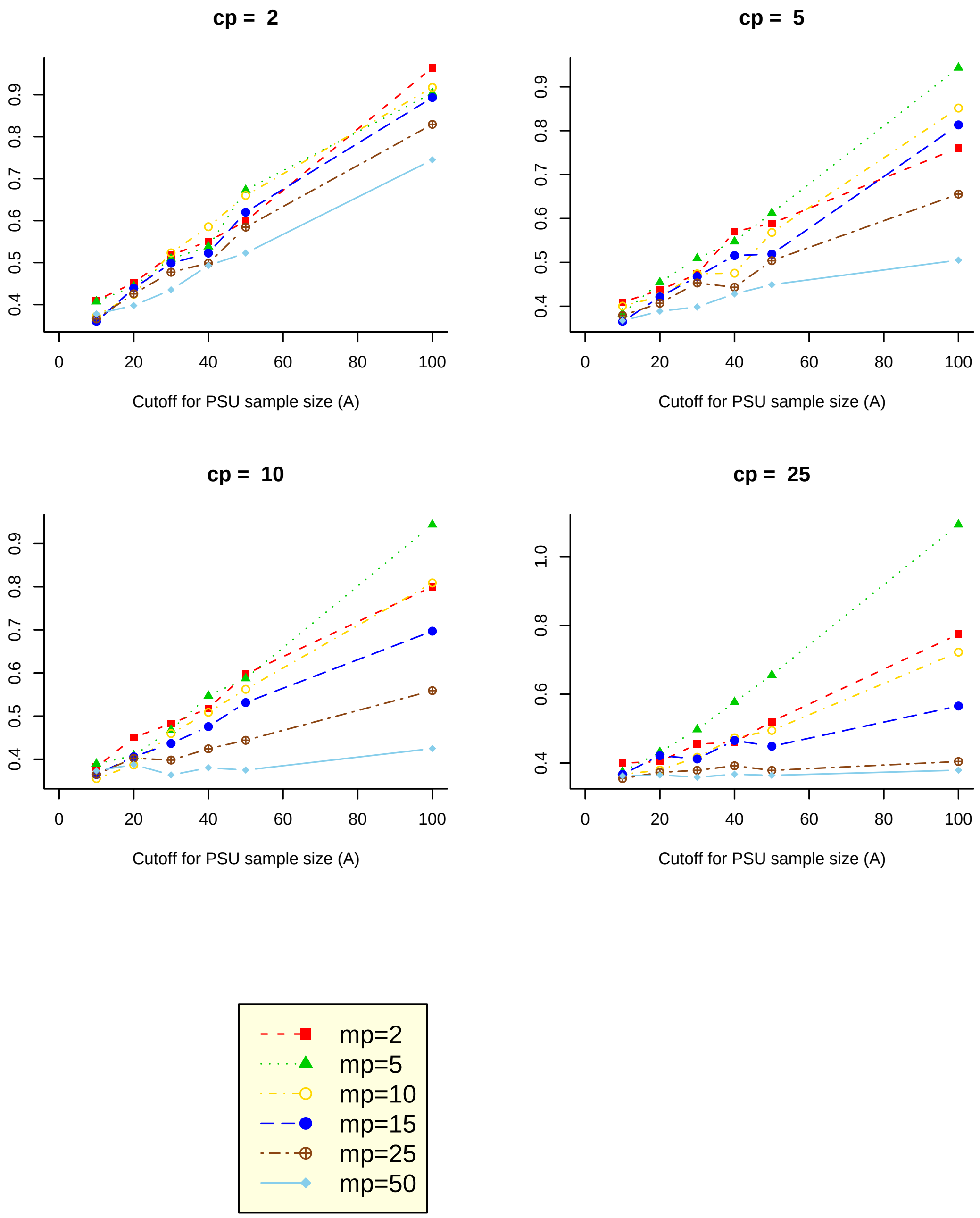


Table E.6: Variance of $\hat{\beta}$, ($\times 10^3$), calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_f=5000, C_1=2 and C_2=1. ρ=0.1)

Pilot		True Variance of $(\hat{\beta})$ for ρ=0.1					
PSUs	Obs	Cutoff for Within-PSU Sample Size (A)					
c_p	m_p	10	20	30	40	50	100
2	2	0.504	0.621	0.747	0.883	0.839	1.593
2	5	0.473	0.614	0.790	0.903	1.033	1.880
2	10	0.484	0.608	0.762	0.858	1.019	1.591
2	15	0.510	0.615	0.660	0.786	0.959	1.458
2	25	0.490	0.597	0.658	0.794	0.819	1.399
2	50	0.492	0.541	0.660	0.692	0.775	1.131
5	2	0.513	0.558	0.739	0.830	0.878	1.353
5	5	0.521	0.600	0.702	0.794	0.966	1.593
5	10	0.459	0.539	0.658	0.649	0.788	1.244
5	15	0.468	0.526	0.584	0.651	0.750	1.044
5	25	0.477	0.480	0.533	0.580	0.604	0.804
5	50	0.469	0.513	0.502	0.514	0.528	0.540
10	2	0.520	0.549	0.636	0.756	0.834	1.167
10	5	0.497	0.544	0.739	0.755	0.834	1.411
10	10	0.455	0.534	0.580	0.569	0.675	1.009
10	15	0.464	0.495	0.521	0.570	0.556	0.676
10	25	0.468	0.453	0.454	0.484	0.495	0.581
10	50	0.443	0.471	0.478	0.479	0.468	0.478
25	2	0.487	0.599	0.633	0.727	0.817	1.247
25	5	0.467	0.539	0.647	0.758	0.820	1.292
25	10	0.456	0.491	0.526	0.552	0.493	0.597
25	15	0.464	0.470	0.484	0.465	0.472	0.521
25	25	0.464	0.456	0.451	0.440	0.489	0.482
25	50	0.456	0.456	0.456	0.456	0.456	0.456

Figure E.10: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=2 and C_2=1, ρ=0.1)

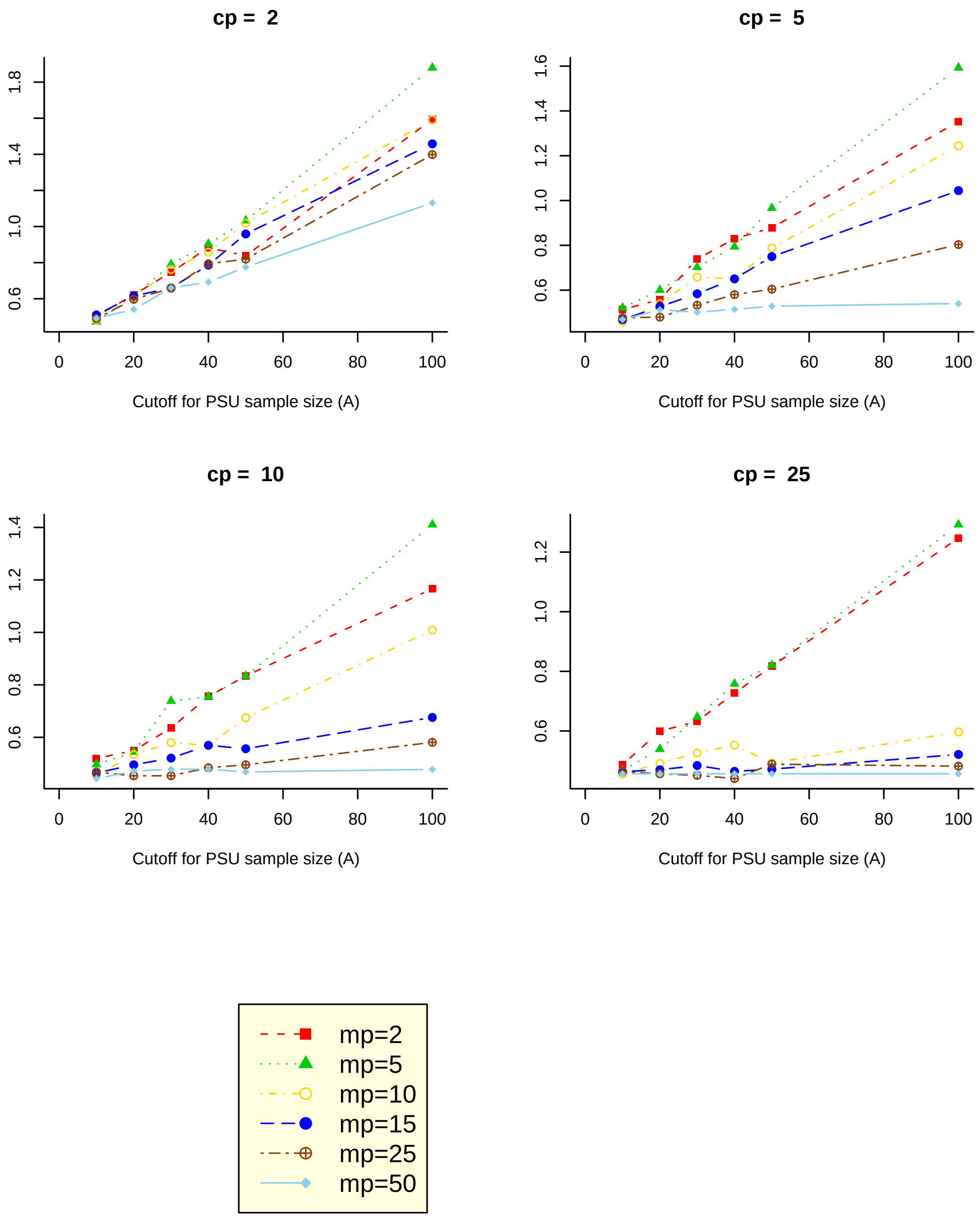


Table E.7: Variance of $\hat{\beta}$, $(\times 10^3)$, calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=10 and C_2=1. ρ=0.01 and 0.025)

Pilot		True Variance of $(\hat{\beta})$ for ρ=0.025						True Variance of $(\hat{\beta})$ for ρ=0.05					
PSUs	Obs	Cutoff for Within-PSU Sample Size (A)						Cutoff for Within-PSU Sample Size (A)					
c_p	m_p	10	20	30	40	50	100	10	20	30	40	50	100
2	2	0.646	0.598	0.602	0.597	0.612	0.625	0.645	0.704	0.704	0.714	0.715	0.825
2	5	0.511	0.465	0.452	0.463	0.468	0.551	0.549	0.544	0.519	0.646	0.590	0.778
2	10	0.486	0.410	0.391	0.414	0.445	0.485	0.547	0.526	0.521	0.554	0.570	0.811
2	15	0.491	0.394	0.407	0.396	0.406	0.444	0.515	0.503	0.484	0.502	0.522	0.693
2	25	0.448	0.402	0.385	0.367	0.401	0.444	0.509	0.484	0.521	0.531	0.537	0.714
2	50	0.456	0.378	0.384	0.373	0.379	0.448	0.506	0.460	0.482	0.483	0.541	0.706
5	2	0.580	0.545	0.549	0.535	0.561	0.588	0.642	0.642	0.653	0.631	0.689	0.797
5	5	0.486	0.424	0.441	0.392	0.434	0.461	0.568	0.506	0.533	0.534	0.535	0.708
5	10	0.452	0.412	0.407	0.373	0.408	0.456	0.514	0.496	0.498	0.499	0.527	0.685
5	15	0.461	0.388	0.386	0.380	0.384	0.423	0.527	0.465	0.492	0.508	0.493	0.692
5	25	0.428	0.371	0.354	0.375	0.392	0.423	0.525	0.484	0.475	0.487	0.525	0.588
5	50	0.444	0.365	0.366	0.383	0.366	0.422	0.498	0.487	0.469	0.512	0.495	0.576
10	2	0.533	0.487	0.463	0.509	0.488	0.577	0.581	0.589	0.592	0.584	0.609	0.723
10	5	0.454	0.399	0.393	0.381	0.384	0.480	0.526	0.497	0.534	0.528	0.531	0.744
10	10	0.458	0.364	0.361	0.380	0.403	0.456	0.535	0.457	0.483	0.540	0.563	0.707
10	15	0.466	0.389	0.390	0.366	0.377	0.464	0.510	0.464	0.474	0.504	0.514	0.649
10	25	0.468	0.382	0.376	0.371	0.366	0.412	0.495	0.463	0.461	0.490	0.491	0.641
10	50	0.454	0.363	0.367	0.354	0.348	0.414	0.503	0.452	0.503	0.512	0.526	0.522
25	2	0.494	0.456	0.441	0.465	0.443	0.482	0.552	0.519	0.550	0.552	0.586	0.674
25	5	0.449	0.409	0.360	0.379	0.399	0.469	0.512	0.469	0.466	0.516	0.557	0.735
25	10	0.441	0.360	0.354	0.360	0.365	0.437	0.514	0.463	0.483	0.508	0.502	0.672
25	15	0.425	0.391	0.361	0.361	0.382	0.449	0.522	0.448	0.463	0.440	0.552	0.645
25	25	0.442	0.365	0.357	0.354	0.366	0.388	0.511	0.481	0.462	0.503	0.475	0.585
25	50	0.457	0.370	0.364	0.364	0.372	0.387	0.510	0.456	0.478	0.448	0.461	0.522

Figure E.11: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=10 and C_2=1, ρ=0.025)

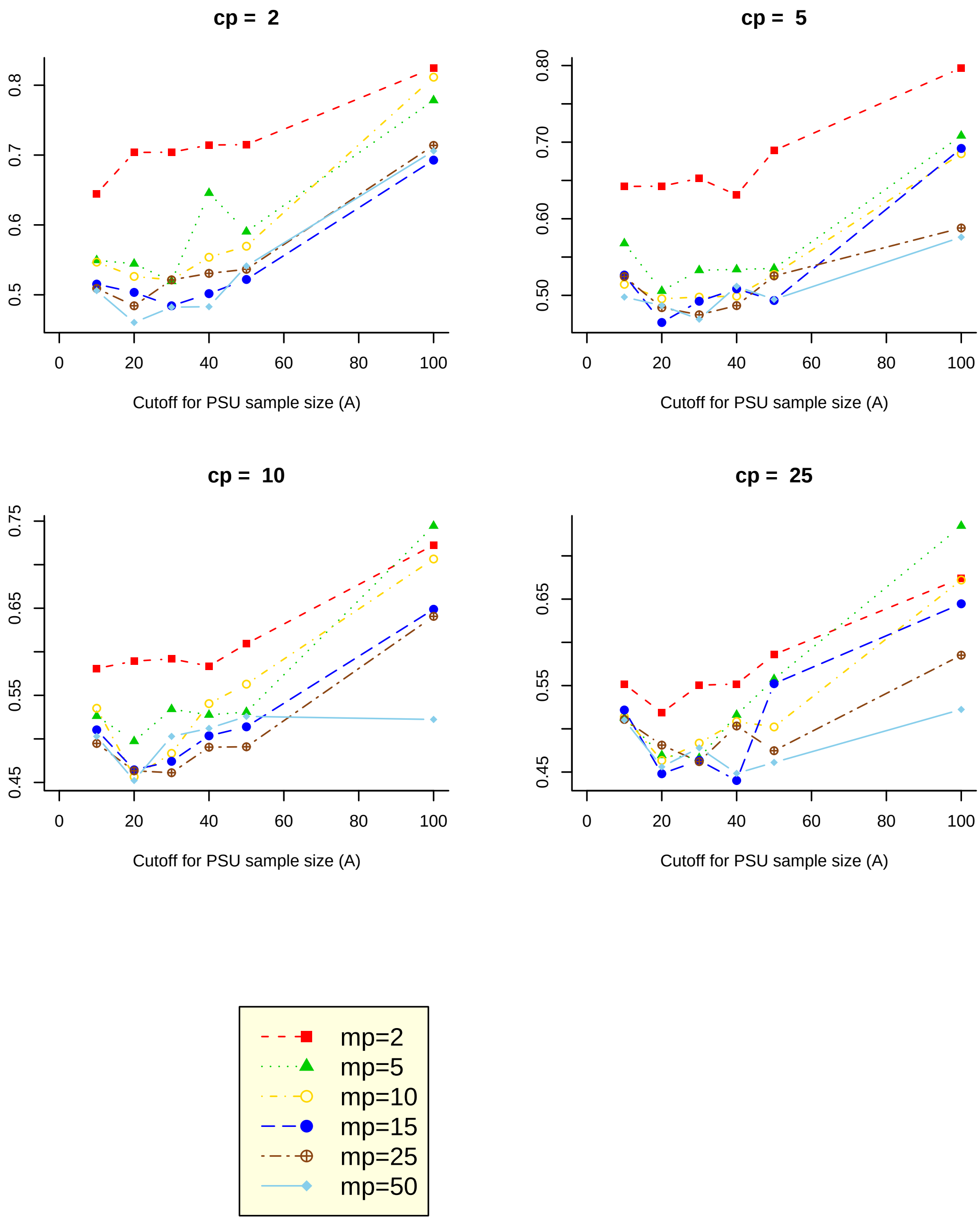
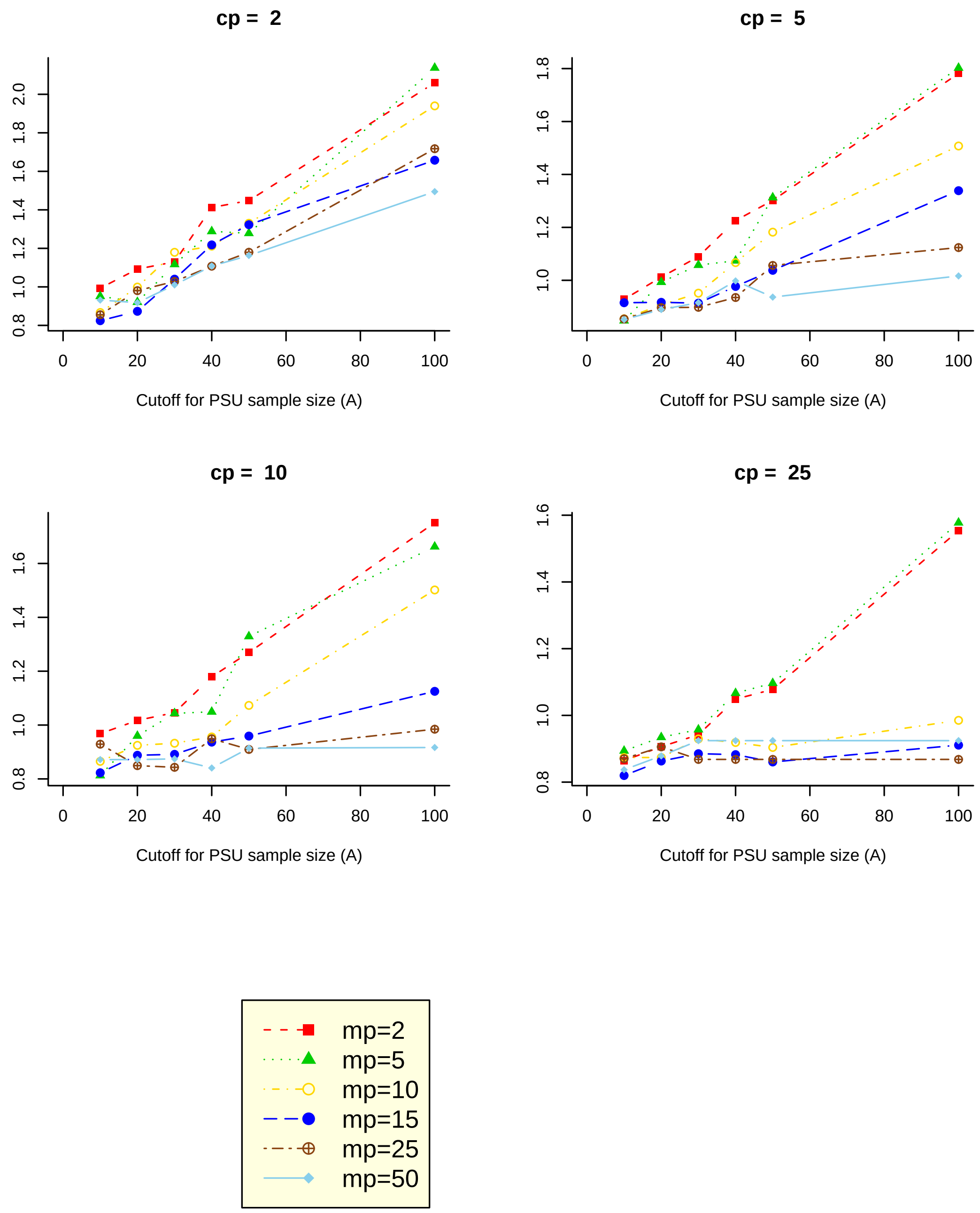


Table E.8: Variance of $\hat{\beta}$, $(\times 10^3)$, calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_f=5000, C_1=2 and C_2=1. ρ=0.1)

Pilot		True Variance of $(\hat{\beta})$ for ρ=0.1					
PSUs	Obs	Cutoff for Within-PSU Sample Size (A)					
c_p	m_p	10	20	30	40	50	100
2	2	0.993	1.093	1.130	1.412	1.448	2.060
2	5	0.951	0.921	1.117	1.288	1.278	2.137
2	10	0.867	0.999	1.180	1.212	1.330	1.940
2	15	0.824	0.873	1.040	1.218	1.323	1.658
2	25	0.855	0.980	1.028	1.108	1.180	1.717
2	50	0.932	0.918	1.010	1.108	1.163	1.495
5	2	0.928	1.012	1.089	1.226	1.302	1.782
5	5	0.847	0.993	1.057	1.074	1.313	1.803
5	10	0.855	0.907	0.951	1.067	1.182	1.508
5	15	0.915	0.917	0.914	0.977	1.038	1.339
5	25	0.854	0.897	0.898	0.935	1.056	1.124
5	50	0.852	0.890	0.915	0.998	0.937	1.017
10	2	0.968	1.017	1.046	1.179	1.269	1.751
10	5	0.813	0.960	1.044	1.049	1.329	1.662
10	10	0.865	0.925	0.933	0.956	1.073	1.502
10	15	0.823	0.888	0.891	0.937	0.959	1.125
10	25	0.929	0.849	0.843	0.949	0.910	0.985
10	50	0.872	0.872	0.874	0.841	0.914	0.917
25	2	0.864	0.906	0.942	1.049	1.078	1.554
25	5	0.894	0.934	0.957	1.067	1.097	1.578
25	10	0.874	0.874	0.929	0.919	0.904	0.985
25	15	0.820	0.863	0.885	0.882	0.860	0.911
25	25	0.870	0.906	0.868	0.868	0.868	0.868
25	50	0.837	0.879	0.924	0.924	0.924	0.924

Figure E.12: Variance of $\hat{\beta}$ calculated from a main survey with budget C_f=5000, designed using a pilot survey (C_1=10, C_2=1, ρ=0.1)



Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Berkhof, J. and Snijders, T. A. B. (2001). Variance component testing in multilevel models. *Journal of Educational and Behavioral Statistics*, 26(2):133–152.

Brooks, S. H. (1955). The estimation of an optimum subsampling number. *Journal of the American Statistical Association*, 50(270):398–415.

Chernoff, H. (1954). On the distribution of the likelihood ratio. *The Annals of Mathematical Statistics*, 25(3):573–578.

Clark, R. G. and Steel, D. G. (2002). The effect of using household as a sampling unit. *International Statistical Review*, 70(2):289–314.

Cochran, W. G. (1977). *Sampling Techniques*. John Wiley and Sons, Inc.

Commenges, D. and Jacqmin, H. (1994). The intraclass correlation coefficient: Distribution-free definition and test. *Biometrics*, 50(2):517 – 526.

Crow, E. L. and Shimizo, K., editors (1988). *Lognormal Distribution: Theory and Application*. Marcel Decker, Inc., New York.

Diggle, P. J., Liang, K. Y., and Zeger, S. L. (1994). *Analysis of Longitudinal Data*. Clarendon Press, New York.

Faes, C., Molenberghs, H., Aerts, M., Verbeke, G., and Kenward, M. G. (2009). The effective sample size and an alternative small-sample degrees-of-freedom method. *The American Statistician*, 63(4):389–399.

Gao, S. and Smith, T. M. F. (1998). A constrained MINQU estimator of correlated response variance from unbalanced data in complex surveys. *Statistica Sinica*, 8:1175–1188.

Goldstein, H. (2003). *Multilevel Statistical Models*. Kendall's Library of Statistics 3. Arnold, London, third edition.

Hansen, M. H. and Hurwitz, W. N. (1943). On the theory of sampling from finite populations. *The Annals of Matematical Statistics*, 14(4):333–362.

Hansen, M. H. and Hurwitz, W. N. (1951). Modern methods in the sampling of human populations, some methods of area sampling in a local community. *American Journal of Public Health*, 41:662–668.

Hansen, M. H., Hurwitz, W. N., and Madow, W. G. (1953). *Sample Survey Methods and Theory*, volume 1, 2. John Wiley and Sons, Inc., New York.

Hocking, R. R. (1996). *Methods and Applications of Linear Modls.* John Wiley and Sons, Inc., New York.

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T. (1999). Bayesian model averaging: a tutorial. *Statistical Science*, 14(4):382–401.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. *Proceedings of the Fifth Berekley Symposium on Mathematical Statistics and Probability, University of California, Berekley*, 11:221–233.

Kalsbeek, W. D., Mendoza, O. M., and Budescu, D. V. (1981). A new cost model for optimum allocation in two-stage sampling. pages 55–60. Proceedings of the American Statistical Association – Section on Survey Research Methods, Alexandria, VA: American Statistical Association.

Kapadia, A. S., Chan, W., and Moyé (2005). *Mathematical Statistics with Applications*. Chapman and Hall/CRC.

Kass, R. E. and Wasserman, L. A. (1995). Bayesian test for nested hypotheses with large samples. *Journal of the American Statistical Association, Theory and Methods*, 90(431):928–934.

Kenward, M. G. and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53(3):983–997.

Killip, S., Mahfoud, Z., and Pearce, K. (2004). What is an intracluster correlation coefficients? crucial concepts for primary case researcher. *Annals of Family Medicine*, 2(3):204–208.

Kish, L. (1965). *Survey Sampling*. John Wiley and Sons, Inc.

Krishnamoorthy, K. and Mathew, T. (2003). Inferences on the means of lognormal distributions using generalized p-values and generalized confidence intervals. *Journal of Statistical Planning and Inference*, 115(1):103–121.

LaHuis, D. M. and Ferguson, M. W. (2009). The accuracy of significance tests for slope variance components in multilevel random coefficient models. *Organizational Research Methods*, 12(3):418–435.

Lancaster, G. A., Dodd, S., and Williamson, P. R. (2004). Design and analysis of pilot studies: recommendations for good practice. *Journal of Evaluation in Clinical Practice*, 10(2):307–312.

Lehtonen, R. and Pahkinen, E. J. (1994). *Practical Methods for Design and Analysis of Complex Surveys*. Chichester, UK: Wiley.

Li, G. and Shi, J. (2010). Application of Bayesian model averaging in modeling long-term wind speed distributions. *Renewable Energy*, 35(6):1192–1202.

Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22.

Longford, N. T. (2008). An alternative analysis of variance. *Statistics and Operations Research Transactions (SORT)*, 32(1):77–92.

Longford, N. T. and Pittau, M. G. (2006). Stability of household income in European countries in the 1990s. *Journal of Computational Statistics and Data Analysis*, 51:1364–1383.

MacKinnon, J. G. and White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29:305–325.

Madigan, D. and Ridgeway, G. (2003). *Bayesian Data Analysis*, chapter 5, pages 104–132. Lawrence Erlbaum, New Jersey.

Muthén, B. O. and Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, 25:267–316.

Niser, J. C. (2010). Study abroad education in new england higher education: a pilot survey. *International Journal of Educational Management*, 24(1):48–55.

Ott, W. R. (1995). *Environmental Statistics and Data Analysis*. Lewis Publishers, Boca Raton, FL.

Patterson, H. D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3):545–554.

Pfeffermann, D. (1993). The role of sampling weights when modeling survay data. *International Statistical Review / Revue Internationale de Statistique*, 61(2):317–337.

Pfeffermann, D., Skinner, C. J., Holmes, D., Goldstein, H., and Rasbash, J. (1998). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 60(1):23–56.

Pfeffermann, D. and Sverchkov, M. (1999). Parametric and semi-parametric estimation of regression models fitted to survey data. *Sankhyā: The Indian Journal of Statistics, Series B*, Special Issue on Sample Surveys, 61(Pt. 1):166–186.

Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-effects Models in S and S-plus*. Springer, New York.

Posada, D. and Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic Biology*, 53(5):793–808.

R Development Core Team (2007). *R: A Language and Environment for*

Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rao, J. N. K. and Wu, C. F. J. (1988). Resampling inference with complex survey data. *Journal of the American Statistical Association*, 83(401):231–241.

Rao, P. S. R. S. (1997). *Variance Components Estimation, Mixed Models, Methodologies and Applications*. Chapman and Hall.

Sahai, H. and Ojeda, M. M. (2004). *Analysis of Variance for Random Models, Balanced Data, Theory, Methods, Applications and Data analysis*, volume 1. Birkhauser, Boston.

Sahai, H. and Ojeda, M. M. (2005). *Analysis of Variance for Random Models, Unbalanced Data, Theory, Methods, Applications and Data Analysis*, volume 2. Birkhauser, Boston.

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika*, 6:309–316.

Scheipl, F., Greven, S., and Küchenhoff, H. (2007). Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models. *Computational Statistics and Data Analysis*, 52(7):3283–3299.

Scott, A. J. and Holt, D. (1982). The effect of two-stage sampling on ordinary least squares methods. *Journal of the American Statistical Association*, 77(380):848–854.

Self, S. G. and Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82(398):605–610.

Sitter, R. R. (1992). A resampling procedure for complex survey data. *Journal of the American Statistical Association*, 87(419):755–765.

Skinner, V., Agho, K., Lee-White, T., and Judy, H. (2007). The development of a tool to assess levels of stress and burnout. *Australian Journal of Advanced Nursing*, 24(4):8–13.

Snijders, T. A. B. (2001). *Sampling*, chapter 11, pages 159–174. John Wiley and Sons, Inc.

Sorenson, D. and Gianola, D. (2002). *Likelihood, Bayesian and MCMC Methods in Quantitative Genetics*. Springer.

SPSS (2007). *SPSS Base 16.0 User´s Guide*. SPSS Inc., 233 South Wacker Drive, 11th Floor, Chicago, USA, IL 60606-6412. Patent No. 7,023,453. ISBN-13: 978-0-13-603601-2, url = http://www.spss.com,.

Steel, D. and Clark, R. G. (2006). Accounting for the uncertainty of information on clustering in the design of a clustered sample. Conference paper, University of Wollongong, Faculty of Informatics, http://ro.uow.edu.au/infopapers/740, Survey Research Methodology Conference, Taiwan.

Stopher, P. R. and Metcalf, H. M. A. (1996). *Methods for Household Travel Surveys, NCHRP Synthesis 236.* Transportation Research Board, Washington, DC.

Stram, D. O. and Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics*, 50(4):1171–1177.

Sugden, R. A. and Smith, T. M. (1984). Ignorable and informative designs in survey sampling inference. *Biometrika*, 71(3):495–506.

Tate, J. E. and Hudgens, M. G. (2007). Estimating population size with two- and three-stage sampling designs. *American Journal of Epidemiology*, 165(11):1314–1320.

Teijlingen, E. R. and Hundley, V. (2002). The importance of pilot studies. *Nursing standard*, 16(40):33–36.

Ukoumunne, O. C. (2002). A comparison of confdence interval methods for the intraclass correlation coeffcient in cluster randomized trials. *Statistics in Medicine*, 21:3757–3774.

Verma, V., Scott, C., and O'Muircheartaigh, C. (1980). Sample design and sampling errors for the world fertility survey. *Journal of the Royal Statistical Society, Series A (General)*, 143(4):431–473.

Visscher, P. M. (2006). A note on the asymptotic distribution of likelihood ratio tests to test variance components. *Twin Research and Human Genetics*, 9(4):490–495.

West, B. T., Welch, K. B., and Galecki, A. T. (2007). *Linear Mixed Model: A Practical Guide Using Statistical Software*. Chapman and Hall/CRC, Boca Raton, Florida.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, 50(1):1–25.

Yuan, Z. and Yang, Y. (2005). Combining linear regression models: when and how? *Journal of the American Statistical Association*, 100(472):1202–1214.

Zabel, J. (1999). Controlling for quality in house price indices. *Journal of Real Estate Finance and Economics*, 3(3):223–241.