2007-07-10

A Comparative Simulation Study of Robust Estimators of Standard Errors

Natalie Johnson
Brigham Young University - Provo

A COMPARATIVE SIMULATION STUDY OF ROBUST ESTIMATORS OF STANDARD ERRORS

by

Natalie Johnson

A project submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Statistics Statistics

Brigham Young University

August 2007

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a project submitted by

Natalie Johnson

This project has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	G. Bruce Schaalje, Chair
Date	Natalie J. Blades
Date	Scott D. Grimshaw

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the project of Natalie Johnson in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date	G. Bruce Schaalje Chair, Graduate Committee

Accepted for the Department

Scott D. Grimshaw
Graduate Coordinator

Accepted for the College

Thomas W. Sederberg
Associate Dean, College of Physical and Mathematical Sciences

ABSTRACT

A COMPARATIVE SIMULATION STUDY OF ROBUST ESTIMATORS OF STANDARD ERRORS

Natalie Johnson

Department of Statistics Statistics

Master of Science

The estimation of standard errors is essential to statistical inference. Statistical variability is inherent within data, but is usually of secondary interest; still, some options exist to deal with this variability. One approach is to carefully model the covariance structure. Another approach is robust estimation. In this approach, the covariance structure is estimated from the data. White (1980) introduced a biased, but consistent, robust estimator. Long et al. (2000) added an adjustment factor to White's estimator to remove the bias of the original estimator. Through the use of simulations, this project compares restricted maximum likelihood (REML) with four robust estimation techniques: the Standard Robust Estimator (White 1980), the Long estimator (Long 2000), the Long estimator with a quantile adjustment (Kauermann 2001), and the `empirical` option of the MIXED procedure in SAS.

The results of the simulation show small sample and asymptotic properties of the five estimators. The REML procedure is modelled under the true covariance structure, and is the most consistent of the five estimators. The REML procedure shows a slight small-sample bias as the number of repeated measures increases. The

REML procedure may not be the best estimator in a situation in which the covariance structure is in question. The Standard Robust Estimator is consistent, but it has an extreme downward bias for small sample sizes. The Standard Robust Estimator changes little when complexity is added to the covariance structure. The Long estimator is unstable estimator. As complexity is introduced into the covariance structure, the coverage probability with the Long estimator increases. The Long estimator with the quantile adjustment works as designed by mimicking the Long estimator at an inflated quantile level. The `empirical` option of the MIXED procedure in SAS works well for homogeneous covariance structures. The `empirical` option of the MIXED procedure in SAS reduces the downward bias of the Standard Robust Estimator when the covariance structure is homogeneous.

ACKNOWLEDGEMENTS

I would like to express my thanks to Dr. Bruce Schaalje whose time, editing, and encouragement has made possible the completion of this project. I would also like to thank Dr. Natalie Blades and Dr. Scott Grimshaw for their contributions both in time and knowledge to this project.

I would like to express my gratitude to my family for their constant love and encouragement. I would also like to thank my roommates, friends, and all others who have supported me through this adventure. Most of all I want to express my gratitude to my Heavenly Father for granting me the opportunity to study at this university and for giving me the strength to push past my own limitations.

CONTENTS

CHAPTER

TABLES

Table

FIGURES

Figure

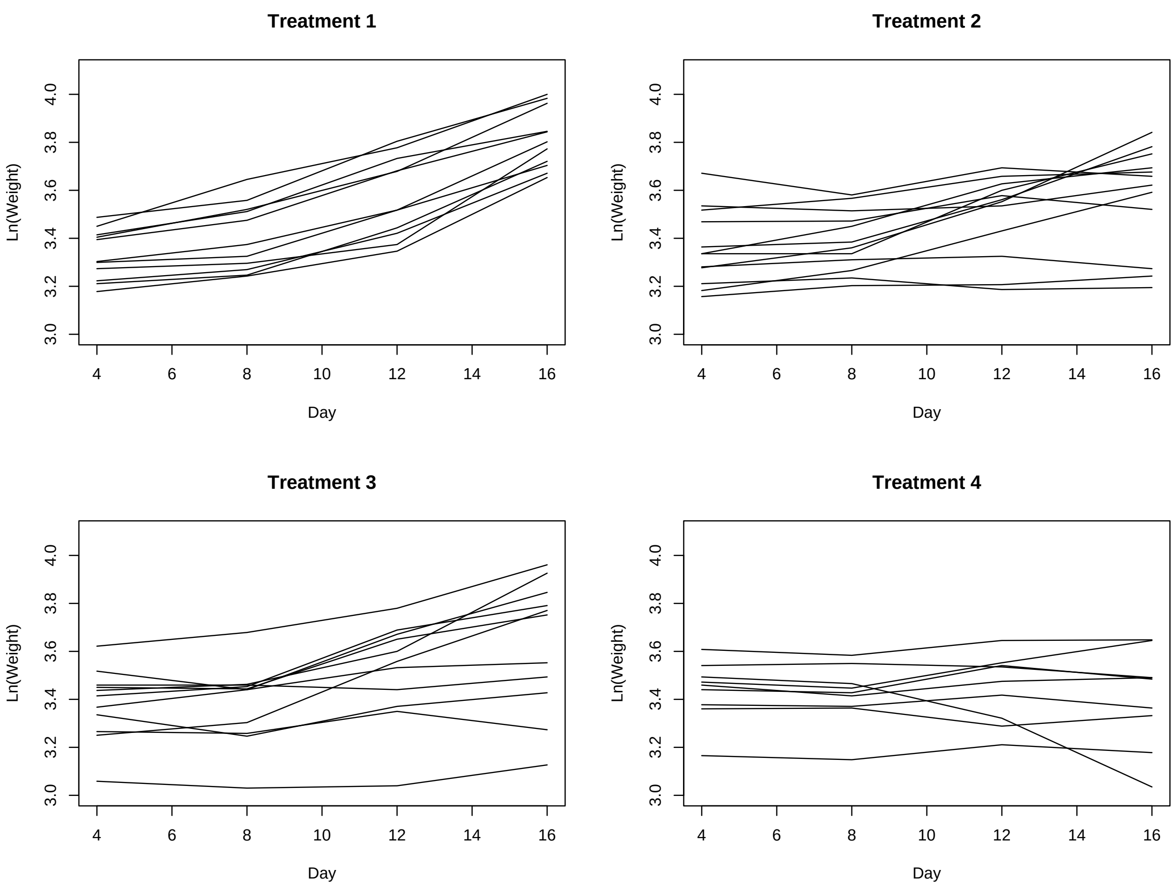
1. INTRODUCTION

The estimation of standard errors is essential to statistical inference. In many situations (e.g. linear models), the primary interest of estimation is in the parameters of the mean structure. Statistical variability itself is usually of secondary interest and is sometimes even considered to be a nuisance (Huber 1972). The efficient estimation of the covariance parameters, however, usually leads to efficient inference about the parameters of the mean structure.

The statistical variability inherent in data and its effects on standard errors of parameter estimates in linear models can be dealt with in several ways. One approach is to fully model the covariance structure of the data, then use parametric methods to estimate the covariance parameters (e.g. restricted maximum likelihood). A poor choice for the specified covariance structure of the data does not usually greatly affect the validity of the parameter estimates, only the efficiency of these estimates (Diggle 1994) and subsequent inferences. Another approach to dealing with statistical variability is robust estimation. In this approach, the covariance structure itself is estimated from the data. The procedure is "robust" in that it is said to produce valid results for many different covariance structures. With robust estimation, the parameter estimates are valid but are not as efficient as when the true covariance structure is known and a parametric method is used.

Because there are many approaches to dealing with statistical variability, deciding which procedure is most appropriate for the situation may be difficult. The use of restricted maximum likelihood (REML) would be appropriate, but the true covariance structure is usually unknown. Several options to model the covariance structure are available, such as spatial power (SP), variance components (VC), unstructured (UN), and compound symmetric (CS). Information criteria such as AIC

Figure 1.1: Weights of 43 mice in a study of teratology of carbon black oil split into four treatments



or BIC may also be used to select a covariance structure. Another option is to use a robust estimator proposed by White (1980) or variations on this estimator (Long 2000, SAS Institute 2003).

As an example, Hansen et al. (2000) studied teratogenic effects of carbon black oil on mice. After receiving one of the four treatments, mouse pups were weighed at each of four preselected ages (4, 8, 12, and 16 days). Figure 1.1 shows the rate of growth for the 43 mice in the study. The log of the weight of the mouse was used as the response variable.

Table 1.1 shows the parameter estimates and the standard errors for several methods used to estimate the covariance structure. The standard errors differed among the estimators in this example. In this situation, the unstructured covariance matrix was selected by AIC. This raises the question: when are each of these procedures (REML plus AIC or the robust estimators) appropriate for use with repeated measures data?

This project is a comparative simulation study of the statistical properties of robust estimators of standard errors for repeated measures data. Specifically, this project compares parametric methods based on REML (Thompson 1969) and robust approaches based on the sandwich estimator (White 1980). This is done using several covariance structures with both homogeneous and heterogeneous variances. Other items of interest that will vary are the amount of correlation within the covariance matrices, the number of repeated measures per unit, and the number of units per treatment. The prevalence with which the parameter is located within the computed 95% confidence interval is measured.

This is an important study because both procedures (REML and the robust estimators) for estimation of standard errors are commonly used, yet only a few simulation studies have been done to compare the estimators, especially in small-sample and repeated measures situations. This study provides valuable and practical information on the effects that the covariance structure and other study features have on the efficiency of the estimates.

Table 1.1: Comparison of the estimates and standard errors of the parameters with respect to various covariance estimation techniques for the control group of the teratogenic study on mice

Method	**Parameter**	**Estimate**	**Standard Error**
REML—Auto-Regressive	Intercept	3.336700	0.070960
	Day	-0.011930	0.011900
	Day^2	0.002612	0.000570
REML—Random Coefficients	Intercept	3.335400	0.141000
	Day	-0.011810	0.032160
	Day^2	0.002612	0.001583
REML—Unstructured*	Intercept	3.335400	0.045710
	Day	-0.011810	0.007293
	Day^2	0.002612	0.000463
REML—Compound Symmetric	Intercept	3.335400	0.082820
	Day	-0.011810	0.015920
	Day^2	0.002612	0.000784
Robust Estimator	Intercept	3.335447	0.021389
	Day	-0.011807	0.009662
	Day^2	0.002612	0.000475
Long Estimator	Intercept	3.335447	0.103007
	Day	-0.011807	0.025508
	Day^2	0.002612	0.001268
Empirical	Intercept	3.336300	0.020550
	Day	-0.012700	0.008426
	Day^2	0.002645	0.000422

* Refers to the covariance structure under REML with the smallest AIC

2. LITERATURE REVIEW

In order to more fully consider each of the estimators discussed in the previous example, the literature detailing these methods will be examined. This chapter is divided into three sections. The first section is an introduction to the notation of the Robust estimator, including some of its asymptotic properties. The second section is an overview of the options for implementing the Robust estimator. The third section is a review of previous simulation studies involving the Robust estimator.

2.1 Notation

Consider the linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations from a dependent variable, $\mathbf{X}$ is an $n \times k$ matrix of observations from an independent variable assumed to be full rank, $\boldsymbol{\epsilon}$ is distributed normally with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{V}$, a positive definite diagonal or block-diagonal matrix. The generalized least squares estimate for $\boldsymbol{\beta}$ and its covariance matrix are

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \tag{2.2}$$

and

$$\text{cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \tag{2.3}$$

These estimates assume that the covariance matrix $\mathbf{V}$ is known. If $\mathbf{V}$ is unknown, neither equation (2.2) nor (2.3) can be calculated.

One approach for estimating the covariance matrix $\mathbf{V}$ is to use maximum likelihood. This approach maximizes the joint likelihood of the original observations.

However, the maximum likelihood approach sometimes yields biased estimates of the variance components (Long 1962). Restricted maximum likelihood (REML) is an approach that is designed to reduce the bias of the estimates of the variance components brought about by maximum likelihood. The REML procedure maximizes the joint likelihood where the sufficient statistics are location-invariant (Long 1962). The REML procedure is an iterative procedure using one of several maximizing algorithms to converge to the global maximum (Burch 2001).

Another possible approach for inference is to use the ordinary least squares estimator for $\boldsymbol{\beta}$,

$$\hat{\hat{\boldsymbol{\beta}}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{2.4}$$

This estimator is not fully efficient, but it no longer has the unknown $\mathbf{V}$ in it; it can always be calculated, and it is unbiased (Rencher 2000, p. 152). The covariance matrix of $\hat{\hat{\boldsymbol{\beta}}}$ for the model in (2.1) is

$$\mathrm{cov}(\hat{\hat{\boldsymbol{\beta}}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}. \tag{2.5}$$

The estimator $\hat{\hat{\boldsymbol{\beta}}}$ avoids the use of $\mathbf{V}$, but its covariance matrix (2.5) still includes the unknown $\mathbf{V}$. Thus, inferences cannot be carried out based on $\hat{\hat{\boldsymbol{\beta}}}$ unless $\mathbf{V}$ can somehow be estimated. Several researchers have investigated estimators of $\mathrm{cov}(\hat{\hat{\boldsymbol{\beta}}})$ and their properties. Eiker (1963) investigated sufficiency and consistency of $\mathrm{cov}(\hat{\hat{\boldsymbol{\beta}}})$, equation (2.5), using the covariance of the errors $\mathrm{cov}(\boldsymbol{\epsilon}\boldsymbol{\epsilon}')$ as the estimate of the $\mathbf{V}$ matrix. Huber (1967) further investigated the asymptotic properties of the covariance matrix using maximum likelihood methods. Huber (1967) used the negative of the information matrix to estimate $\mathbf{V}$. Robust estimation is based on the claim that $\mathrm{cov}(\hat{\hat{\boldsymbol{\beta}}})$, equation (2.5), is somewhat insensitive to the quality of the estimate of $\mathbf{V}$ (Diggle 1994, p. 69). Under fairly broad conditions, this robust estimator is consistent (Kauermann 2001).

A more general analog of $\hat{\hat{\boldsymbol{\beta}}}$, equation (2.4), is

$$\hat{\hat{\boldsymbol{\beta}}}_W = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{y}, \tag{2.6}$$

where $\mathbf{W}$ is a "working" covariance matrix. This analog is unbiased for estimating $\hat{\hat{\boldsymbol{\beta}}}$, and the covariance matrix for $\hat{\hat{\boldsymbol{\beta}}}_W$ is

$$\text{cov}(\hat{\hat{\boldsymbol{\beta}}}_W) = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{V}\mathbf{W}\mathbf{X}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}. \tag{2.7}$$

2.2 Estimation of $\text{cov}(\hat{\hat{\boldsymbol{\beta}}})$

White (1980) introduced a consistent estimator for the situation in which $\mathbf{V}$ is diagonal but possibly heteroscedastic. This method uses the estimate

$$\widehat{\text{cov}}(\hat{\hat{\boldsymbol{\beta}}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}, \tag{2.8}$$

where

$$\hat{\mathbf{V}} = \begin{pmatrix} (y_1 - x_1'\hat{\hat{\boldsymbol{\beta}}})^2 & 0 & \cdots & 0 \\ 0 & (y_2 - x_2'\hat{\hat{\boldsymbol{\beta}}})^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (y_n - x_n'\hat{\hat{\boldsymbol{\beta}}})^2 \end{pmatrix}. \tag{2.9}$$

Notice that this estimate of $\mathbf{V}$ does not allow for correlation among observations.

This estimate has several names, including the sandwich estimator, the consistent variance estimator, the Robust estimator, and the White estimator. In this project, it will be referred to as the Standard Robust Estimator. The Standard Robust Estimator is biased downward, but it is consistent (Kauermann 2001). The consistency of the estimator is what makes it appealing. One of the benefits of this estimator is that it does not rely on a specific formal model of the heteroscedasticity for its consistency (White 1980).

Several programs offer the option of using the Robust estimator. In SAS, the REG procedure uses the `acov` option to compute the weighted least squares sandwich estimate (SAS Institute 2003). The `acov` option uses

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'(\mathrm{diag}(e_i^2))\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}, \tag{2.10}$$

where

$$e_i = y_i - \mathbf{x}_i'\hat{\hat{\boldsymbol{\beta}}}. \tag{2.11}$$

Some efficiency is lost when the Robust estimator is used, but the estimate is still consistent (Agresti 2002, p. 472).

The Robust estimator is computed using ordinary least squares residuals, which tend to be too small (White 1985). Kauermann (2001) provides an option to reduce bias when using the Robust estimator; namely, substitution of

$$\tilde{\boldsymbol{\epsilon}}_i = \frac{\hat{\boldsymbol{\epsilon}}_i}{(1-h_{ii})^{1/2}}, \tag{2.12}$$

where h_{ii} is the ith diagonal element of the hat matrix

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \tag{2.13}$$

for

$$\hat{\boldsymbol{\epsilon}}_i = (y_i - \mathbf{X}_i\hat{\hat{\boldsymbol{\beta}}})(y_i - \mathbf{X}_i\hat{\hat{\boldsymbol{\beta}}})'. \tag{2.14}$$

A generalization of the Robust estimator is often used when $\mathbf{V}$ is block-diagonal, with blocks $(\mathbf{V}_i)$ of equal size. One can use the least squares estimator (2.4) with

$$\widehat{\mathrm{cov}}(\hat{\hat{\boldsymbol{\beta}}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}_b\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}, \tag{2.15}$$

where

$$\hat{\mathbf{V}}_{\mathbf{b}} = \begin{pmatrix} (\mathbf{y}_1 - \mathbf{X}_1\hat{\hat{\boldsymbol{\beta}}})(\mathbf{y}_1 - \mathbf{X}_1\hat{\hat{\boldsymbol{\beta}}})' & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{y}_2 - \mathbf{X}_2\hat{\hat{\boldsymbol{\beta}}})(\mathbf{y}_2 - \mathbf{X}_2\hat{\hat{\boldsymbol{\beta}}})' & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & (\mathbf{y}_n - \mathbf{X}_n\hat{\hat{\boldsymbol{\beta}}})(\mathbf{y}_n - \mathbf{X}_n\hat{\hat{\boldsymbol{\beta}}})' \end{pmatrix}. \tag{2.16}$$

This is similar to the estimator that SAS uses in the `empirical` option of the MIXED procedure (SAS Institute 2003). The `empirical` option uses the estimator

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\hat{\mathbf{V}}_{\mathbf{b}}^{-1}\mathbf{X})^{-1}(\Sigma_{i=1}^{S}\mathbf{X}_i'\hat{\mathbf{V}}_{\mathbf{b}i}^{-1}\hat{\boldsymbol{\epsilon}}_i\hat{\boldsymbol{\epsilon}}_i'\hat{\mathbf{V}}_{\mathbf{b}i}^{-1}\mathbf{X}_i)(\mathbf{X}'\hat{\mathbf{V}}_{\mathbf{b}}^{-1}\mathbf{X})^{-1}, \tag{2.17}$$

where

$$\hat{\boldsymbol{\epsilon}}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\hat{\boldsymbol{\beta}}}. \tag{2.18}$$

In this situation, S is the number of subjects in the study. The matrices with the subscript i are those for the i-th subject (SAS Institute 2003).

2.3 Simulation Studies of the Robust Estimator

Many simulation studies have been done to investigate the properties of the REML estimator (e.g. Thompson 1969, Swallow 1984, Burch 2001, Schaalje 2002). Gomez (2004) used simulation to study properties of the AIC- and BIC-selected REML estimator. However, only a few studies have been done to understand the properties of the Robust estimator.

2.3.1 Kesselman et al. 1999

Kesselman et al. (1999) conducted a simulation study to compare Type I error rates and power for REML-based tests and the Welch-James (WJ) test, a test based on a robust variance estimator similar to the Robust estimator. The researchers were interested in testing the hypothesis

$$H_0 : \mathbf{C}\boldsymbol{\mu} = \mathbf{0}, \tag{2.19}$$

where $\mathbf{C}$ is a set of independent linear contrasts in factorial studies with heterogeneous non-independent covariance structures. Kesselman et al. (1999) generated data from both Normal and Lognormal distributions using covariance structures of Unstructured (UN), Heterogeneous First-order Autoregressive (ARH), or Random Coefficients (RC). All of the covariance structures had within-subject heterogeneity, as well as equal and unequal between-subject variances.

The researchers compared the WJ test, the REML-based test using the correct covariance structure, and the REML-based test using the covariance structure selected by the Akaike Information Criteria (AIC). They let the AIC choose between eleven possible covariance structures: Compound Symmetric (CS), Unstructured (UN), First-order Autoregressive (AR), Huynh-Feldt (HF), Heterogeneous Compound Symmetric (CSH), Heterogeneous First-order Autoregressive (ARH), Random Coefficients (RC), UN_j, HF_j, AR_j, and RC_j. The subscript j represents unequal between-subject covariance. The researchers used total sample sizes of 30, 45, and 60. They also used positive and negative pairings of sample sizes with covariance matrices. A positive pairing is a situation in which the largest subsamples are associated with the groups with the largest variances. A negative pairing is a situation in which the largest subsamples are associated with the groups with the smallest variances.

The WJ test is given by

$$T_{WJ} = (\mathbf{C}\bar{\mathbf{Y}})'(\mathbf{C}\mathbf{S}\mathbf{C}')^{-1}(\mathbf{C}\bar{\mathbf{Y}}), \tag{2.20}$$

where $\bar{\mathbf{Y}} = (\bar{\mathbf{Y}}_1', \cdots, \bar{\mathbf{Y}}_j')'$, with $E(\bar{\mathbf{Y}}) = \boldsymbol{\mu}$ (Kesselman 1999). The sample covariance matrix of $\bar{\mathbf{Y}}$ is $\mathbf{S} = \mathrm{diag}(\mathbf{S}_1/n_1, \cdots, \mathbf{S}_j/n_j)$, where $\mathbf{S}_j$ is the sample covariance matrix for the j-th level of the between-subject grouping factor (Kesselman 1999). The WJ test was divided by a constant to become distributed as an approximate F (Kesselmann 1999). The WJ test is similar to the Robust estimator, but it assumes that individuals within a group have the same covariance matrix.

For tests of the main effects, the Type I error rate was always lower for the WJ test than for the AIC-chosen covariance structure. The AIC-based test exceeded the target error rate when the error structure included negative pairings. The WJ test only slightly exceeded the target error rate when the structure included negative pairings. Tests of interaction effects showed little consistency as to when the error rate exceeded the target level. The AIC-based test and the WJ test often seemed to overshoot the error rate, revealing a bias toward allowing more error than desired. The WJ procedure, however, seemed less affected by heteroscedasticity. The results of the non-Normal data were similar to the results of the Normal data.

The power comparison between the REML-based test and the WJ test showed that the WJ test was consistently less powerful than the F-test. However, the power was similar enough that the WJ test could be considered for use when the covariance structure is unknown.

2.3.2 Long et al. 2000

White and MacKinnon (1985) proposed three alternatives to reduce the bias of the Standard Robust Estimator (2.8). They did this after noticing that the original robust estimator introduced by White (1980) did not take into account that ordinary least squares residuals tended to be too small. This problem with the residuals resulted in the downward bias of this robust estimator (White 1985). The Standard Robust Estimator and three alternatives introduced to reduce bias are

$$\mathrm{cov}_0(\hat{\hat{\boldsymbol{\beta}}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathrm{diag}(\mathrm{e}_i^2))\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}, \tag{2.21}$$

$$\mathrm{cov}_1(\hat{\hat{\boldsymbol{\beta}}}) = \frac{\mathrm{n}}{\mathrm{n}-\mathrm{k}}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathrm{diag}(\mathrm{e}_i^2))\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}, \tag{2.22}$$

$$\mathrm{cov}_2(\hat{\hat{\boldsymbol{\beta}}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathrm{diag}(\frac{\mathrm{e}_i^2}{1-\mathrm{h}_{ii}}))\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}, \tag{2.23}$$

and

$$\mathrm{cov}_3(\hat{\hat{\boldsymbol{\beta}}}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathrm{diag}(\frac{\mathrm{e}_i^2}{(1-\mathrm{h}_{ii})^2}))\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}, \tag{2.24}$$

where

$$\mathrm{e}_i = y_i - \mathbf{X}_i\hat{\hat{\boldsymbol{\beta}}}, \tag{2.25}$$

and h_{ii} is the i-th diagonal element of the hat matrix (2.13). Each of these three alternatives is designed to reduce the downward bias of $\hat{\mathbf{V}}_{\mathbf{b}}$ (2.15). The properties of these three alternatives are described by White and MacKinnon and applied to data from various experiments. Long et al. (2000) ran a simulation study to compare the Robust estimator (2.21) to the three alternatives (2.22, 2.23, and 2.24). They also tested the ordinary least squares estimator with the usual estimate for the covariance matrix

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = \frac{\Sigma e_i^2}{n-k}(\mathbf{X}'\mathbf{X})^{-1}. \tag{2.26}$$

The covariance matrix for the ordinary least squares estimator (2.26) is the estimator with the worst coverage probability (a minimum bound) when the true covariance structure is heterogeneous. Long et al. (2000) tested each estimator under varying degrees of heteroscedasticity. They also varied the sample sizes from 25 to 1000.

Long et al. (2000) generated 100,000 observations for each independent variable. They next generated random errors with the desired error structure. Then, with these errors, they computed the dependent variables. Random samples of the proper size were drawn without replacement. Each of these combinations were run using the

Standard Robust Estimator (2.21), the three alternatives to the Standard Robust Estimator (2.22, 2.23, and 2.24), and the ordinary least squares estimator (2.26).

Since the ordinary least squares estimator is designed for homoscedastic data, it had the best properties for homoscedastic data. The properties for the $\text{cov}_3(\hat{\hat{\boldsymbol{\beta}}})$, (2.24), were nearly the same as the properties of the ordinary least squares estimator, even at the smallest sample size ($n = 25$). The tests for the other three estimators had varying degrees of distortion for sample sizes $n \leq 100$. All of the tests had nearly the same properties for $n \geq 250$.

Under milder forms of heteroscedasticity, the ordinary least squares estimator worked well for all sample sizes. With more extreme forms of heteroscedasticity, the ordinary least squares estimator performed increasingly worse as the sample size increased. When $n \geq 500$, the four different estimators performed similarly. Long et al. (2000) concluded that the estimator (2.24) worked better than all of the other estimators under both heteroscedasticity and homoscedasticity. Long et al. (2000) proposed that the estimator in $\text{cov}_3(\hat{\hat{\boldsymbol{\beta}}})$ (2.24) should always be used.

2.3.3 Kauermann et al. 2001

Kauermann et al. (2001) studied the efficiency of the sandwich estimator using $\text{cov}_2(\hat{\hat{\boldsymbol{\beta}}})$ (2.23). The matrix $\hat{\mathbf{V}}$ is calculated using $\hat{\mathbf{V}}_{\mathbf{b}}$ (2.16). By replacing $\hat{\boldsymbol{\epsilon}}_i$ (2.18) with $\tilde{\boldsymbol{\epsilon}}_i$ (2.12), the Robust estimator is shown to be less biased.

Kauermann et al. (2001) found that the Robust estimator leads to a substantial loss of efficiency. Using the unbiased Robust estimator (2.23), they tested the coverage probability of parameters in simple linear regression. They also compared confidence intervals based on corrected quantiles from a Standard Normal distribution and t-distribution quantiles, and a Robust estimator based on the Jackknife procedure. The confidence intervals of the corrected quantiles were computed by increasing the quantile (e.g. from 95% to 99%) to compensate for the downward bias of the Robust estimator. The t-distribution quantile confidence intervals used quantiles from a t-distribution instead of quantiles from a Standard Normal distribution. Kauermann et al. (2001) claimed that the Robust estimator based on the Jackknife was mathematically equivalent to $\text{cov}_2(\hat{\hat{\boldsymbol{\beta}}})$ (2.23). These estimators were compared across models with varying degrees of heteroscedasticity.

The corrected quantiles for the Robust estimator outperformed the estimator using t-distribution quantiles and the Robust estimator (based on the Jackknife procedure) in data from Uniform, Normal, and Laplace distributions. Neither the t-distribution quantile estimator nor the Robust estimator (based on the Jackknife procedure) achieved the 95% coverage level. The bias-adjusted estimator with the corrected quantiles—a less biased Robust estimator—exceeded the 95% coverage level only when the distribution of the data was Normal. This estimator achieved the 95% coverage level when the sample size was 20 only for the homoscedastic model. Also, this estimator achieved the 95% coverage level when the sample size was 40 for both the homoscedastic and heteroscedastic models.

2.3.4 Summary of Simulation Studies of the Robust Estimator

The original Robust estimator as introduced by White (1980) is a biased estimator. Several authors have considered this estimator to be a reasonable estimator when heteroscedasticity is present as long as an adjustment factor is included in the estimator. Kesselman et al. (1999) showed that a variation of the Robust estimator, the WJ test (2.20), is frequently more accurate than the AIC-based test. Long et al. (2000) tested three variations on the original Robust estimator (2.22, 2.23, and 2.24) in order to determine which variation would best reduce the downward bias of the Robust estimator. They concluded that $\text{cov}_3(\hat{\hat{\boldsymbol{\beta}}})$ (2.24) reduces the bias of the original estimator and they suggested that this equation should always be used to estimate the covariance matrix of any analysis. Kauermann et al. (2001) compared quantile adjustments to compensate for the downward bias of the Robust estimator. They concluded that the best adjustment is made by increasing the quantile for the Standard Normal distribution in the confidence intervals.

3. METHODS

This chapter details the techniques employed to generate the data, compute the estimators and their corresponding confidence intervals, and analyze the results. The data were generated with varying numbers of repeated measurements for each subject. The number of subjects per data set was varied to investigate the small- and large-sample properties of each estimator. The covariance structure was compound symmetric for both the homogeneous and heterogeneous data.

3.1 Estimators

The small- and large-sample properties of five estimators were compared. The five estimators are the REML procedure (using the correct covariance structure), the Robust estimator using the `empirical` option of the MIXED procedure in SAS (2.17), the Standard Robust Estimator as introduced by White (2.15), the bias-reducing adjustment to the Robust estimator that Long (2000) introduced (2.24), and the corrected quantiles adjustment (Kauermann 2001) in conjunction with the Long estimator (2.24).

The REML procedure was modelled under the true compound symmetric covariance structure. When the correct covariance structure was modelled, the REML procedure remained extremely close to the target coverage probability. The REML estimator in the MIXED procedure provided the most accurate estimates under the true covariance structure for both the point estimates and the standard errors because the true covariance structure was known and not just estimated from the data.

The Standard Robust Estimator provided the least accurate standard errors because of its inherent downward bias. The `empirical` option of the MIXED procedure in SAS differed from the Standard Robust Estimator in that it included a working covariance matrix in its estimation. The Long adjustment to the Standard Robust Estimator and the corrected quantiles addition to this estimator provided further insight into the consistency of the Standard Robust Estimator when it has been adjusted to reduce bias. The quantile adjustment to the Long estimator was computed with 99% coverage quantiles instead of the 95% coverage quantiles with which the other estimators were computed.

3.2 Data Generation

The covariance structure used in this simulation study was compound symmetric. Table 3.1 shows the compound symmetric structure for both the homogeneous and heterogeneous covariance matrices (SAS Institute 2003).

Table 3.1: Compound symmetric structure for both the homogeneous and heterogeneous covariance matrices

Compound Symmetric	**Heterogeneous Compound Symmetric**
$\begin{pmatrix} \sigma^2 & \rho & \rho \\ \rho & \sigma^2 & \rho \\ \rho & \rho & \sigma^2 \end{pmatrix}$	$\begin{pmatrix} \sigma_1^2 & \rho\sigma_{12} & \rho\sigma_{13} \\ \rho\sigma_{12} & \sigma_2^2 & \rho\sigma_{23} \\ \rho\sigma_{13} & \rho\sigma_{23} & \sigma_3^2 \end{pmatrix}$

The identity matrix was the basic homoscedastic matrix. The variances for the basic heterogeneous matrix were odd beginning with 1 and increasing by two for every additional repeated measure. Added to these base covariance matrices was a small, moderate, or high level of correlation. The low level of correlation for the homogeneous covariance matrices was 0.0, the low level of correlation for the heterogeneous covariance matrices was 0.1. A small amount of correlation was added to the heterogeneous covariance matrices because without any correlation SAS returned an error that the covariance matrix was not positive definite. The homogeneous and

Table 3.2: Covariance matrices for 3, 4, and 5 repeated measures

Homogeneous Covariance Matrices			
Repeated Measure	Correlation		
	0	0.6	0.9
3	$\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0.6 & 0.6\\0.6 & 1 & 0.6\\0.6 & 0.6 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0.9 & 0.9\\0.9 & 1 & 0.9\\0.9 & 0.9 & 1\end{pmatrix}$
4	$\begin{pmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0.6 & 0.6 & 0.6\\0.6 & 1 & 0.6 & 0.6\\0.6 & 0.6 & 1 & 0.6\\0.6 & 0.6 & 0.6 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0.9 & 0.9 & 0.9\\0.9 & 1 & 0.9 & 0.9\\0.9 & 0.9 & 1 & 0.9\\0.9 & 0.9 & 0.9 & 1\end{pmatrix}$
5	$\begin{pmatrix}1 & 0 & 0 & 0 & 0\\0 & 1 & 0 & 0 & 0\\0 & 0 & 1 & 0 & 0\\0 & 0 & 0 & 1 & 0\\0 & 0 & 0 & 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0.6 & 0.6 & 0.6 & 0.6\\0.6 & 1 & 0.6 & 0.6 & 0.6\\0.6 & 0.6 & 1 & 0.6 & 0.6\\0.6 & 0.6 & 0.6 & 1 & 0.6\\0.6 & 0.6 & 0.6 & 0.6 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0.9 & 0.9 & 0.9 & 0.9\\0.9 & 1 & 0.9 & 0.9 & 0.9\\0.9 & 0.9 & 1 & 0.9 & 0.9\\0.9 & 0.9 & 0.9 & 1 & 0.9\\0.9 & 0.9 & 0.9 & 0.9 & 1\end{pmatrix}$

heterogeneous covariance matrices had 0.6 and 0.9 for the moderate and high levels of correlation, respectively. Table 3.2 lists the homogeneous variance-covariance matrices for 3, 4, and 5 repeated measures, respectively. Tables 3.3, 3.4, and 3.5 list the heterogeneous variance-covariance matrices for 3, 4, and 5 repeated measures by their level of correlation.

These covariance matrices were used to generate the data. Random noise was generated from a Standard Normal distribution and then multiplied by the Cholesky decomposition of the covariance matrix that corresponded to the combination of factors currently being generated. The statistical model (2.1) was used to generate this data. The design matrix, $\mathbf{X}$, in Table 3.6, corresponds to the number of repeated measures in the combination. The appropriate design matrix was multiplied by the true vector of coefficients

Table 3.3: Covariance matrices for 3, 4, and 5 repeated measures with correlation 0.1

Heterogeneous Covariance Matrices	
Repeated Measure	Correlation 0.1
3	$\begin{pmatrix} 1 & 0.1\sqrt{1\cdot 3} & 0.1\sqrt{1\cdot 5} \\ 0.1\sqrt{1\cdot 3} & 3 & 0.1\sqrt{3\cdot 5} \\ 0.1\sqrt{1\cdot 5} & 0.1\sqrt{3\cdot 5} & 5 \end{pmatrix}$
4	$\begin{pmatrix} 1 & 0.1\sqrt{1\cdot 3} & 0.1\sqrt{1\cdot 5} & 0.1\sqrt{1\cdot 7} \\ 0.1\sqrt{1\cdot 3} & 3 & 0.1\sqrt{3\cdot 5} & 0.1\sqrt{3\cdot 7} \\ 0.1\sqrt{1\cdot 5} & 0.1\sqrt{3\cdot 5} & 5 & 0.1\sqrt{5\cdot 7} \\ 0.1\sqrt{1\cdot 7} & 0.1\sqrt{3\cdot 7} & 0.1\sqrt{5\cdot 7} & 7 \end{pmatrix}$
5	$\begin{pmatrix} 1 & 0.1\sqrt{1\cdot 3} & 0.1\sqrt{1\cdot 5} & 0.1\sqrt{1\cdot 7} & 0.1\sqrt{1\cdot 9} \\ 0.1\sqrt{1\cdot 3} & 3 & 0.1\sqrt{3\cdot 5} & 0.1\sqrt{3\cdot 7} & 0.1\sqrt{3\cdot 9} \\ 0.1\sqrt{1\cdot 5} & 0.1\sqrt{3\cdot 5} & 5 & 0.1\sqrt{5\cdot 7} & 0.1\sqrt{5\cdot 9} \\ 0.1\sqrt{1\cdot 7} & 0.1\sqrt{3\cdot 7} & 0.1\sqrt{5\cdot 7} & 7 & 0.1\sqrt{7\cdot 9} \\ 0.1\sqrt{1\cdot 9} & 0.1\sqrt{3\cdot 9} & 0.1\sqrt{5\cdot 9} & 0.1\sqrt{7\cdot 9} & 9 \end{pmatrix}$

$$\boldsymbol{\beta} = \begin{pmatrix} 1 & 1 \end{pmatrix}. \tag{3.1}$$

With the design matrix and the true vector of coefficients, the random noise, multiplied by the Cholesky decomposition of the proper covariance matrix, was added to $\mathbf{X}\boldsymbol{\beta}$. This was done for sample sizes of 5, 10, 25, and 50 subjects within each data set. Point estimates and standard errors were computed for the intercept and slope of each of the five estimators. A 95% confidence interval was calculated for each of the estimators. The Kenward-Roger adjustment (Kenward and Roger 1997) was used to compute the confidence interval for the REML procedure.

Each combination of factors represented a data set. Each data set was replicated 10,000 times. A marker identified which of the confidence intervals contained the true parameter value. These were totalled and converted into percentages. The

Table 3.4: Covariance matrices for 3, 4, and 5 repeated measures with correlation 0.6

Heterogeneous Covariance Matrices	
Repeated Measure	Correlation 0.6
3	$\begin{pmatrix} 1 & 0.6\sqrt{1\cdot 3} & 0.6\sqrt{1\cdot 5} \\ 0.6\sqrt{1\cdot 3} & 3 & 0.6\sqrt{3\cdot 5} \\ 0.6\sqrt{1\cdot 5} & 0.6\sqrt{3\cdot 5} & 5 \end{pmatrix}$
4	$\begin{pmatrix} 1 & 0.6\sqrt{1\cdot 3} & 0.6\sqrt{1\cdot 5} & 0.6\sqrt{1\cdot 7} \\ 0.6\sqrt{1\cdot 3} & 3 & 0.6\sqrt{3\cdot 5} & 0.6\sqrt{3\cdot 7} \\ 0.6\sqrt{1\cdot 5} & 0.6\sqrt{3\cdot 5} & 5 & 0.6\sqrt{5\cdot 7} \\ 0.6\sqrt{1\cdot 7} & 0.6\sqrt{3\cdot 7} & 0.6\sqrt{5\cdot 7} & 7 \end{pmatrix}$
5	$\begin{pmatrix} 1 & 0.6\sqrt{1\cdot 3} & 0.6\sqrt{1\cdot 5} & 0.6\sqrt{1\cdot 7} & 0.6\sqrt{1\cdot 9} \\ 0.6\sqrt{1\cdot 3} & 3 & 0.6\sqrt{3\cdot 5} & 0.6\sqrt{3\cdot 7} & 0.6\sqrt{3\cdot 9} \\ 0.6\sqrt{1\cdot 5} & 0.6\sqrt{3\cdot 5} & 5 & 0.6\sqrt{5\cdot 7} & 0.6\sqrt{5\cdot 9} \\ 0.6\sqrt{1\cdot 7} & 0.6\sqrt{3\cdot 7} & 0.6\sqrt{5\cdot 7} & 7 & 0.6\sqrt{7\cdot 9} \\ 0.6\sqrt{1\cdot 9} & 0.6\sqrt{3\cdot 9} & 0.6\sqrt{5\cdot 9} & 0.6\sqrt{7\cdot 9} & 9 \end{pmatrix}$

percentages represented the coverage probabilities for each combination of factors.

Table 3.5: Covariance matrices for 3, 4, and 5 repeated measures with correlation 0.9

Heterogeneous Covariance Matrices	
Repeated Measure	Correlation 0.9
3	$\begin{pmatrix} 1 & 0.9\sqrt{1\cdot 3} & 0.9\sqrt{1\cdot 5} \\ 0.9\sqrt{1\cdot 3} & 3 & 0.9\sqrt{3\cdot 5} \\ 0.9\sqrt{1\cdot 5} & 0.9\sqrt{3\cdot 5} & 5 \end{pmatrix}$
4	$\begin{pmatrix} 1 & 0.9\sqrt{1\cdot 3} & 0.9\sqrt{1\cdot 5} & 0.9\sqrt{1\cdot 7} \\ 0.9\sqrt{1\cdot 3} & 3 & 0.9\sqrt{3\cdot 5} & 0.9\sqrt{3\cdot 7} \\ 0.9\sqrt{1\cdot 5} & 0.9\sqrt{3\cdot 5} & 5 & 0.9\sqrt{5\cdot 7} \\ 0.9\sqrt{1\cdot 7} & 0.9\sqrt{3\cdot 7} & 0.9\sqrt{5\cdot 7} & 7 \end{pmatrix}$
5	$\begin{pmatrix} 1 & 0.9\sqrt{1\cdot 3} & 0.9\sqrt{1\cdot 5} & 0.9\sqrt{1\cdot 7} & 0.9\sqrt{1\cdot 9} \\ 0.9\sqrt{1\cdot 3} & 3 & 0.9\sqrt{3\cdot 5} & 0.9\sqrt{3\cdot 7} & 0.9\sqrt{3\cdot 9} \\ 0.9\sqrt{1\cdot 5} & 0.9\sqrt{3\cdot 5} & 5 & 0.9\sqrt{5\cdot 7} & 0.9\sqrt{5\cdot 9} \\ 0.9\sqrt{1\cdot 7} & 0.9\sqrt{3\cdot 7} & 0.9\sqrt{5\cdot 7} & 7 & 0.9\sqrt{7\cdot 9} \\ 0.9\sqrt{1\cdot 9} & 0.9\sqrt{3\cdot 9} & 0.9\sqrt{5\cdot 9} & 0.9\sqrt{7\cdot 9} & 9 \end{pmatrix}$

Table 3.6: Design matrices for 3, 4, and 5 repeated measures

Repeated Measures		
3	4	5
$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix}$

4. RESULTS

This chapter details the results of the simulation study. Complete tables of the results are given in Appendix B.

4.1 Intercept

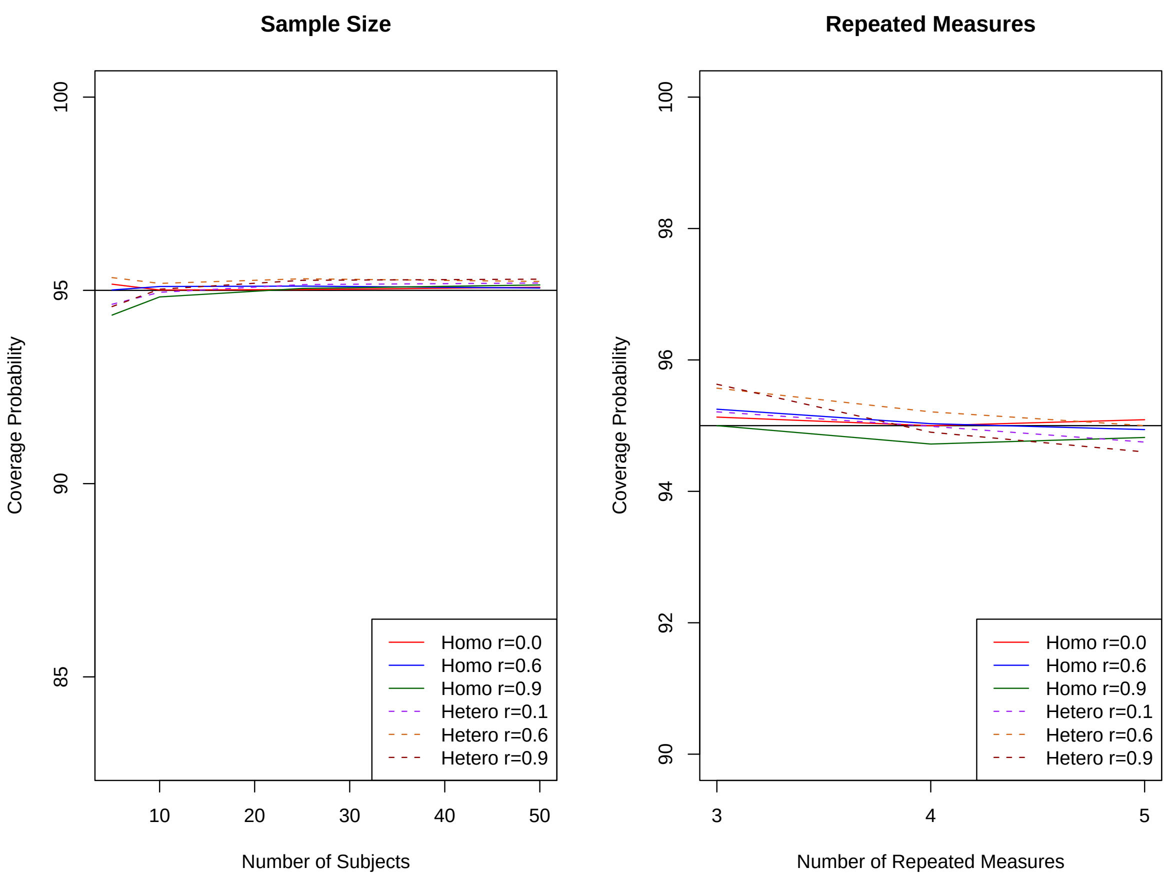
The intercept is a stationary column. Certain properties of the estimators are enhanced when examining the intercept. Each estimator is examined individually, then the estimators are compared.

4.1.1 REML

The small sample and asymptotic properties of an estimator can be determined by looking at how increasing sample size affects the properties of an estimator. Because the REML procedure used the correct covariance structure, it was expected that this estimator would have excellent properties. Figure 4.1 contains plots of the coverage percentages in the REML procedure. The REML procedure never varied more than 0.64% from the target 95% coverage probability. The largest deviation for the REML procedure occurred under a heterogeneous covariance structure with a correlation of 0.9 and a sample size of 5. This plot shows that even the REML procedure modelled under the proper covariance structure had some small-sample bias; the bias was slight, however, which was evidence of the consistency of this estimator. Regardless of correlation or number of repeated measures, the heterogeneous covariance structure converged to just above the 95% target coverage value with a sample size of 25. The homogeneous covariance structure with its various correlations converged to 95% with a sample size of 10.

Figure 4.1 shows a downward trend in coverage rates as the number of repeated

Figure 4.1: Coverage probability of confidence intervals for the intercept associated with the REML estimator by sample size and number of repeated measures



measures increased. The REML procedure did not vary far from the target 95% coverage rate. There was bias, however, as shown by a 1% change from three repeated measures to five repeated measures. The heterogeneous covariance structure with its various correlations had a steeper slope in general than its homogeneous counterpart. It appears that the sample size required to obtain the 95% coverage probability grows larger as the number of repeated measures increases. This may be due to the Kenward-Roger adjustment.

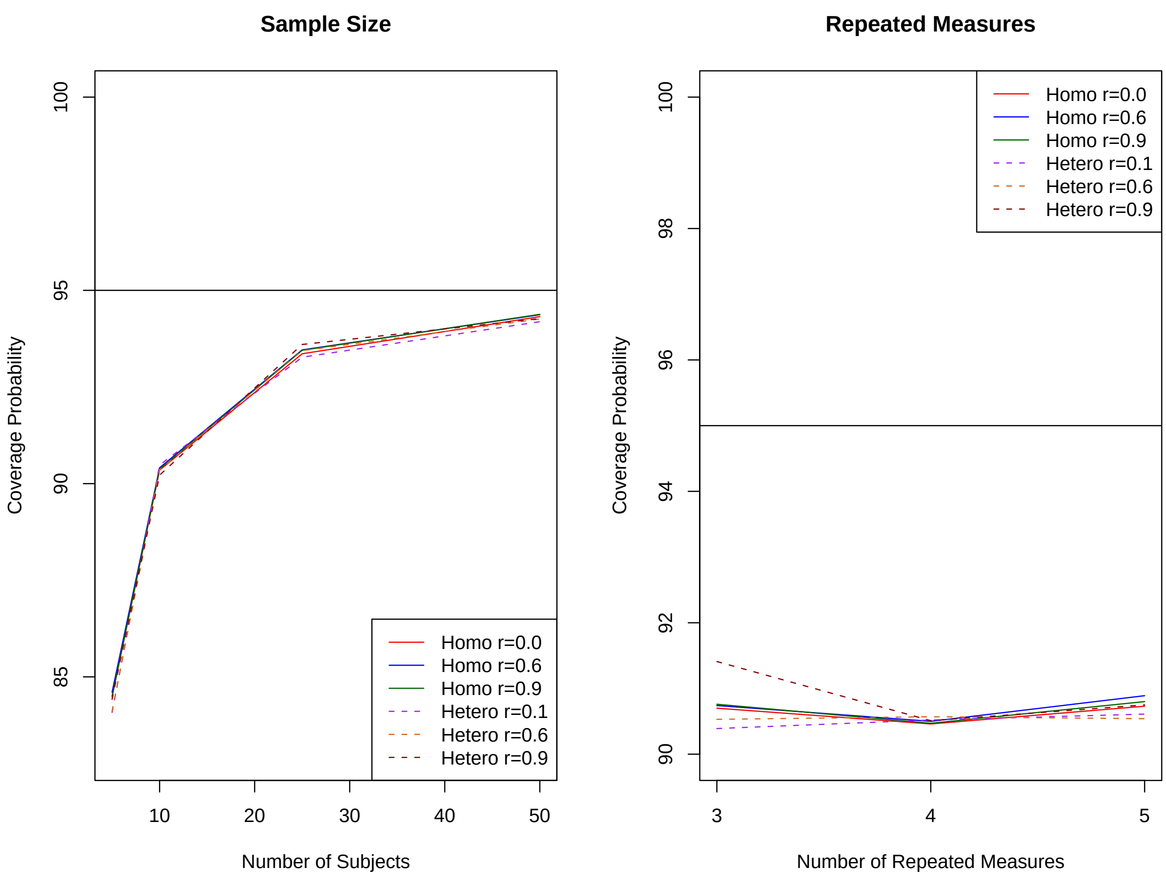
4.1.2 Standard Robust Estimator

The Standard Robust Estimator is the basic form of the four robust estimators that are being examined. Because it has no adjustment to deal with its inherent downward bias, it had the smallest average coverage probability of all of the estimators. This downward bias was extreme with small sample sizes. For example, with the identity matrix as its covariance structure, when $n = 5$ the Standard Robust Estimator yielded a coverage rate of 84.49%, more than 10% below the target level. Under a sample size of 50, the Standard Robust Estimator peaked at 94.42% coverage.

Figure 4.2 shows how consistent the Standard Robust Estimator was. The difference between the six possible covariance structures for each sample size was at most 0.52% with a sample size of five. This estimator produced extremely consistent results regardless of the covariance structure or the amount of correlation within a given covariance structure.

The downward bias of the Standard Robust Estimator was only affected by sample size and not the number of repeated measures. Figure 4.2 illustrates that while there was more variation among the covariance structures and their respective correlations when considering the number of repeated measures, this estimator is reasonably static.

Figure 4.2: Coverage probability of confidence intervals for the intercept associated with the Standard Robust Estimator by sample size and number of repeated measures

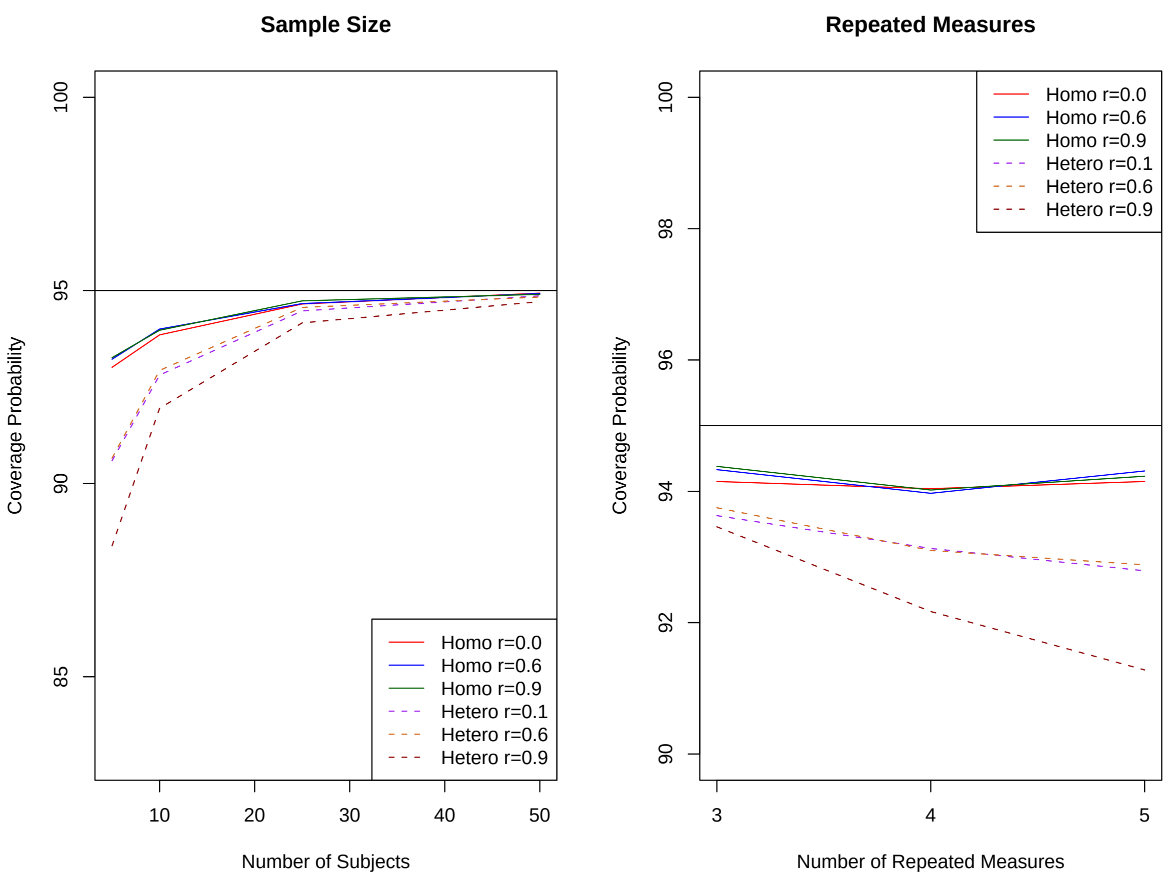


4.1.3 `Empirical` Option

The `empirical` option in the MIXED procedure of SAS is the Standard Robust Estimator with a working covariance matrix included in the equation. Coverage rates for this estimator had a shape similar to the Standard Robust Estimator, but there were a few differences between the two estimators. Figure 4.3 shows the plots of the `empirical` option coverage rates for sample sizes and repeated measures. The `empirical` option had a less significant small-sample bias, regardless of the covariance structure. The data sets with the homogeneous covariance structure, regardless of the level of correlation, were not as biased downward as the data sets with the heterogeneous covariance structure. The homoscedastic models produced nearly equivalent results. The heteroscedastic model with the high level of correlation had the most extreme small-sample bias for this estimator, falling to a coverage level of 88.39%. The minimal and moderate levels of correlation produced almost identical results. All of the covariance structures converged to just below the 95% target coverage value. This estimator was not as consistent across the changing covariance structure as the Standard Robust Estimator. The working covariance matrix used by the `empirical` option reduced the downward bias of the Robust estimator, but this reduction introduced a sensitivity to heteroscedasticity.

The same pattern of the reduction in the downward bias adding a sensitivity to heteroscedasticity held true for the relationship the `empirical` option had with the number of repeated measures. As the number of repeated measures increased from three to five, the heterogeneous covariance matrix with a correlation of 0.9 exhibited decreasing coverage probability from 93.46% to 91.28%. For heterogeneous covariance matrices with minimal and moderate levels of correlation, coverage probability decreased at almost equivalent rates, decreasing by 0.84% and 0.87%, respectively. The three homogeneous covariance structures produced static results, changing 0.36% or less. This estimator, though not as consistent as the Standard Robust Estimator,

Figure 4.3: Coverage probability of confidence intervals for the intercept associated with the `empirical` option of the MIXED procedure in SAS by sample size and number of repeated measures



decreased the downward bias inherent in its base estimator.

4.1.4 Long Estimator

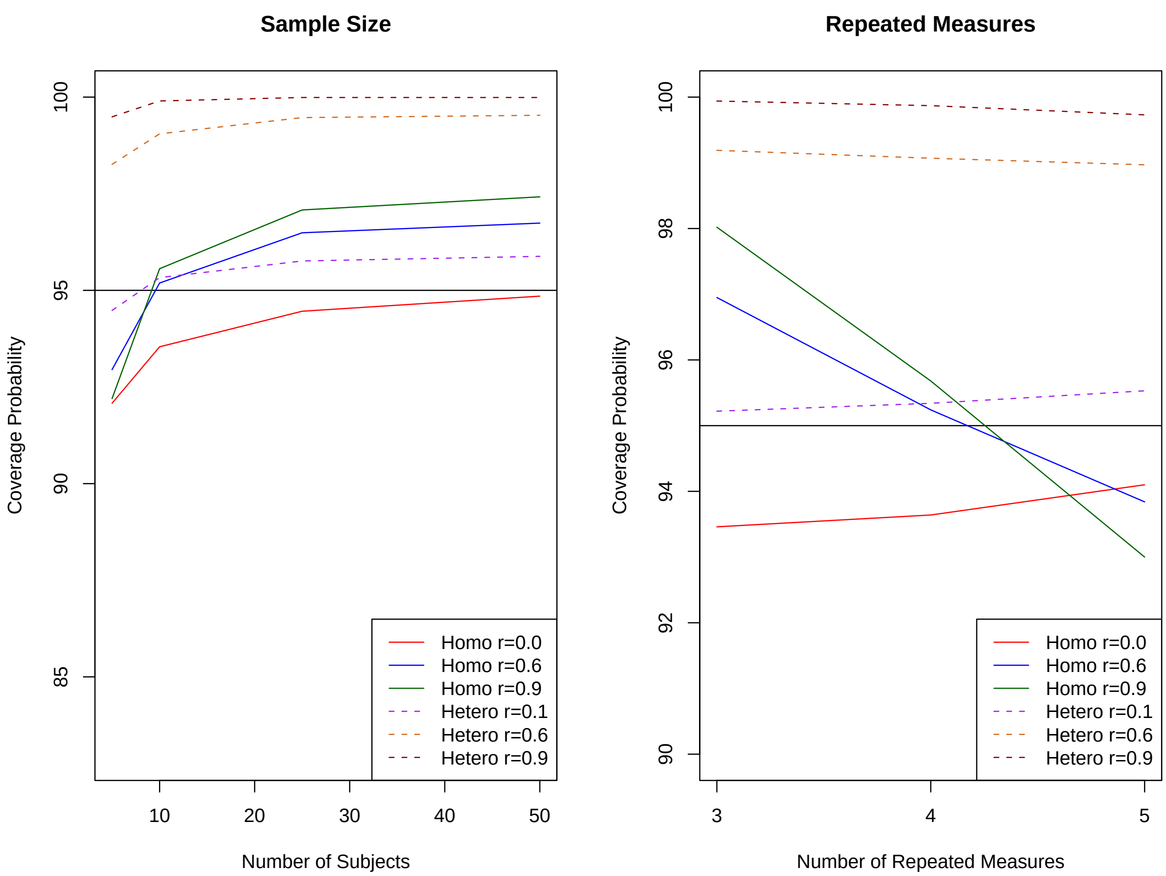
The Long estimator is the Standard Robust Estimator, except the covariance matrix uses standardized residuals instead of ordinary residuals. This change in the residuals considerably reduced the downward bias of the Standard Robust Estimator (Figure 4.4). However, the results were inconsistent among covariance structures. When the identity was the covariance structure, there was small-sample bias, but coverage rates converged to 95%. When nonzero correlation was added to this base structure, the coverage probability increased with respect to the amount of correlation added to the covariance matrix. The base heteroscedastic matrix converged to 95.88%. Adding correlation to this model increased the coverage probability and reduced the effect of sample size. This suggests the need for a different method of standardization in this estimator.

With respect to the number of repeated measures in the covariance matrix, the coverage probabilities that the Long estimator exhibited were somewhat unpredictable. The repeated measures plot in Figure 4.4 illustrates this instability. The homoscedastic covariance structure with correlations of 0.6 and 0.9 failed to maintain a constant coverage probability as the number of repeated measures increased. Coverage rates for these two structures fell from 96.95% and 98.02% to 93.84% and 93.00%, respectively. All of the other structures exhibited stable coverage rates. Coverage rates for the equivalent heterogeneous structures were very near 100% coverage probability and did show a slight downward bias.

4.1.5 Long Estimator with Quantile Adjustment

Coverage rates for the Long estimator with the quantile adjustment should mimic those for the Long estimator, but should be about 4% higher than the Long

Figure 4.4: Coverage probability confidence intervals for the intercept associated with the Long estimator by sample size and number of repeated measures



estimator. The Long estimator with the quantile adjustment performed as expected, except when the Long estimator had more than 96% coverage probability.

Figure 4.5 shows the sample size and repeated measures plots for the Long estimator with the quantile adjustment. The Long estimator with the quantile adjustment mimicked the Long estimator in regards to repeated measures.

4.1.6 Summary of the Intercept

Table 4.1 lists coverage rates for the intercept relative to sample size, and Figure 4.1 contains plots of each estimator at each sample size for the covariance structures. The REML procedure had a slight small-sample effect, but it was the most consistent of the five estimators. The Standard Robust Estimator had the smallest coverage probability of all the estimators. This estimator had an extreme downward bias with small sample sizes. The small-sample downward bias that was inherent in the Robust estimator was less apparent for the `empirical` option, especially when the data were from a homogeneous covariance structure. Having a heterogeneous covariance structure seemed to force the Long estimator to be more consistent but only because the coverage rates were close to 100%. The dramatic increase in the coverage probability of the Long estimator was unexpected because this estimator was slightly downward biased with sample sizes much larger than 50 (Long 2000). This suggests that a more elaborate method of standardization might be needed. The Long estimator with the quantile adjustment mimicked the Long estimator with coverage rates, staying about 4% above the Long estimator.

Table 4.2 shows the coverage probabilities of each of the estimators across the various levels of correlation for the homogeneous and heterogeneous covariance structures, as well as across numbers of repeated measures. Figure 4.7 contains plots of coverage rates of each estimator at each value of repeated measures for the respective covariance structures in terms of the intercept.

Figure 4.5: Coverage probability of confidence intervals for the intercept associated with the Long estimator with the quantile adjustment by sample size and number of repeated measures

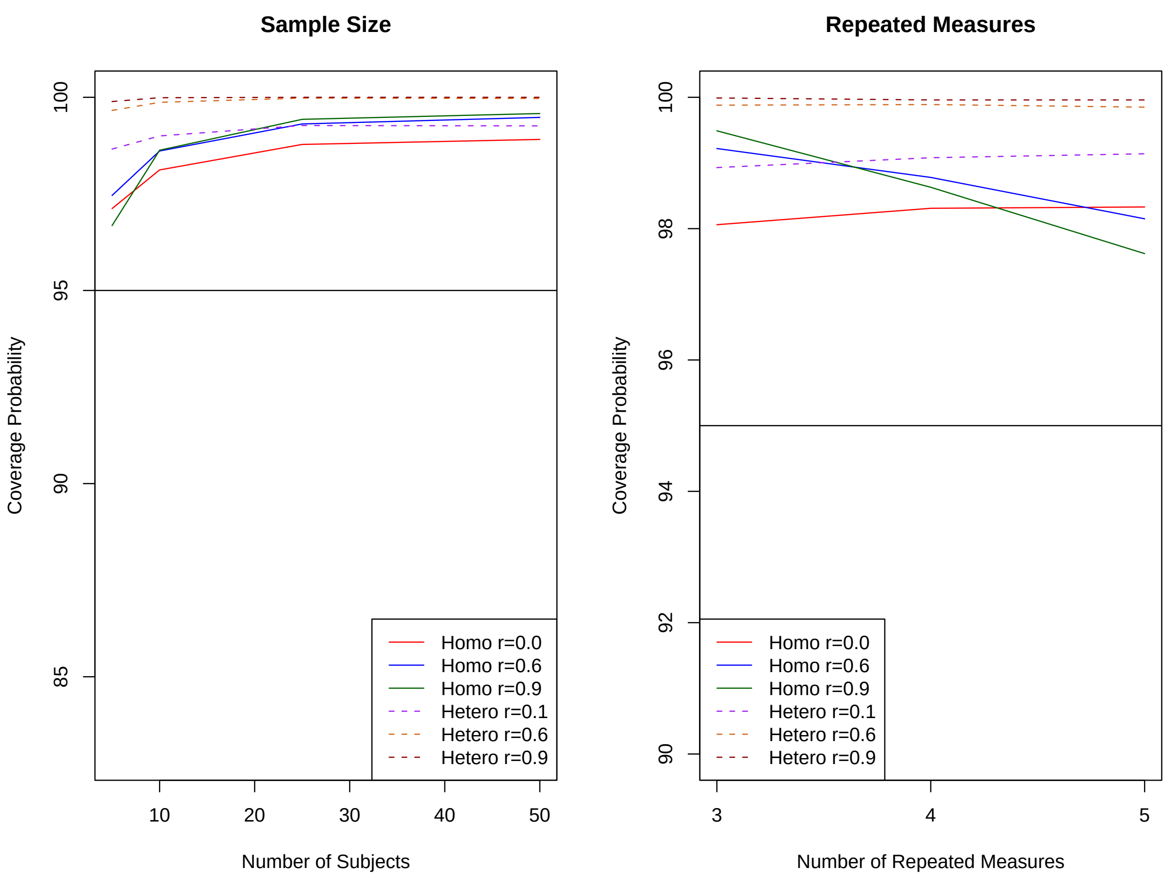


Figure 4.6: Coverage probabilities of each estimator for the intercept with respect to sample size

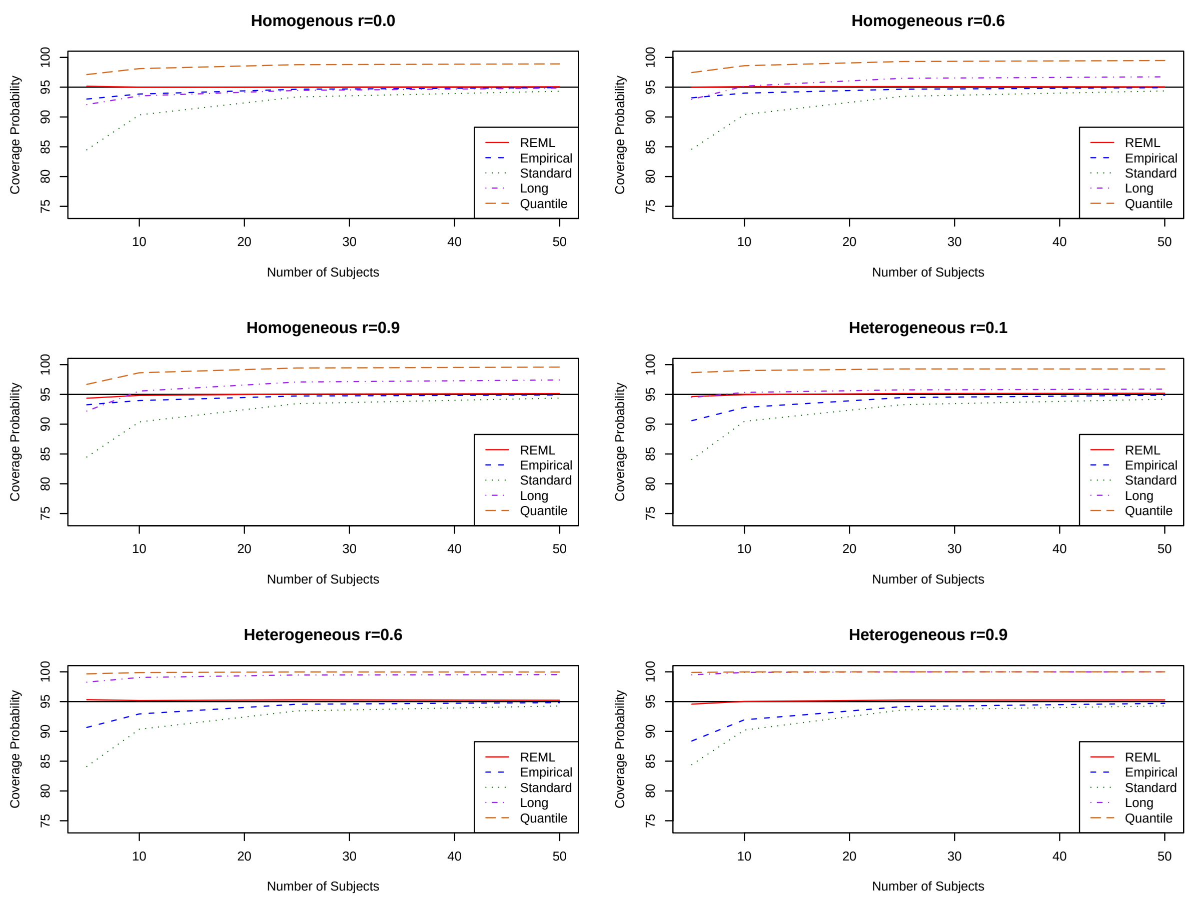
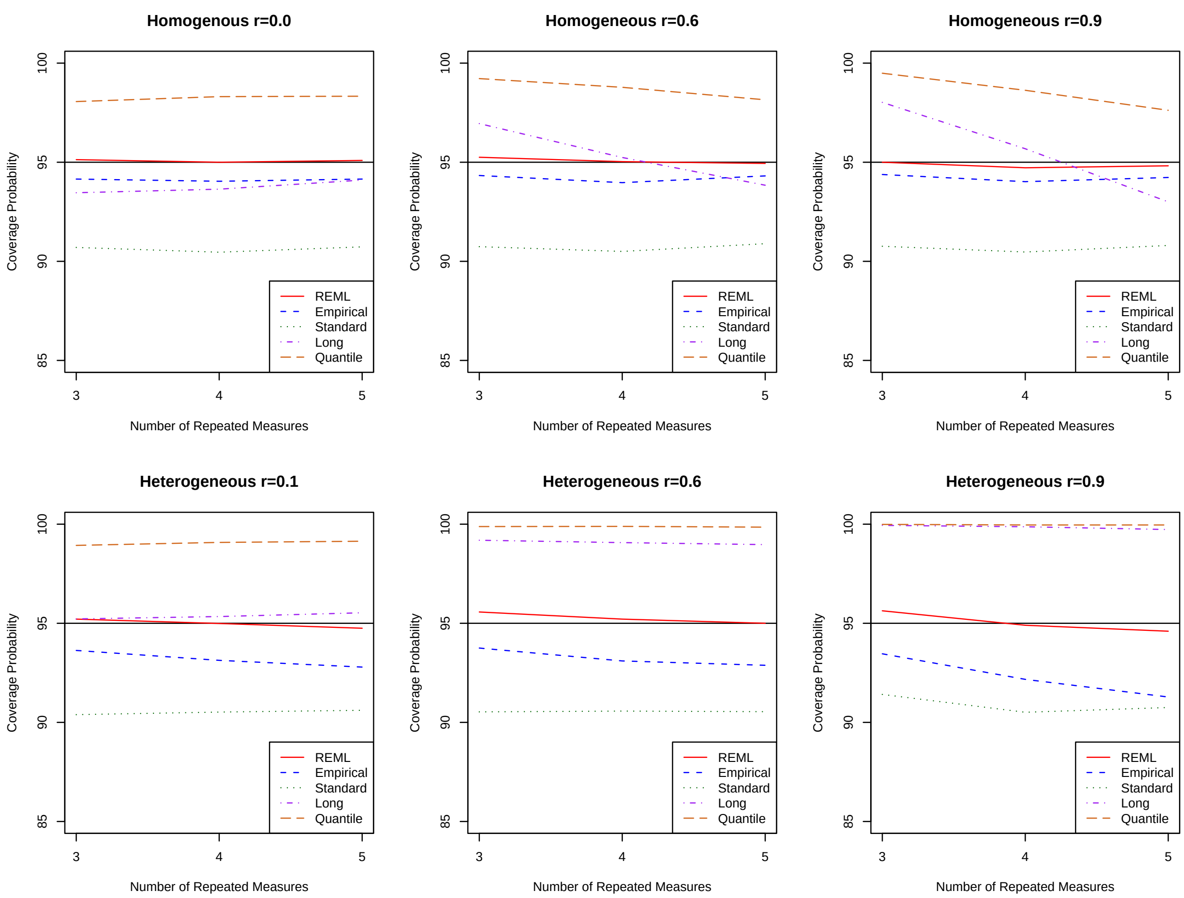


Figure 4.7: Coverage probabilities of each estimator for the intercept with respect to the number of repeated measures



The REML procedure was consistent across repeated measures. The Standard Robust Estimator was consistent across repeated measures, but it had the lowest coverage probability of all five of the estimators. The `empirical` option of the MIXED procedure in SAS had a slight downward trend as the number of repeated measures increased. The Long estimator was biased downward as the number of repeated measures increased. The quantile adjustment to the Long estimator mimicked the Long estimator.

4.2 Slope

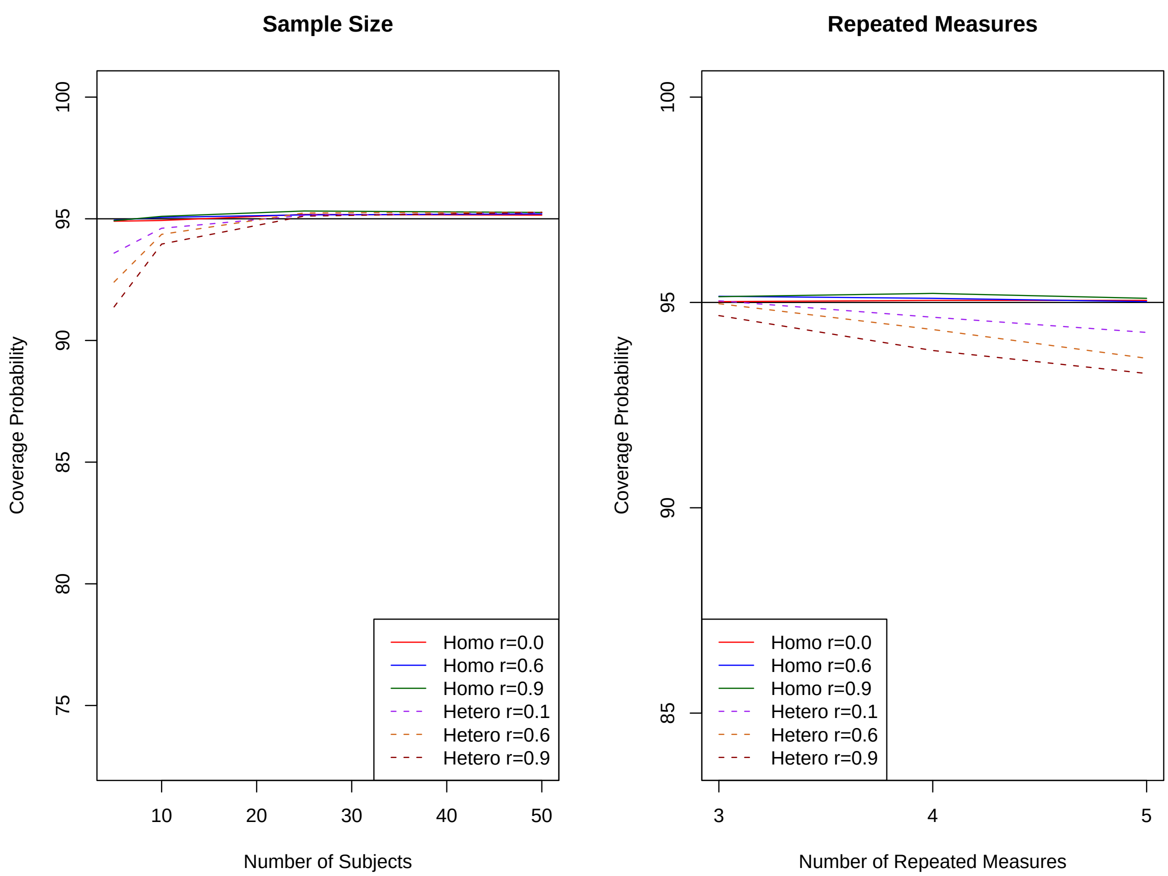
The slope is a changing column. Properties of the estimators are enhanced when examining the slope that may not have been seen when examining the intercept. Each estimator is examined individually, then the estimators are compared.

4.2.1 REML

Figure 4.8 shows plots of each of the six covariance structures by sample and repeated measures for the REML procedure. The REML procedure showed a more distinct small-sample bias for the slope than it did for the intercept. With a sample size of 25, all six covariance structures had converged to just above the target 95% coverage probability. The heterogeneous covariance structures had a much greater small-sample effect than the homogeneous covariance structures. The smallest coverage rate was for the heterogeneous covariance structure with a correlation of 0.9, which reached a coverage level of 91.38%. The lowest coverage probability for any homogeneous structure was 94.90%.

With respect to the number of repeated measures in a data set, the REML procedure remained above the 95% coverage rate for the three homogeneous structures. The three heterogeneous structures were not able to maintain the same coverage probability as the three homogeneous structures. However, even with a correlation of 0.9

Figure 4.8: Coverage probability of confidence intervals for the slope associated with the REML procedure by sample size and number of repeated measures



and five repeated measures in the heterogeneous covariance structure, the REML procedure had a coverage rate of 93.27%. This estimator was stable, but lost efficiency as the number of repeated measures increased.

4.2.2 Standard Robust Estimator

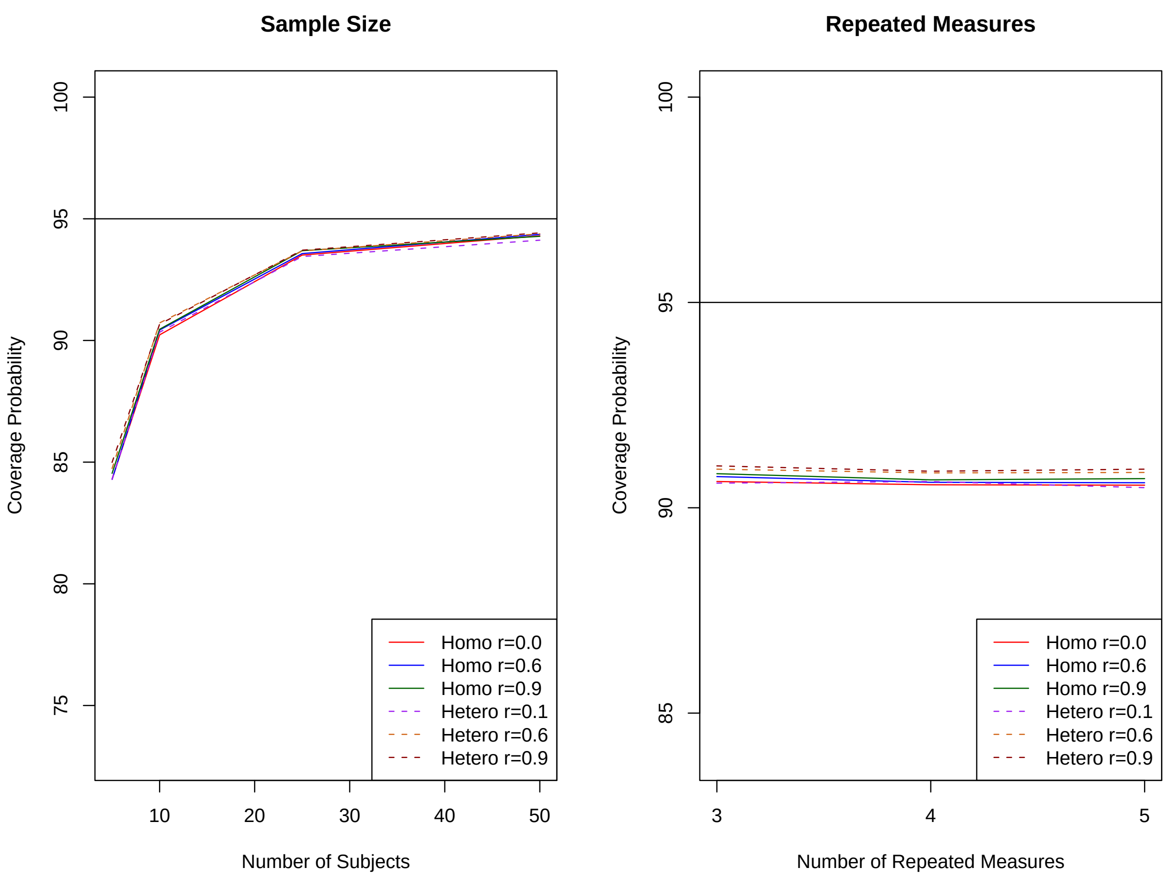
The Standard Robust Estimator was an extremely consistent estimator. As shown in Figure 4.9, the estimates for the six covariance structures lie almost on top of each other when calculated according to sample size. There was a small-sample bias, with the lowest coverage rate reaching 84.29%. The highest coverage average reached 94.42%.

As the number of repeated measures increased, the six covariance structures were static. However, the estimators did not lie on top of each other as they did with the increase in sample size. As correlation and heteroscedasticity were increased, the coverage probability slightly increased. The coverage rate remained between 90.49% and 91.02%. The number of repeated measures did not significantly impact the ability of the Standard Robust Estimator to reach its asymptotic coverage probability.

4.2.3 `Empirical` option

The `empirical` option of the MIXED procedure in SAS yielded results for the slope similar to the results for the intercept. Figure 4.10 shows coverage rates of confidence intervals for the six covariance structures associated with the `empirical` option by the sample size and the number of repeated measures. With respect to the sample size, the homogeneous covariance structures lay on top of each other. They had a small-sample bias that decreased the coverage probability to 87.25%. The homogeneous covariance matrices converged to 94.54%. The heterogeneous covariance matrices converged to 94%; however, the small-sample bias was not consistent among the three heterogeneous covariance matrices. The most biased of the three

Figure 4.9: Coverage probability of confidence intervals for the slope associated with the Standard Robust Estimator by sample size and number of repeated measures



heteroscedastic structures was the structure with a correlation of 0.9, for which the coverage probability reached to only 74.52% when the sample size was 5.

The number of repeated measures in the covariance matrix decreased the ability of the `empirical` option to reach the target 95% coverage level. The three homogeneous covariance structures lay on top of each other. There was a slight downward bias as the number of repeated measures increased. The slope for the three heterogeneous covariance structures steepened as the number of repeated measures increased. The lowest coverage rate was 84.32%, more than 10% below the target 95% coverage level.

4.2.4 Long Estimator

The Long estimator had no consistency about the covariance structures. Figure 4.11 illustrates this inconsistency with respect to the sample size and the number of repeated measures. When the sample sizes were small, the homogeneous and heterogeneous covariance structures at the low level of correlation were the only structures with coverage rates below the target 95% level. Adding correlation to these structures increased the coverage rate. The coverage probability for the homogeneous covariance structures with moderate and high correlation had higher coverage probabilities than their heterogeneous counterparts. This estimator tended to have a higher coverage rate than the target 95% rate.

The number of repeated measures also had an impact on this estimator. When the covariance structure was homogeneous with moderate and high levels of correlation, the coverage probability was essentially stationary around 99.25% and 99.96%, respectively. The heterogeneous covariance structure with moderate and high levels of correlation had a downward bias. With moderate correlation, the heterogeneous covariance structure changed coverage probabilities from 97.85% to 96.69% as the number of repeated measures increased. The high level of correlation for the hetero-

Figure 4.10: Coverage probability of confidence intervals for the slope associated with the `empirical` option in the MIXED procedure of SAS by sample size and number of repeated measures

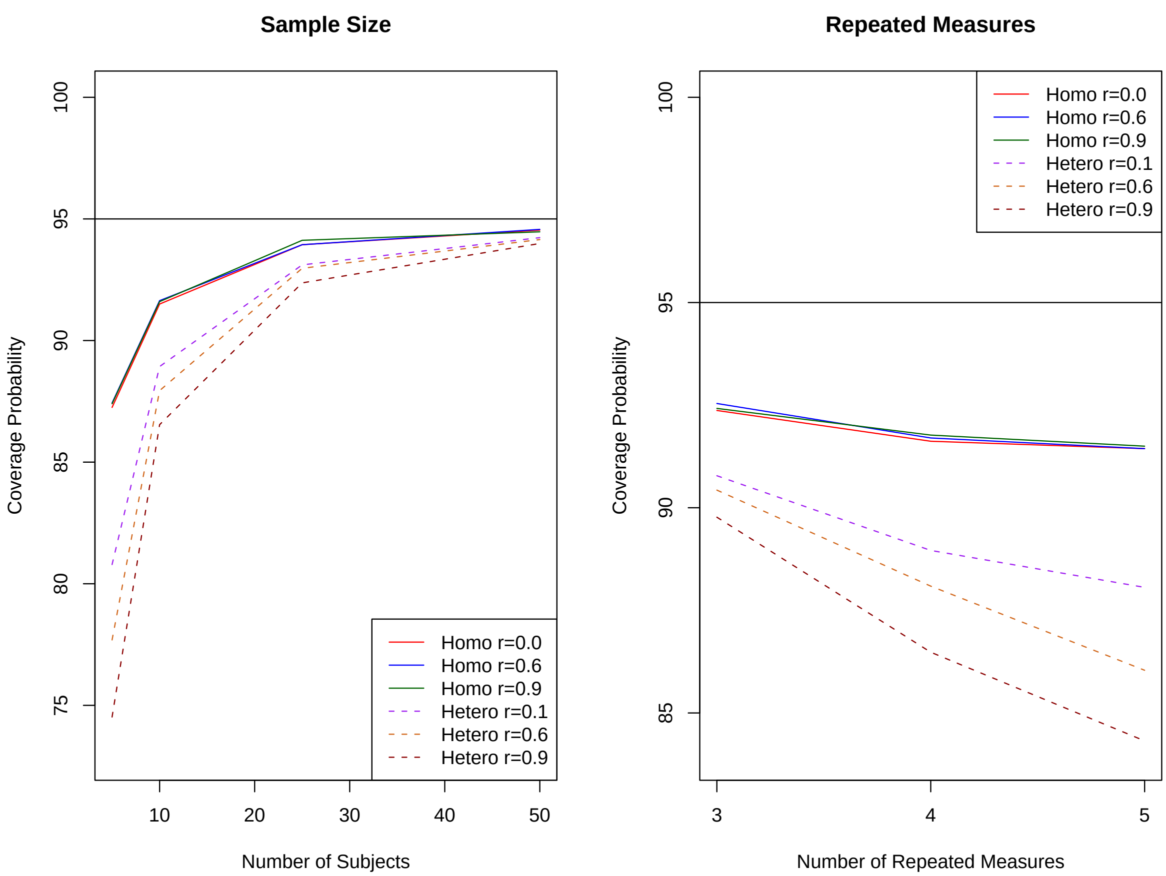
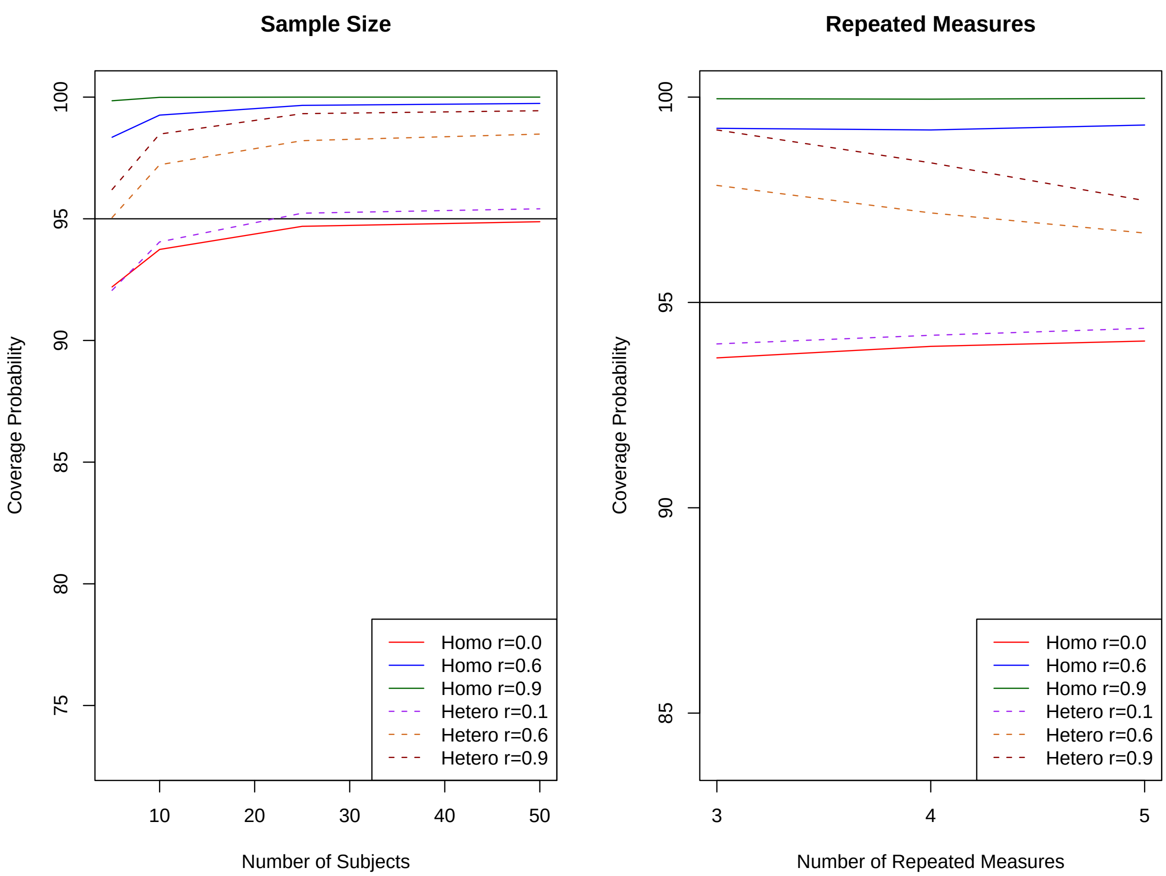


Figure 4.11: Coverage probability of confidence intervals for the slope associated with the Long estimator by sample size and number of repeated measures



geneous covariance structure decreased coverage probability from 99.20% to 97.48% as the number of repeated measures increased. The homogeneous and heterogeneous covariance structures at the low level of correlation had a slight upward bias, increasing by 0.41% and 0.38%, respectively. This is evidence of the unpredictability of the coverage rate of the Long estimator.

4.2.5 Long Estimator with Quantile Adjustment

The Long estimator with the quantile adjustment mimicked the Long estimator. As shown in Figure 4.12, no covariance structure dropped below 97.20%. This adjustment worked as designed, with about a 4% increase in coverage rate.

4.2.6 Summary of the Slope

Table 4.3 displays the results for each of the estimators by sample size over the six covariance structures. Figure 4.13 illustrates the effect of the six covariance structures with respect to the sample size.

The homogeneous covariance structure with a correlation of zero showed that the five estimators converged to the 95% level when the sample size was 50. The quantile adjustment to the Long estimator and the REML procedure both achieved the 95% coverage level when the sample size was small. The remaining three estimators (the Long estimator, the `empirical` option in SAS, and the Standard Robust Estimator) showed a small-sample effect. These estimators eventually converged to just below the target 95% coverage level. The REML procedure was stationary around the 95% target rate.

Table 4.4 lists each estimator and their coverage probabilities across the varying levels of correlation and numbers of repeated measures. As the number of repeated measures increased, there did not seem to be much change within most of the estimators.

Figure 4.12: Coverage probability of confidence intervals for the slope associated with the Long estimator with the quantile adjustment by sample size and number of repeated measures

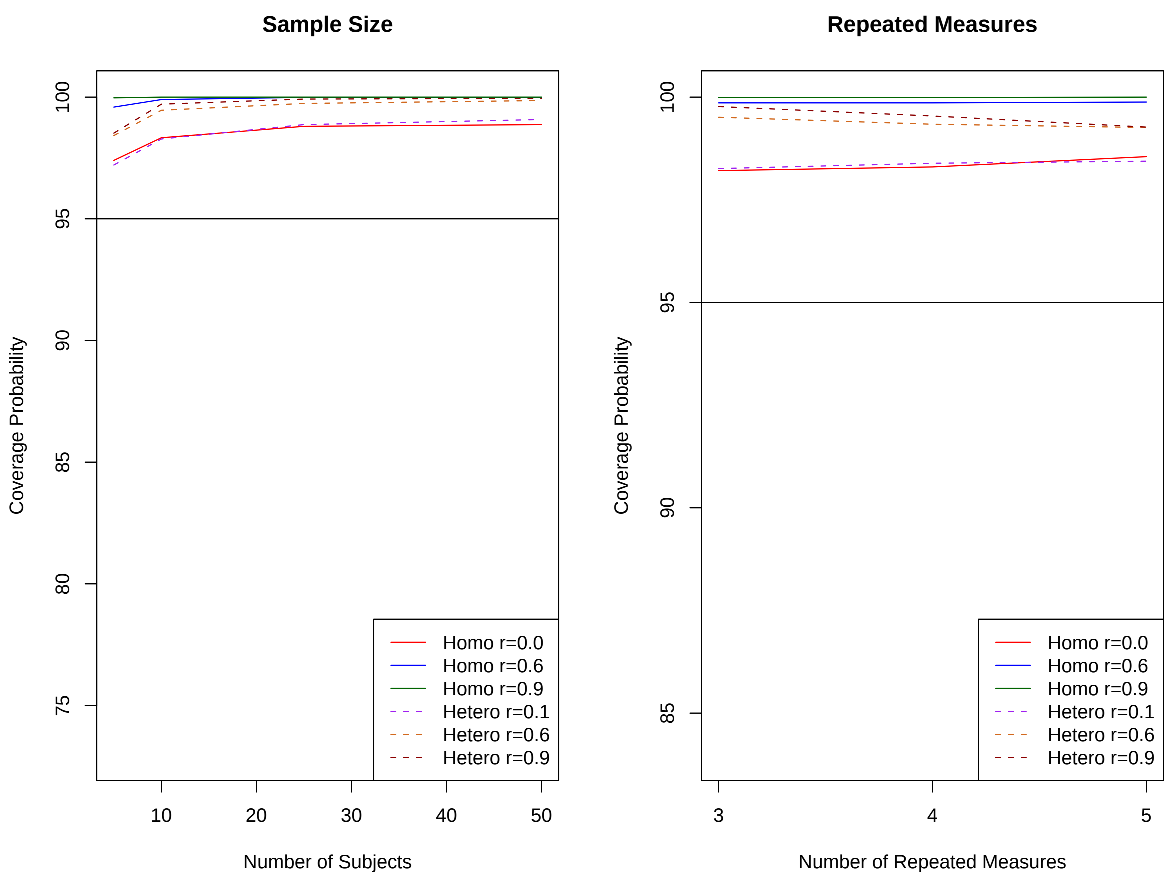


Figure 4.13: Coverage probabilities for the slope with respect to sample Size

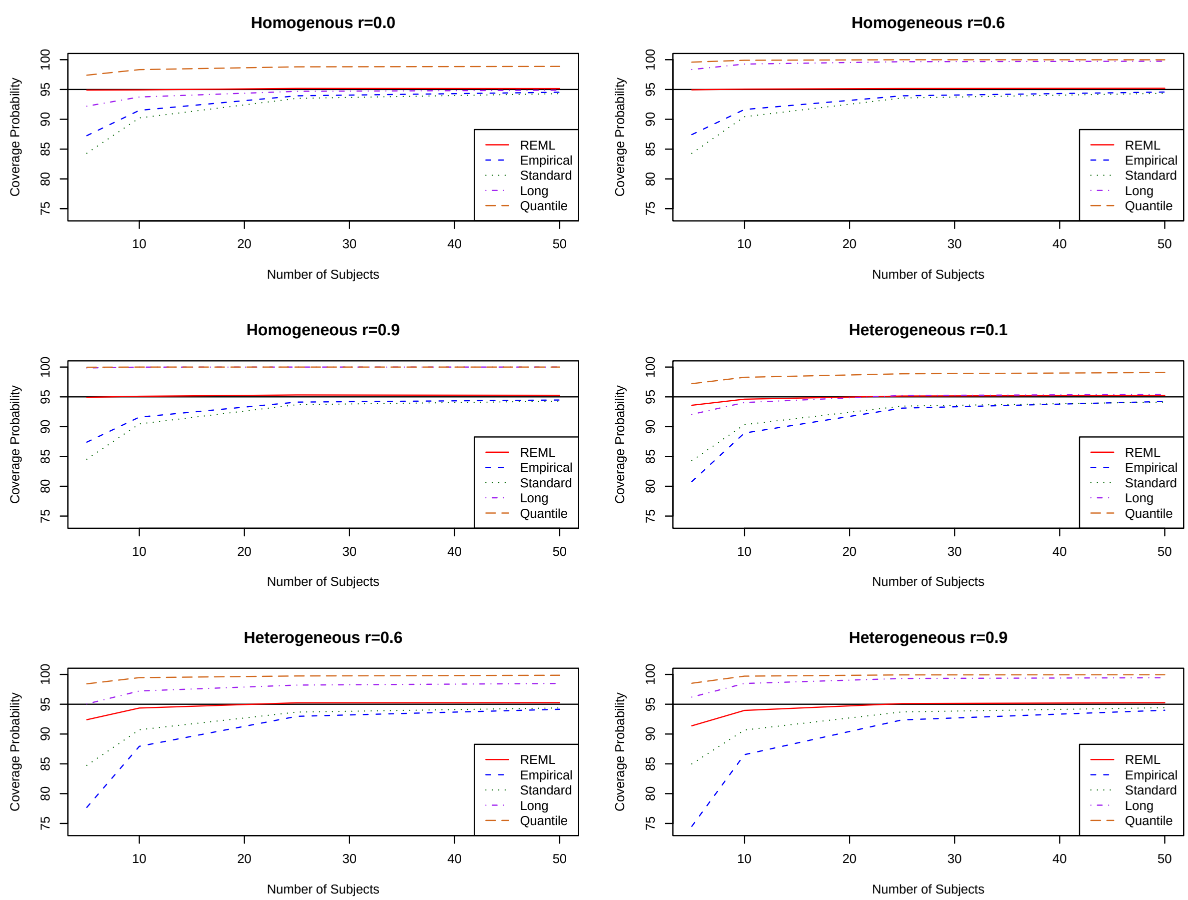


Figure 4.14 shows the effect of the six covariance structures on the coverage rates of the five estimators with an increasing number of repeated measures. When the correlation was at its lowest level, the estimators were fairly flat. The quantile adjustment to the Long estimator and the REML procedure were the only two estimators to achieve the 95% coverage level. The `empirical` estimator performed worse than the Standard Robust Estimator when the covariance structure was heterogeneous. The Standard Robust Estimator, which slightly surpassed the 90% coverage probability level, could not maintain the coverage rates of the other estimators.

Within the homogeneous covariance structure, all five of the estimators remained reasonably static. They did not change much when the number of repeated measures was increased. As the correlation and number of repeated measures increased within the heterogeneous covariance structure, four of the five estimators developed a downward bias. The fifth estimator, the Standard Robust Estimator, remained static regardless of changes to the covariance structure or the number of repeated measures.

Figure 4.14: Coverage probabilities for the slope with respect to the number of repeated measures

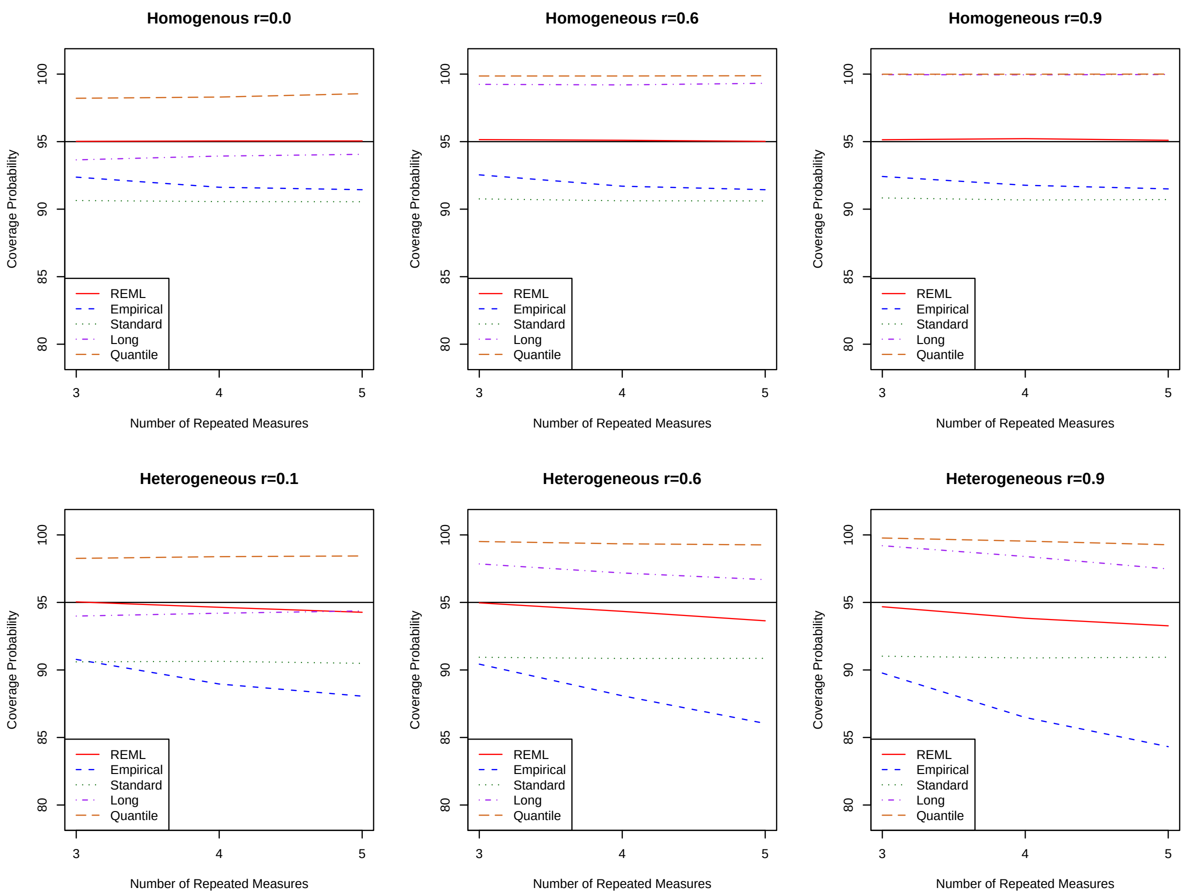


Table 4.1: Coverage rates for the intercept averaged over the number of repeated measures

Method	Covariance Structure: Type	Covariance Structure: Correlation	Sample Size: 5	Sample Size: 10	Sample Size: 25	Sample Size: 50
REML	Homogeneous	0.0	95.16	95.01	95.02	95.08
		0.6	95.01	95.10	95.11	95.06
		0.9	94.36	94.83	95.05	95.14
	Heterogeneous	0.1	94.64	94.95	95.15	95.19
		0.6	95.33	95.18	95.30	95.23
		0.9	94.58	95.03	95.26	95.29
Standard Robust	Homogeneous	0.0	84.49	90.35	93.36	94.32
		0.6	84.60	90.41	93.46	94.37
		0.9	84.49	90.38	93.45	94.38
	Heterogeneous	0.1	84.08	90.48	93.27	94.19
		0.6	84.11	90.36	93.44	94.26
		0.9	84.42	90.22	93.60	94.27
Empirical	Homogeneous	0.0	93.01	93.85	94.65	94.93
		0.6	93.22	94.00	94.66	94.92
		0.9	93.26	93.97	94.73	94.90
	Heterogeneous	0.1	90.59	92.81	94.47	94.86
		0.6	90.66	92.93	94.56	94.83
		0.9	88.39	91.95	94.16	94.71
Long Adjustment	Homogeneous	0.0	92.08	93.54	94.46	94.85
		0.6	92.95	95.19	96.49	96.74
		0.9	92.20	95.56	97.08	97.42
	Heterogeneous	0.1	94.48	95.33	95.76	95.88
		0.6	98.26	99.05	99.47	99.53
		0.9	99.49	99.90	99.99	99.99
Long Quantile Adjustment	Homogeneous	0.0	97.12	98.12	98.78	98.91
		0.6	97.46	98.61	99.31	99.48
		0.9	96.68	98.63	99.43	99.58
	Heterogeneous	0.1	98.66	99.00	99.27	99.26
		0.6	99.66	99.87	99.98	99.97
		0.9	99.89	99.99	100.00	100.00

Table 4.2: Coverage rates for the intercept averaged over the sample size

	Covariance Structure		Repeated Measure		
Method	Type	Correlation	3	4	5
REML	Homogeneous	0.0	95.13	95.00	95.09
		0.6	95.25	95.03	94.94
		0.9	95.00	94.72	94.82
	Heterogeneous	0.1	95.21	94.99	94.75
		0.6	95.57	95.21	95.00
		0.9	95.63	94.90	94.60
Standard Robust	Homogeneous	0.0	90.70	90.46	90.73
		0.6	90.74	90.50	90.89
		0.9	90.76	90.47	90.80
	Heterogeneous	0.1	90.39	90.52	90.61
		0.6	90.53	90.57	90.54
		0.9	91.41	90.51	90.75
Empirical	Homogeneous	0.0	94.15	94.04	94.15
		0.6	94.33	93.97	94.31
		0.9	94.38	94.02	94.23
	Heterogeneous	0.1	93.63	93.13	92.79
		0.6	93.75	93.10	92.88
		0.9	93.46	92.17	91.28
Long Adjustment	Homogeneous	0.0	93.46	93.64	94.10
		0.6	96.95	95.24	93.84
		0.9	98.02	95.68	93.00
	Heterogeneous	0.1	95.22	95.34	95.53
		0.6	99.19	99.07	98.97
		0.9	99.94	99.87	99.73
Long Quantile Adjustment	Homogeneous	0.0	98.06	98.31	98.33
		0.6	99.22	98.78	98.15
		0.9	99.49	98.63	97.62
	Heterogeneous	0.1	98.93	99.08	99.14
		0.6	99.88	99.89	99.85
		0.9	99.99	99.96	99.96

Table 4.3: Coverage rates for the Slope averaged over the number of repeated measures

	Covariance Structure		Sample Size			
Method	Type	Correlation	5	10	25	50
REML	Homogeneous	0.0	94.90	94.93	95.18	95.15
		0.6	94.94	95.06	95.16	95.20
		0.9	94.92	95.10	95.32	95.26
	Heterogeneous	0.1	93.59	94.61	95.14	95.25
		0.6	92.40	94.36	95.24	95.27
		0.9	91.38	93.96	95.11	95.26
Standard Robust	Homogeneous	0.0	84.30	90.23	93.51	94.30
		0.6	84.29	90.43	93.57	94.36
		0.9	84.54	90.46	93.69	94.28
	Heterogeneous	0.1	84.29	90.34	93.45	94.12
		0.6	84.74	90.72	93.68	94.38
		0.9	84.99	90.67	93.71	94.42
Empirical	Homogeneous	0.0	87.25	91.50	93.94	94.54
		0.6	87.43	91.64	93.94	94.57
		0.9	87.40	91.60	94.12	94.47
	Heterogeneous	0.1	80.79	88.93	93.11	94.23
		0.6	77.69	87.94	92.97	94.15
		0.9	74.52	86.54	92.37	93.99
Long Adjustment	Homogeneous	0.0	92.20	93.74	94.69	94.88
		0.6	98.35	99.26	99.66	99.74
		0.9	99.85	99.99	100.00	100.00
	Heterogeneous	0.1	92.06	94.05	95.23	95.41
		0.6	95.05	97.22	98.21	98.48
		0.9	96.20	98.48	99.32	99.44
Long Quantile Adjustment	Homogeneous	0.0	97.40	98.33	98.80	98.87
		0.6	99.59	99.90	99.99	99.98
		0.9	99.97	100.00	100.00	100.00
	Heterogeneous	0.1	97.21	98.28	98.87	99.08
		0.6	98.42	99.46	99.74	99.86
		0.9	98.52	99.71	99.92	99.95

Table 4.4: Coverage rates for the slope averaged over sample size

	Covariance Structure		Repeated Measure		
Method	Type	Correlation	3	4	5
REML	Homogeneous	0.0	95.02	95.05	95.05
		0.6	95.15	95.10	95.02
		0.9	95.14	95.22	95.10
	Heterogeneous	0.1	95.04	94.64	94.27
		0.6	94.97	94.34	93.64
		0.9	94.68	93.83	93.27
Standard Robust	Homogeneous	0.0	90.64	90.56	90.55
		0.6	90.76	90.62	90.61
		0.9	90.83	90.68	90.71
	Heterogeneous	0.1	90.60	90.64	90.49
		0.6	90.94	90.85	90.86
		0.9	91.02	90.89	90.94
Empirical	Homogeneous	0.0	92.37	91.62	91.44
		0.6	92.54	91.70	91.44
		0.9	92.42	91.77	91.50
	Heterogeneous	0.1	90.78	88.96	88.06
		0.6	90.43	88.09	86.04
		0.9	89.77	86.48	84.32
Long Adjustment	Homogeneous	0.0	93.65	93.93	94.06
		0.6	99.24	99.20	99.32
		0.9	99.96	99.95	99.97
	Heterogeneous	0.1	93.99	94.20	94.37
		0.6	97.85	97.18	96.69
		0.9	99.20	98.40	97.48
Long Quantile Adjustment	Homogeneous	0.0	98.21	98.30	98.55
		0.6	99.86	99.86	99.88
		0.9	99.99	99.99	100.00
	Heterogeneous	0.1	98.26	98.39	98.44
		0.6	99.51	99.34	99.26
		0.9	99.77	99.54	99.27

5. CONCLUSION

5.1 Estimators

Each estimator reacts differently to increases in sample size, repeated measures, correlation, and homoscedasticity This section summarizes the properties each estimator exhibits. It also includes a recommendation based on the properties of the estimators.

The small-sample results for the intercept and the slope for the REML procedure are similar. The small-sample properties of the REML procedure demonstrate the ability of this procedure to quickly converge to the expected coverage probability. When the sample size is only 5, the REML procedure shows only a slight bias for all covariance structures. When the sample size is 10, coverage rates for the homoscedastic matrices converge to the 95% coverage probability level. For the sample size of 25, results for the heterogeneous covariance matrices converge to the target probability level. The REML procedure stays extremely close to the target coverage level once it converges. However, as the number of repeated measures increases, the REML procedure develops a downward trend, possibly because of the Kenward-Roger adjustment. This means that a different sample size is needed to attain convergence than is needed when there are fewer repeated measures. Regardless of the number of repeated measures, the REML procedure converges under the true covariance structure by the time the sample size reaches 25. The REML procedure works extremely well in this situation because the true covariance structure is known and can be used in the simulation. The REML procedure, however, may not be the best estimator in a situation in which the covariance structure is in question.

The Standard Robust Estimator has an extreme downward bias when the sample size is small. This downward bias decreases as the sample size increases, and seems to converge by the time $n = 50$; however, the Standard Robust Estimator converges to just below the 95% coverage probability level. Despite this downward bias, this estimator is extremely predictable across covariance structures. Little change is seen when the covariance structure is changed or when correlation is added to the covariance structure. The number of repeated measures has little influence on the Standard Robust Estimator; no more than a 1% change in coverage probability occurs when the number of repeated measures is changed.

The `empirical` option of the MIXED procedure in SAS is an enhancement of the consistent estimator that White introduced (SAS Institute 2003). The Standard Robust Estimator and the `empirical` option do not mimic each other in coverage rates. The `empirical` option adds the REML estimate of the covariance matrix as a working covariance matrix. This variation of the Standard Robust Estimator changes the consistency of the Robust estimator with the smaller sample sizes. The `empirical` option with the small sample sizes is much more variable than the Standard Robust Estimator. The homogeneous covariance structures have a larger coverage probability than the heterogeneous covariance structures. As the correlation increases, the downward bias of the estimator increases—especially with the heteroscedastic structure. Overall, the `empirical` option has a higher coverage probability than the Standard Robust Estimator when the covariance structure is homogeneous. The Standard Robust Estimator has a higher coverage probability when the covariance structure is heterogeneous.

The Long estimator uses standardized residuals in the computation of the covariance matrix instead of the residuals used in the ordinary least squares computation of the covariance matrix. This change in residuals reduces the downward bias of the Standard Robust Estimator. Although this change does reduce the downward bias, it is not able to maintain the same properties as the Standard Robust Estimator. The Long estimator yields some very unexpected results. With a homogeneous covariance structure and no correlation in the model, the Long estimator appears similar to the Standard Robust Estimator, but has a smaller downward bias. However, when correlation is added to the model, the coverage probability increases by 2% or more. When heteroscedasticity is added to the model, a similar result occurs. The higher the correlation, especially when heteroscedasticity is present, the more extreme the increase in coverage over the Standard Robust Estimator. This increase quickly reaches the point where the Long estimator has a coverage probability of more than 99%. If the expected length of the interval does not increase much, a coverage rate greater than 95% is adequate. While looking at the number of repeated measures per subject, the Long estimator is static when a heterogeneous covariance structure is used. Conversely, when a homogeneous covariance structure is used, especially with correlation, the Long estimator exhibits a downward bias. These results are quite interesting, but they make it difficult to predict the extremity of the bias even though the direction of bias (upward or downward) is more easily ascertained.

The quantile adjustment to the Long estimator is computed by using a higher quantile than the desired level in order to compensate for the downward bias of the estimator. The quantile adjustment mimics the Long estimator, but more often produces coverage rates near 100% (the Long estimator produces rates over 99%).

No estimator is universally preferable. The Standard Robust Estimator works well for large sample sizes under all covariance structures. The Long estimator works well for all situations, but the expected length of the confidence intervals may get too large under heterogeneous covariance structures. The `empirical` estimator works well for all homogeneous covariance structures.

When the covariance structure is known, my recommendation is to use the REML procedure. This procedure is accurate even with sample sizes as small as 10. The Long estimator should not be used unless the size of the confidence interval does not matter. The `empirical` option in the MIXED procedure of SAS should only be used with homoscedastic data when a large sample size is available.

5.2 Example

Now that the properties of the estimators have been detailed, recall the teratology example from the introduction. Mice pups were split into 4 treatments. Treatment 1 includes 11 mice. This treatment is used in Table 1.1 to compare the different estimators.

The first four estimators are the REML procedure under four different covariance structures. Without knowing the actual underlying covariance structure, the AIC is used to select the appropriate structure. In this example, the smallest AIC occurs with an unstructured covariance structure. If the underlying covariance in this example is truly unstructured, the REML procedure using this structure provides the most accurate estimates of the standard errors that the REML procedure can provide.

Since the AIC chose the unstructured covariance structure, it is reasonable to believe that there is heterogeneity within this data. The Standard Robust Estimator is known to be downward biased. With only 11 mice pups in the sample, this estimator is expected to produce standard errors smaller than they actually are. The same is expected of the `empirical` option of the MIXED procedure in SAS. The Standard Robust Estimator and the `empirical` option provide similar standard errors. The Long estimator, as shown previously, reduces the downward bias of the Standard Robust Estimator, but it also inflates the standard errors of the estimates when heterogeneity and/or correlation is present within the data.

The question is how which of the known biases is more acceptable in the standard errors? This depends on the researcher. For this example, my recommendation would be to use the Standard Robust Estimator or the `empirical` option, as they provide similar results. These two robust estimators also provide results similar to the AIC-chosen REML model. Each analysis is different and no estimator should be exclusively used to estimate standard errors. A proper examination of the data should be performed in order to determine the best estimator to use.

5.3 Further Research

During the course of this project, many ideas were suggested by myself as well as my committee. Only a few of these ideas were implemented, because the rest were outside the scope of this project. These ideas are introduced here as possible further research into the topic of robust estimators.

When a covariance structure was set up, only positive correlation was used. Some estimators may react differently to negative correlation. It may be interesting to study the estimators using both positive and negative levels of correlation.

This study used only a compound symmetric covariance matrix. These estimators may not be as consistent under a different covariance structure, such as auto-regressive, unstructured, Toeplitz, or spatial power.

In this study, the REML procedure was modelled under the true covariance structure. The true covariance structure is often unknown. One option is to use AIC to choose a reasonable covariance structure for the data. The REML procedure may not be as consistent under an AIC-chosen model. The AIC-chosen structure may not be the true covariance structure. If this is the case, the REML procedure may perform worse than the other estimators. The REML procedure could be modelled under the true covariance structure for comparison. The Standard Robust Estimator with a quantile adjustment is a viable option to compare to the AIC-chosen REML model.

Also, a more elaborate Long-type standardization is of interest in order to reduce the bias of the Standard Robust Estimator. A new standardization could yield more consistent coverage rates. This investigation should include an comparison of the length of the Long-type intervals to confidence intervals from more consistent robust estimators.

Each of these considerations has application throughout the field of statistics. Like the teratogenic study with the mice, all estimators do not provide the same standard errors. Discretion is needed when using any of the options that are available.

BIBLIOGRAPHY

Agresti, A. (2002), *Categorical Data Analysis*, John Wiley & Sons, Inc., 2nd ed.

Burch, B. D. and Harris, I. R. (2001), "Closed-Form Approximations to the REML Estimator of a Variance Ratio (or Heritability) in a Mixed Linear Model," *Biometrics*, 57, 1148–1156.

Diggle, P., Liang, K.-Y., and Zeger, S. (1994), *Analysis of Longitudinal Data*, Clarendon Press.

Eicker, F. (1963), "Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions," *The Annals of Mathematical Statistics*, 34, 447–456.

Gomez, E. V. (2004), "Performance of the Kenward-Roger Method when the Covariance Structure is Selected Using AIC and BIC," Master's thesis, Brigham Young University.

Hansen, J. M., Reynolds, P. R., Booth, G. M., Schaalje, G. B., and Seegmiller, R. E. (2000), *Teratology*, 62, 227–232.

Huber, P. J. (1967), "The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*.

— (1972), "The 1972 Wald Lecture Robust Statistics: A Review," *The Annals of Mathematical Statistics*, 43, 1041–1067.

Kauermann, G. and Carroll, R. (2001), "A Note on the Efficiency of Sandwich Covariance Matrix Estimation," *Journal of the American Statistical Association*, 96, 1387–1396.

Kenward, M. G. and Roger, J. H. (1997), "Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood," *Biometrics*, 53, 983–997.

Keselman, H., Algina, J., Kowalchuk, R. K., and Wolfinger, R. D. (1999), "The Analysis of Repeated Measurements: A Comparison of Mixed-Model Satterthwaite F Tests and a Nonpooled Adjusted Degrees of Freedom Multivariate Test," *Communications in Statistics: Part A Theory and Methods*, 28, 2967–2999.

Long, J. S. and Ervin, L. H. (2000), "Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model," *The American Statistician*, 54, 217–224.

Rencher, A. (2000), *Linear Models in Statistics*, John Wiley and Sons, Inc.

SAS Institute, I. (2003), *SAS/STAT User's Guide*, SAS Version 9, Cary, North Carolina.

Schaalje, G. B., McBride, J. J., and Fellingham, G. W. (2002), "Adequacy of Approximations to Distributions of Test Statistics in Complex Mixed Linear Models," *Journal of Agricultural, Biological and Environmental Statistics*, 7, 512–524.

Swallow, W. H. and Monahan, J. F. (1984), "Monte Carlo Comparison of ANOVA, MIVQUE, REML, and ML Estimators of Variance Components," *Technometrics*, 26, 47–57.

Thompson, R. (1969), "Iterative Estimation of Variance Components for the Non-Orthogonal Data," *Biometrics*, 25, 767–773.

Thompson, W. (1962), "The Problem of Negative Estimates of Variance Components," *The Annals of Mathematical Statistics*, 33, 273–289.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817–838.

White, H. and MacKinnon, J. G. (1985), “Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties,” .

1. TERATOGENIC STUDY CODE

```
options ls=80 formdlim="*";
PROC IMPORT OUT= WORK.all
            DATAFILE=
"C:\Documents and Settings\Natalie\Desktop\mouse.xls"
            DBMS=EXCEL2000 REPLACE;
     SHEET="Sheet1";
     GETNAMES=YES;
RUN;
data allm;
    set all;
    drop w4--afw;
    wt=w4; logwt=log(wt); day=4; daysq=day**2; output;
    wt=w8; logwt=log(wt); day=8; daysq=day**2; output;
    wt=w12; logwt=log(wt);day=12; daysq=day**2; output;
    wt=w16; logwt=log(wt);day=16; daysq=day**2; output;
    wt=tbw; logwt=log(wt);day=18; daysq=day**2; output;
run;

symbol c=black i=join l=1 repeat=100;
proc gplot data=allm;
    by t;
    plot logwt*day=id;
run;
```

```
proc mixed data=allm maxit=500;
    where day lt 18;
    class t id;
    model logwt= t t*day t*daysq/noint solution;
    random intercept day daysq / subject=id(t) type=un;
run;
proc mixed data=allm maxit=500;
    where day lt 18;
    class t id;
    model logwt=t t*day t*daysq/noint solution;
    repeated /subject=id(t) type=cs;
run;
proc mixed data=allm maxit=500;
    where day lt 18;
    class t id;
    model logwt=t t*day t*daysq/noint solution;
    repeated /subject=id(t) type=un;
run;
proc mixed data=allm maxit=500;
    where day lt 18;
    class t id;
    model logwt=t t*day t*daysq/noint solution;
    random id(t);
    repeated /subject=id(t) type=sp(pow)(day);
run;
proc mixed data=allm maxit=500;
```

```
    where day lt 18;
    class t id;
    model logwt=t t*day t*daysq/noint solution;
    repeated /subject=id(t) type=vc;
run;
proc mixed data=allm maxit=500 empirical;
    where day lt 18;
    class t id;
    model logwt=t t*day t*daysq/noint solution;
    repeated /subject=id(t) type=un;
run;
data allinc;
    set allm;
    where day lt 18;
run;
data allinc;
    set allinc;
    where t=1;
run;
proc iml;
    use allinc;
    read all var{logwt} into y;
    read all var{t day daysq} into x;

    ij = (j(4,1,1)||j(4,10,0))//(j(4,1,0)||j(4,1,1)||j(4,9,0))//
         (j(4,2,0)||j(4,1,1)||j(4,8,0))//(j(4,3,0)||j(4,1,1)||
         j(4,7,0))//(j(4,4,0)||j(4,1,1)||j(4,6,0))//(j(4,5,0)||
```

```
    j(4,1,1)||j(4,5,0))//(j(4,6,0)||j(4,1,1)||j(4,4,0))//
    (j(4,7,0)||j(4,1,1)||j(4,3,0))//(j(4,8,0)||j(4,1,1)||
    j(4,2,0))//(j(4,9,0)||j(4,1,1)||j(4,1,0))//(j(4,10,0)||
    j(4,1,1));

beta_hat = (x`*x)**-1*x`*y;
y_hat = x*beta_hat;
e_hat=y-y_hat;
e_hat2=e_hat#ij;
d_hat_b = e_hat2*e_hat2`;
hat = x*(x`*x)**-1*x`;

*Cov beta hat for robust estimator;
covb_hat = (x`*x)**-1*x`*d_hat_b*x*(x`*x)**-1;

*Cov beta hat for robust estimator with long adjustment;
mhat=i(44)-hat;
mhat2=mhat*mhat`;
long=d_hat_b/mhat2;
v3=diag(long);
covb_hat_l = (x`*x)**-1*x`*v3*x*(x`*x)**-1;

*se for beta for robust estimator;
se_beta_hat = (vecdiag(covb_hat))##(1/2);

*se for beta for robust estimator with long adjustment;
se_beta_hat_l = (vecdiag(covb_hat_l))##(1/2);
```

```
    param = {"Intercept","Day","Daysq"};
    print param beta_hat se_beta_hat se_beta_hat_l;
quit;
```

2. TABLES OF COVERAGE RATES FOR THE INTERCEPT AND SLOPE

Table B.1: Coverage rates for the intercept

Method	Covariance Structure Type	Correlation	Repeated Measure	Sample Size 5	10	25	50
REML	Homogeneous	0.0	3	95.29	94.97	95.08	95.17
			4	94.88	95.10	94.89	95.11
			5	95.32	94.97	95.08	94.97
		0.6	3	95.37	95.13	95.19	95.29
			4	94.83	95.17	95.02	95.10
			5	94.84	94.99	95.13	94.80
		0.9	3	94.64	94.65	95.26	95.43
			4	93.98	95.00	94.78	95.13
			5	94.45	94.83	95.11	94.87
	Heterogeneous	0.1	3	95.27	94.98	95.23	95.37
			4	94.60	94.97	95.10	95.28
			5	94.05	94.90	95.12	94.92
		0.6	3	96.06	95.43	95.46	95.34
			4	95.33	95.21	95.23	95.08
			5	94.61	94.91	95.22	95.26
		0.9	3	95.86	95.55	95.54	95.56
			4	94.57	94.96	95.08	94.99
			5	93.32	94.59	95.16	95.33

Table B.1: Coverage rates for the intercept (ctnd.)

Method	Covariance Structure Type	Correlation	Repeated Measure	Sample Size 5	10	25	50
Standard Robust	Homogeneous	0.0	3	84.72	90.17	93.42	94.47
			4	84.37	90.14	93.10	94.24
			5	84.38	90.74	93.55	94.26
		0.6	3	84.83	89.94	93.58	94.61
			4	84.30	90.42	93.08	94.20
			5	84.66	90.87	93.73	94.30
		0.9	3	84.69	90.19	93.58	94.57
			4	84.26	90.15	93.14	94.33
			5	84.52	90.81	93.62	94.24
	Heterogeneous	0.1	3	83.98	90.16	93.32	94.09
			4	84.06	90.32	93.36	94.32
			5	84.19	90.97	93.14	94.15
		0.6	3	84.03	90.04	93.83	94.20
			4	83.96	90.46	93.43	94.41
			5	84.35	90.57	93.07	94.16
		0.9	3	84.35	89.89	93.77	94.49
			4	83.98	90.33	93.54	94.18
			5	84.93	90.44	93.48	94.15
Empirical	Homogeneous	0.0	3	93.10	93.72	94.70	95.08
			4	93.02	93.81	94.40	94.94
			5	92.91	94.03	94.86	94.78
		0.6	3	93.44	94.00	94.70	95.16
			4	92.99	93.77	94.34	94.78
			5	93.23	94.22	94.95	94.83
		0.9	3	93.72	93.90	94.81	95.09
			4	92.84	93.97	94.43	94.85
			5	93.21	94.03	94.94	94.75

Table B.1: Coverage rates for the intercept (ctnd.)

	Covariance Structure		Repeated	Sample Size			
Method	Type	Correlation	Measure	5	10	25	50
Empirical	Heterogeneous	0.1	3	91.57	93.06	94.75	95.13
			4	90.50	92.80	94.35	94.87
			5	89.70	92.58	94.30	94.59
		0.6	3	91.99	93.24	94.76	95.02
			4	90.27	93.04	94.40	94.69
			5	89.72	92.50	94.52	94.78
		0.9	3	90.98	93.02	94.71	95.12
			4	88.12	92.21	93.96	94.40
			5	86.08	90.62	93.81	94.61
Long Adjustment	Homogeneous	0.0	3	91.37	93.06	94.47	94.94
			4	91.83	93.74	94.23	94.77
			5	93.04	93.81	94.69	94.84
		0.6	3	94.84	96.66	98.04	98.25
			4	92.50	95.29	96.32	96.83
			5	91.51	93.61	95.10	95.14
		0.9	3	95.67	97.98	99.12	99.30
			4	92.06	95.82	97.11	97.71
			5	88.86	92.89	95.00	95.26
	Heterogeneous	0.1	3	93.98	95.08	95.93	95.88
			4	94.25	95.32	95.68	96.12
			5	95.21	95.60	95.67	95.65
		0.6	3	98.38	99.18	99.57	99.64
			4	98.22	99.08	99.40	99.59
			5	98.18	98.89	99.43	99.36
		0.9	3	99.77	99.97	100.00	100.00
			4	99.58	99.89	99.99	100.00
			5	99.12	99.83	99.99	99.98

Table B.1: Coverage rates for the intercept (ctnd.)

	Covariance Structure		Repeated	Sample Size			
Method	Type	Correlation	Measure	5	10	25	50
Long Quantile Adjustment	Homogeneous	0.0	3	96.69	97.93	98.73	98.88
			4	97.20	98.26	98.80	98.97
			5	97.48	98.18	98.80	98.87
		0.6	3	98.23	99.22	99.65	99.77
			4	97.49	98.69	99.36	99.56
			5	96.67	97.91	98.92	99.10
		0.9	3	98.45	99.61	99.92	99.98
			4	96.60	98.73	99.51	99.67
			5	94.98	97.55	98.87	99.08
	Heterogeneous	0.1	3	98.30	98.82	99.32	99.29
			4	98.74	99.10	99.16	99.30
			5	98.95	99.07	99.34	99.19
		0.6	3	99.64	99.88	100.00	99.98
			4	99.70	99.87	100.00	100.00
			5	99.64	99.87	99.94	99.94
		0.9	3	99.95	100.00	100.00	100.00
			4	99.87	99.98	100.00	100.00
			5	99.83	100.00	100.00	100.00

Table B.2: Coverage rates for the slope

	Covariance Structure		Repeated	Sample Size			
Method	Type	Correlation	Measure	5	10	25	50
REML	Homogeneous	0.0	3	94.80	94.71	95.33	95.23
			4	94.83	94.95	95.01	95.41
			5	95.07	95.12	95.19	94.82
		0.6	3	94.97	94.73	95.53	95.38
			4	94.89	95.20	95.11	95.21
			5	94.96	95.25	94.83	95.02
		0.9	3	94.95	94.89	95.42	95.30
			4	94.82	95.26	95.52	95.28
			5	94.98	95.16	95.03	95.21
	Heterogeneous	0.1	3	94.76	94.64	95.37	95.37
			4	93.38	94.61	95.21	95.34
			5	92.62	94.59	94.85	95.03
		0.6	3	94.47	94.90	95.39	95.13
			4	92.33	94.25	95.37	95.41
			5	90.40	93.92	94.97	95.26
		0.9	3	93.73	94.47	95.32	95.18
			4	91.05	93.91	95.04	95.33
			5	89.36	93.50	94.96	95.27

Table B.2: Coverage rates for the slope (ctnd.)

Method	Covariance Structure Type	Correlation	Repeated Measure	Sample Size 5	10	25	50
Standard Robust	Homogeneous	0.0	3	84.56	89.92	93.65	94.44
			4	83.93	90.38	93.57	94.36
			5	84.40	90.40	93.30	94.09
		0.6	3	84.32	90.37	93.87	94.47
			4	84.21	90.47	93.36	94.45
			5	84.34	90.44	93.49	94.15
		0.9	3	84.70	90.37	93.89	94.36
			4	84.35	90.51	93.62	94.25
			5	84.56	90.49	93.55	94.23
	Heterogeneous	0.1	3	84.29	90.05	94.01	94.03
			4	84.31	90.50	93.38	94.35
			5	84.51	90.48	92.97	93.98
		0.6	3	85.18	90.51	93.64	94.43
			4	84.47	90.63	93.87	94.42
			5	84.57	91.02	93.53	94.30
		0.9	3	85.37	90.62	93.77	94.30
			4	84.58	90.75	93.70	94.54
			5	85.03	90.63	93.65	94.43
Empirical	Homogeneous	0.0	3	88.61	91.79	94.30	94.79
			4	86.61	91.46	93.88	94.53
			5	86.54	91.26	93.65	94.31
		0.6	3	88.84	92.10	94.44	94.79
			4	86.99	91.52	93.66	94.63
			5	86.45	91.31	93.72	94.29
		0.9	3	88.64	91.99	94.43	94.63
			4	86.96	91.56	94.11	94.44
			5	86.61	91.25	93.81	94.34

Table B.2: Coverage rates for the slope (ctnd.)

Method	Covariance Structure Type	Covariance Structure Correlation	Repeated Measure	Sample Size 5	Sample Size 10	Sample Size 25	Sample Size 50
Empirical	Heterogeneous	0.1	3	84.61	90.15	93.86	94.51
			4	79.82	88.67	93.09	94.24
			5	77.94	87.98	92.39	93.94
		0.6	3	83.83	90.11	93.60	94.19
			4	77.15	87.55	93.18	94.48
			5	72.10	86.15	92.14	93.78
		0.9	3	81.62	89.75	93.40	94.29
			4	73.71	85.96	92.19	94.04
			5	68.22	83.92	91.51	93.64
Long Adjustment	Homogeneous	0.0	3	91.53	93.38	94.80	94.87
			4	92.15	93.86	94.53	95.17
			5	92.92	93.97	94.75	94.60
		0.6	3	98.27	99.29	99.63	99.77
			4	98.15	99.23	99.69	99.73
			5	98.64	99.27	99.65	99.72
		0.9	3	99.85	100.00	100.00	100.00
			4	99.82	99.98	100.00	100.00
			5	99.88	100.00	100.00	100.00
	Heterogeneous	0.1	3	91.48	93.61	95.54	95.34
			4	91.93	94.08	95.15	95.63
			5	92.77	94.46	94.99	95.27
		0.6	3	95.96	97.71	98.75	98.98
			4	94.76	97.27	98.15	98.54
			5	94.44	96.68	97.73	97.91
		0.9	3	97.88	99.28	99.78	99.86
			4	96.11	98.57	99.38	99.54
			5	94.61	97.58	98.80	98.93

Table B.2: Coverage rates for the slope (ctnd.)

	Covariance Structure		Repeated	Sample Size			
Method	Type	Correlation	Measure	5	10	25	50
Long Quantile Adjustment	Homogeneous	0.0	3	97.13	98.11	98.78	98.80
			4	97.23	98.36	98.71	98.88
			5	97.83	98.52	98.91	98.92
		0.6	3	99.57	99.89	100.00	99.98
			4	99.54	99.91	99.99	99.99
			5	99.66	99.90	99.97	99.98
		0.9	3	99.97	100.00	100.00	100.00
			4	99.96	100.00	100.00	100.00
			5	99.99	100.00	100.00	100.00
	Heterogeneous	0.1	3	96.89	98.07	98.90	99.16
			4	97.19	98.33	98.93	99.09
			5	97.55	98.43	98.79	99.00
		0.6	3	98.69	99.61	99.82	99.92
			4	98.32	99.39	99.80	99.86
			5	98.25	99.39	99.59	99.80
		0.9	3	99.27	99.85	99.98	99.99
			4	98.50	99.74	99.93	99.97
			5	97.78	99.54	99.85	99.89

```
options formdlim="*";
\%let beta0=1;
\%let beta1=1;
\%macro simulated3a(nsub, nit, nmeas);
proc iml;
    nsub=&nsub;
    nit=&nit;
    nmeas=&nmeas;
    np1 = nit*nsub*nmeas+1;
    *d2a = {1 0, 0 1};
    d = {1 0 0, 0 1 0, 0 0 1};
    *d4a = {1 0 0 0, 0 1 0 0, 0 0 1 0, 0 0 0 1};
    *d5a = {1 0 0 0 0, 0 1 0 0 0, 0 0 1 0 0, 0 0 0 1 0, 0 0 0 0 1};

    *d2b = {1 0.6, 0.6 1};
    *d3b = {1 0.6 0.6, 0.6 1 0.6, 0.6 0.6 1};
    *d4b = {1 0.6 0.6 0.6, 0.6 1 0.6 0.6, 0.6 0.6 1 0.6,
            0.6 0.6 0.6 1};
    *d5b = {1 0.6 0.6 0.6 0.6, 0.6 1 0.6 0.6 0.6,
            0.6 0.6 1 0.6 0.6, 0.6 0.6 0.6 1 0.6,
            0.6 0.6 0.6 0.6 1};

    *d2c = {1 0.9, 0.9 1};
    *d3c = {1 0.9 0.9, 0.9 1 0.9, 0.9 0.9 1};
```

```
*d4c = {1 0.9 0.9 0.9, 0.9 1 0.9 0.9, 0.9 0.9 1 0.9,
        0.9 0.9 0.9 1};
*d5c = {1 0.9 0.9 0.9 0.9, 0.9 1 0.9 0.9 0.9,
        0.9 0.9 1 0.9 0.9, 0.9 0.9 0.9 1 0.9,
        0.9 0.9 0.9 0.9 1};


*d2d = {1 sqrt(3), sqrt(3) 1};
*d3d = {1 sqrt(3) sqrt(5), sqrt(3) 3 sqrt(3*5),
        sqrt(5) sqrt(3*5) 5};
*d4d = {1 sqrt(3) sqrt(5) sqrt(7),
        sqrt(3) 3 sqrt(3*5) sqrt(3*7),
        sqrt(5) sqrt(3*5) 5 sqrt(5*7),
        sqrt(7) sqrt(3*7) sqrt(5*7) 7};
*d5d = {1 sqrt(3) sqrt(5) sqrt(7) sqrt(9),
        sqrt(3) 3 sqrt(3*5) sqrt(3*7) sqrt(3*9),
        sqrt(5) sqrt(3*5) 5 sqrt(5*7) sqrt(5*9),
        sqrt(7) sqrt(3*7) sqrt(5*7) 7 sqrt(7*9),
        sqrt(9) sqrt(3*9) sqrt(5*9) sqrt(7*9) 9};


*d2e = {1 0.6*sqrt(3), 0.6*sqrt(3) 1};
*d3e = {1 0.6*sqrt(3) 0.6*sqrt(5), 0.6*sqrt(3) 3 0.6*sqrt(3*5),
        0.6*sqrt(5) 0.6*sqrt(3*5) 5};
*d4e = {1 0.6*sqrt(3) 0.6*sqrt(5) 0.6*sqrt(7),
        0.6*sqrt(3) 3 0.6*sqrt(3*5) 0.6*sqrt(3*7),
        0.6*sqrt(5) 0.6*sqrt(3*5) 5 0.6*sqrt(5*7),
        0.6*sqrt(7) 0.6*sqrt(3*7) 0.6*sqrt(5*7) 7};
*d5e = {1 0.6*sqrt(3) 0.6*sqrt(5) 0.6*sqrt(7) 0.6*sqrt(9),
```

```
        0.6*sqrt(3) 3 0.6*sqrt(3*5) 0.6*sqrt(3*7) 0.6*sqrt(3*9),
        0.6*sqrt(5) 0.6*sqrt(3*5) 5 0.6*sqrt(5*7) 0.6*sqrt(5*9),
        0.6*sqrt(7) 0.6*sqrt(3*7) 0.6*sqrt(5*7) 7 0.6*sqrt(7*9),
        0.6*sqrt(9) 0.6*sqrt(3*9) 0.6*sqrt(5*9) 0.6*sqrt(7*9) 9};

*d2f = {1 0.9*sqrt(3), 0.9*sqrt(3) 1};
*d3f = {1 0.9*sqrt(3) 0.9*sqrt(5),
        0.9*sqrt(3) 3 0.9*sqrt(3*5),
        0.9*sqrt(5) 0.9*sqrt(3*5) 5};
*d4f = {1 0.9*sqrt(3) 0.9*sqrt(5) 0.9*sqrt(7),
        0.9*sqrt(3) 3 0.9*sqrt(3*5) 0.9*sqrt(3*7),
        0.9*sqrt(5) 0.9*sqrt(3*5) 5 0.9*sqrt(5*7),
        0.9*sqrt(7) 0.9*sqrt(3*7) 0.9*sqrt(5*7) 7};
*d5f = {1 0.9*sqrt(3) 0.9*sqrt(5) 0.9*sqrt(7) 0.9*sqrt(9),
        0.9*sqrt(3) 3 0.9*sqrt(3*5) 0.9*sqrt(3*7) 0.9*sqrt(3*9),
        0.9*sqrt(5) 0.9*sqrt(3*5) 5 0.9*sqrt(5*7) 0.9*sqrt(5*9),
        0.9*sqrt(7) 0.9*sqrt(3*7) 0.9*sqrt(5*7) 7 0.9*sqrt(7*9),
        0.9*sqrt(9) 0.9*sqrt(3*9) 0.9*sqrt(5*9) 0.9*sqrt(7*9) 9};



*x2i = {1 1, 1 2};
xi = {1 1, 1 2, 1 3};
*x4i = {1 1, 1 2, 1 3, 1 4};
*x5i = {1 1, 1 2, 1 3, 1 4, 1 5};
beta={&beta0, &beta1};

beta_hatm={.,.};
```

```
ll_m={.,.};
ul_m={.,.};
cover={.,.};
coverm={.,.};
ll_m_l={.,.};
ul_m_l={.,.};
cover_l={.,.};
coverm_l={.,.};
ll_m_lq={.,.};
ul_m_lq={.,.};
cover_lq={.,.};
coverm_lq={.,.};
ll_m_k={.,.};
ul_m_k={.,.};
cover_k={.,.};
coverm_k={.,.};

nn=nsub*nmeas;
rootcovd = (root(d))`; *rootcovd=1;
seed=1;
y=.;
x={. .};
w=.;
item = {1,2};
itm = item;
do it=1 to nit;
    do sub=1 to nsub;
```

```
    z=normal(repeat(seed,nmeas,1));
    yi=rootcovd*z;
    yi=xi*beta+yi;
    wi=(sub@j(nmeas,1,1));
    w=w//wi;
    y=y//yi;
    x=x//xi;
end;
its=it@j(nn,1,1);
iter=iter//its;

ycurr=y[nn*(it-1)+2:nn*it+1,1];
xcurr=x[nn*(it-1)+2:nn*it+1,];
xpxi = (xcurr`*xcurr)**-1;

beta_hat=xpxi*xcurr`*ycurr;
y_hat = xcurr*beta_hat;

itm = itm//item;

e_hat=ycurr-y_hat;
e_hat2= e_hat#(i(nsub)@j(nmeas,1,1));

d_hat_b = e_hat2*e_hat2`;

hat = xcurr*xpxi*xcurr`;
```

```
*Cov beta hat for robust estimator;
covb_hat = xpxi*xcurr`*d_hat_b*xcurr*xpxi;

*Cov beta hat for robust estimator with long adjustment;
mhat=i(nn)-hat;
mhat2=mhat*mhat`;
long=d_hat_b/mhat2;
v3=diag(long);
covb_hat_l = xpxi*xcurr`*v3*xcurr*xpxi;

*Cov beta hat for robust estimator with Kauermann adjustment;
kau=d_hat_b/mhat;
v2=diag(kau);
covb_hat_k = xpxi*xcurr`*v2*xcurr*xpxi;



*95\% CI for beta for robust estimator;
ll_beta_hat = beta_hat-1.96*(vecdiag(covb_hat))##(1/2);
ul_beta_hat = beta_hat+1.96*(vecdiag(covb_hat))##(1/2);

*CI contain beta for robust estimator;
if ll_beta_hat[1]<beta[1] & ul_beta_hat[1]>beta[1]
    then cover[1]=1;
else cover[1]=0;

if ll_beta_hat[2]<beta[2] & ul_beta_hat[2]>beta[2]
    then cover[2]=1;
```

```
else cover[2]=0;

*95\% CI for beta for robust estimator with long adjustment;
ll_beta_hat_l = beta_hat-1.96*(vecdiag(covb_hat_l))##(1/2);
ul_beta_hat_l = beta_hat+1.96*(vecdiag(covb_hat_l))##(1/2);

*CI contain beta for robust estimator with long adjustment;
if ll_beta_hat_l[1]<beta[1] & ul_beta_hat_l[1]>beta[1]
     then cover_l[1]=1;
else cover_l[1]=0;

if ll_beta_hat_l[2]<beta[2] & ul_beta_hat_l[2]>beta[2]
     then cover_l[2]=1;
else cover_l[2]=0;

*95\% CI for beta for robust estimator with kauermann
    adjustment;
ll_beta_hat_k = beta_hat-1.96*(vecdiag(covb_hat_k))##(1/2);
ul_beta_hat_k = beta_hat+1.96*(vecdiag(covb_hat_k))##(1/2);

*CI contain beta for robust estimator with kauermann
    adjustment;
if ll_beta_hat_k[1]<beta[1] & ul_beta_hat_k[1]>beta[1]
    then cover_k[1]=1;
else cover_k[1]=0;

if ll_beta_hat_k[2]<beta[2] & ul_beta_hat_k[2]>beta[2]
```

```
    then cover_k[2]=1;
else cover_k[2]=0;


*95\% CI for beta for robust estimator with long adjustment
    *with corrected quantiles (99\%);
ll_beta_hat_lq = beta_hat-2.576*(vecdiag(covb_hat_l))##(1/2);
ul_beta_hat_lq = beta_hat+2.576*(vecdiag(covb_hat_l))##(1/2);

*CI contain beta for robust estimator with long adjustment;
if ll_beta_hat_lq[1]<beta[1] & ul_beta_hat_lq[1]>beta[1]
     then cover_lq[1]=1;
else cover_lq[1]=0;

if ll_beta_hat_lq[2]<beta[2] & ul_beta_hat_lq[2]>beta[2]
     then cover_lq[2]=1;
else cover_lq[2]=0;

beta_hatm=beta_hatm//beta_hat;
ll_m=ll_m//ll_beta_hat;
ul_m=ul_m//ul_beta_hat;
coverm=coverm//cover;
ll_m_l=ll_m_l//ll_beta_hat_l;
ul_m_l=ul_m_l//ul_beta_hat_l;
coverm_l=coverm_l//cover_l;
ll_m_lq=ll_m_lq//ll_beta_hat_lq;
ul_m_lq=ul_m_lq//ul_beta_hat_lq;
```

```
    coverm_lq=coverm_lq//cover_lq;
    ll_m_k=ll_m_k//ll_beta_hat_k;
    ul_m_k=ul_m_k//ul_beta_hat_k;
    coverm_k=coverm_k//cover_k;
end;

    *w is subject;
    w=w[2:np1,1];
    y=y[2:np1,1];
    x=x[2:np1,1:2];
    beta_hatm=beta_hatm[3:nit*2+2,1];
    ll_m=ll_m[3:nit*2+2,1];
    ul_m=ul_m[3:nit*2+2,1];
    coverm=coverm[3:nit*2+2,1];
    ll_m_l=ll_m_l[3:nit*2+2,1];
    ul_m_l=ul_m_l[3:nit*2+2,1];
    coverm_l=coverm_l[3:nit*2+2,1];
    ll_m_lq=ll_m_lq[3:nit*2+2,1];
    ul_m_lq=ul_m_lq[3:nit*2+2,1];
    coverm_lq=coverm_lq[3:nit*2+2,1];
    ll_m_k=ll_m_k[3:nit*2+2,1];
    ul_m_k=ul_m_k[3:nit*2+2,1];
    coverm_k=coverm_k[3:nit*2+2,1];
    itm=itm[3:nit*2+2,1];

    x1 = x[,1];
    x2 = x[,2];
```

```
create data1 var { y x1 x2 w iter };
append;

create mysim var {beta_hatm ll_m ul_m coverm ll_m_l ul_m_l
  coverm_l ll_m_lq ul_m_lq coverm_lq ll_m_k ul_m_k coverm_k itm};
append; quit;

*REML;
ods listing close;
proc mixed data=data1;
    class w;
    by iter;
    model y = x2 / solution cl ddfm=kenwardroger;
    repeated / subject=w type=cs;
    ods output SolutionF=beta_reml;
run;

*Whites estimator;
proc mixed data=data1 empirical;
    class w;
    by iter;
    model y = x2 / solution cl ddfm=kenwardroger;
    repeated /subject=w type=cs;
    ods output SolutionF=beta_white;
run;
ods listing;
```

```
*REML;
data beta_reml;
    set beta_reml;
    if lower<&beta0 and upper>&beta0 and effect="Intercept"
        then cover=1;
    else if lower<&beta1 and upper>&beta1 and effect="X2"
        then cover=1;
    else cover=0;
run;
proc sort data=beta_reml;
    by effect;
run;
proc freq data=beta_reml;
    by effect;
    table cover;
run;
proc capability data=beta_reml noprint;
    var estimate;
    histogram / kernel (k=Normal c=MISE color=RED l=1 w=3)
        cfill=BLUE;
    by effect;
run;

*White;
data beta_white;
    set beta_white;
```

```
    if lower<&beta0 and upper>&beta0 then cover=1;
    else if lower<&beta1 and upper>&beta1 and effect="X2"
        then cover=1;
    else cover=0;
run;
proc sort data=beta_white;
    by effect;
run;
proc freq data=beta_white;
    by effect;
    table cover;
run;
proc capability data=beta_white noprint;
    var estimate;
    histogram / kernel (k=Normal c=MISE color=RED l=1 w=3)
        cfill=BLUE;
    by effect;
run;

*Values from IML computations;
proc sort data=mysim;
    by itm;
run;
proc freq data=mysim;
    by itm;
    table coverm;
run;
```

```
proc freq data=mysim;
    by itm;
    table coverm_l;
run;
proc freq data=mysim;
    by itm;
    table coverm_lq;
run;
proc freq data=mysim;
    by itm;
    table coverm_k;
run;
proc capability data=mysim noprint;
    var beta_hatm;
    histogram / kernel (k=Normal c=MISE color=RED l=1 w=3)
        cfill=BLUE;
    by itm;
run;
\%mend;

\%macro simulated3f(nsub, nit, nmeas);
proc iml;
    nsub=&nsub;
    nit=&nit;
    nmeas=&nmeas;
    np1 = nit*nsub*nmeas+1;
    *d2a = {1 0, 0 1};
```

```
*d3a = {1 0 0, 0 1 0, 0 0 1};
*d4a = {1 0 0 0, 0 1 0 0, 0 0 1 0, 0 0 0 1};
*d5a = {1 0 0 0 0, 0 1 0 0 0, 0 0 1 0 0, 0 0 0 1 0, 0 0 0 0 1};

*d2b = {1 0.6, 0.6 1};
*d3b = {1 0.6 0.6, 0.6 1 0.6, 0.6 0.6 1};
*d4b = {1 0.6 0.6 0.6, 0.6 1 0.6 0.6, 0.6 0.6 1 0.6,
        0.6 0.6 0.6 1};
*d5b = {1 0.6 0.6 0.6 0.6, 0.6 1 0.6 0.6 0.6,
        0.6 0.6 1 0.6 0.6, 0.6 0.6 0.6 1 0.6,
        0.6 0.6 0.6 0.6 1};

*d2c = {1 0.9, 0.9 1};
*d3c = {1 0.9 0.9, 0.9 1 0.9, 0.9 0.9 1};
*d4c = {1 0.9 0.9 0.9, 0.9 1 0.9 0.9, 0.9 0.9 1 0.9,
        0.9 0.9 0.9 1};
*d5c = {1 0.9 0.9 0.9 0.9, 0.9 1 0.9 0.9 0.9,
        0.9 0.9 1 0.9 0.9, 0.9 0.9 0.9 1 0.9,
        0.9 0.9 0.9 0.9 1};

*d2d = {1 sqrt(3), sqrt(3) 1};
*d3d = {1 sqrt(3) sqrt(5), sqrt(3) 3 sqrt(3*5),
        sqrt(5) sqrt(3*5) 5};
*d4d = {1 sqrt(3) sqrt(5) sqrt(7),
        sqrt(3) 3 sqrt(3*5) sqrt(3*7),
        sqrt(5) sqrt(3*5) 5 sqrt(5*7),
        sqrt(7) sqrt(3*7) sqrt(5*7) 7};
```

```
*d5d = {1 sqrt(3) sqrt(5) sqrt(7) sqrt(9),
        sqrt(3) 3 sqrt(3*5) sqrt(3*7) sqrt(3*9),
        sqrt(5) sqrt(3*5) 5 sqrt(5*7) sqrt(5*9),
        sqrt(7) sqrt(3*7) sqrt(5*7) 7 sqrt(7*9),
        sqrt(9) sqrt(3*9) sqrt(5*9) sqrt(7*9) 9};

*d2e = {1 0.6*sqrt(3), 0.6*sqrt(3) 1};
*d3e = {1 0.6*sqrt(3) 0.6*sqrt(5), 0.6*sqrt(3) 3 0.6*sqrt(3*5),
        0.6*sqrt(5) 0.6*sqrt(3*5) 5};
*d4e = {1 0.6*sqrt(3) 0.6*sqrt(5) 0.6*sqrt(7),
        0.6*sqrt(3) 3 0.6*sqrt(3*5) 0.6*sqrt(3*7),
        0.6*sqrt(5) 0.6*sqrt(3*5) 5 0.6*sqrt(5*7),
        0.6*sqrt(7) 0.6*sqrt(3*7) 0.6*sqrt(5*7) 7};
*d5e = {1 0.6*sqrt(3) 0.6*sqrt(5) 0.6*sqrt(7) 0.6*sqrt(9),
        0.6*sqrt(3) 3 0.6*sqrt(3*5) 0.6*sqrt(3*7) 0.6*sqrt(3*9),
        0.6*sqrt(5) 0.6*sqrt(3*5) 5 0.6*sqrt(5*7) 0.6*sqrt(5*9),
        0.6*sqrt(7) 0.6*sqrt(3*7) 0.6*sqrt(5*7) 7 0.6*sqrt(7*9),
        0.6*sqrt(9) 0.6*sqrt(3*9) 0.6*sqrt(5*9) 0.6*sqrt(7*9) 9};

*d2f = {1 0.9*sqrt(3), 0.9*sqrt(3) 1};
d = {1 1.558845727 2.01246118, 1.558845727 3 3.485685012,
     2.01246118 3.485685012 5};
*d4f = {1 0.9*sqrt(3) 0.9*sqrt(5) 0.9*sqrt(7),
        0.9*sqrt(3) 3 0.9*sqrt(3*5) 0.9*sqrt(3*7),
        0.9*sqrt(5) 0.9*sqrt(3*5) 5 0.9*sqrt(5*7),
        0.9*sqrt(7) 0.9*sqrt(3*7) 0.9*sqrt(5*7) 7};
*d5f = {1 0.9*sqrt(3) 0.9*sqrt(5) 0.9*sqrt(7) 0.9*sqrt(9),
```

```
        0.9*sqrt(3) 3 0.9*sqrt(3*5) 0.9*sqrt(3*7) 0.9*sqrt(3*9),
        0.9*sqrt(5) 0.9*sqrt(3*5) 5 0.9*sqrt(5*7) 0.9*sqrt(5*9),
        0.9*sqrt(7) 0.9*sqrt(3*7) 0.9*sqrt(5*7) 7 0.9*sqrt(7*9),
        0.9*sqrt(9) 0.9*sqrt(3*9) 0.9*sqrt(5*9) 0.9*sqrt(7*9) 9};


*x2i = {1 1, 1 2};
xi = {1 1, 1 2, 1 3};
*x4i = {1 1, 1 2, 1 3, 1 4};
*x5i = {1 1, 1 2, 1 3, 1 4, 1 5};
beta={&beta0, &beta1};

beta_hatm={.,.};
ll_m={.,.};
ul_m={.,.};
cover={.,.};
coverm={.,.};
ll_m_l={.,.};
ul_m_l={.,.};
cover_l={.,.};
coverm_l={.,.};
ll_m_lq={.,.};
ul_m_lq={.,.};
cover_lq={.,.};
coverm_lq={.,.};
ll_m_k={.,.};
ul_m_k={.,.};
```

```
cover_k={.,.};
coverm_k={.,.};

nn=nsub*nmeas;
rootcovd = (root(d))`; *rootcovd=1;
seed=1;
y=.;
x={. .};
w=.;
item = {1,2};
itm = item;
do it=1 to nit;
    do sub=1 to nsub;
        z=normal(repeat(seed,nmeas,1));
        yi=rootcovd*z;
        yi=xi*beta+yi;
        wi=(sub@j(nmeas,1,1));
        w=w//wi;
        y=y//yi;
        x=x//xi;
    end;
    its=it@j(nn,1,1);
    iter=iter//its;

    ycurr=y[nn*(it-1)+2:nn*it+1,1];
    xcurr=x[nn*(it-1)+2:nn*it+1,];
    xpxi = (xcurr`*xcurr)**-1;
```

```
beta_hat=xpxi*xcurr`*ycurr;
y_hat = xcurr*beta_hat;

itm = itm//item;

e_hat=ycurr-y_hat;
e_hat2= e_hat#(i(nsub)@j(nmeas,1,1));

d_hat_b = e_hat2*e_hat2`;

hat = xcurr*xpxi*xcurr`;

*Cov beta hat for robust estimator;
covb_hat = xpxi*xcurr`*d_hat_b*xcurr*xpxi;

*Cov beta hat for robust estimator with long adjustment;
mhat=i(nn)-hat;
mhat2=mhat*mhat`;
long=d_hat_b/mhat2;
v3=diag(long);
covb_hat_l = xpxi*xcurr`*v3*xcurr*xpxi;

*Cov beta hat for robust estimator with Kauermann adjustment;
kau=d_hat_b/mhat;
v2=diag(kau);
covb_hat_k = xpxi*xcurr`*v2*xcurr*xpxi;
```

```
*95\% CI for beta for robust estimator;
ll_beta_hat = beta_hat-1.96*(vecdiag(covb_hat))##(1/2);
ul_beta_hat = beta_hat+1.96*(vecdiag(covb_hat))##(1/2);

*CI contain beta for robust estimator;
if ll_beta_hat[1]<beta[1] & ul_beta_hat[1]>beta[1]
    then cover[1]=1;
else cover[1]=0;

if ll_beta_hat[2]<beta[2] & ul_beta_hat[2]>beta[2]
    then cover[2]=1;
else cover[2]=0;


*95\% CI for beta for robust estimator with long adjustment;
ll_beta_hat_l = beta_hat-1.96*(vecdiag(covb_hat_l))##(1/2);
ul_beta_hat_l = beta_hat+1.96*(vecdiag(covb_hat_l))##(1/2);

*CI contain beta for robust estimator with long adjustment;
if ll_beta_hat_l[1]<beta[1] & ul_beta_hat_l[1]>beta[1]
    then cover_l[1]=1;
else cover_l[1]=0;

if ll_beta_hat_l[2]<beta[2] & ul_beta_hat_l[2]>beta[2]
    then cover_l[2]=1;
```

```
else cover_l[2]=0;


*95\% CI for beta for robust estimator with long adjustment
    *with corrected quantiles (99\%);
ll_beta_hat_lq = beta_hat-2.576*(vecdiag(covb_hat_l))##(1/2);
ul_beta_hat_lq = beta_hat+2.576*(vecdiag(covb_hat_l))##(1/2);

*95\% CI for beta for robust estimator with kauermann
    adjustment;
ll_beta_hat_k = beta_hat-1.96*(vecdiag(covb_hat_k))##(1/2);
ul_beta_hat_k = beta_hat+1.96*(vecdiag(covb_hat_k))##(1/2);

*CI contain beta for robust estimator with kauermann
    adjustment;
if ll_beta_hat_k[1]<beta[1] & ul_beta_hat_k[1]>beta[1]
    then cover_k[1]=1;
else cover_k[1]=0;

if ll_beta_hat_k[2]<beta[2] & ul_beta_hat_k[2]>beta[2]
    then cover_k[2]=1;
else cover_k[2]=0;

*CI contain beta for robust estimator with long adjustment;
if ll_beta_hat_lq[1]<beta[1] & ul_beta_hat_lq[1]>beta[1]
    then cover_lq[1]=1;
else cover_lq[1]=0;
```

```
    if ll_beta_hat_lq[2]<beta[2] & ul_beta_hat_lq[2]>beta[2]
        then cover_lq[2]=1;
    else cover_lq[2]=0;

    beta_hatm=beta_hatm//beta_hat;
    ll_m=ll_m//ll_beta_hat;
    ul_m=ul_m//ul_beta_hat;
    coverm=coverm//cover;
    ll_m_l=ll_m_l//ll_beta_hat_l;
    ul_m_l=ul_m_l//ul_beta_hat_l;
    coverm_l=coverm_l//cover_l;
    ll_m_lq=ll_m_lq//ll_beta_hat_lq;
    ul_m_lq=ul_m_lq//ul_beta_hat_lq;
    coverm_lq=coverm_lq//cover_lq;
    ll_m_k=ll_m_k//ll_beta_hat_k;
    ul_m_k=ul_m_k//ul_beta_hat_k;
    coverm_k=coverm_k//cover_k;
end;

    *w is subject;
    w=w[2:np1,1];
    y=y[2:np1,1];
    x=x[2:np1,1:2];
    beta_hatm=beta_hatm[3:nit*2+2,1];
    ll_m=ll_m[3:nit*2+2,1];
    ul_m=ul_m[3:nit*2+2,1];
```

```
        coverm=coverm[3:nit*2+2,1];
        ll_m_l=ll_m_l[3:nit*2+2,1];
        ul_m_l=ul_m_l[3:nit*2+2,1];
        coverm_l=coverm_l[3:nit*2+2,1];
        ll_m_lq=ll_m_lq[3:nit*2+2,1];
        ul_m_lq=ul_m_lq[3:nit*2+2,1];
        coverm_lq=coverm_lq[3:nit*2+2,1];
        ll_m_k=ll_m_k[3:nit*2+2,1];
        ul_m_k=ul_m_k[3:nit*2+2,1];
        coverm_k=coverm_k[3:nit*2+2,1];
        itm=itm[3:nit*2+2,1];

        x1 = x[,1];
        x2 = x[,2];

create data1 var { y x1 x2 w iter };
append;

create mysim var {beta_hatm ll_m ul_m coverm ll_m_l ul_m_l
    coverm_l ll_m_lq ul_m_lq coverm_lq ll_m_k ul_m_k coverm_k itm};
append;
quit;

*REML;
ods listing close;
proc mixed data=data1;
    class w;
```

```
    by iter;
    model y = x2 / solution cl ddfm=kenwardroger;
    repeated / subject=w type=csh;
    ods output SolutionF=beta_reml;
run;


*Whites estimator;
proc mixed data=data1 empirical;
    class w;
    by iter;
    model y = x2 / solution cl ddfm=kenwardroger;
    repeated /subject=w type=csh;
    ods output SolutionF=beta_white;
run;
ods listing;

*REML; data beta_reml;
    set beta_reml;
    if lower<&beta0 and upper>&beta0 and effect="Intercept"
        then cover=1;
    else if lower<&beta1 and upper>&beta1 and effect="X2"
        then cover=1;
    else cover=0;
run;
proc sort data=beta_reml;
    by effect;
```

```
run;
proc freq data=beta_reml;
    by effect;
    table cover;
run;
proc capability data=beta_reml noprint;
    var estimate;
    histogram / kernel (k=Normal c=MISE color=RED l=1 w=3)
        cfill=BLUE;
    by effect;
run;

*White;
data beta_white;
    set beta_white;
    if lower<&beta0 and upper>&beta0 then cover=1;
    else if lower<&beta1 and upper>&beta1 and effect="X2"
        then cover=1;
    else cover=0;
run;
proc sort data=beta_white;
    by effect;
run;
proc freq data=beta_white;
    by effect;
    table cover;
run;
```

```
proc capability data=beta_white noprint;
    var estimate;
    histogram / kernel (k=Normal c=MISE color=RED l=1 w=3)
        cfill=BLUE;
    by effect;
run;

*Values from IML computations;
proc sort data=mysim;
    by itm;
run;
proc freq data=mysim;
    by itm;
    table coverm;
run;
proc freq data=mysim;
    by itm;
    table coverm_l;
run;
proc freq data=mysim;
    by itm;
    table coverm_lq;
run;
proc freq data=mysim;
    by itm;
    table coverm_k;
run;
```

```
proc capability data=mysim noprint;
    var beta_hatm;
    histogram / kernel (k=Normal c=MISE color=RED l=1 w=3)
        cfill=BLUE;
    by itm;
run;
\%mend;

*\%simulatedNL(nsub, nit, nmeas);

*Runs 1-4: da matrix, nmeas=3, nsub={5, 10, 25, 50};
title 'Run 1: Identity r=0 m=3 n=5';
\%simulated3a(5, 1000, 3);
title 'Run 2: Identity r=0 m=3 n=10';
\%simulated3a(10, 1000, 3);
title 'Run 3: Identity r=0 m=3 n=25';
\%simulated3a(25, 1000, 3);
title 'Run 4: Identity r=0 m=3 n=50';
\%simulated3a(50, 1000, 3);

*Runs 21-24: df matrix, nmeas=3, nsub={5, 10, 25, 50};
title 'Run 21: Hetero r=0.9 m=3 n=5';
\%simulated3f(5, 10000, 3);
title 'Run 22: Hetero r=0.9 m=3 n=10';
\%simulated3f(10, 10000, 3);
title 'Run 23: Hetero r=0.9 m=3 n=25';
\%simulated3f(25, 10000, 3);
```

```
title 'Run 24: Hetero r=0.9 m=3 n=50';

\%simulated3f(50, 10000, 3);
```