Analysis of Longitudinal Data in Perinatal Trials when the Length of Follow-up is Informative

Tessa Longstaff

Thesis submitted for the degree of

Master of Philosophy

in

Statistics

at

The University of Adelaide

(Faculty of Engineering, Computer and Mathematical Sciences)

School of Mathematical Sciences

THE UNIVERSITY
of ADELAIDE

May 14, 2016

Contents

List of Tables

List of Figures

Abbreviations

ARGEE	Autoregressive Generalised Estimating Equation
CWGEE	Cluster Weighted Generalised Estimating Equation
EGEE	Exchangeable Generalised Estimating Equation
GEE	Generalised Estimating Equation
ICS	Informative Cluster Size
IEE	Independence Estimating Equation
MM	Mixed Model
NICS	Non Informative Cluster Size
POPPET	Providing Optimal Protein for Prems via Enteral Tubes
SD	Standard Deviation of the Estimates
SE	Model Based Standard Error
WCR	Within Cluster Resampling

Signed Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

SIGNED: ..... DATE: .....

Acknowledgements

I would like to begin by thanking my two supervisors, Associate Professor Gary Glonek and Dr. Lisa Yellend, for their help and guidance throughout my Masters.

I am grateful to the staff at the Women's and Children's Health Research Institute who were involved in conducting the POPPET trial and for sharing the data.

Of course I couldn't do this without the help of my family: my dad, Simon, for always being interesting in what I was doing and my mother, Jacqui, for always encouraging me.

Finally, I would like to thanks my partner, Brock, for his ongoing support and always being there for me.

Abstract

Background: Most commonly used statistical methods assume that the data consist of independent observations. Clustered data occur in many settings, such as longitudinal studies, where outcomes are repeatedly measured over time on each subject. Observations from the same subject are dependent and hence form a cluster. Two commonly used methods of analysis for clustered data are mixed models and generalised estimating equations (GEEs).

Additional complexity arises when analysing clustered data where the cluster size is informative; that is, where the cluster size is related to the outcome. Most methods of analysis for clustered data, including mixed models and GEEs, generally assume non informative cluster size and hence may not be suitable when the cluster size is informative.

Aim: The aim of this thesis is to compare methods for analysing longitudinal data when the cluster size (length of follow up) is informative.

Methods: Both real and simulated data were used to compare methods for analysing clustered data with informative cluster size. A range of methods were considered including: GEEs with independent, autoregressive or exchangeable working correlation structures; cluster weighted GEEs; and mixed models. The real data come from a perinatal trial (the POPPET trial), which investigated the effect of high versus standard protein content human milk fortifier on the growth of 60 preterm infants. This dataset was used to investigate different methods of analysis for estimating the effect of treatment on infant growth when informative cluster size was suspected.

As real data cannot be used to show which methods of analysis are performing best in general, a simulation study was conducted to compare methods when the true parameter values were known. The data were simulated based on the POPPET trial. Different treatment effects, sample sizes, and correlations between the cluster size and the outcomes were considered.

Results: For the POPPET trial, evidence of informative cluster size was found. Different methods of analysis produced quite different parameter estimates but similar conclusions about the effect of the intervention.

The simulation results showed that when cluster size was non informative, all methods performed very well. When cluster size was informative, mixed models and autoregressive GEEs always performed well. However, the independence, exchangeable and cluster weighted GEEs often produced low coverage probabilities and model based standard errors that differed from the standard deviation of the parameter estimates. These methods generally performed better when the trial size was larger and when there was no correlation between individual growth trajectories and cluster size.

Conclusions: It is recommended that mixed models or autoregressive GEEs be used to analyse longitudinal data with informative cluster size in general, including the POPPET trial data. Independence, exchangeable and cluster weighted GEEs should only be used when the sample size is large and there is no correlation between individual growth trajectories and cluster size.

Chapter 1

Introduction

1.1 Motivating Example

The motivation for this research project came from the recently completed POPPET Trial, where POPPET stands for Providing Optimal Protein for Prems via Enteral Tubes (Registration number; AC- TRN12611001275954). This perinatal trial was conducted to assess the safety and efficacy of supplementing expressed breast milk and direct breast feeds with a high protein content human milk fortifier compared to a standard protein content human milk fortifier, in infants born preterm. Gestational age and weight were recorded at birth, where the infants' gestational ages ranged from 28 to 32 weeks. Gestational age was divided into two categories; infants born at less than 30 weeks and infants born at greater than or equal to 30 weeks. Infants were then randomly allocated to one of the two treatment groups (high protein or standard protein), with equal numbers of infants assigned to each treatment group within each of the two gestational age categories. In other words, the randomisation was stratified by gestational age. Weight measurements were taken daily from trial entry until the nasogastric tube was removed, where the timing of removal varied between infants depending on their medical needs. There were 60 infants enrolled in the trial, with 42 remaining at 5 weeks and only 3 at 10 weeks. The number of days spent in the trial (from entry until removal of the nasogastric tube) indicates the length of follow up and will be referred to as duration. One aim of the trial was to estimate the effect of treatment (i.e. the high protein content human milk fortifier) on weight.

The Poppet Trial data are longitudinal due to the repeated measurements taken on each infant. In addition, there are some twins and triplets in the trial. This creates two levels of clustering in the data, infants within mothers and repeated measurements within infants. Any analysis of the POPPET data needs to consider both levels of clustering.

In longitudinal studies multiple measurements of some variable are taken on individuals at various times. In the simplest case, the measurements are taken at regular intervals for a fixed period of time, so that all individuals contribute equally to the analysis. In some studies however, the length of follow up is not fixed but rather depends on the characteristics of the individual. When the length of follow up is not fixed it may be related to the outcomes of interest. This phenomenon is known as informative cluster size (Seaman et al., 2014a). For example, in perinatal trials where treatments are given to the mother and/or her infant to try and improve the health of the infant, measurements are often taken on the infant from birth until some milestone, such as hospital discharge or removal of the nasogastric tube in the case of the POPPET Trial. The timing of this milestone will generally depend on the health of the infant, meaning that sicker infants take longer to reach the milestone and therefore contribute more data to the analysis. This can introduce bias into the estimated treatment effects (Hoffman et al., 2001).

As will be discussed in Chapter 2, methods are available for analysing clustered data when cluster size is informative but it is unclear if they are suitable for longitudinal data.

1.2 Research Aims

The broad aim of this thesis is to identify suitable methods of analysis for longitudinal data when the cluster size or length of follow up is informative. This will allow the POPPET data to be analysed to estimate the effect of treatment on weight, and more general recommendations to be made regarding the use of different methods for analysing longitudinal data when cluster size is informative.

More specifically, this project entails answering the following research questions:

- What methods of analysis are available for analysing clustered data with informative cluster size?
- Are these methods suitable for analysing longitudinal data?
- If multiple methods are suitable, which one is better to use?
- How can simulations be used to help understand the statistical properties of methods for analysing longitudinal data when cluster size is informative?
- Finally, what is the estimated effect of treatment on weight for the POPPET data and how does this vary between analysis methods?

1.3 Thesis Outline

The remainder of this thesis is organised as follows. Chapter 2 provides an overview of the problem of informative cluster size. Different analysis methods for longitudinal data are discussed, including those that do and do not assume non informative cluster size. In Chapter 3, data from the POPPET trial are analysed to estimate the effect of treatment on weight and to compare methods for analysing these type of data. A simulation study based on the POPPET trial is conducted in Chapter 4 to compare analysis methods more generally. Three extensions to the simulation study are considered in Chapter 5 to answer key questions which arose based on the results of the Chapter 4 simulation study. Finally, a summary of the key results and recommendations regarding methods for analysing longitudinal data with informative cluster size are provided in Chapter 6.

Chapter 2

Methods for Analysing Clustered Data

In Chapter 2, the problem of informative cluster size is discussed in detail. The literature is reviewed to identify methods for analysing clustered data. Standard methods of analysis for clustered data and approaches for dealing with informative cluster size are considered.

2.1 Clustered Data

Definition 1 (Cluster)**.** A cluster *is a group of observations that are not independent. Each group of dependent observations forms a cluster.*

Clustered data occur in many settings. Longitudinal data are a special type of clustered data, where each cluster corresponds to a set of repeated measurements on a single individual. For example, consider a study which measures height on individuals every week for a year. Heights from the same individual over time will be dependent, so these data will be clustered.

Clustered data can also occur in cross sectional studies, where each measurement comes from a different subject but there is some dependence present between certain subjects. For example, consider a study which measures blood pressure of patients at several different clinics. There may be dependence between measurements taken at the same clinic due

to similarities in characteristics of patients treated at the same clinic, creating clustered data. In this case, each cluster corresponds to the set of measurements taken on patients at a single clinic.

As clustered data violate the assumption of independent observations, they cannot be analysed using standard methods such as Ordinary Least Squares Regression. Alternative methods that accommodate the dependence within clusters are required.

2.2 Standard Methods of Analysis for Clustered Data

Two commonly used methods for analysing clustered data are mixed models (MMs) and generalised estimating equations (GEEs).

To motivate the development of MMs and GEEs, consider first the ordinary least squares method for $i = 1, ..., K$ independent observations (Dunlop, 1994). Let $\mathbf{Y}$ be a $K \times 1$ vector of outcomes, let $\boldsymbol{e}$ be a $K \times 1$ vector of residuals, let $\boldsymbol{\beta}$ be a $p \times 1$ vector of unknown population parameters and let X be a $K \times p$ matrix of covariates. Then:

$$\mathbf{Y} = X^T\boldsymbol{\beta} + \boldsymbol{e}, \quad \text{where} \quad E(\boldsymbol{e}) = \mathbf{0} \quad \text{and} \quad \text{Var}(\boldsymbol{e}) = \sigma^2 I$$

$$\text{and let} \quad \boldsymbol{\mu} = E(\mathbf{Y}) = X^T\boldsymbol{\beta}.$$

The estimates, $\hat{\boldsymbol{\beta}}$, of $\boldsymbol{\beta}$ are the solutions to the equation:

$$X^T(\mathbf{Y} - \boldsymbol{\mu}) = \mathbf{0}.$$

Both MMs and GEEs arise as extensions of this well known method of analysis.

2.2.1 Mixed Model

The mixed model (MM) is a popular method of analysis that can be used when the assumption of independence between observations is violated (Laird and Ware, 1982). The model includes both fixed and random terms. A fixed term is one that is assumed to be constant across the population, whereas a random term is one that is allowed to vary

between clusters. In general, both the fixed and random effects may be of direct interest. However, in the context of this thesis only the fixed effects will be of direct interest and the random effects will be used only to model the correlations structure. The random terms allow for dependency in the data to be explicitly taken into account. For example, in longitudinal data a random effect can give each individual a different intercept, hence accounting for the dependence among the repeated measurements within each individual. Random slopes are also common for outcomes that vary linearly with time.

MMs have been studied extensively and can be implemented using a variety of different programs, such as R using the lme4 package.

Mixed Model Formulation

Let the clusters be labelled $i = 1, ..., M$ and let the number of members for cluster i be N_i

Let $\mathbf{Y_i}$ be an $N_i \times 1$ vector of outcomes for the ith cluster, let $\boldsymbol{\beta}$ be a $p \times 1$ vector of unknown population parameters, let X_i be an $N_i \times p$ matrix of covariates, let $\boldsymbol{a_i}$ be a $w \times 1$ vector of unknown individual effects for cluster i, let Z_i be an $N_i \times w$ matrix of covariates associated with $\boldsymbol{a_i}$ and let $\mathbf{e_i}$ be an $N_i \times 1$ vector of residuals.

Then for each cluster i the form of the linear mixed model is

$$\mathbf{Y_i} = X_i\boldsymbol{\beta} + Z_i\boldsymbol{a_i} + \boldsymbol{e_i},$$

$$\text{where} \quad E(\mathbf{Y_i}) = X_i\boldsymbol{\beta} \quad \text{and} \quad \text{Var}(\mathbf{Y_i}) = R_i + Z_iDZ_i^T.$$

In this model the $\boldsymbol{\beta}$ are the fixed effects and the $\boldsymbol{a_i}$ are the random effects. Estimates for these parameters can be obtained using MLE or REML ((Harville, 1974) and (Harville, 1976)).

The assumptions of the MM are as follows (Laird and Ware, 1982).

- $\boldsymbol{e_i}$ are distributed $N(\mathbf{0}, R_i)$ where R_i is a $N_i \times N_i$ positive definite covariance matrix
- $\boldsymbol{a_i}$ are distributed $N(\mathbf{0}, D)$ where D is an $w \times w$ positive definite covariance matrix
- $\boldsymbol{e_i}$ are independent random vectors
- $\boldsymbol{a_i}$ are independent random vectors
- $\boldsymbol{a_i}$ are independent of the $\boldsymbol{e_i}$

2.2.2 Generalised Estimating Equations

An alternative method for analysing clustered data is the generalised estimating equations (GEEs) approach proposed by Liang and Zeger (Liang and Zeger, 1986) (Zeger and Liang, 1986). It arises from the Generalised Linear Model (GLM), which is an extension of Generalised Least Squares.

Generalised Least Squares

The assumptions of equal variance and independence made in the ordinary least squares method, described at the beginning of Section 2.2, can be relaxed to obtain the generalised least squares method (Dunlop, 1994):

$$\mathbf{Y} = X^T\boldsymbol{\beta} + \boldsymbol{e} \quad \text{where} \quad E(\boldsymbol{e}) = \mathbf{0}, \qquad \text{Var}(\mathbf{Y}) = \sigma^2 V,$$

$$\text{and } V \text{ is a known, symmetric, positive definite matrix.}$$

The estimates of $\hat{\boldsymbol{\beta}}$ are now the solutions to the equation

$$X^T V^{-1}(\mathbf{Y} - \boldsymbol{\mu}) = \mathbf{0} \quad \text{where} \quad \boldsymbol{\mu} = E(\mathbf{Y}) = X^T\boldsymbol{\beta}.$$

Generalised Linear Model (GLM)

An extension of the generalised least squares method is the the generalised linear model, where the outcomes Y_i are generated from a distribution in the exponential family, e.g.

binomial, Poisson or gamma. A link function g is used to relate $E(Y_i)$ to the linear predictors. For generalised linear models, the $Y_i s$ are assumed to have known variance which is a function of μ_i.

The mean model for GLMs have the form

$$g(\mu_i) = {X_i}^T \boldsymbol{\beta}$$

where

$$E(Y_i) = \mu_i = g^{-1}(X_i^T \boldsymbol{\beta}),$$

$$\text{Var}(Y_i) = \sigma^2 V(\mu_i)$$

and V is an $K \times K$ positive definite, symmetric matrix.

The assumption of independence between observations is still made, which means that the covariance matrix of Y is a diagonal matrix, i.e. $\sigma^2 V = \sigma^2 \,\text{diag}[V(\mu_1), ..., V(\mu_K)]$.

It can be shown (McCullagh and Nelder, 1989) that the MLEs of $\hat{\boldsymbol{\beta}}$, in vector notation, are the solutions to the equation:

$$D^T V^{-1}(\mathbf{Y} - \boldsymbol{\mu}) = \mathbf{0}, \qquad \text{where} \quad D = \frac{\delta \boldsymbol{\mu}}{\delta \boldsymbol{\beta}}$$

As generalised linear models assume the observations to be independent, they are unsuitable for use with clustered data. Generalised linear models were extended by Liang and Zeger to allow for dependence amongst the observations (Liang and Zeger, 1986) using the method of GEEs.

GEE Formulation

Consider M clusters and let N_i be the number of observations in cluster i.

For cluster i, let $\mathbf{Y_i}$ be the $N_i \times 1$ vector of outcomes, X_i be the $N_i \times p$ matrix of covariates and $\mathbf{Y} = [\mathbf{Y_1}, ..., \mathbf{Y_M}]^T$ be the $\sum(N_i) \times 1$ vector of outcomes.

For cluster i, the mean model is

$$g(\boldsymbol{\mu_i}) = X_i^T \boldsymbol{\beta}.$$

The covariance matrix of $\mathbf{Y_i}$ can be written as $\sigma^2 V_i$, where

$$V_i = (\text{diag}[V(\mu_{i1}), ..., V(\mu_{iN_i})])^{1/2} R_i (\text{diag}[V(\mu_{i1}), ..., V(\mu_{iN_i})])^{1/2},$$

$V(\mu_{ij}) = \text{Var}(Y_{ij})$ and R_i is the $N_i \times N_i$ correlation matrix for Y_i. In practice, R_i is unknown and a working correlation structure must be specified. In what follows, R_i will be taken to be the working correlation matrix. Observations within clusters may be dependent but observations across clusters are assumed independent so that the covariance matrix for $\mathbf{Y}$ is $\sigma^2 V$, where the block diagonal matrix V is:

$$V = \begin{bmatrix} V_1 & 0 & \dots & & 0 \\ 0 & V_2 & & & \vdots \\ \vdots & & \ddots & & \\ & & & V_{N-1} & 0 \\ 0 & \dots & & 0 & V_N \end{bmatrix}.$$

A number of different working correlation structures are commonly assumed. The simplest is the independent working correlation structure, in which no correlation is present: $R_i = I$. An exchangeable working correlation is one in which all correlations (ρ) are equal:

$$R_i = \begin{bmatrix} 1 & \rho & \dots & & \rho \\ \rho & 1 & & & \vdots \\ \vdots & & \ddots & & \\ & & & 1 & \rho \\ \rho & \dots & & \rho & 1 \end{bmatrix}$$

Another possible structure is an autoregressive correlation structure, which can be used with longitudinal data and models a diminishing correlation over time. For example, an autoregressive correlation structure with order one has the form:

$$R_i = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \dots & & \rho^{n-1} \\ \rho & 1 & & & & & \rho^{n-2} \\ \rho^2 & & \ddots & & & & \rho^{n-3} \\ \vdots & & & & & & \vdots \\ & & & & & 1 & \rho \\ \rho^{n-1} & \dots & & & & \rho & 1 \end{bmatrix}$$

It is also possible to consider an unstructured working correlation structure. However, with large cluster sizes it is not always possible to apply this method. Unstructured working correlation structures involve many parameters leading to computational and statistical difficulties.

A key property of GEEs is that they are robust to misspecification of the working correlation structure and will lead to consistent parameter estimates, provided the mean model is correctly specified (Liang and Zeger, 1986).

The estimates of $\boldsymbol{\beta}$ are solutions to the equation:

$$\sum_{i=1}^{N} D_i^T V_i^{-1} (\mathbf{Y_i} - \boldsymbol{\mu_i}) = \mathbf{0},$$

$$\text{where} \quad D_i = \frac{\delta \boldsymbol{\mu_i}}{\delta \boldsymbol{\beta}}.$$

These equations are called the generalised estimating equations (Liang and Zeger, 1986). Let $\hat{\boldsymbol{\beta}}$ denote the estimate for $\boldsymbol{\beta}$ found by solving the GEE. The equation has the same form as the Generalised Linear Model but the matrix V can now include off-diagonal terms to allow for dependence within clusters.

The robust or sandwich variance estimate (Σ) is used for variance estimation of parameter estimates and has the form

$$\Sigma = \sum_{i=1}^{N} \left(D_i^T V_i^{-1} D_i \right)^{-1} \left(\sum_{i=1}^{N} D_i^T V_i^{-1} \operatorname{cov}(\mathbf{Y_i}) V_i^{-1} D_i \right) \sum_{i=1}^{N} \left(D_i^T V_i^{-1} D_i \right)^{-1},$$

where cov($\mathbf{Y_i}$) is a sample covariance.

When $R_i = I$, the GEE is reduced to the maximum likelihood equation for a generalised linear model. It will therefore give the same parameter estimates as GLM but different standard errors, because GEEs use robust variance estimation (Zeger and Liang, 1986).

An assumption of GEEs is that observations within clusters are correlated but observations between clusters are independent. In the POPPET Trial, there are sets of twins that create dependence amongst observations across individuals from the same mother. In this case of perfectly nested clusters (measurements are nested within infants who are nested within mothers), accounting for clustering on only the top level cluster (mothers) is required (Miglioretti and Heagerty, 2006). Such structures can thus be accommodated in the GEE framework.

2.2.3 Comparison

The key difference between MMs and GEEs is the way in which they handle the dependency between measurements taken from the same cluster. MMs explicitly account for the dependence by the addition of the random effects terms, whereas GEEs implicitly take into account the dependency through the working correlation matrix, R_i.

The type of model that should be used to analyse clustered data depends on whether the aim of the analysis is to estimate population-averages or cluster specific parameters ((Hu et al., 1998); (J. M. Neuhaus, 1991); (Ritz and Spiegelman, 2004)). GEEs are population averaged models, as they are used to estimate population averaged effects. Population averaged models specify a marginal mean function in term of the covariates, hence GEEs are also known as marginal models. MMs are an example of a cluster-specific model, as they are used to estimate cluster-specific effects. Cluster-specific models specify a conditional mean function in terms of a random effect and the covariates, hence MMs are also known as conditional models. For the linear model, marginal and conditional

parameters coincide so the choice between MMs and GEEs is less important here than for other link functions (Ritz and Spiegelman, 2004). Both MMs and GEEs will be considered for the POPPET data as a linear model will be assumed.

2.3 Informative Cluster Size

2.3.1 Definition

Definition 2 (Informative Cluster Size (ICS))**.** Informative cluster size *occurs when the cluster size is not independent of the outcome. That is (Seaman et al., 2014a):*

$$E[Y|\mathbf{X} = \mathbf{x}, N] \neq E[Y|\mathbf{X} = \mathbf{x}] \quad \textit{for some } x \tag{2.1}$$

where Y is the outcome variable, $\mathbf{X}$ is the covariate vector for a random member of a cluster and N is the cluster size.

2.3.2 Examples

To understand how informative cluster size (ICS) arises, we consider several examples of clustered data.

A Study on Periodontal Disease

Consider a study on periodontal disease, that measures whether each tooth of an individual is diseased or not (Hoffman et al., 2001). The outcome is binary and takes the value 1 if the tooth has periodontal disease and 0 otherwise. Every tooth from each individual is included in the study and hence contributes to the analysis. As outcomes of teeth from the same individual are dependent, this study produces clustered data in which the clusters are the individuals and the cluster members are the teeth. Not all individuals will have the same number of teeth, so cluster size is not fixed and may be related to the outcome. In fact, factors which lead to periodontal disease may also lead to tooth loss, so those with the disease may have already lost teeth as a result. This would cause a negative relationship between the number of teeth and the disease status. That is, the

less teeth an individual has the higher the chance of disease on their teeth. Hence, as the outcome is not independent of the cluster size, this is an example of ICS.

An outcome measured at repeated visits to the doctor

Consider an outcome, such as general heath, that is measured each time an individual visits the doctor. Each individuals' measurements will form a cluster, as measurements from the same individual will be dependent. The cluster size will be the number of visits to the doctor. As cluster size is not fixed it could be related to the outcome of interest. For example, sicker individuals may visit the doctor more often and hence contribute more to the analysis. This could result in a negative relationship between the cluster size and the health of the individual. That is, the larger the cluster size the worse, the general health outcome is. This is again an example of ICS, as the cluster size is not independent of the outcome.

A study on blood pressure

Consider a study on the blood pressure of blood donors, in which the blood pressure is measured each time the individual attends the dontation clinic. Each indiviuals' measurements will form a cluster, as measurements from the same individual will be dependent. The cluster size will be the number of visits to the donation clinic. Cluster size is not fixed but it will not be related to the outcome of interest in this case, because the number of blood donations a donor makes will not be related to their health or blood pressure. Hence, this is an example of non informative cluster size.

2.3.3 Issues Arising for Data with ICS

Non-informative cluster size (NICS) is an assumption in general for the standard methods of analysis for clustered data, including MMs and GEEs. For MMs, one of the assumptions is that the random effects $\mathbf{a_i}$ are independent and hence independent of N_i. If $\mathbf{a_i}$ is related to N_i this suggests ICS is present ((Seaman et al., 2014b), (Neuhaus and Mcculloch, 2011), (Chen et al., 2011)). For GEEs, it is generally assumed that the same mean and variance models apply within each cluster, irrespective of size. As such, this method

does not allow for ICS.

When ICS is present, standard methods of analysis may result in bias in the parameter estimates. Very different results can occur depending on which method is chosen (Hoffman et al., 2001). If a GEE is used when ICS is present, choosing a different working correlation structure can lead to substantially different parameter estimates (Hoffman et al., 2001).

Example

To illustrate that different methods can produce very different results when ICS is present, consider again the periodontal example discussed in Section 2.3.2.

The periodontal study data were analysed using two different methods (Hoffman et al., 2001): a GEE with an independent working correlation structure (IEE) and a GEE with an exchangeable working correlation structure (EGEE). The analysis model included a single predictor variable indicating whether or not the participant brushed their teeth at least twice a day. The results of the analysis were as follows:

	est. IEE	est. EGEE
Intercept	0.0685	-0.389
Brush at least twice	-0.786	-0.499

Table 2.1: Coefficients for the IEE and EGEE methods

	est. IEE	est. EGEE
Dont Brush at least twice	0.52	0.40
Brush at least twice	0.33	0.29

Table 2.2: Probability of disease for the IEE and EGEE methods

Tables 2.1 and 2.2 show that the two analysis methods produce quite different results. For

example, the probability of disease for those who don't brush their teeth twice a week was estimated as 0.40 for EGEE and 0.52 for IEE. GEEs are robust to the misspecification of the working correlation, so presumably the discrepancy is the result of ICS. Hence, in the presence of ICS the choice of analysis method makes a substantial difference to the results for this data set. This is problematic, as it is unclear which method should be used to draw conclusions from the data.

2.3.4 Relationship with Missing Data

A common misconception is that ICS is really just a missing data problem. Missing data refer to observations that are missing but should have taken values ((Schafer and Graham, 2002) and (Cummings, 2013)). Missing data are a different problem to ICS. Although ICS can arise when cluster sizes vary due to missing data, it can also occur in the absence of missing data. For example, consider a study involving singletons and twins with no missing measurements, giving a maximum cluster size of two. If ICS was a missing data problem, clusters of size 1 would be considered to have missing values. This does not make sense, as this implies all the singletons would be missing a twin. As this is not the case it is not a missing data problem but ICS may be present if the outcome of interest varies between singletons and twins, which is often the case in practice. This thesis focuses on methods for dealing with ICS, and will not consider methods for dealing with missing data.

2.4 Approaches for Dealing with ICS

When cluster size is informative, standard methods of analysis for clustered data may result in bias and hence alternative methods are required to provide valid inference ((Hoffman et al., 2001), (Benhin et al., 2005), (Seaman et al., 2014a)). Alternative methods have been proposed, but their applicability to longitudinal data is unclear.

In this section, methods of analysis that allow ICS will be considered. These include: within cluster resampling; including cluster size as a covariate; joint models for the out-

come and the cluster size; GEEs with independent correlation structure; and cluster weighted GEEs.

2.4.1 Within Cluster Resampling

Within Cluster Resampling (WCR) is a method that provides valid inference even when the cluster size is informative (Hoffman et al. 2001). For a dataset with M clusters, the method is as follows:

1. Randomly sample one observation from each of the M clusters. This is the resampled data set and its observations are independent. Repeat this process K times, where K is large.

2. Analyse each of the K samples to obtain K valid estimates of the parameter of interest, denoted $\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}, k = 1, ..., K$. The estimate, $\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}$, is found by solving the estimating equation from the Generalised Linear Model (since observations in each resampled dataset are independent) (Section 2.2.1).

3. The WCR estimator, $\bar{\beta}$, is then the average of the K estimates obtained:

$$\bar{\boldsymbol{\beta}} = K^{-1} \sum_{k=1}^{K} \hat{\boldsymbol{\beta}}_{\boldsymbol{k}}$$

This method can be used even when cluster size is informative because it weights clusters equally, unlike GEEs. Only one observation from each cluster is used each time the parameters of interest are estimated, so the effects of ICS do not affect the results found using WCR. Since all clusters are weighted equally, the parameters will have a cluster level interpretation, i.e. parameters apply to a randomly selected member from a randomly selected cluster.

Note that under regularity conditions, as $M \to \infty$ $M^{1/2}(\bar{\boldsymbol{\beta}} - \boldsymbol{\beta}) \to N(0, \Sigma)$, where Σ is a finite positive definite matrix and $\boldsymbol{\beta}$ is the true parameter for a random member of a random cluster.

Each estimate $\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}$ comes from a dataset where all observations are independent, but the K datasets are correlated (because they are all samples drawn from the same dataset), hence the K estimates are dependent. This dependence needs to be taken into account in the variance formula of $\bar{\boldsymbol{\beta}}$. The asymptotic variance estimator is:

$$\hat{\Sigma} = \hat{\text{Var}}(M^{1/2}(\bar{\boldsymbol{\beta}} - \boldsymbol{\beta})) = M(K^{-1}\sum_{k=1}^{K}\hat{\Sigma}_k - \sum_{k=1}^{K}K^{-1}(\hat{\boldsymbol{\beta}}_{\boldsymbol{k}} - \bar{\boldsymbol{\beta}})(\hat{\boldsymbol{\beta}}_{\boldsymbol{k}} - \bar{\boldsymbol{\beta}})^T)$$

where $\hat{\Sigma}_k$ is the covariance matrix estimated from the kth resample.

Within cluster resampling is a very simple method and involves using only the generalised linear model, which is very easy to implement. However, it is a computationally inefficient method, and the results will also depend on the random samples chosen (Hoffman et al., 2001). For these reasons, the approach is rarely used in practice and will not be considered further in this thesis.

2.4.2 Joint Models

For mixed models, one way to relax the assumption that the random effects are independent of the cluster size, is to jointly model the outcome and the cluster size ((Dunson et al., 2003), (Gueorguieva, 2005), (Seaman et al., 2014a) and (Chen et al., 2011)). This is achieved by combining a mixed model for the outcome, with a model for the distribution of the cluster size.

One option for a joint model is a shared parameter model. This jointly models the cluster size (N_i) and the outcome (Y_i), by requiring that the N_i and Y_i are conditionally independent given random effects ($\mathbf{a_i}$), hence $\mathbf{a_i}$ is the shared parameter.

This method is theoretically appealing and it would be plausible to implement this method. However, this method is beyond the scope of this thesis and will therefore not be considered further.

2.4.3 Include Cluster Size as a Covariate

If cluster size is included as a covariate in either a MM or a GEE both become valid model choices. Recall the definition of NICS (Seaman et al., 2014a) :

$$E[Y|\mathbf{X} = \mathbf{x}, N] = E[Y|\mathbf{X} = \mathbf{x}] \quad \forall x \tag{2.2}$$

where Y is the outcome for a random member of a random cluster and $\mathbf{X}$ is the vector of covariates for a random member of a random cluster. Including cluster size as a covariate ensures that this relationship holds for all clusters and hence cluster size is non informative.

However, it is not always appropriate to include cluster size as a covariate in the analysis. For example, consider a clinical trial where the exposure of interest is the randomised treatment. Any covariate determined after treatment is given (i.e. after randomisation) should not be adjusted for, as including any such covariates can remove part of the effect of the treatment on the outcome. This is because these covariates could themselves be affected by the treatment. Hence if cluster size is determined after treatment is given, it is not appropriate to include it as a covariate (the Committee for Proprietary Medicinal Products, 2004).

If cluster size cannot be adjusted for directly, an alternative is to adjust for baseline covariates that are expected to affect the cluster size. This means that covariates that might influence the cluster size, but were determined before the exposure of interest, are included instead. This does not always guarantee NICS, as it does not directly ensure that equation 2.2 holds. Further testing is then needed to determine whether the inclusion of baseline covariates has resulted in NICS and if so then standard methods, such as MMs and GEEs, become appropriate.

2.4.4 GEE Based Approaches

In general, GEEs do not provide valid inference when cluster size is informative. However, two specific types of GEEs have been show to give valid inference: cluster weighted GEEs (CWGEE) and independence estimating equations (IEE) (Williamson et al., 2003). These

two approaches estimate different associations of potential interest when ICS is present.

Cluster Weighted Generalised Estimating Equation Approach (CWGEE)

The CWGEE approach is used to describe the association between Y and X for a random member of a random cluster. ((Benhin et al., 2005), (Yelland et al., 2015)). It involves fitting a GEE with an independent working correlation matrix and weighting each cluster inversely by cluster size, such that each cluster contributes equally to the analysis. Weightings for clusters is an optional input for most GEE functions, and is easily implemented in a variety of statistical programs.

This Cluster Weighted GEE (CWGEE) approach is asymptotically equivalent to WCR. That is, as K approaches infinity, the WCR estimate $\bar{\beta}$ approaches the solution to the CWGEE. However, CWGEE is much simpler as it only requires a single model and is recommended in preference to WCR (Williamson et al., 2003)

Independence Estimating Equation Approach (IEE)

The IEE approach is used to describes the association between Y and X among all cluster members (Yelland et al., 2015). It involves fitting a GEE with an independence working correlation matrix, where no weights are specified and hence each observation contributes equally to the analysis. This independence estimating equation (IEE) approach makes more sense with cross sectional clusters than with longitudinal data, since interest is often in cluster level effects with longitudinal data. Hence, weighting each cluster equally in the analysis using CWGEE is more appropriate.

Comparison between CWGEE and IEE

The CWGEE and IEE methods differ in the parameter they estimate and are based on different marginal models, both of which may be of interest in practice. To understand the difference between these two approaches, consider examples by simplifying the POPPET trial data in two ways.

Example 1- Cross Sectional Study

Consider removing the longitudinal aspect of the trial and only taking one measurement at a fixed time point for each infant within mother. In this case, clusters correspond to mothers, with one observation for each infant. The size of each cluster is the number of infants born to each mother. To illustrate, we select three mothers from the POPPET trial with a total of 5 infants. Two mothers had singletons with durations of 26 and 22 days, and gestational ages of 30 weeks. The third mother had triplets with durations of 47, 44 and 43 days, and a gestational age of 30 weeks. At day 22, (the last day when all infants remained in the trial) the weights for the five infants were: 2545, 2390, 1915, 2225 and 2100 grams respectively.

If we fit an IEE to the weights on day 22, then each infant will be given equal weighting and the sample mean will be:

$$\bar{y} = (2545 + 2390 + 1915 + 2225 + 2100) * \frac{1}{5} = 2235g.$$

In contrast, if we fit a CWGEE then the singletons will each be given a weight of one and the triplets will each be given a weight of $\frac{1}{3}$, so their mother has a total weight of 1. The sample mean will then be:

$$\bar{y} = (2350 + 2390 + (1915 + 2225 + 2100) * \frac{1}{3}) * \frac{1}{3} = 2273.3g.$$

These two approaches result in different estimated means $\bar{y}$, with the CWGEE giving a mean that is 38.3 grams larger than IEE. Hence, it is important to choose the correct method depending on the parameter of interest. In this case, an IEE may make more

sense because the interest will likely lie in the average infant weight, rather than the average infant weight across mothers.

Example 2 - Longitudinal Study

Consider now simplifying the POPPET trial to remove the multiple births and leave only one level of clustering (the longitudinal aspect). In this case, the infant is the cluster and the size of the cluster is the number of measurements taken on the infant. To illustrate, again consider only 5 infants. These were chosen from the POPPET singleton infants and had durations of 14, 72, 30, 14 and 58 days and sum of daily weight measurements over these durations of 32570, 140220, 67410, 35304, 105815 grams respectively.

If we fit an IEE then each repeated measurement within each infant will be given equal weighting and the sample mean will be:

$$\bar{y} = (32570 + 140220 + 67410 + 35304 + 105815) * \frac{1}{14 + 72 + 30 + 14 + 58} = 2028.3g.$$

Alternatively, if we fit a CWGEE, each infant (cluster) will be inversely weighted by the duration (cluster size), giving a sample mean of:

$$\bar{y} = (32570 * \frac{1}{14} + 140220 * \frac{1}{72} + 67410 * \frac{1}{30} + 35304 * \frac{1}{14} + 105815 * \frac{1}{58}) * \frac{1}{5} = 2173.4g.$$

Again, these result in different estimated means $\bar{y}$, where the CWGEE has a mean that is 145.1 grams larger than IEE.

Both these examples illustrate how different the estimates based on IEE and CWGEE can be when cluster size in informative, and the importance of specifying the parameter of interest when choosing between these methods.

2.4.5 Approaches for the POPPET Trial

GEE Approaches for the POPPET Trial

The POPPET trial has two types of clustering: cross sectional clustering (infants within mothers) and also longitudinal clusters (repeated measurements within infants). Due to this complexity, it is not clear whether an IEE or a CWGEE is more appropriate. Hence, both methods will be considered when analysing the POPPET trial. The IEE, which weights cluster members equally, would result in all measurements from all infants having equal weighting in the analysis. The CWGEE, which weights clusters equally, would result in weights inversely proportional to duration for singletons and inversely proportional to the sum of the durations for multiple births.

Mixed Model Approaches for the POPPET Trial

MMs will also be considered for the POPPET trial. The two levels of clustering will be directly taken into account by the inclusion of random effects for each. The method of including baseline covariates will also be considered for a MM in an attempt to remove ICS and hence make the MM an appropriate choice for analysis.

2.5 Summary of Methods

When ICS is absent, clustered data can be analysed using MMs or GEEs. However, when cluster size is informative, an alternative method is needed for analysis. Some alternatives have been discussed in this chapter including: WCR, including cluster size as a covariate, joint models, IEE and CWGEE. It is, however, unclear which method should be chosen for longitudinal data with ICS. This issue will be addressed by first considering a real data set in Chapter 3 and then via simulation in Chapter 4 and 5.

Chapter 3

POPPET Data Analysis

In Chapter 2, the concept of clustered data was explored. Extensions to the simple linear model were discussed that allow for the dependence within clusters to be taken into account in the analysis. This included methods of analysis which can be used when cluster size is informative.

In Chapter 3, a real clustered data set will be analysed using the methods described in Chapter 2, and the property of ICS for this data set will be explored. All analysis and simulations in this chapter and throughout this thesis will be performed using R.

The data are from the POPPET Trial, introduced in Section 1.1. The trial was conducted to assess the effect of high protein content human milk fortifier on the growth of 60 preterm infants, where measurements were taken daily from trial entry until the removal of the nasogastric tube. Recall, the number of days spent in the trial (from entry till removal of the nasogastric tube) will be referred to as duration. To determine whether treatment affected growth, these data will be analysed to estimate the effect of treatment group on the daily weight measurements.

The POPPET Trial data set has two levels of clustering. The data are longitudinal due to the repeated measurements taken on each infant. There is also cross sectional clustering present due to some mothers having a multiple birth. This creates two levels of nested

clustering; infants within mothers and repeated measurements within infants.

Due to the two levels of clustering, the total cluster size here is the total number of observations from infants within the same mother. That is, for multiple births the cluster size is the sum of the infants' durations and for singletons it is just their duration. Alternatively, the cluster size at each level of the clustering can be considered separately. At the mother level, the cluster size is the number of infants each mother has and at the infant level it is each infant's duration. In this thesis the cluster size will be considered at each level separately when examining the data for evidence of ICS.

ICS is a potential a problem for this data set. Recall that ICS occurs when the cluster size is related to the outcome. If ICS is present, one of the methods discussed in Section 2.4 may be needed for analysis. We therefore investigate whether ICS is present in Section 3.2. Methods assuming ICS and NICS will be applied to the POPPET data and the results compared.

Misspecification of the correlation matrix is also a potential problem for the POPPET Trial data. This is because the correlation between twins will be stronger if they are monozygotic (one egg fertilised by one sperm that splits) vs dizygotic (two eggs fertilised by two sperm) and also monochorionic (shared placenta) vs dichorionic (separate placenta). As we do not have the information regarding zygosity or chorionicity, we cannot model the correlation matrix properly. Instead, simplifying assumptions such as equal correlation for all sets of twins must be made.

In addition, there are a number of missing weight measurements in this dataset. Not all infants have weight measurements recorded for each day they were in the study and the analysis was performed using the available data. The missing data are expected to have little impact on the results, due to the large durations and small number of missing observations. There are 80 missing measurements out of 2411, which is less than 3.4%.

3.1 Variables in the POPPET Data Set

The variables from the POPPET data set that will be considered in this thesis are listed in Table 3.1. The table gives the name of the variable, which will be used throughout, and a description of the variable.

Variable Name	Description
Duration	The number of days spent in the trial from entry until removal of the nasogastric tube (length of follow up).
Plurality	The number of infants born to the mother. Given values 1 for singletons and 2 for multiple births.
Time	Time in weeks since trial entry.
Gestational Age	Gestational age at birth measured in weeks.
Group	The randomised treatment group.
Sex	The gender of the infant.
Birthweight	The birthweight of the infant in grams.
Weight	The weight of the infant measured daily in grams (outcome).

Table 3.1: Variables to be used in the analysis of the POPPET data set

3.2 ICS for the POPPET Data

For the POPPET data set ICS could be present and it is simplest to investigate this by considering the relationship between cluster size and the outcome separately for each level of clustering.

The number of infants each mother has may be related to their weights. It is very common for infants from multiple births to have lower weights compared to singleton infants. This indicates the cluster size (at the mother level) is expected to be related to the outcome, which would create ICS. However, in this case ICS at the mother level will not be a problem, because plurality can be included in the model as it is a baseline covariate and

cannot be influenced by treatment. If the cluster size is included as a covariate, then ICS cannot be present according to the definition (See Section 2.4.3).

As the length of follow up in this trial is not fixed, it may be related to the outcome of interest. That is, the duration of the infant's time in the trial may be related to their weight. If this is the case, then ICS is present (at the infant level). Problematically, ICS at the infant level can not easily be accounted for, since duration is a post randomisation variable that could be influenced by treatment. Thus, it cannot be included in the analysis model to remove ICS and ICS is potential a problem for the POPPET data.

To investigate informally whether ICS is present for the POPPET data, consider Figure 3.1. The y axis is the weight of the infant in grams and on the x-axis is time, where 0 represents a full term pregnancy of 40 weeks. Hence, $x = -84$ would represent a trial entry time for an infant born at 28 weeks. Each line on the plot is an individual infant's trajectory showing their weight from when they entered the study until the day they exited the study. As gestational ages vary in this trial, not all trajectories begin at the same point on the time axis.

An interesting aspect of this plot is that there is an apparent truncation of the lines at the top of the plot. This arises because discharge weights were quite similar for all infants. If all infants are being discharged at roughly the same weight, then low birthweight infants will remain in the trial for longer, as they take longer to reach this discharge weight than high birthweight infants. This pattern suggests that the duration is related to the weight and hence ICS is present.

ICS can also be investigated using a MM. For a MM the assumption of NICS is that the random effects are independent of the cluster size. By fitting a MM to the POPPET data, the correlation between the random effects and duration can be calculated. If it is non zero, there is evidence to suggest that ICS is present. Investigating potential ICS using a MM will be considered in Section 3.3.

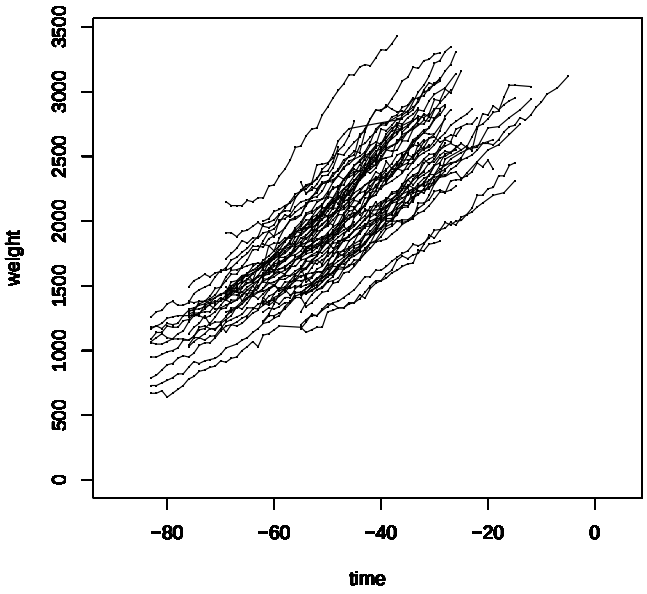

Figure 3.1: Scatterplot of weight against time for each infant

3.3 Mixed Model Analysis

The MM, described in Section 2.2, was fit to the POPPET data. To begin with, a model building procedure was applied to ensure that the usual assumptions of the MM were met. These assumptions include normality of the residuals and random effects, linearity and homoscedasticity. As well as this, is the assumption of non informative cluster size, which is that the random effects are independent of the cluster size. This was not assessed until the final model. If all of these assumptions hold, the MM should provide valid inference.

3.3.1 Model Developement

For the POPPET data, a MM could be used to estimate the effect of treatment on weight from trial entry until removal of the nasogastric tube. Recall the MM includes both fixed and random terms, where the random terms are used to account for the clustering explicitly. In this case, there are two levels of clustering and each can be allowed for with an appropriate random effect. A random effect for each infant is needed to account for the dependence in the outcomes between repeated measurements. This will give a different

intercept for each infant. As weight varies with time we will also allow for the slope to vary between infants by including a random slope. For simplicity the infant level intercept, slope and quadratic random effects are assumed to be independent. While such a model may not necessarily be a good fit to the data, nevertheless provides a realistic basis for our simulation studies. A random effect for each mother is also needed to account for the dependence in the outcomes between infants with the same mother. This will give a different intercept for each mother. Fixed effects for group and time will be included in the model, as well as stratification variables of plurality and sex.

Treatment group will be included in the model in the form of an interaction term between group and time but a main effect for group is not considered. *The Principle of Marginality* (McCullagh and Nelder, 1989) states that higher order effects or interactions are marginal to the main effect. This means that all main effects and lower order effects would usually be included in the model if the higher order effect is included. This principle is not applicable in this case because it does not make sense to include a main effect for treatment in this trial. If a main effect were included, then a difference between the two groups at time zero could occur. As it was a randomised trial, no difference between the two groups can occur at time zero, except by chance. On the other hand the difference between growth rates can be modelled by including an interaction without a main effect. Consider that treatment group is a factor which is either a 0 or a 1. If this is multiplied by the time variable, one group will have values which are zero at all time points and the other will have values which change over time, depending on the estimated coefficient. Note, a main effect of time will be included as well as this interaction between group and time. This will allow the two groups to grow apart over time as one group will have an extra term in the model, which changes over time. These ideas are illustrated in Figure 3.2. The left part shows how the two groups can have a difference at zero if a main effect is included and the right part shows how the groups can grow at different rates by including an interaction term but no main effect for treatment.

For each infant i the MM takes the form:

$$\mathbf{Y_i} = X_i\boldsymbol{\beta} + Z_i\boldsymbol{a_i} + \mathbf{e_i} \tag{3.1}$$

where X_i consists of time, gestational age, sex, plurality and group $\times$ time effects and Z_i consists of a random effect for mother, a random intercept for infant and a random slope for infant. The $\boldsymbol{e_i}$ are distributed $N(\mathbf{0}, R_i)$ where R_i is a $N_i \times N_i$ positive definite covariance matrix and the $\boldsymbol{a_i}$ are distributed $N(\mathbf{0}, D)$ where D is an $w \times w$ positive definite covariance matrix.

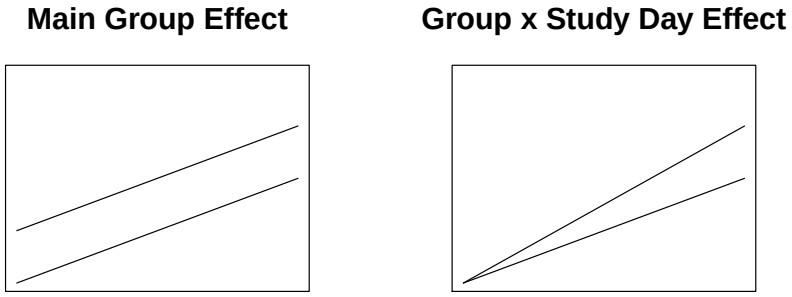

Figure 3.2: Plots of main group effect vs group x time interaction, with no main effect

Model 3.1 was fit to the POPPET data, where here it is assumed that $R_i = I$. The assumptions of the MM were considered to determine whether the model is adequate. The assumptions of linearity and homoscedasticity can be checked with a scatterplot of the residuals vs the fitted values, given in Figure 3.3. It is noted that the residuals used do not include the random effects. Figure 3.3 shows there is an obvious upward curve to the data. This means there is not random scatter or even spread about the zero line, so the assumptions of linearity and homoscedasticity do not hold. To attempt to account for this curvature, a fixed quadratic term in time and an interaction between time squared and treatment group was added to model 3.1. Time was not centered before its inclusion as a quadratic term. This new model was fit and the coefficient of the quadratic time terms were found to be highly statistically significant ($p < 0.0001$). Again, the assumption of linearity and homoscedasticity was checked and the residuals vs fitted plot of the

quadratic model is shown in Figure 3.4. In this case, there is no obvious pattern and the upward curvature seen in Figure 3.3 is no longer present. This indicates the assumptions of linearity and homoscedasticity are satisfied for the quadratic model.

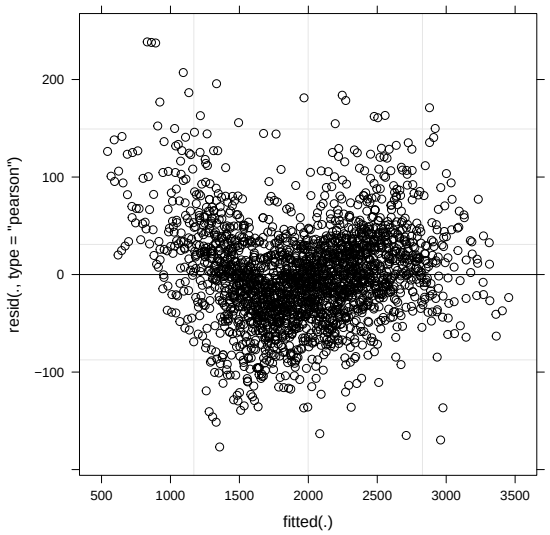

Figure 3.3: Scatterplot of residuals vs fitted values

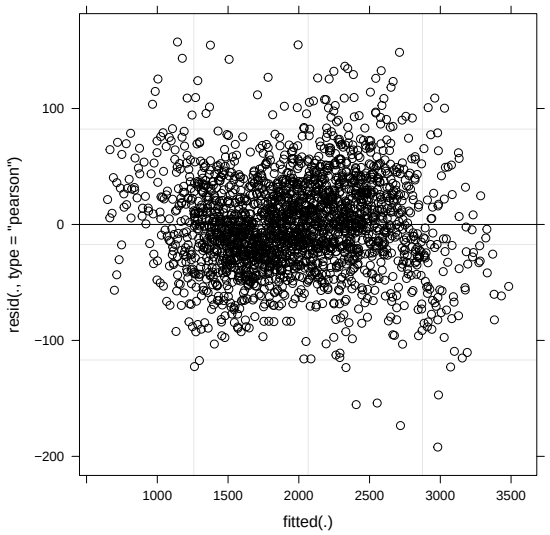

Figure 3.4: Scatterplot of residuals vs fitted values for quadratic model

To check the normality of the residuals and of the random effects, normal quantile plots are used. Figure 3.5 and 3.6 give the normal quantile plots of the residuals and the ran-

dom effects respectively. The normal quantile plot for the random effects appears roughly linear. The normal quantile plot for the residuals is also linear except at the tails where it deviates slightly.

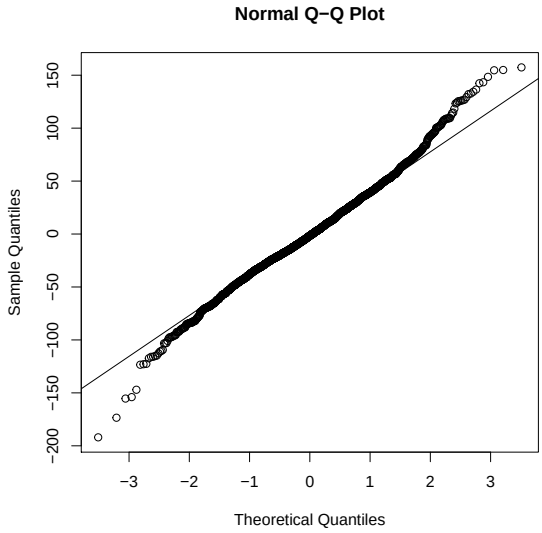

Figure 3.5: Normal quantile plot of residuals

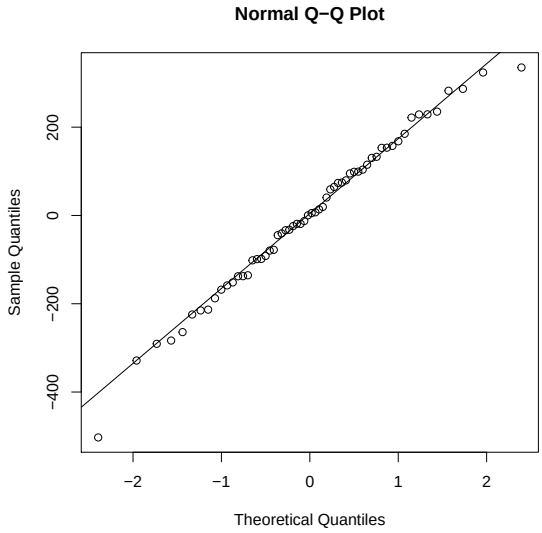

Figure 3.6: Normal quantile plot of random effects

As a quadratic fixed effect is included in the model, a quadratic random effect for each infant (different curve for each infant) could also be considered in this case. Model 3.1

was modified to include a random quadratic effect. To compare this model to the model without a quadratic random effect, a likelihood ratio test of the nested models is used. This tests whether the difference in the residual sum of squares is statistically significant and can only be used for nested models. In this case, the two models are nested so the test can be used and is performed using the `anova` function in R. The p-value is: $p < 0.0001$ with 3 degrees of freedom, which indicates there is a statistically significant difference between the models. Hence, the model including the quadratic random effect should be used.

We use the estimated random effects from this model to investigate potential ICS in these data.

3.3.2 Informative Cluster Size

For the MM the assumption of NICS implies the random effects are independent of the cluster size. There are four random effects to consider: random infant intercept, slope and quadratic effect, and random mother intercept. In this case, the assumption of NICS implies the infant level random effects are independent of duration and the mother level random effect is independent of plurality. As plurality is included in the model there cannot be ICS on the mother level, hence it does not need to be considered. The assumption of NICS for the infant level does need to be considered. Plots of duration vs the infant level random effects are given in Figure 3.7. The plots show a negative relationship between duration and each of the infant level random effects. That is, as the random effects decrease the duration increases, where the sample correlations are $r = -0.710, -0.345, -0.208$ for the infant intercept, slope and quadratic effect respectively. These show that the correlation is non zero and is very large for the random infant intercept. This represents strong evidence of ICS in the POPPET data.

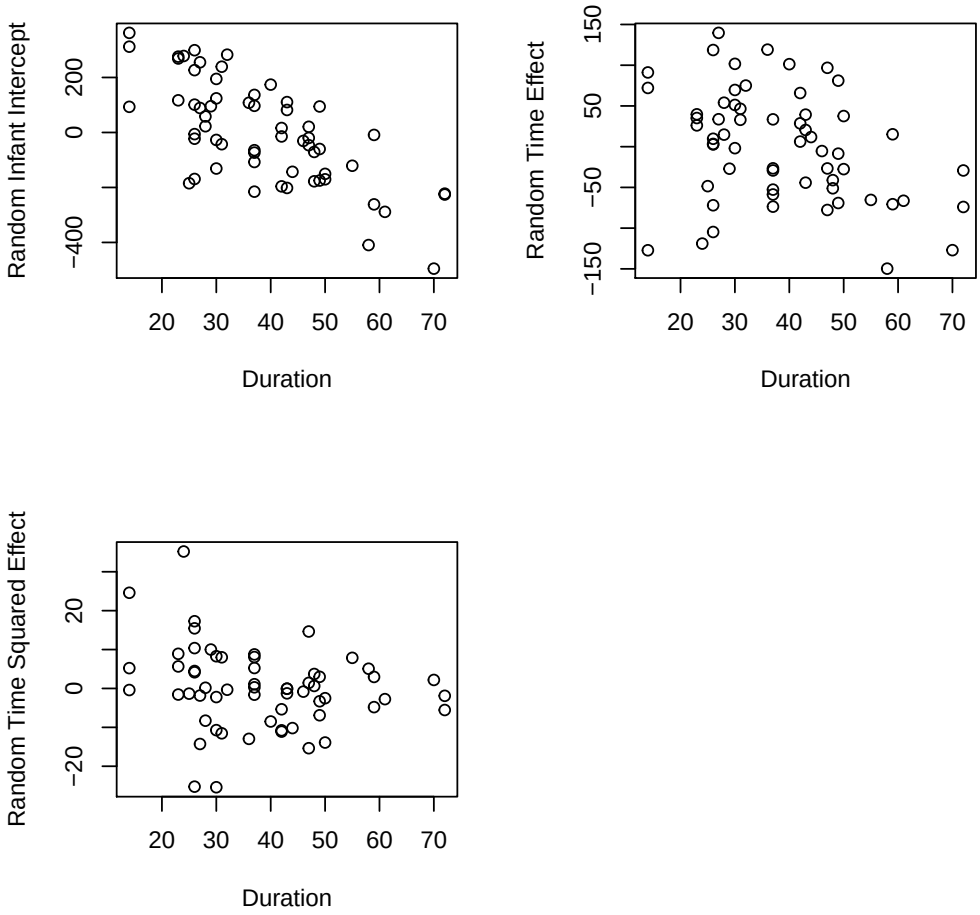

Figure 3.7: Scatterplot of random effects vs duration

3.3.3 Reducing Informative Cluster Size by Adjusting for Baseline Covariates

Recall from Chapter 2 that if cluster size is included as a covariate, there cannot be ICS. However, it is not always possible to include cluster size, in X_R, directly. In this case, cluster size (duration) is a response variable that is potentially influenced by the treatment. Thus, it should not be included as a covariate, as adjusting for a variable that may have been influenced by treatment can distort the estimated treatment effect (CPMP, 2004).

As duration cannot be accounted for directly, an alternative approach is to consider

baseline covariates, that might influence duration. To investigate possible covariates, consider a simple linear model with log(duration) as the outcome and baseline covariates that may affect duration as predictors. This model is:

$$\log(\mathbf{N}) = X^T\boldsymbol{\beta} + \mathbf{e}$$

where X includes group, sex, plurality, gestational age and birthweight.

This model was fit in R and the results are summarised in Table 3.2. Using backwards selection, based on a p-value criteria, group was removed from the model first with a p-value of 0.93. After refitting, plurality was removed with a p-value of 0.38 and finally sex was removed with a p-value of 0.40. Thus, the final model included terms of birthweight and gestational age and is given in Table 3.3. This implies that duration is significantly affected by these two terms. Adding these terms to the model may reduce NICS. However NICS is not guaranteed, as would be the case for duration, because it does not ensure equation 2.2 holds.

	estimate	s.e	t-value	p-value	ci lower	ci upper
Intercept	8.067	0.83	9.698	< 0.0001	6.40	9.73
Group	-0.0056	0.0597	-0.094	0.93	-0.13	0.11
Sex(Male)	-0.066	0.066	-1.00	0.32	-0.20	0.066
Plurality(2)	0.044	0.051	0.87	0.39	-0.057	0.15
Gestational Age	-0.13	0.029	-4.47	< 0.0001	-0.19	-0.072
Birthweight	-0.00037	$9.80 \times 10^{-}5$	-3.83	0.00034	-0.0006	-0.00018

Table 3.2: Summary of initial duration model

The MM for weight including birthweight and gestational age in weeks as fixed effects was fit in an attempt to remove ICS. It should be noted that there is potential for collinearity between the gestational age and birthweight terms. Removing one term could change the

	estimate	s.e	t-value	p-value	ci lower	ci upper
Intercept	7.81	0.69	11.33	< 0.0001	6.43	9.18
Gestational Age	-0.12	0.025	-4.70	< 0.0001	-0.17	-0.068
Birthweight	-0.00041	0.000090	-4.60	< 0.0001	-0.0006	-0.00023

Table 3.3: Summary of final duration model

estimates but as both terms were significant they were retained in the model. Again, ICS was examined using the correlations between cluster size and each of the three infant level random effects. Plots of duration vs the infant level random effects are given in Figure 3.8. The first plot between duration and the random infant intercept shows no obvious pattern. This is very different to Figure 3.7, where it was found to have a very strong negative relationship before fixed effects for birthweight and gestational age were included in the model. The other two plots remain mainly unchanged by the inclusion of these new terms. In this case, the sample correlations were found to be $r = -0.028, -0.367, -0.152$ for the intercept, slope and quadratic effects respectively. These show there is still a negative linear relationship between the random infant slope and duration and a weaker relationship between the random quadratic effect and duration.

Including birthweight and gestational age in the MM has removed the correlation between the random infant intercept and duration (sample correlation changed from -0.710 to -0.028), but has not had the same effect for the other random infant effects. Hence, including baseline covariates that influence cluster size has reduced but not completely removed ICS. This means the MM may still not provide valid inference.

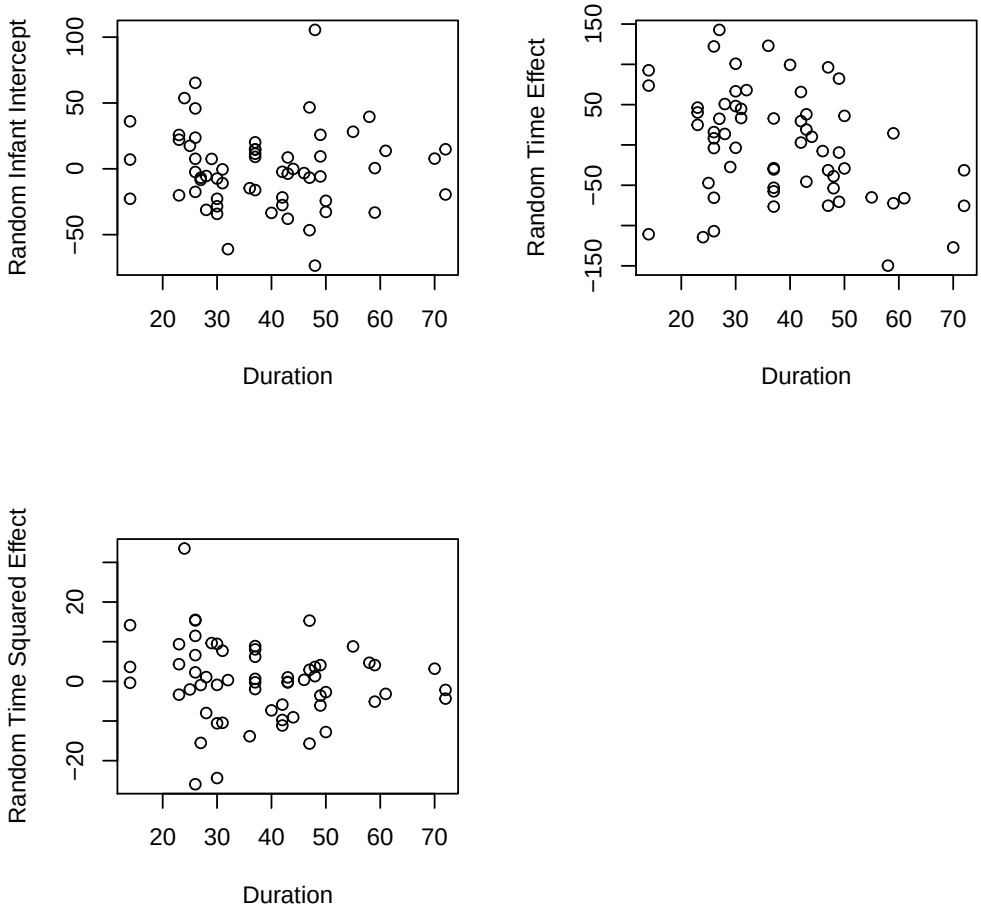

Figure 3.8: Scatterplot of random effects vs duration for model with baseline covariates

3.3.4 Mixed Model Analysis Results

For comparison with the other methods, the results of the final MM are summarised in Tables 3.4 and 3.5. These results need to be interpreted with caution as the NICS assumption appears to be violated. However, using this model we can consider whether treatment has an effect on weight. That is, whether the treatment significantly improves the growth rate of preterm infants. This is described by the two interaction terms, which have estimates of -5.0237 and 3.9817 respectively. Neither of these estimates are significant. An F-test was performed to determine if the terms are jointly different from zero, using the `anova.lme` function in R. The p-value from the Wald F-test was found to be 0.2721. Hence, there is not sufficient evidence to conclude that treatment has an effect

Fixed Effect	estimate	s.e	t-value	p-value	ci lower	ci upper
Intercept	-549.86	341.74	-1.61	0.11	-1219.67	119.94
Time	208.10	12.82	16.23	< 0.0001	182.95	233.21
Sex(Male)	13.20	22.66	0.58	0.57	-31.21	57.62
Plurality(2)	-8.63	29.49	-0.29	0.77	-66.42	49.16
Gestational Age	28.13	12.09	2.33	0.025	4.44	51.83
Birthweight	0.70	0.037	18.80	< 0.0001	0.63	0.78
Time Squared	6.60	2.09	3.15	0.0016	2.50	10.71
Group*Time	-5.02	18.00	-0.28	0.78	-40.32	30.27
Group*Time Squared	3.98	2.93	1.36	0.17	-1.76	9.73

Table 3.4: Summary of fixed effects for mixed model

Group	Name	Variance	Standard Deviation
Infant	Intercept	2105.6	45.89
	Time (Slope)	4950.4	70.36
	Time Squared	118.8	10.90
Mother	Intercept	6762.2	82.23
Residual		1343.8	36.66

Table 3.5: Summary of random effects for mixed model

on the growth of preterm infants.

It is relevant also to consider how well the MM fits the POPPET data. To compare the final model to the actual data, a plot of actual weight and estimated weight against time for each infant is given in Figure 3.9. In black is the infants actual trajectory and in red is their predicted trajectory from the MM. Figure 3.9 shows an extremely good fit with the red lines following the black lines very closely. This can be attributed to two factors. Firstly, the MM produces a reasonable fit for the data. Secondly, due to its structure, the MM allows for good fit by design. The random infant intercept in the model allow for

each infant to have their own intercept and the random linear and quadratic terms allow for each infants trajectories to differ (linearly and quadratically).

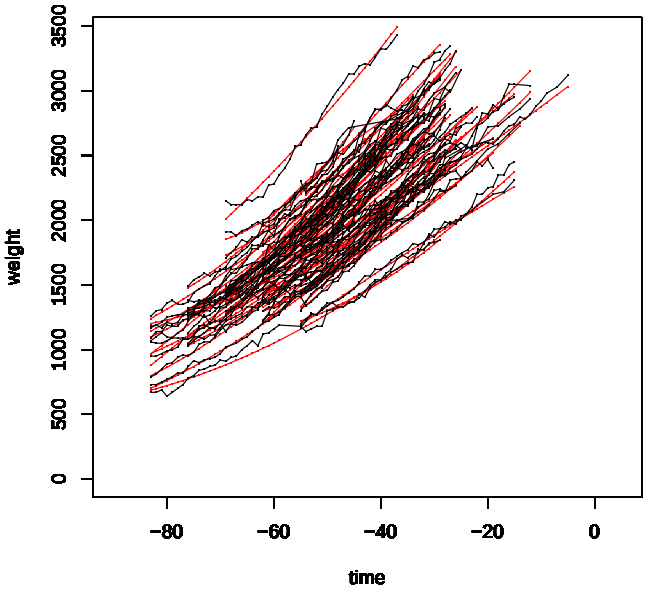

Figure 3.9: Actual (in black) and expected (in red) weight against time for each infant

The close agreement between the predicted and actual growth trajectories can be seen more clearly by only considering a selection of the infants from the trial, as the full plot is hard to distinguish with so many trajectories. Consider now the trajectories for a randomly selected 6 infants, given in Figure 3.10. In this plot, the close agreement between the estimated weights and the true weights is very clear.

In addition to considering individual growth trajectories, it is interesting to consider the data in terms of the observed mean weight over time. This can be done by comparing the mean on each study day from the raw data to the predicted mean on each day, calculated as the average fixed effects on each study day for Group 1 and Group 2 separately. This plot is given in Figure 3.11, where the black indicates the sample mean for Group 1, the blue indicates the sample mean for Group 2 and the green and red indicates the estimated means for Group 1 and 2 respectively. In this plot, time corresponds to the study day,

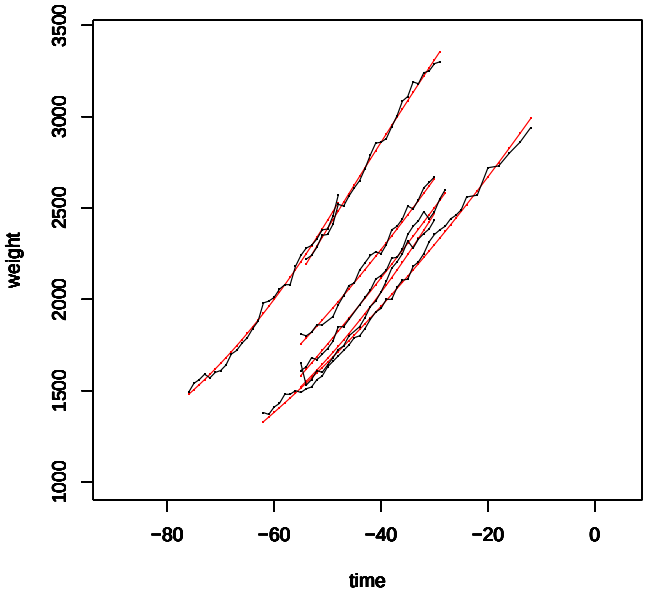

Figure 3.10: Scatterplot of weight against time for every 10th infant with a plot of fitted values against time for every 10th infant (in red)

which can range between 1 and 72. In all cases the observed mean weight drops below the estimated mean weight after approximately 30 weeks, with the gap increasing over time. This means the modelled growth continues fairly constantly, almost linear, whereas the data shows a drop off. This drop off is because the smaller sicker infants tend to stay in hospital longer and thus are more likely to contribute to the raw mean at later time points than larger healthier infants.

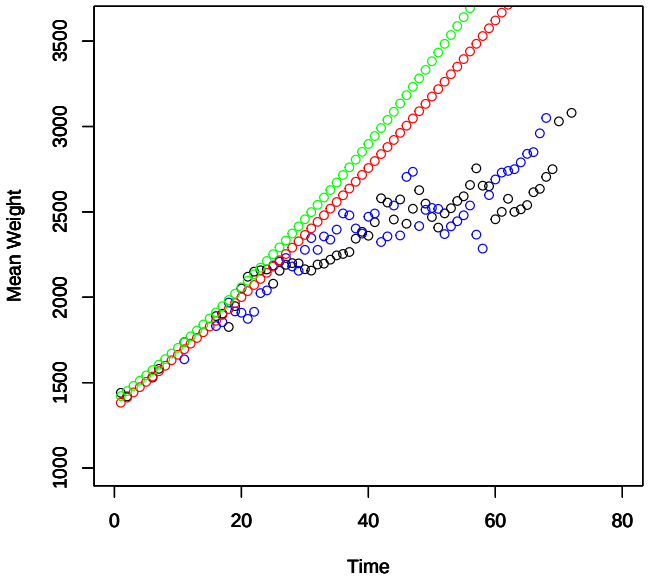

Figure 3.11: Scatterplot of mean weight for Group 1 (black), mean weight for Group 2 (blue) and mean fixed effects for each group (Group 1 is green and Group 2 is red) vs time

To see why this is the case, consider the plot of the fitted values for each infant extrapolated out over the max 72 day study period, given in Figure 3.12. These fitted values are in pink for when the infant was in the study and are green once the infant has left the study. The reason the MM means are so high is because they are based on these green lines. The MM means continue the projection of the path for infants even after they have left the study. On the other hand, the raw means only consider the infants while they are in the study, given by the pink lines. Toward the end of the 72 day period, the actual data gradually flattens off and most of the remaining pink lines are not very steep, resulting in a low sample mean weight. It is worth noting that this discrepancy is not a model defect.

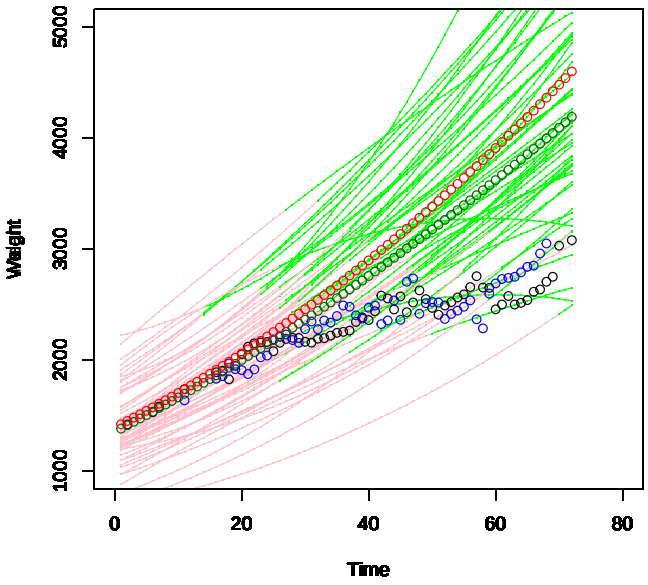

Figure 3.12: Scatterplot of mean weight for Group 1 (black), mean weight for Group 2 (blue) and mean fixed effects for each group (Group 1 is green and Group 2 is red) vs time with a plot of the fitted values for each infant for the entire duration of 72 days. It is pink while the infant was in the study and then light green once the infant was discharged.

3.3.5 Attempt to Remove ICS by Having a Fixed Trial Length

One way to remove the problem of ICS is to use a fixed trial length, which forces all cluster sizes to be equal. This approach removes the problem of ICS because if the cluster size is fixed, it cannot be related to the outcomes of interest. For the POPPET data the shortest duration across all infants is 14 days. ICS can therefore be removed by only considering the data from all infants up to day 14. This formulation also results in no missing values. However, it means discarding a large amount of data that could provide important information about the effect of treatment on growth.

The same MM from the previous section was applied again but only using the measurements for all infants from the first 14 days of the study. The results are summarised in Tables 3.6 and 3.7, and show a similar pattern to the previous results with baseline terms having similar estimates and standard errors. However, the terms which vary over time

do have different standard errors. For example, both time and time squared have larger standard errors in the fixed trial length model (18 and 7.97 for the fixed trial length model compared to 12.82 and 2.09 for the non fixed trial length model). This could be expected since the number of time points has been greatly reduced.

Fixed Effect	estimate	s.e	t-value	p-value	ci lower	ci upper
Intercept	-149.56	327.95	-0.46	0.65	-792.13	493.00
Time	134.82	18.00	7.50	< 0.0001	99.53	170.11
Sex(Male)	-6.25	23.87	-0.26	0.80	-53.04	40.55
Plurality(2)	8.46	26.11	0.32	0.75	-42.70	59.63
Gestational Age	13.53	11.32	1.19	0.24	-8.69	35.72
Birthweight	0.77	0.036	21.14	< 0.0001	0.70	0.84
Time Squared	32.10	7.97	4.027	0.0001	16.48	47.7
Group*Time	-5.05	3.65	-1.38	0.17	-85,41	14.75
Group*Time Squared	0.30	0.23	1.31	0.19	-7.36	37.02

Table 3.6: Summary of fixed effects for fixed trial length model

Group	Name	Variance	Standard Deviation
Infant	Intercept	3141.1	56.05
	Time (Slope)	8497.2	92.18
	Time Squared	1642.1	40.52
Mother	Intercept	4209.4	64.88
Residual		452.8	21.28

Table 3.7: Summary of random effects for fixed trial length model

In summary, creating a fixed trial length does remove the problem of ICS but it may also remove a lot of information. For this reason, it is generally not considered to be a suitable method of analysis.

3.4 GEE Analysis

For the POPPET data, a GEE could be used to estimate the effect of treatment on weight from trial entry until removal of the nasogastric tube. For infant i, the mean model is:

$$g(\boldsymbol{\mu_i}) = X_i^T \boldsymbol{\beta},$$

where X_i comprises a treatment $\times$ time interaction, treatment $\times$ time squared interaction, gestational age, sex, plurality, time, time squared and birthweight. These were the fixed effect terms included in the final MM and hence will be used as the starting point for the GEE model. The outcomes will be assumed to be generated from a Gaussian distribution, so the link function will be the identity.

Recall from Chapter 2, that the GEE takes into account the dependence in the data through the working correlation matrix. Many different working correlation matrices can be used and several will be considered here for comparison: an exchangeable (EGEE), an autoregressive of order 1 (ARGEE) and an independence (IEE) working correlation structure. A cluster weighted GEE will also be considered. This uses an independence working correlation structure but each cluster is now inversely weighted by cluster size (CWGEE), as described in Section 2.4.4. For the CWGEE the weighting is based on the total number of observations for each cluster (mother). For twins, the weight is the inverse of the total number of measurements taken on the two infants. For singletons, the weight is the inverse of the number of measurements taken on the infant.

The results of the four GEE methods are summarised in Tables 3.8 - 3.11. For the EGEE, the value of the correlation parameter in R_i was estimated to be $\rho = 0.6621$ and for the ARGEE, the value of the correlation parameter in R_i was estimated to be $\rho = 0.984$.

The results from the four GEE fits show that using a different working correlation structure can produce different results when ICS is present. There are some large differences between the results for the EGEE, ARGEE, IEE and CWGEE. The IEE and CWGEE have the most similar results and have very similar estimates for most of the terms. These methods also result in the same terms being statistically significant ($p < 0.05$) as the MM

	estimate	s.e	Wald	p-value	ci lower	ci upper
Intercept	-424	706	0.36	0.548	-1808	960.2
Time	249.86	12.63	391.66	< 0.0001	225	275
Sex(Male)	32.85	61.18	0.288	0.591	-87.06	152.8
Plurality(2)	35.15	51.51	0.466	0.495	-65.81	136.1
Gestational Age	2.245	27.38	0.007	0.935	-51.42	55.91
Birthweight	1.10	0.1235	79.30	< 0.0001	0.858	1.342
Time Squared	-2.56	1.76	2.125	0.145	-6.01	0.89
Group*Time	-8.72	14.69	0.352	0.553	-37.5	20.1
Group*Time Squared	2.23	2.08	1.155	0.282	-1.85	6.31

Table 3.8: Summary of the EGEE model

	estimate	s.e	Wald	p-value	ci lower	ci upper
Intercept	532.38	1304.29	0.17	0.683	-2024	3089
Time	275.63	32.69	71.10	< 0.0001	212	340
Sex(Male)	-106.13	83.06	1.63	0.201	-269	56.7
Plurality(2)	40.99	64.4	0.40	0.525	-85.3	167
Gestational Age	-32.9162	54.21	0.37	0.544	-139	73.3
Birthweight	1.20	0.27	19.90	< 0.0001	0.675	1.73
Time Squared	-5.89	4.15	2.01	0.156	-14	2.24
Group*Time	-31.34	33.84	0.86	0.354	-97.7	35
Group*Time Squared	8.46	4.38	3.73	0.054	-0.125	17

Table 3.9: Summary of the ARGEE model

did. That is, the intercept, time, time squared, gestational age and birthweight. The ARGEE and the EGEE only result in two significant variables, time and birthweight. Hence, the largest change between the models can be seen in the time squared and gestational age terms. The standard error for time squared increased for the ARGEE method (4.18 for the ARGEE compared to 1.89 and 2.603 for the IEE and CWGEE) and for the

	estimate	s.e	Wald	p-value	ci lower	ci upper
Intercept	-1293	535	5.85	0.0156	-2342	-244
Time	263	13.5	379.08	< 0.0001	237	289
Sex(Male)	29.78	48.3	0.38	0.54	-64.9	124
Plurality(2)	17.44	47.7	0.13	0.715	-76.1	111
Gestational Age	44.6	19.3	5.36	0.0206	6.77	82.4
Birthweight	0.82	0.0739	123.53	< 0.0001	0.677	0.966
Time Squared	-5.81	1.89	9.42	0.0021	-9.51	-2.11
Group*Time	4.39	21.0	0.040	0.835	-36.8	45.5
Group*Time Squared	0.768	3.20	0.060	0.810	-5.5	7.04

Table 3.10: Summary of the IEE model

	estimate	s.e	Wald	p-value	ci lower	ci upper
Intercept	-1375	474	8.42	0.0037	-2304	-446
Time	262.10	14.97	306.43	< 0.0001	233	291
Sex(Male)	20.00	41.00	0.24	0.6267	-60.4	100
Plurity(2)	18.00	42.9	0.18	0.6743	-66.1	102
Gestational Age	49.5	17.2	8.26	0.0041	15.8	83.2
Birthweight	0.780	0.0630	153.52	< 0.0001	0.657	0.903
Time Squared	-5.161	2.603	3.93	0.0474	-10.3	-0.0591
Group*Time	17.304	21.149	0.67	0.4132	-24.1	58.8
Group*Time Squared	-1.773	3.685	0.23	0.6305	-9	5.45

Table 3.11: Summary of the CWGEE model

EGEE the estimate decreased while the standard error remained similar to the IEE and CWGEE (estimate of -2.56 compared to -5.89, -5.81 and -5.161 for the ARGEE, IEE and CWGEE respectively). Similarly, for gestational age the standard error was substantially larger for the EGEE and ARGEE models (27.38 and 54.21 for the EGEE and ARGEE compared to 19.3 and 17.2 for the IEE and CWGEE respectively). Both, birthweight and

time are fairly constant across all methods and although there is a very large difference in the estimated interaction terms, all had large standard errors.

None of the methods resulted in a significant treatment effect, as all had non significant interaction terms. To test whether the interaction terms jointly differ from zero, a model with the interactions terms was compared to a model without them. The `anova` function in R was used to test whether the models were significantly different. That is, testing whether both interaction terms were jointly significantly different from zero. For all four Chi Squared Tests there were 2 degrees of freedom and the p-values were 0.51, 0.048, 0.65, 0.64 for the EGEE, ARGEE, IEE and CWGEE respectively. This implies that for the EGEE, IEE and CWGEE, the interaction terms are not significantly different from zero. For the ARGEE, the p-value is less than 0.05 but only marginally. Consistent with the MM analysis, the results of the GEE analysis for the EGEE, IEE and CWGEE methods also provide no evidence that the use of high protein milk fortifier improves growth for preterm infants compared with the regular protein milk fortifier. The ARGEE model did suggest that the interaction terms were significantly different from zero but only marginally, providing some evidence of a treatment effect. However, significance is probably exaggerated by considering many tests.

To investigate how well the GEE models fit the data, a plot of actual and estimated weight against time is given in Figures 3.13 - 3.16 for the EGEE, ARGEE, IEE and CWGEE respectively. Comparing the GEE methods, the ARGEE fits very similarly to the EGEE, perhaps a little better for some infants. The IEE and CWGEE show very similar fit and both appear to fit better than the EGEE and ARGEE. In comparison to the fit for the MM, the GEE fits appear worse. This is expected, because a GEE only includes fixed terms and models dependence using a working correlation structure, hence each infant does not have their own intercept and slope terms. Infant trajectories are only separated according to the birthweight, gestational age, sex, plurality and group (through the interaction terms). For a fair comparison, consider trajectories from the MM using only fixed effects, given in Figure 3.17. This plot shows fit which is better than the ARGEE and EGEE and is similar to the fit of the IEE and CWGEE methods.

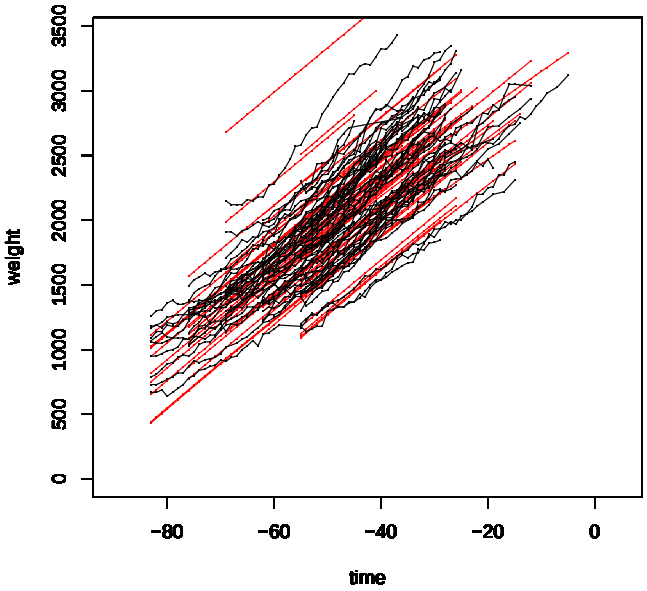

Figure 3.13: Actual (in black) and expected (in red) weight against time for each infant for the EGEE

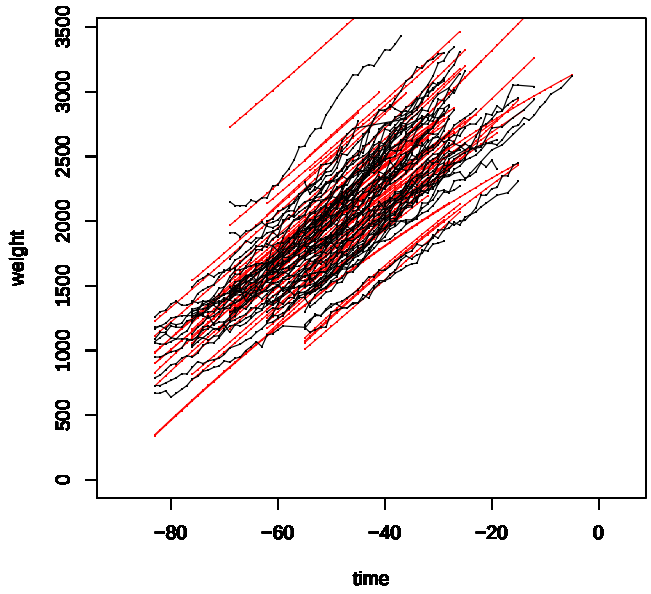

Figure 3.14: Actual (in black) and expected (in red) weight against time for each infant for the ARGEE

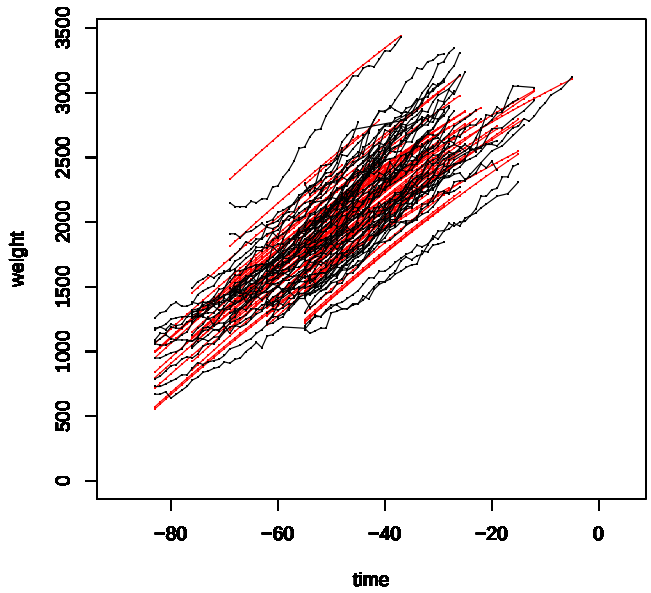

Figure 3.15: Actual (in black) and expected (in red) weight against time for each infant for the IEE

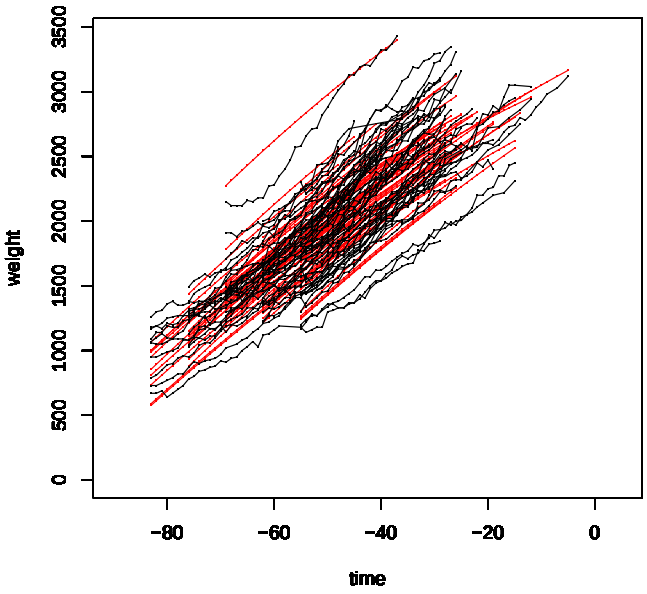

Figure 3.16: Actual (in black) and expected (in red) weight against time for each infant for the CWGEE

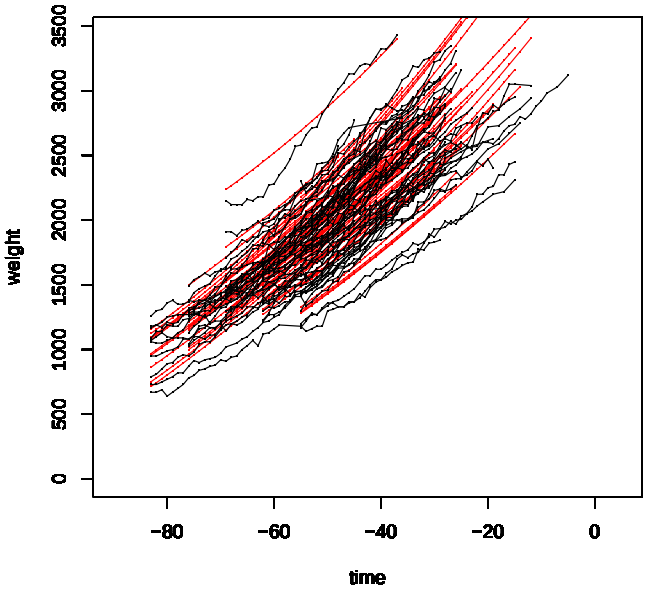

Figure 3.17: Actual (in black) and expected (in red) weight against time for each infant for the MM

Again, we will now consider the data in terms of the overall mean weight on each study day as opposed to considering each infant separately. This was investigated by plotting the sample mean based on the raw data and the estimated mean on each study day for Group 1 and Group 2 separately against time for each GEE method. These plots are given in Figure 3.18, where the black indicates the sample mean for Group 1, the light green indicates the sample mean for Group 2 and the blue and pink lines are for the estimated means for Group 1 and 2 respectively. Figure 3.18 shows that the EGEE, ARGEE and the IEE indicate some difference in the mean of each group by the end of the 72 day period, which is not significantly different except for the ARGEE. In contrast, the CWGEE does not show a visible change in the mean of the groups by the end of the study period. In all cases the observed mean weight drops below the estimated mean weight after approximately 30 weeks, with the gap increasing over time. This means the modelled growth continues fairly constantly, almost linear, whereas the data shows a drop off. This drop off is because the smaller sicker infants tend to stay in hospital longer and thus are more likely to contribute to the raw mean at later time points than larger healthier infants.

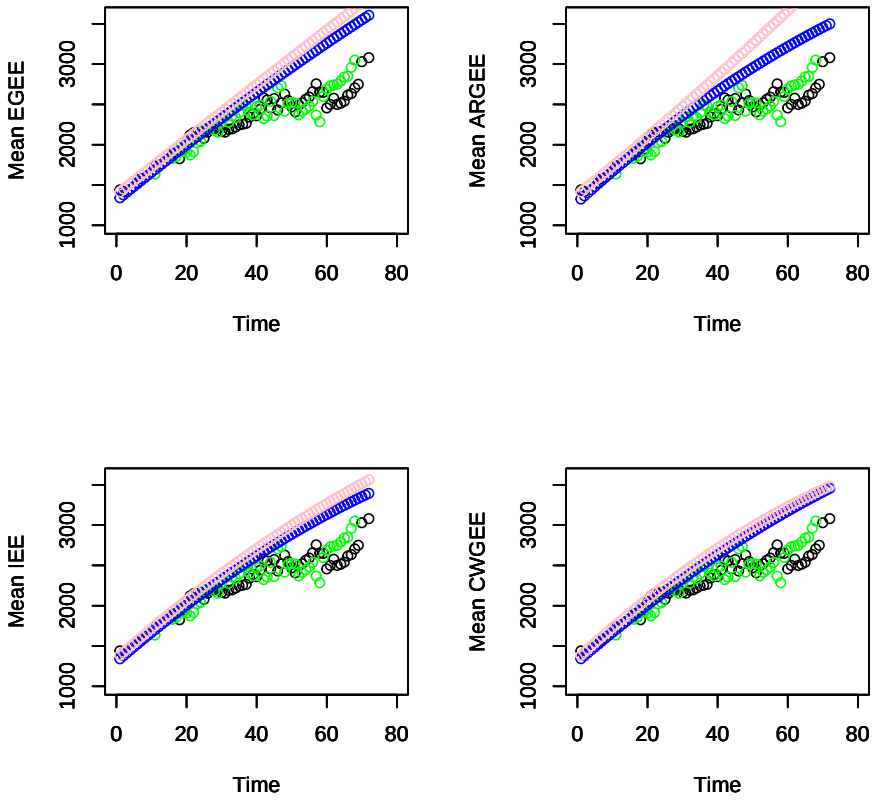

Figure 3.18: Scatterplot of mean weight for Group 1 (black), for Group 2 (light green) and mean fitted values from GEE for each group (Group 1 is blue and Group 2 is pink) vs time for each of the GEEs

3.5 Comparison of Methods

The aim of the POPPET trial data was to determine the effect of treatment on the growth of preterm infants. Five different methods of analysis were used to analyse the POPPET data and the results for the estimated treatment effect from each are summarised in Tables 3.12 and 3.13.

	estimate	s.e	Wald	p-value	ci lower	ci upper
MM	-5.02	18.00	-0.28	0.78	-40.32	30.27
EGEE	-8.72	14.69	0.352	0.553	-37.5	20.1
ARGEE	-31.34	33.84	0.86	0.354	-97.7	35
IEE	4.39	21.0	0.040	0.835	-36.8	45.5
CWGEE	17.304	21.149	0.67	0.4132	-24.1	58.8

Table 3.12: $Group \times Time$ estimates for each model

	estimate	s.e	Wald	p-value	ci lower	ci upper
MM	3.98	2.93	1.36	0.17	-1.76	9.73
EGEE	2.23	2.08	1.155	0.282	-1.85	6.31
ARGEE	8.46	4.38	3.73	0.054	-0.125	17
IEE	0.768	3.20	0.060	0.810	-5.5	7.04
CWGEE	-1.773	3.685	0.23	0.6305	-9	5.45

Table 3.13: $Group \times Time^2$ estimates for each model

Tables 3.12 and 3.13 show that the estimates for both terms vary between the methods. Considered separately, coefficients appear different, but in combination the coefficients describe a similar treatment effect over time, although with the exception of group $\times$ time squared for the ARGEE, none are significantly different from zero. For all but the ARGEE, it was also found that the terms were jointly not significantly different from zero, with the ARGEE only finding a slightly significant effect (p=0.048). In that respect the models are all reaching a very similar conclusion.

The combined effect of the two parameters describing the treatment effect can be seen by considering outcomes on a specific day. Table 3.14 gives the predicted difference in weight based on each group allocation for day 45, from the formula $\beta_8 * Time + \beta_9 * Time^2$. This example shows that the difference in the estimates depending on group allocation is similar in magnitude for each method, with Group 2 always resulting in a higher weight.

The ARGEE results in the largest difference in this example of 148g with the MM having a similar difference of 132g and both having similar standard errors. The EGEE, IEE and CWGEE resulted in similar differences when the large standard errors are considered. These differences were slightly smaller than the differences for the MM and ARGEE methods. By considering the treatment effect from the combined interaction terms, it is much clearer that all methods are roughly describing the same effect.

	Difference (Group 2 - Group 1)	Standard Error
MM	132	98
EGEE	36	57.3
ARGEE	148	108
IEE	60	64.7
CWGEE	38	74.8

Table 3.14: Trajectories for an infant given each group allocation

In summary, the results of the analysis for the POPPET data did not indicate that the treatment was effective at improving the growth of preterm infants. Four out of five methods indicated the treatment effect wasn't significantly different from zero, and the last only found a marginally significant difference. Again, the significance could be exaggerated by considering many methods. However, the appropriateness of some of these methods is questionable when ICS is present. Their performance can be tested through simulation before any final conclusions can be made for the POPPET trial, and this will be addressed in Chapter 4 and 5.

Chapter 4

Simulation Study

In Chapter 3, the POPPET data set was analysed using a variety of methods to estimate the effect of treatment on weight. As ICS was found to be present for these data, the appropriateness of some of the methods is questionable. In Chapter 4, a simulation study is conducted to investigate the performance of these methods when ICS is present.

The aim of the simulation study is to help understand which methods are most appropriate for analysing longitudinal data when the length of follow up is informative. The simulated data will be analysed using each of the methods described in Chapter 2. This will allow the methods to be compared when the true parameter values are known.

4.1 Generating Simulated Data

Data were simulated based on the POPPET data set with two simplifications. Firstly, we consider only one level of clustering, that is longitudinal data, and assume all mothers give birth to single infants. Secondly, only a linear relationship between time and weight was considered and the quadratic effect seen in the POPPET data was not included.

Each simulation initially included the same number of individuals as the POPPET trial. That is, weights for 60 infants were generated for each simulated dataset. The data were generated from the following MM:.

$$
\begin{aligned}
Weight =& \beta_0 + \beta_1 Treatment : Time + \beta_2 GestationalAge + \beta_3 Birthweight \\
&+ \beta_4 Time + b_0 + b_1 Time \\
&+ Error
\end{aligned}
\tag{4.1}
$$

where b_0 is the random infant intercept and b_1 is the random infant slope. The $Error$ is distributed $N(\mathbf{0}, R)$ where R is a $N \times N$ positive definite covariance matrix and the $\boldsymbol{b} = [b_0, b_1]$ are distributed $N(\mathbf{0}, D)$ where D is an $w \times w$ positive definite covariance matrix.

To generate weight measurements for each individual using model 4.1, values for the coefficients and for each predictor variable needed to be determined. The approach used to determine each of these quantities will now be discussed.

4.1.1 Gestational Age in Weeks

Gestational age in weeks was the first variable to be generated, as several of the other variables depend on gestational age. For the POPPET trial, infants' gestational ages ranged between 28 and 32 weeks. The distribution of gestational age for POPPET is shown in Figure 4.1 and appears roughly uniform. Gestational age was therefore generated from a discrete uniform distribution unif(28,32). In R this can be achieved using the sample function with replacement.

4.1.2 Birthweight

The distribution of the birthweight for POPPET is given in Figure 4.2 and shows a roughly normal distribution for the birth weights with a mean of 1516 grams and a standard deviation of 413.7 grams. However, mean birthweight increases with gestational age, as shown in Figure 4.3.

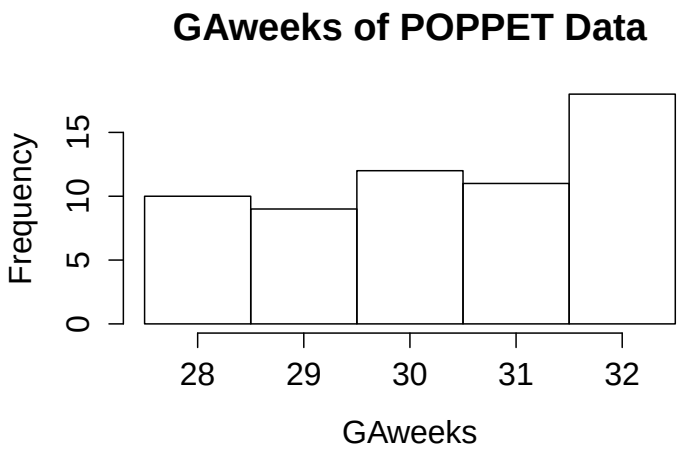

Figure 4.1: Histogram of gestational age in weeks

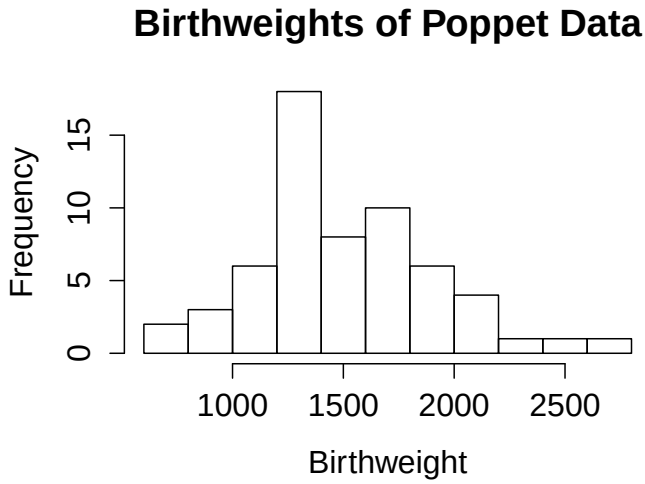

Figure 4.2: Histogram of birthweight in grams

Fitting a linear regression of gestational age on birthweight showed that gestational age has a significant effect on birthweight (Birthweight$= -3889.9 + 178.4\times$Gestational Age). For the simulation study, it was concluded that birth weights should be generated separately within each gestational age based on a different normal distribution. To obtain the mean and standard deviation for the normal distributions, the POPPET data were not used, as within each gestational age the sample size was so small that reasonable estimates of these quantities could not be obtained. Instead, the means and standard deviations were obtained from a published study (Roberts and Lancaster, 1999), which

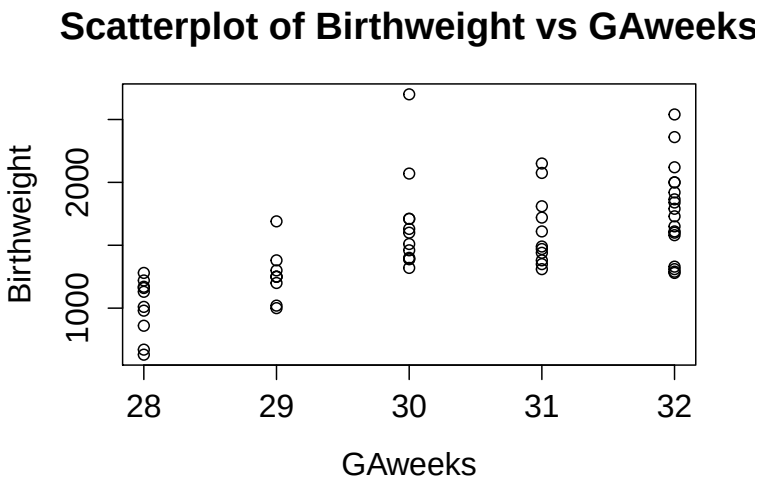

Figure 4.3: Scatterplot of birthweight vs gestational age

reported means and standard deviations for infants' birthweight by gestational age, using much larger sample sizes than were available in POPPET. As a check to see whether the variances used from the published study reflected the POPPET data, the variance estimate from the POPPET data (grouped by gestational age) can be compared to the published variances. The variance estimates from POPPET are 51432, 41250, 150057, 83250 and 128724; and the 5 chosen variances are 57600, 68121, 97969, 96721 and 142884 for gestational ages 28-32 weeks, respectively.

4.1.3 Treatment Group

In the POPPET trial, the gestational age was divided into two categories (as described in Section 1.1); less than 30 weeks and greater than or equal to 30 weeks. Infants were then randomly allocated to one of the two treatment groups, with equal numbers of infants assigned to each treatment within each gestational age category. That is, the randomisation was stratified by gestational age. Hence, in the simulated data within these two gestational age groups, equal numbers of infants were allocated to each treatment. If there was an even number, 2N, of infants in a gestational age category then N were allocated to each group. This was done using the sample function in R (without replacement) sampling from N ones and N zeros, where zero represents allocation to Group 1 and one represents

allocation to Group 2. If there was an odd number of infants in a gestational age category, then a one or zero was chosen randomly to determine which treatment group would have the extra infant, using the sample function. The remaining infants were then allocated to treatment groups as described previously.

4.1.4 Coefficients, Error and Random Effects

The coefficients, $\beta_0 - \beta_4$, and the standard deviation of the random effects and error terms, were determined by fitting model 4.1 to the POPPET data, with output shown in Tables 4.1 and 4.2. The error and random effects were then generated from a normal distribution (using `rnorm` in R) with a mean of zero and the obtained standard deviations.

Fixed Effect	estimate	s.e	t-value	p-value	ci lower	ci upper
Intercept	-1042	309.9	-3.36	0.0008	-1649.47	-434.59
Time	244	7.473	32.72	< 0.0001	229.84	259.13
Gestational Age	41.96	11.38	3.69	0.0002	19.66	64.27
Birthweight	0.721	0.0404	17.86	< 0.0001	0.64	0.80
Group*Time	13.99	10.72	1.31	0.19	-7.02	35.01

Table 4.1: Summary of fixed effects for simulation study mixed model

Group	Name	Variance	Standard Deviation
Infant	Intercept	9531	97.63
	Time (Slope)	1664	40.79
Mother	Intercept	2720	52.15
Residual		1343.8	36.66

Table 4.2: Summary of random effects for simulation study mixed model

4.1.5 Infant Weights Over Time

To generate the weights for each infant in the study over time, initially the duration was set to be a fixed value of 180 days, for every infant. The predictor variables and coefficients were then applied in model 4.1 to obtain the weights for each infant on each day. Once the cluster size was determined (as discussed in Section 4.2), the remaining measurements (out of the 180 calculated) were removed to produce a trial with unequal follow up lengths, between infants, resulting in a maximum possible cluster size of 180.

4.2 Method for Determining Cluster Size

The main aim of the simulation study is to investigate the performance of analysis methods when informative cluster size is present. A method of determining cluster size in such a way that it induces informative cluster size is therefore needed.

In a previous simulation study investigating the problem of informative cluster size, Neuhaus and Mcculloch generated cluster size from the random effects poisson distribution Poisson$(e^{\gamma_0+\gamma_1 b_0+\gamma_2 b_1}) + N_{min}$, where b_0 is the random intercept, b_1 is the random slope and N_{min} is the minimum cluster size (Neuhaus and Mcculloch, 2011).

Recall that for the MM, cluster size is non informative when the random effects are independent of the cluster size. The above method thus induces informative cluster size by making the cluster size depend on the random effects.

As the simulated data are based on the POPPET trial, the distribution of the cluster size in the simulation study should be as similar as possible to POPPET. Cluster size (duration), has a mean of 38.7 days, a variance of 191.59 days and a distribution as shown in Figure 4.4. If the durations in POPPET followed a Poisson distribution then the variance would be equal to the mean, which is clearly not the case for the POPPET data. In that, the POPPET data has mean of 38 but a variance of 192. This means that the POPPET

data are over dispersed and not Poisson distributed. However, Neuhaus's method of generating cluster size from a Poisson distribution will still be used for simplicity. The value of N_{min} will be set to 14 to match the minimum duration present in the POPPET data set.

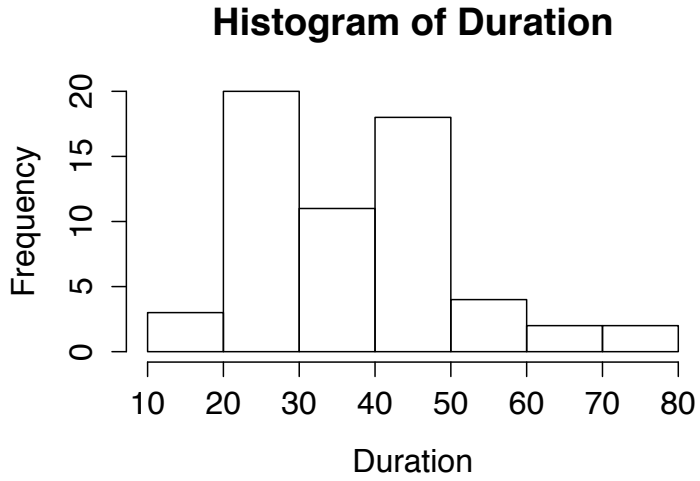

Figure 4.4: Histogram of duration

The mean and standard deviation of the cluster size distribution was chosen to match the POPPET trial. This was achieved by equating the expectation and variance of Neuhaus's distribution to the POPPET trial mean and variance and solving for the required values of $\gamma_0, \gamma_1, \gamma_2$. The calculations are given in Section 4.2.1.

4.2.1 Expected Value and Variance of Cluster Size Distribution

Conditionally on the random effects, the cluster size N follows a Poisson distribution. That is,

$$N|b_0, b_1 \sim Po(e^{\gamma_0+\gamma_1 b_0+\gamma_2 b_1}) + N_{\min},$$

which has mean and variance

$$E[N|b_0, b_1] = e^{\gamma_0+\gamma_1 b_0+\gamma_2 b_1} + N_{\min}$$

$$\text{and } \text{Var}[N|b_0, b_1] = e^{\gamma_0+\gamma_1 b_0+\gamma_2 b_1}.$$

The joint distribution of the random effect is bivariate normal. This has the form $N(\boldsymbol{\mu}, \Sigma)$ where $\boldsymbol{\mu}$ is the mean vector and Σ is the covariance matrix. For the random effects,

$$\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{0,0} & \sigma_{0,1} \\ \sigma_{1,0} & \sigma_{1,1} \end{pmatrix} \right),$$

where $\sigma_{i,j}$ is the covariance between b_i and b_j.

$$\text{Let} \quad Z = \gamma_0 + \gamma_1 b_0 + \gamma_2 b_1.$$

The expected value and variance of Z can be determined as follows:

$$\begin{aligned} E[Z] &= E[\gamma_0 + \gamma_1 b_0 + \gamma_2 b_1] \\ &= \gamma_0 + \gamma_1 E[b_0] + \gamma_2 E[b_1] \\ &= \gamma_0 \end{aligned}$$

$$\begin{aligned} \operatorname{Var}(Z) &= \operatorname{Var}(\gamma_0 + \gamma_1 b_0 + \gamma_2 b_1) \\ &= \gamma_1^2 \operatorname{Var}(b_0) + \gamma_2^2 \operatorname{Var}(b_1) + 2\gamma_1\gamma_2 \operatorname{Cov}(b_0, b_1) \\ &= \gamma_1^2 \sigma_{0,0}^2 + \gamma_2^2 \sigma_{1,1}^2 + 2\gamma_1\gamma_2\sigma_{0,1} \end{aligned}$$

Hence, Z follows a normal distribution with mean γ_0 and variance $\gamma_1^2\sigma_{0,0}^2 + \gamma_2^2\sigma_{1,1}^2 + 2\gamma_1\gamma_2\sigma_{0,1}$, that is,

$$Z \sim N(\gamma_0, \gamma_1^2\sigma_{0,0}^2 + \gamma_2^2\sigma_{1,1}^2 + 2\gamma_1\gamma_2\sigma_{0,1}).$$

$$\text{Let} \quad W = e^Z$$

The expected value and variance of W can be determined as follows:

$$\begin{aligned} E[W] = E[e^Z] = E[e^{Zt}]|_{t=1} &= M_Z(1) \\ &= e^{E[Z]+\frac{1}{2}Var(Z)} \\ &= e^{\gamma_0+\frac{1}{2}(\gamma_1^2\sigma_{0,0}^2+\gamma_2^2\sigma_{1,1}^2+2\gamma_1\gamma_2\sigma_{0,1})}, \end{aligned}$$

where $M_Z(t)$ is the moment generating function of Z and

$$\begin{aligned} \mathrm{Var}[W] = Var[e^Z] &= E[e^{2Z}] - E[e^Z]^2 \\ &= E[e^{Zt}]|_{t=2} - E[e^Z]^2 \\ &= M_Z(2) - E[e^Z]^2 \\ &= e^{2\gamma_0+2\gamma_1^2\sigma_{0,0}^2+2\gamma_2^2\sigma_{1,1}^2+4\gamma_1\gamma_2\sigma_{0,1}} \\ &- e^{2\gamma_0+\gamma_1^2\sigma_{0,0}^2+\gamma_2^2\sigma_{1,1}^2+2\gamma_1\gamma_2\sigma_{0,1}}. \end{aligned}$$

Finally the expectation and the variance of the cluster size can be calculated as follows:

$$\begin{aligned} E[N] &= E\big[E[N|b_0,b_1]\big] \\ &= E[e^{\gamma_0+\gamma_1 b_0+\gamma_2 b_1} + N_{\min}] \\ &= E[W] + N_{\min} \\ &= e^{\gamma_0+\frac{1}{2}(\gamma_1^2\sigma_{0,0}^2+\gamma_2^2\sigma_{1,1}^2+2\gamma_1\gamma_2\sigma_{0,1})} + N_{\min} \end{aligned} \tag{4.2}$$

and

$$\begin{aligned} Var(N) &= E\big(Var[N|b_0,b_1]\big) + Var\big(E[N|b_0,b_1]\big) \\ &= E(e^{\gamma_0+\gamma_1 b_0+\gamma_2 b_1}) + Var(e^{\gamma_0+\gamma_1 b_0+\gamma_2 b_1} + N_{\min}) \\ &= E(W) + Var(W) \\ &= e^{\gamma_0+\frac{1}{2}(\gamma_1^2\sigma_{0,0}^2+\gamma_2^2\sigma_{1,1}^2+2\gamma_1\gamma_2\sigma_{0,1})} + e^{2\gamma_0+2\gamma_1^2\sigma_{0,0}^2+2\gamma_2^2\sigma_{1,1}^2+4\gamma_1\gamma_2\sigma_{0,1}} \\ &- e^{2\gamma_0+\gamma_1^2\sigma_{0,0}^2+\gamma_2^2\sigma_{1,1}^2+2\gamma_1\gamma_2\sigma_{0,1}}. \end{aligned} \tag{4.3}$$

4.2.2 Determine Parameters to Match POPPET Trial Cluster Size Distribution

It is desired that the simulated cluster size distribution be similar to that for the POPPET trial. The parameters γ_0, γ_1 and γ_2 were therefore determined to match the expected value and variance of POPPET duration, that is, $\bar{N} = 38.36667$, $N_{min} = 14$ and $S_N^2 = 191.5921$. There is no correlation between the random effects (i.e $\sigma_{0,1} = 0$) and the random effects will be standardised for simplicity of calculation.

The sample mean of POPPET was equated to Equation 4.2:

$$\begin{aligned}
\bar{N} &= e^{\gamma_0 + \frac{1}{2}(\gamma_1^2\sigma_{0,0}^2 + \gamma_2^2\sigma_{1,1}^2 + 2\gamma_1\gamma_2\sigma_{0,1})} + N_{\min} \\
38.3667 &= e^{\gamma_0 + \frac{1}{2}(\gamma_1^2 + \gamma_2^2)} + 14 \\
\ln(24.3667) &= \gamma_0 + \frac{1}{2}(\gamma_1^2 + \gamma_2^2) \\
2(\ln(24.3667) - \gamma_0) &= \gamma_1^2 + \gamma_2^2
\end{aligned} \tag{4.4}$$

and the sample variance of POPPET was equated to Equation 4.3:

$$\begin{aligned}
S_N^2 &= e^{\gamma_0 + \frac{1}{2}(\gamma_1^2\sigma_{0,0}^2 + \gamma_2^2\sigma_{1,1}^2 + 2\gamma_1\gamma_2\sigma_{0,1})} + e^{2\gamma_0 + 2\gamma_1^2\sigma_{0,0}^2 + 2\gamma_2^2\sigma_{1,1}^2 + 4\gamma_1\gamma_2\sigma_{0,1}} \\
&\quad - e^{2\gamma_0 + \gamma_1^2\sigma_{0,0}^2 + \gamma_2^2\sigma_{1,1}^2 + 2\gamma_1\gamma_2\sigma_{0,1}} \\
191.5922 &= e^{\gamma_0 + \frac{1}{2}(\gamma_1^2 + \gamma_2^2)} + e^{2\gamma_0 + 2\gamma_1^2 + 2\gamma_2^2} - e^{2\gamma_0 + \gamma_1^2 + \gamma_2^2}
\end{aligned}$$

Substituting in Equation 4.4 gives

$$
\begin{aligned}
191.5922 &= e^{\gamma_0 + ln(24.3667) - \gamma_0} + e^{2\gamma_0 + 4ln(24.3667) - 4\gamma_0} - e^{2\gamma_0 + 2ln(24.3667) - 2\gamma_0} \\
191.5922 &= 24.3667 + e^{-2\gamma_0 + 4ln(24.3667)} - 24.3667^2 \\
ln(760.9616) &= -2\gamma_0 + 4ln(24.3667) \\
\gamma_0 &= \frac{4ln(24.3667) - ln(760.96)}{2} \\
&= 3.0691.
\end{aligned}
$$

Substituting this value for γ_0 into Equation 4.4 gives

$$
\begin{aligned}
2(ln(24.3667) - 3.0691) &= \gamma_1^2 + \gamma_2^2 \\
0.2482 &= \gamma_1^2 + \gamma_2^2.
\end{aligned}
$$

Therefore, to obtain the same expectation and variance as in the POPPET trial, γ_1 and γ_2 can be taken to be any value on the circle with equation $0.248 = \gamma_1^2 + \gamma_2^2$.

To confirm that this method has created informative cluster size in the simulated data, some test simulations were run and a MM was fit to each data set. Informative cluster size was then tested for by finding the correlation between the cluster size and the random effects (here a random intercept and a random slope). It was determined that, depending on the choice of γ_1 and γ_2, a range of different correlation strengths could be achieved, and hence the method could successfully be used to induce informative cluster size.

Both γ_1 and γ_2 needed to be negative to obtain a negative correlation between the random intercept and duration and the random slope and duration. This is desired because these correlations were negative for the POPPET data. Hence, only the quarter of the circle where both γ_1 and γ_2 are negative was considered.

Five combinations of values for γ_1 and γ_2 equally spaced around the negative quarter of the circle were chosen for the simulation study, as these covered a range of possible

correlations on the random effects. The combinations are given in Table 4.3. In Case 1, only the random intercept is related to duration. Similarly, in Case 2, only the random slope is related to duration. For Case 3, duration is equally correlated with the random intercept and the random slope, while for Cases 4 and 5, there was stronger correlation for the random intercept and the random slope, respectively.

Case	γ_1	γ_2	$cor(b_0, N)$	$cor(b_1, N)$
1	-0.50	0	-0.86	0
2	0	-0.50	0	-0.87
3	-0.35	-0.35	-0.62	-0.62
4	-0.46	-0.19	-0.80	-0.34
5	-0.19	-0.46	-0.34	-0.81

Table 4.3: Values of γ_1 and γ_2 and resulting correlations

4.3 Simulation Scenarios

When model 4.1 was fit to the POPPET data, the estimated treatment effect was 14 grams per week, which was found not to be significant (see Table 4.1). In this simulation study, varying levels of treatment effect will be explored.

By increasing the treatment effect in steps by 7 grams a week (starting from 0) and using the treatment effect together with model 4.1 to obtain weight, it was determined that the maximum treatment effect that would be reasonable was 28 grams a week. When the treatment effect is 28 grams a week, then the predicted weights at 40 weeks are roughly 4000 grams. Anything over this would be large and unrealistic. Hence, for the simulation study five different treatment effects were considered, that is: 0, 7, 14, 21 and 28 grams a week. For each of these treatment effects, the data were simulated using the five sets of values of γ_1 and γ_2 given in Table 4.3, resulting in 25 different simulation scenarios to consider.

In addition to these 25 simulation scenarios where informative cluster size is present, scenarios with non informative cluster size were also of interest. This will give us the ability to compare how methods perform when informative cluster size ir or is not present. NICS can be achieved by either having a fixed cluster size or allowing it to vary, in a way that is unrelated to the outcome. The first option was achieved by setting the cluster size to be fixed at the mean of the POPPET trial cluster size, which is 38 days. To allow the cluster size to vary but not depend on the outcome, the method described in Section 4.2 can be used where γ_1 and γ_2 are set to zero. Then the cluster size only depends on some constant and so the random effects are independent of the cluster size, resulting in NICS. To maintain the same expected cluster size as POPPET, the calculations in Section 4.2.2 need to be repeated with the requirement $\gamma_1 = \gamma_2 = 0$. This results in $\gamma_0 = \ln(24.36667) = 3.19$.

These 2 settings with non informative cluster size will be considered for each of the 5 treatment effects, resulting in 10 additional simulation scenarios. This brings the total number of simulation scenarios to 35, which are listed in Table 4.4. For each of these 35 simulation scenarios 10000 datasets were generated. Each simulated dataset was analysed using a mixed model (MM), an exchangeable generalised estimating equation (EGEE), an autoregressive generalised estimating equation (ARGEE), an independent generalised estimating equation (IEE) and a cluster weighted generalised estimating equation (CWGEE). These methods were described in Chapter 2.

The five different analysis approaches were compared based on 3 standard properties. First, the bias was defined as the difference between the estimate and the true value, which should be close to zero if the method is performing well. Second, the standard deviation of the estimates was compared to the model based standard error. The difference between these two values should be close to zero if the model is performing well. Third, the Coverage Probability is the proportion of times the 95% confidence interval contains the true value. The coverage probability should be close to the expected value of 0.95 if the method is performing well. In fact, for 10000 simulations the coverage probability should be between 0.9457 and 0.9542.

Scenario	Cluster Size(days)	γ_0	γ_1	γ_2	Treatment Effect(g/week)
1	38	NA	NA	NA	28
2	Varying	3.19	0	0	28
3	Varying	3.07	-0.50	0	28
4	Varying	3.07	0	-0.50	28
5	Varying	3.07	-0.35	-0.35	28
6	Varying	3.07	-0.46	-0.19	28
7	Varying	3.07	-0.19	-0.46	28
8	38	NA	NA	NA	21
92	Varying	3.19	0	0	21
10	Varying	3.07	-0.50	0	21
11	Varying	3.07	0	-0.50	21
12	Varying	3.07	-0.35	-0.35	21
13	Varying	3.07	-0.46	-0.19	21
14	Varying	3.07	-0.19	-0.46	21
15	38	NA	NA	NA	14
16	Varying	3.19	0	0	14
17	Varying	3.07	-0.50	0	14
18	Varying	3.07	0	-0.50	14
19	Varying	3.07	-0.35	-0.35	14
20	Varying	3.07	-0.46	-0.19	14
21	Varying	3.07	-0.19	-0.46	14
22	38	NA	NA	NA	7
23	Varying	3.19	0	0	7
24	Varying	3.07	-0.50	0	7
25	Varying	3.07	0	-0.50	7
26	Varying	3.07	-0.35	-0.35	7
27	Varying	3.07	-0.46	-0.19	7
28	Varying	3.07	-0.19	-0.46	7
29	38	NA	NA	NA	0

30	Varying	3.19	0	0	0
31	Varying	3.07	-0.50	0	0
32	Varying	3.07	0	-0.50	0
33	Varying	3.07	-0.35	-0.35	0
34	Varying	3.07	-0.46	-0.19	0
35	Varying	3.07	-0.19	-0.46	0

Table 4.4: Simulation Scenarios

4.4 Results

Each column in the results tables give the average of the 10000 estimates of a model parameter followed by the average of the 10000 model based standard errors (SE), and the standard deviation of the 10000 estimates (SD). In the results tables β_1 is the estimated coefficient for the treatment $\times$ time interaction effect, β_2 is the estimated coefficient for gestational age, β_3 is the estimated coefficient for birthweight, β_4 is the estimated coefficient for time and β_0 is the intercept. For a treatment effect of 28 grams a week, the true parameter values are: $\beta_1 = 28$, β_2=42, β_3=0.721, β_4=244, β_0=-1042. The value of β_1 will differ for different treatment effects. These values were used to generate the simulated data and were determined in Section 4.1.4. Primary interest lies in the interaction term (β_1), as this summarises the treatment effect. For this reason, the methods are compared in terms of their performance in estimating this parameter.

4.4.1 NICS Scenarios

The results for the NICS simulation scenarios with a treatment effect of 28 grams a week are given in Tables 4.5 and 4.6.

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.93	10.47	10.73	42.04	11.95	12.23	0.72	0.04	0.04	244.50	7.44	7.47	-1043.64	320.97	328.90
EGEE	27.91	10.38	10.65	41.93	17.31	18.32	0.72	0.06	0.07	244.52	7.32	7.43	-1040.82	466.64	491.60
IEE	27.87	12.19	12.74	41.94	17.34	18.42	0.72	0.06	0.07	244.54	8.00	8.23	-1040.99	466.84	493.70
CWGEE	27.87	12.19	12.74	41.94	17.34	18.42	0.72	0.06	0.07	244.54	8.00	8.23	-1040.99	466.84	493.70
ARGEE	27.93	10.81	11.26	41.96	17.56	18.66	0.72	0.06	0.07	244.51	7.63	7.82	-1041.84	473.26	500.66

Table 4.5: Scenario 1 (fixed trial length of 38 days, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.91	10.48	10.82	41.80	11.93	12.42	0.72	0.04	0.04	244.51	7.45	7.63	-1036.97	320.62	333.65
EGEE	27.94	10.96	11.44	41.58	17.43	18.69	0.72	0.06	0.07	244.50	7.71	8.10	-1030.93	469.45	500.80
IEE	27.88	12.59	13.31	41.56	17.79	19.22	0.72	0.06	0.07	244.53	8.94	9.56	-1030.42	478.66	514.59
CWGEE	27.88	12.34	12.95	41.57	17.45	18.78	0.72	0.06	0.07	244.53	8.67	9.17	-1030.59	469.60	502.78
ARGEE	27.93	10.93	11.43	41.52	17.77	19.10	0.72	0.06	0.07	244.49	7.71	8.08	-1029.09	478.31	511.79

Table 4.6: Scenario 2 ($\gamma_0 = 3.19, \gamma_1 = 0, \gamma_2 = 0$, treatment effect = 28 g/week)

Bias

Consider Tables 4.5 and 4.6, which give the results of the two simulation scenarios with NICS. Both show minimal bias in the estimated treatment effect. This can also been seen in Figure 4.5, which shows boxplots by method of the 10000 estimates from the simulation scenario in Table 4.6. The red line on the plot gives the true treatment effect. All methods have similar spread, with the IEE and CWGEE having slightly larger spread than the other methods. In addition, for all methods the median estimate is very close to the red line, implying there is no bias present. This is what we expect to see, as when cluster size is non informative all methods should provide valid inference. A similar pattern was seen in the boxplot of the treatment effect estimates for the other NICS scenario (see Appendix B.1 Figure B.2).

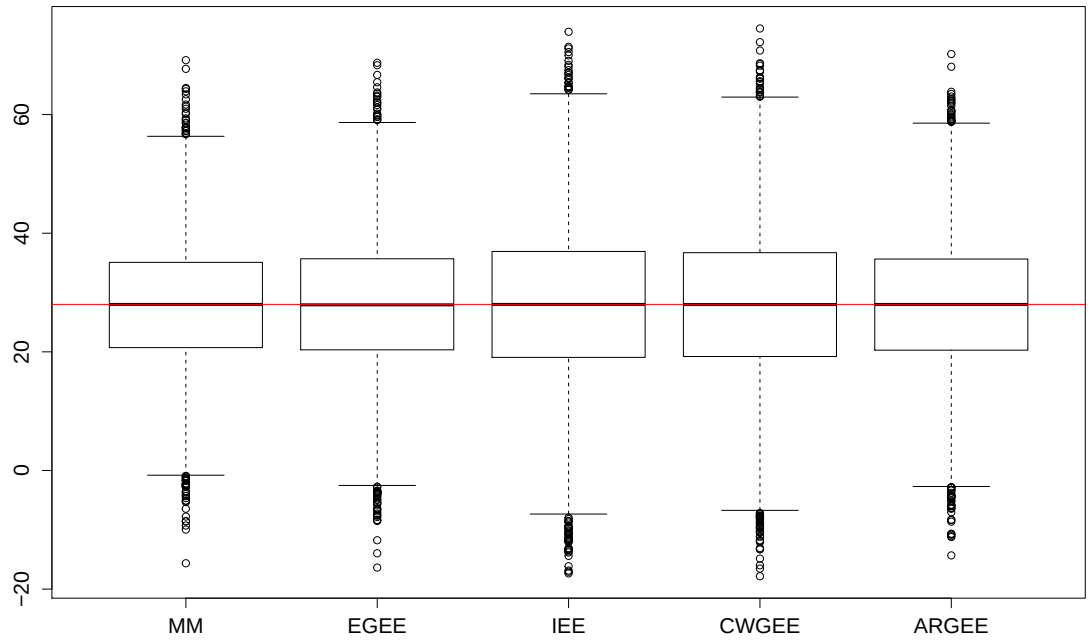

Figure 4.5: Boxplots of Interaction Effect estimates for scenario from Table 4.6

Standard Errors of Interaction Term

For the interaction effect, the standard error of the estimates and the model based standard error are roughly equal for each method. This can be seen in Figure 4.6, which shows boxplots of the ratio of SE for each of the 10000 simulated datasets divided by SD, for the simulation scenario in Table 4.6. It shows that there is only a slight difference between SE and SD, with the median ratio always falling slightly below 1. A similar pattern was seen in the boxplots of the SE/SD for the other NICS scenario (see Appendix B.1 Figure B.9). As seen in Table 4.5, the IEE and CWGEE have the largest SE, but for all the methods, the difference between SE and SD is small. This is what we expect to see, as all methods should perform well when ICS is absent.

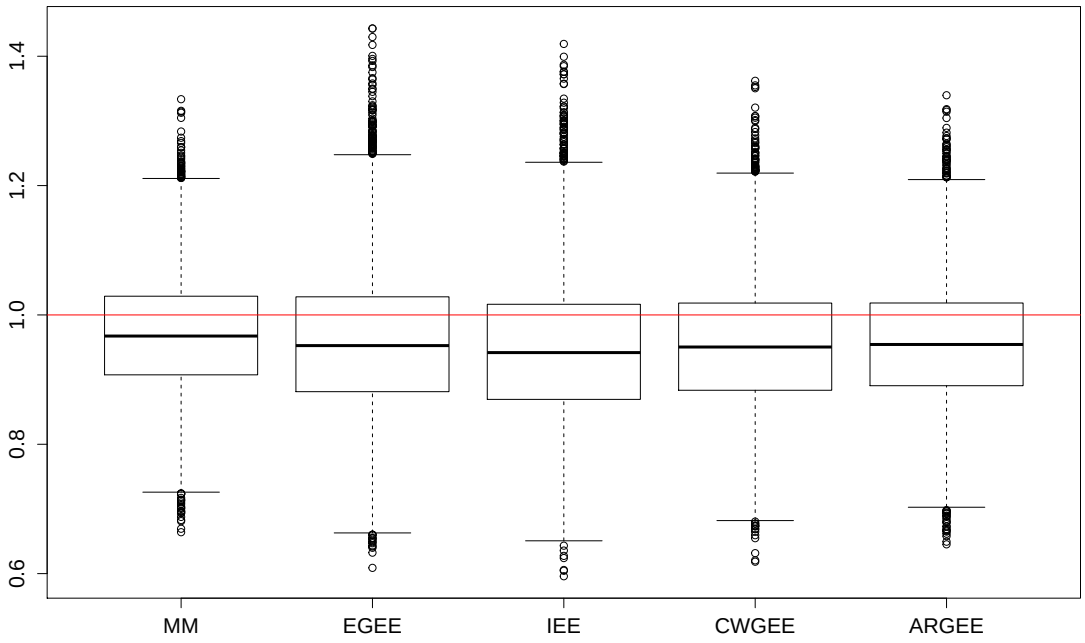

Figure 4.6: Boxplots of Interaction Effect standard error / standard deviation for simulation 1

Coverage Probability

For the NICS simulation scenarios, the coverage probability of the Wald type confidence interval for the interaction effect are shown in Table 4.7. These are all lower than the expected value of 0.95 but not by very much.

Simulation Scenario	1	2
MM	0.938	0.941
EGEE	0.933	0.940
IEE	0.933	0.937
CWGEE	0.932	0.937
ARGEE	0.933	0.939

Table 4.7: Coverage probabilities for NICS scenarios

Summary

In summary, for the NICS scenarios, there is no bias in the estimated treatment effect, the SE is roughly equal to the SD and the coverage probabilities are all close to 0.95. This is what is expected when the methods are performing well.

4.4.2 ICS Scenarios

The results for the ICS simulation scenarios with a treatment effect of 28 grams a week are given in Tables 4.8 to 4.12.

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.99	10.56	10.76	42.12	11.95	12.38	0.72	0.04	0.04	245.03	7.50	7.50	-1047.11	321.14	332.70
EGEE	27.63	14.40	17.61	42.18	17.73	18.94	0.72	0.06	0.07	244.78	9.97	12.43	-1047.55	477.78	507.73
IEE	27.73	14.20	17.72	42.30	18.95	21.57	0.72	0.07	0.08	226.08	12.13	16.51	-1020.56	511.03	578.20
CWGEE	27.89	12.79	14.21	42.13	16.52	17.79	0.72	0.06	0.06	223.83	11.00	12.73	-987.43	444.97	476.84
ARGEE	27.96	11.81	12.53	42.02	18.02	19.47	0.72	0.06	0.07	244.77	8.19	8.60	-1040.56	485.32	522.08

Table 4.8: Scenario 3 ($\gamma_1 = -0.50$, $\gamma_2 = 0$, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.99	10.44	10.70	41.84	12.00	12.25	0.72	0.04	0.04	243.64	7.42	7.61	-1037.92	322.26	329.33
EGEE	27.78	16.14	23.33	41.83	16.51	17.51	0.72	0.06	0.06	206.65	10.94	14.93	-966.41	444.69	471.39
IEE	27.77	13.22	18.91	42.02	17.12	20.13	0.72	0.06	0.07	181.66	11.55	17.04	-923.75	461.43	542.24
CWGEE	27.85	12.58	14.86	41.84	15.50	16.70	0.72	0.06	0.06	192.99	11.50	12.87	-928.36	417.67	449.53
ARGEE	28.06	12.57	13.31	41.75	18.17	19.52	0.72	0.06	0.07	234.93	8.46	9.81	-1033.89	489.19	524.94

Table 4.9: Scenario 4 ($\gamma_1 = 0, \gamma_2 = -0.50$, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.09	10.51	10.75	42.02	11.94	12.40	0.72	0.04	0.04	244.24	7.46	7.61	-1043.79	320.95	331.96
EGEE	27.99	15.26	21.01	41.85	16.33	17.73	0.72	0.06	0.06	217.60	10.43	13.89	-986.85	440.03	474.24
IEE	27.93	12.45	17.69	41.95	15.91	19.28	0.72	0.06	0.07	186.73	10.87	16.47	-933.98	429.23	516.92
CWGEE	27.98	11.58	13.63	41.87	13.96	15.37	0.72	0.05	0.06	193.29	10.64	12.68	-918.87	376.24	411.28
ARGEE	27.97	13.46	14.22	41.92	17.67	19.14	0.72	0.06	0.07	240.88	8.77	9.68	-1033.65	475.64	511.67

Table 4.10: Scenario 5 ($\gamma_1 = -0.35, \gamma_2 = -0.35$, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.82	10.53	10.74	42.08	11.97	12.38	0.72	0.04	0.04	244.87	7.48	7.56	-1045.29	321.65	332.29
EGEE	27.74	14.63	18.41	41.95	16.99	18.13	0.72	0.06	0.06	230.15	10.04	12.82	-1013.01	457.59	487.61
IEE	27.74	13.18	17.06	42.05	17.20	20.10	0.72	0.06	0.07	203.26	11.32	16.24	-968.88	463.70	541.05
CWGEE	27.78	11.98	13.37	41.93	14.97	16.15	0.72	0.05	0.06	205.68	10.61	12.54	-943.64	403.01	433.95
ARGEE	27.80	12.70	13.33	41.91	17.78	19.16	0.72	0.06	0.07	242.63	8.49	9.19	-1034.20	478.37	514.49

Table 4.11: Scenario 6 ($\gamma_1 = -0.46, \gamma_2 = -0.19$, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.79	10.48	10.73	41.93	11.98	12.34	0.72	0.04	0.04	244.03	7.45	7.51	-1040.58	321.70	332.50
EGEE	27.53	15.91	22.85	42.15	16.10	17.35	0.72	0.06	0.06	209.71	10.79	14.63	-979.75	433.65	465.57
IEE	27.63	12.49	18.31	42.29	15.83	19.14	0.72	0.06	0.07	179.41	10.98	16.75	-927.32	426.74	514.37
CWGEE	27.74	11.84	14.15	42.11	14.19	15.47	0.72	0.05	0.06	189.06	11.01	12.78	-921.01	382.28	415.99
ARGEE	27.70	13.34	14.09	42.04	17.93	19.46	0.72	0.06	0.07	238.50	8.70	9.84	-1040.07	482.14	520.97

Table 4.12: Scenario 7 ($\gamma_1 = -0.19, \gamma_2 = -0.46$, treatment effect = 28 g/week)

Bias

All scenarios with ICS show very little bias in the treatment effect for any method. This is shown in Figure 4.7, which shows boxplots by method of the 10000 estimates of the treatment effect from the simulation scenario in Table 4.8. The median estimate for each method is very close to the true value given by the red line. Hence, for the ICS cases there is no evidence of bias in the estimated treatment effect. A similar pattern was apparent in the boxplots of the treatment effect estimates for the other ICS scenarios (see Appendix B.1 Figures B.3 : B.7).

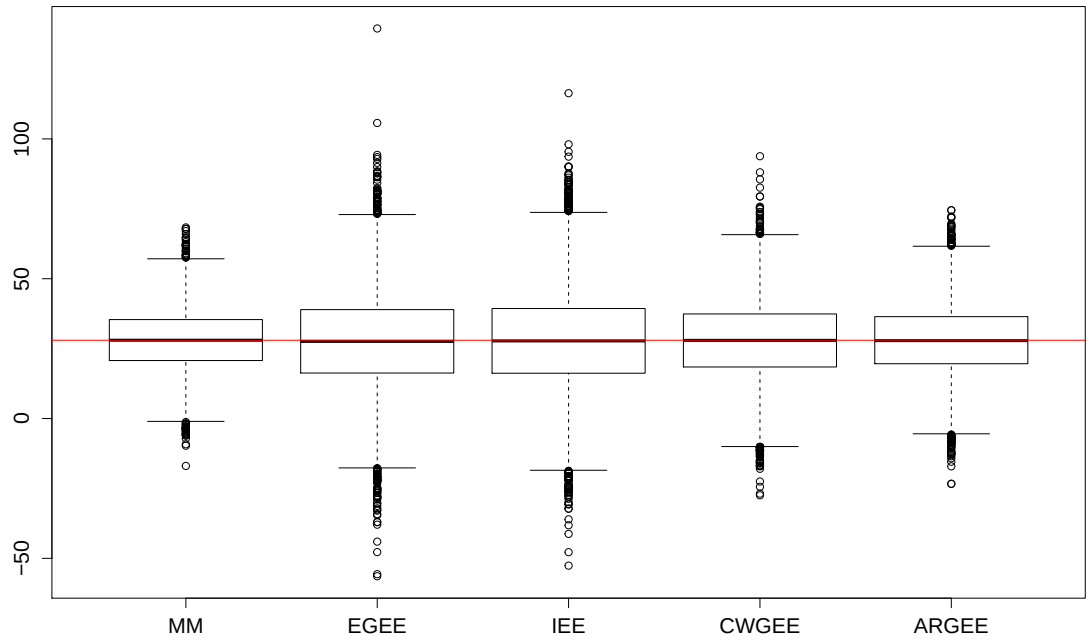

Figure 4.7: Boxplots of Interaction Effect estimates for scenario from Table 4.8

Standard Errors of Interaction Term

In the ICS scenarios, the mean model based standard error is no longer always roughly equal to the standard deviation of the interaction effect estimates for some methods. The difference occurs to varying degrees in different scenarios. Generally, the EGEE and the IEE show the largest difference, the CWGEE and ARGEE show a small difference and the MM shows very little difference. In all cases, the mean of the model based standard error is smaller than the standard deviation of the estimates. This implies that the standard error is being underestimated by these methods.

Table 4.13 gives the ratio of the mean model based standard error and the standard deviation of the estimates, by method for the five simulation scenarios with ICS. Generally for the EGEE, IEE and CWGEE, the worst results can be seen to occur in simulation scenarios 4 and 7, where duration is correlated with the random slope more than with the random intercept. The best performance occurs when duration is highly correlated with the random intercept but not the random slope, seen in simulation scenarios 3 and 6.

Simulation Scenario	3	4	5	6	7
MM	0.98	0.98	0.98	0.98	0.98
EGEE	0.82	0.69	0.73	0.79	0.69
IEE	0.80	0.70	0.70	0.77	0.68
CWGEE	0.90	0.85	0.85	0.90	0.84
ARGEE	0.94	0.94	0.95	0.95	0.95

Table 4.13: Ratio of mean SE divided by SD

The more duration is correlated with the random slope, the larger the difference between SE and SD becomes for the EGEE, IEE and CWGEE. The MM and ARGEE do not appear to vary between simulation scenarios. In general, the ratios are furthest from 1 for the EGEE and IEE and the ratios are moderately different to 1 for the CWGEE, the ratio is slightly different from 1 for the ARGEE and the MM appears to do well in all cases.

To further understand how the different analysis methods compare, consider Figure 4.8. This gives the ratio of SE for each of the 10000 simulated datasets divided by SD for each method when the data from all of the simulation scenarios with ICS are combined. Three outliers with values above 6 lie above the plot limits for the ARGEE and hence are not shown in the plot. Figure 4.8 shows that the EGEE and IEE are substantially underestimating the model based standard error and the the CWGEE and ARGEE are slightly underestimating it. The MM is the only method which doesn't underestimate the standard error. It should also be noted that there is a large amount of variability present, especially for the ARGEE. Separate plots for each simulation scenario showed similar results to the combined plot (see Appendix B.1 Figures B.10 : B.14).

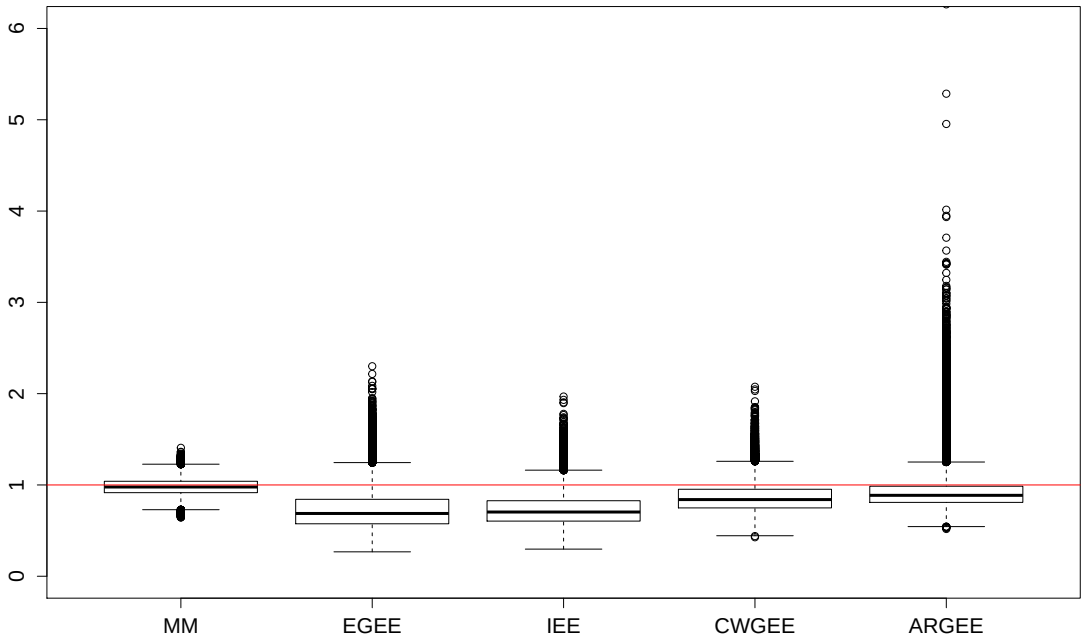

Figure 4.8: Boxplots of Interaction Effect standard error / standard deviation for all ICS data

Coverage Probabilities

The coverage probabilities of the Wald type confidence intervals for the interaction effect for the scenarios with ICS, are given in Table 4.14. The MM and ARGEE have coverage probabilities which are close to but slightly below 0.95 and do not vary much between simulation scenarios. The EGEE and IEE have the worst coverage probabilities, which all fall below 0.9, while the CWGEE has coverage probabilities of around 0.9. The EGEE and IEE methods have coverage probabilities which vary quite a lot between simulation scenarios and the coverage probabilities vary a little for the CWGEE. When the strength of the correlation between the random slope and duration is greater than the strength of the correlation between the random intercept and duration (scenario 4 and 7), the coverage probabilities are lower. This again suggests that the EGEE, IEE and CWGEE perform worse when the random slope is correlated with duration.

Simulation Scenario	3	4	5	6	7
MM	0.942	0.941	0.94	0.94	0.938
EGEE	0.887	0.832	0.861	0.886	0.842
IEE	0.877	0.822	0.829	0.865	0.813
CWGEE	0.915	0.895	0.897	0.912	0.891
ARGEE	0.93	0.931	0.943	0.938	0.941

Table 4.14: Coverage probabilities for ICS scenarios

Summary

In summary, these results show for a treatment effect of 28g/week that when cluster size is non informative, all methods provide valid inference. There is no bias, the model based standard error is being estimated well and the coverage probability is close to 0.95. This is what we expect to see, as all methods should perform well in this case. Notably, the IEE and CWGEE are the most inefficient methods. For simulation scenarios with ICS, no method produced substantial bias in the treatment effect. However, the model based standard error differed from the standard deviation of the estimates for all methods except the MM. In fact, the EGEE and IEE are severely underestimating the model based standard error and the CWGEE is slightly underestimating it. The ARGEE only had a very minor difference between SE and SD. However, there is a large amount of variability in the ratio of median SE divided by SD for the ARGEE. Comparing the variability of the ARGEE to the smallest variability method, the MM, the ratio of SE divided by SD ranged from 0.52 to 16.82 for the ARGEE and 0.64 to 1.40 for the MM, which mean the ARGEE has misleading SE estimates. The coverage probability is also severely low for the EGEE and IEE and slightly low for the CWGEE. For all methods except the MM and ARGEE, the strength of the correlation between duration and the random effects influences how badly these methods preform. When duration is correlated with the random slope only, the difference between SE and SD is greater and the coverage probabilities are lower. The best results occur when duration is not correlated with the random slope. The MM performs well in all circumstances considered and the ARGEE performed nearly as well, but there was a large variability in the ratio of mean SE divided by SD.

The results for the simulation scenarios using the other treatment effects of 0, 7, 14 and 21 g/week, gave similar results to those discussed in this section. The value of the treatment effect estimate changed accordingly but the analysis concerning bias, estimates of SEs and coverage probabilities were unaffected. Full results tables for these scenarios can be found in Appendix A.1 Tables A.1 : A.35.

4.5 Simulation Study Conclusions

The simulation study was conducted to compare methods of analysis for longitudinal data with ICS when the true parameter values are known. This allowed the methods to be compared by considering the bias, the difference between SE and SD, and the coverage probabilities for each method.

The results of the simulation study showed that when there was NICS, all methods performed well, as expected. When ICS was present however, not all methods performed well. The MM performed the best, followed by the ARGEE. The EGEE, IEE and CWGEE all performed poorly, with low coverage probabilities and large differences between SE and SD. The best results for these methods occurred when duration was correlated with the random intercept only, with no relationship between the cluster size and the random slope. It is worth noting that simulations were based on the MM which could be causing it to be favoured in the analysis. This will be explored in Chapter 5.

Overall, the results from the simulation study suggest that the MM and ARGEE are the most suitable methods for analysing longitudinal data when the length of followup is informative. Before a final conclusion is made however, the reasons why the other methods may have performed poorly needs to be explored. This will be the topic of Chapter 5.

Chapter 5

Simulation Study Extensions

In Chapter 4, a simulation study was conducted to compare methods of analysis for longitudinal data with ICS. Three issues arose from these results that are worthy of further investigation:

1. The distribution of the cluster size was chosen to give the same mean and variance as the POPPET data in these simulations. It is relevant to consider the possible effect of different distributions of cluster size on the conclusions.

2. It was observed that a large difference between the model based standard error and the standard deviation of the estimates could occur for various methods under certain ICS scenarios. It is relevant to investigate whether this was due to the relatively small number of clusters considered (n=60).

3. It was observed that the mixed model performed well in all cases, while the performance of the GEE approaches varied for data simulated under a mixed model. It is relevant to consider whether this is true more generally.

In Chapter 5, three extensions to the original simulation study were considered to investigate these issues. In Extension 1, the distribution of the cluster size was varied by increasing the standard deviation to investigate the effect on the performance of the methods. In Extension 2, the sample size of the trial was increased to investigate whether the difference between the model based standard error (SE) and the standard deviation of the estimate (SD) would be reduced for larger studies. This extension was conducted

on both the original distribution of the cluster size and the Extension 1 distribution of the cluster size. In Extension 3, the data were simulated from a multivariate distribution not equivalent to the mixed model, to investigate the relative performance of the methods under a different simulation model. This extension was conducted on both the original and larger sample size.

5.1 Extension 1: Larger Standard Deviation of Cluster Size

The distribution of the cluster size in the Chapter 4 simulation study was chosen to match the POPPET trial data. In the first extension to the Chapter 4 simulation study, the variability of the cluster size was increased by doubling the standard deviation of the cluster size from 14 to 28. It is expected that increasing the variability will increase any problems caused by ICS.

New values for γ_0, γ_1 and γ_2 are required to produce the desired standard deviation. Repeating the calculation in Section 4.2.2 with the new standard deviation gives $\gamma_0 = 2.787813$ and $0.81 = \gamma_1^2 + \gamma_2^2$. Values of γ_1 and γ_2 were then chosen to cover the same five correlation settings used in the original simulation study, as summarised in Table 5.1. This produced a total of 25 new simulation scenarios when combined with 5 different levels of weekly weight gain, which are listed in Table 5.2.

Case	γ_1	γ_2	$cor(b_0, N)$	$cor(b_1, N)$
1	-0.90	0	-0.86	0
2	0	-0.90	0	-0.87
3	-0.64	-0.64	-0.62	-0.62
4	-0.83	-0.34	-0.80	-0.34
5	-0.34	-0.83	-0.34	-0.81

Table 5.1: Extension 1: Values of γ_1 and γ_2 and resulting correlations

Scenario	γ_0	γ_1	γ_2	Treatment Effect(g/week)
36	2.79	-0.90	0	28
37	2.79	0	-0.90	28
38	2.79	-0.64	-0.64	28
39	2.79	-0.83	-0.34	28
40	2.79	-0.34	-0.83	28
41	2.79	-0.90	0	21
42	2.79	0	-0.90	21
43	2.79	-0.64	-0.64	21
44	2.79	-0.83	-0.34	21
45	2.79	-0.34	-0.83	21
46	2.79	-0.90	0	14
47	2.79	0	-0.90	14
48	2.79	-0.64	-0.64	14
49	2.79	-0.83	-0.34	14
50	2.79	-0.34	-0.83	14
51	2.79	-0.90	0	7
52	2.79	0	-0.90	7
53	2.79	-0.64	-0.64	7
54	2.79	-0.83	-0.34	7
55	2.79	-0.34	-0.83	7
56	2.79	-0.90	0	0
57	2.79	0	-0.90	0
58	2.79	-0.64	-0.64	0
59	2.79	-0.83	-0.34	0
60	2.79	-0.34	-0.83	0

Table 5.2: Simulation Scenarios for Extension 1

5.1.1 Results for Extension 1

Recall that in the results tables β_1 is the estimated coefficient for the treatment $\times$ time interaction effect, β_2 is the estimated coefficient for gestational age, β_3 is the estimated coefficient for birthweight, β_4 is the estimated coefficient for time and β_0 is the intercept. The results for the simulation scenarios with a treatment effect of 28 grams a week are given in Tables 5.3 - 5.7.

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.06	10.69	10.87	41.95	12.02	12.31	0.72	0.04	0.04	245.44	7.60	7.66	-1043.43	322.78	331.19
EGEE	27.96	17.25	28.00	41.94	18.16	19.52	0.72	0.06	0.07	244.61	11.59	19.93	-1041.49	488.95	523.95
IEE	27.97	15.72	26.31	42.08	23.38	29.97	0.72	0.08	0.11	228.53	12.85	24.49	-1032.30	630.61	806.55
CWGEE	28.08	14.48	18.73	41.93	16.23	17.80	0.72	0.06	0.06	223.03	12.65	18.02	-981.20	437.11	478.01
ARGEE	28.09	12.56	13.27	41.95	17.78	19.17	0.72	0.06	0.07	246.22	8.65	9.14	-1037.23	478.70	514.17

Table 5.3: Scenario 36 (γ_1=-0.90, γ_2=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.08	10.42	10.69	42.01	11.99	12.39	0.72	0.04	0.04	242.49	7.40	7.64	-1040.62	322.14	332.88
EGEE	27.71	14.21	31.44	41.92	14.70	15.83	0.72	0.05	0.06	184.12	9.73	20.53	-928.41	395.96	423.92
IEE	27.88	9.53	21.32	41.97	15.87	20.92	0.72	0.06	0.07	162.45	7.57	19.05	-891.52	428.02	563.49
CWGEE	27.90	11.37	17.06	41.92	13.44	14.81	0.72	0.05	0.05	173.23	10.27	15.63	-899.67	362.11	398.23
ARGEE	28.04	13.33	14.07	41.81	17.50	18.82	0.72	0.06	0.07	232.64	8.85	9.59	-1034.62	471.11	505.41

Table 5.4: Scenario 37 (γ_2=-0.90, γ_1=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.91	10.59	10.64	41.80	12.02	12.39	0.72	0.04	0.04	243.88	7.53	7.67	-1036.60	322.98	334.80
EGEE	27.93	15.85	29.77	41.87	14.38	15.81	0.72	0.05	0.06	202.11	10.69	19.98	-960.35	387.34	425.76
IEE	27.93	10.39	21.06	42.28	14.96	21.00	0.72	0.05	0.07	175.50	8.66	19.02	-935.78	404.19	566.27
CWGEE	27.89	10.37	15.09	41.91	10.75	12.20	0.72	0.04	0.04	179.31	9.49	14.57	-899.13	289.72	329.03
ARGEE	27.83	13.88	14.30	41.73	16.53	17.97	0.72	0.06	0.06	239.27	9.08	9.49	-1026.04	445.16	483.97

Table 5.5: Scenario 38 (γ_1=-0.64, γ_2=-0.64, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.23	10.65	10.94	42.04	12.00	12.29	0.72	0.04	0.04	244.60	7.57	7.77	-1043.13	322.48	329.58
EGEE	27.81	16.89	29.22	41.76	16.30	17.26	0.72	0.06	0.06	221.90	11.33	20.13	-992.33	439.00	462.73
IEE	27.84	13.51	24.45	41.67	19.32	25.52	0.72	0.07	0.09	198.93	11.15	22.02	-964.23	521.62	686.85
CWGEE	27.99	12.46	17.01	41.74	13.23	14.41	0.72	0.05	0.05	197.89	11.07	16.16	-926.03	356.49	387.32
ARGEE	28.21	13.22	13.83	41.88	16.94	18.01	0.72	0.06	0.07	242.03	8.83	9.40	-1029.13	455.91	483.74

Table 5.6: Scenario 39 (γ_1=-0.83, γ_2=-0.34, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.97	10.48	10.62	41.93	12.01	12.30	0.72	0.04	0.04	243.07	7.45	7.62	-1038.93	322.62	330.28
EGEE	28.24	14.76	31.03	41.70	13.66	14.87	0.72	0.05	0.05	188.50	10.04	20.56	-930.58	368.10	399.64
IEE	28.12	8.38	19.89	41.63	13.18	18.51	0.72	0.05	0.07	162.53	6.87	18.42	-888.39	355.42	499.40
CWGEE	28.11	9.87	15.33	41.75	10.85	12.17	0.72	0.04	0.04	170.49	9.16	15.02	-882.98	292.37	328.16
ARGEE	28.15	13.86	14.34	41.90	16.81	18.28	0.72	0.06	0.07	235.77	9.08	9.46	-1033.34	452.14	490.29

Table 5.7: Scenario 40 (γ_2=-0.83, γ_1=-0.34, treatment effect = 28 g/week)

Bias

In Extension 1, there is very little bias in the estimated treatment effect for any of the analysis methods. This can been seen in Figure 5.1, which shows boxplots by method of the 10000 estimates of the treatment effect from the simulation scenario in Table 5.3. The median estimated treatment effect for each method is very close to the true value (given by the red line), which is similar to the ICS scenarios in the Chapter 4 simulation study. The MM has the smallest spread and the EGEE and IEE have the largest. A similar pattern was seen in the boxplots of the treatment effect estimates for the other ICS scenarios (see Appendix B.2 Figures B.15 : B.19).

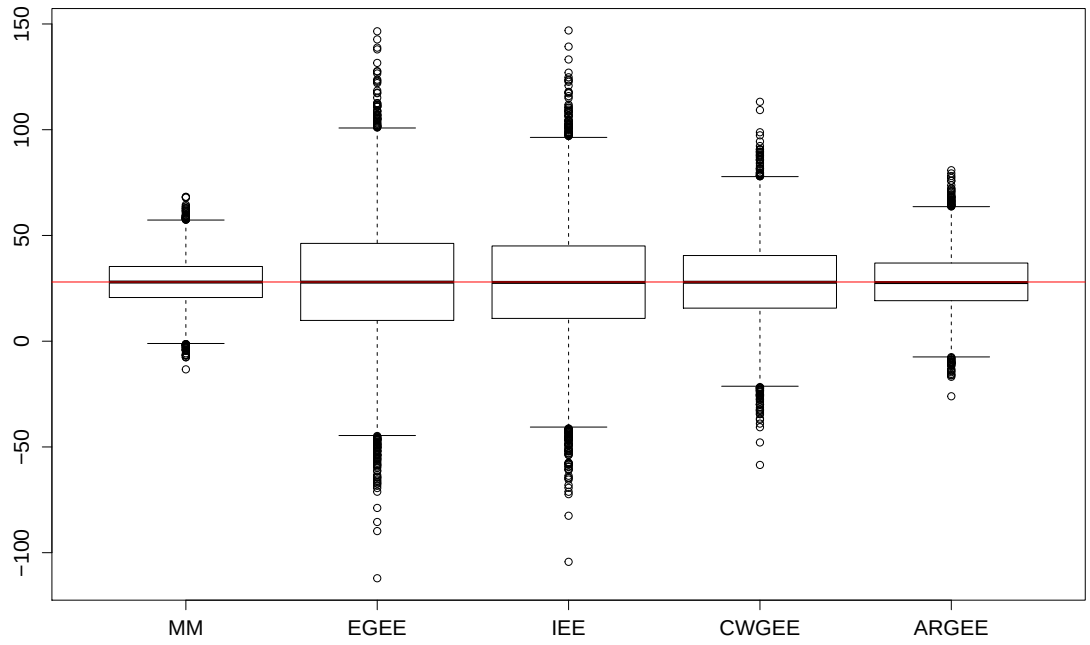

Figure 5.1: Extension 1: Boxplots of Interaction Effect estimates for Scenario 36 in Table 5.3

Standard Errors of Interaction Term

As seen in the Chapter 4 simulation study, for the interaction effect there is still a difference between the model based standard error and the standard deviation of the estimates. In these scenarios, the difference now appears to be exacerbated. Table 5.8 gives the ratio of the mean model based standard error divided by the standard deviation of the estimates, by method for the five simulation scenarios in the extension study with treatment effect 28g/week. It appears that the ratios are similar to those seen in the Chapter 4 simulation study for the MM and the ARGEE, but are different for the EGEE, IEE and CWGEE.

For example, the EGEE ratio in simulation scenario 4 was 0.69, for the Chapter 4 simulation study compared to 0.45 for simulation scenario 37 in Extension 1. In fact, the difference between SE and SD is now so large that the difference is larger than the standard error itself. These results show that the EGEE and IEE are now severely underestimating

the standard error and the CWGEE is also underestimating it quite substantially. Hence, increasing the spread of the distribution of the cluster size appears to have reduced the ability of the EGEE, IEE and CWGEEs methods to correctly estimate the standard error but has had little effect on the MM and ARGEE. The EGEE, IEE and CWGEE methods had ratios closer to 1 when duration was correlated with the random intercept only, similar to the results seen in the Chapter 4 simulation study.

Simulation Scenario	36	37	38	39	40
MM	0.98	0.97	1.00	0.97	0.99
EGEE	0.62	0.45	0.53	0.58	0.48
IEE	0.62	0.45	0.49	0.55	0.42
CWGEE	0.77	0.67	0.69	0.73	0.64
ARGEE	0.95	0.95	0.97	0.96	0.97

Table 5.8: Extension 1: Ratios of mean SE divided by SD

To summarise the difference between SE and SD for the 5 methods, consider Figure 5.2, which gives the ratio of the SE for each of the 10000 simulated datasets divided by the SD for each method when the data from all of the simulation scenarios with ICS are combined. This shows that across all the simulation scenarios when ICS is present, the EGEE and IEE are severly underestimating the model based standard error, the CWGEE is moderately underestimating it and the MM and ARGEE are estimating it quite well. The general pattern is similar to the results found for the original simulations. However, in this case the amount by which each method has underestimated SE is much larger than for the original simulations. It is again observed that there is a large amount of volatility, which is again worst for the ARGEE. Separate plots for each simulation scenario showed similar results to the combined plot (see Appendix B.2 Figures B.20 : B.24).

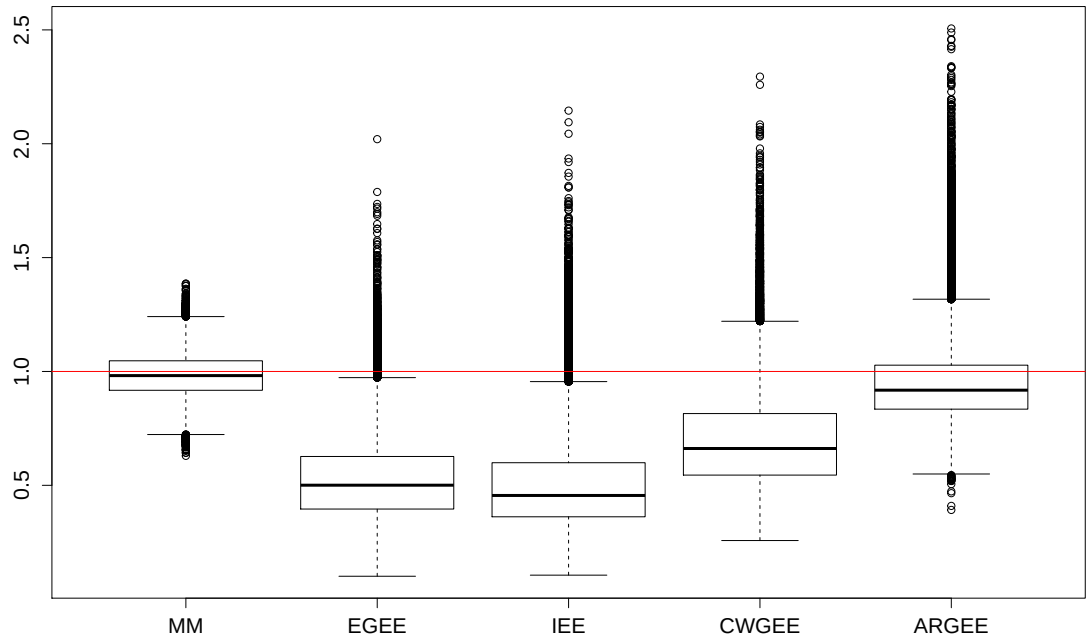

Figure 5.2: Extension 1: Boxplots of Interaction Effect standard error / standard deviation for all ICS data

Coverage Probabilities

The coverage probabilities of the Wald type confidence interval for the interaction effect are given in Table 5.9. For the MM and the ARGEE these are similar to the Chapter 4 simulation study and are close to but slightly below the expected value of 0.95. The EGEE and IEE have the worst coverage probabilities and the CWGEE is marginally better, which is also consistent with the Chapter 4 simulation study. The EGEE, IEE and CWGEE methods have coverage probabilities that vary substantially between simulation scenarios. When the correlation between the random slope and duration is stronger than between the random intercept and duration, the coverage probabilities are lower. The major difference here is that the EGEE, IEE and CWGEE all have much lower coverage probabilities than those found in the original simulations. This agrees with the larger differences between SE and SD for these methods, hence indicating that the EGEE, IEE and CWGEE methods are performing worse when the cluster size is more variable compared to the original simulations.

Simulation Scenario	36	37	38	39	40
MM	0.942	0.942	0.946	0.941	0.943
EGEE	0.755	0.594	0.687	0.732	0.627
IEE	0.746	0.584	0.644	0.711	0.557
CWGEE	0.865	0.789	0.814	0.845	0.777
ARGEE	0.931	0.935	0.945	0.941	0.943

Table 5.9: Extension 1: Coverage probabilities

Summary

Overall, increasing the variability in the cluster sizes has had an effect on the simulated results for some methods. The performance of the MM and ARGEE is largely unaffected by the change in cluster size distribution and they continue to perform well. On the other hand, for the EGEE, IEE and CWGEE methods the difference between SE and SD is greater and the coverage probabilities are lower compared with the original simulation results. In general, it appears that these methods all perform worse when the variability in the cluster sizes is increased.

The results for the simulation scenarios using the other treatment effects of 0, 7, 14 and 21 g/week gave similar results to those discussed above. The value of the treatment effect estimate changed accordingly but the analysis concerning bias, estimates of SEs and coverage probabilities were unaffected. Full results for these scenarios can be found in Appendix A.2, Tables A.36 : A.60.

5.2 Extension 2: Larger Sample Size

In the Chapter 4 simulation study, the standard deviation of the treatment effect estimates differed from the model based standard error and the coverage probabilities were low for the EGEE, IEE and CWGEE methods. This could have occurred because the

study size was not large enough for the models to estimate the standard error correctly. GEEs are known to perform poorly when the number of clusters is small (Mancl and DeRouen, 2001). The simulation study in Chapter 4 only included 60 infants (n=60), so perhaps the models would obtain a more reasonable estimate of the standard error if there were more infants.

In Extension 2, a new simulation study was conducted which included 600 infants (n=600) in each simulated dataset. Both the 35 simulation scenarios from the original simulation study and the 25 from Extension 1 were considered. Note that only 1000 simulations were completed for each simulation scenario, compared to 10000 in the original simulation study, due the computational time required to simulate and analyse data from a larger study.

5.2.1 NICS Scenario Results for Extension 2

The results for the NICS simulation scenarios with a treatment effect of 28 grams a week are given in Tables 5.10 and 5.11.

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.05	3.35	3.43	41.73	3.73	3.77	0.72	0.01	0.01	244.37	2.37	2.41	-1035.37	100.46	101.64
EGEE	28.05	3.35	3.43	41.91	5.62	5.68	0.72	0.02	0.02	244.37	2.37	2.41	-1040.00	151.23	152.66
IEE	27.90	3.97	4.06	41.91	5.62	5.69	0.72	0.02	0.02	244.44	2.60	2.66	-1040.03	151.24	152.75
CWGEE	27.90	3.97	4.06	41.91	5.62	5.69	0.72	0.02	0.02	244.44	2.60	2.66	-1040.03	151.24	152.75
ARGEE	28.00	3.52	3.60	41.93	5.70	5.80	0.72	0.02	0.02	244.40	2.48	2.54	-1040.65	153.33	155.94

Table 5.10: Extension 2 Scenario 1 (fixed trial length=38, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.07	3.36	3.21	41.92	3.74	3.84	0.72	0.01	0.01	244.42	2.37	2.34	-1041.55	100.57	103.23
EGEE	28.08	3.58	3.51	42.07	5.69	5.80	0.72	0.02	0.02	244.42	2.54	2.55	-1045.55	153.14	155.58
IEE	28.07	4.15	4.08	42.10	5.84	5.97	0.72	0.02	0.02	244.44	3.00	3.05	-1046.70	157.15	160.31
CWGEE	28.07	4.04	3.96	42.07	5.69	5.80	0.72	0.02	0.02	244.44	2.87	2.92	-1045.63	153.14	155.74
ARGEE	28.07	3.57	3.46	42.08	5.81	5.91	0.72	0.02	0.02	244.43	2.53	2.54	-1046.11	156.18	158.79

Table 5.11: Extension 2 Scenario 2 (γ_0=log(24.3667),γ_1=0, γ_2=0, treatment effect = 28 g/week)

Bias

For the two simulation scenarios with NICS, the results are very similar to those discussed for the Chapter 4 simulation study. There is very little bias in the estimated treatment effect, as seen in Figure 5.3, which shows boxplots by method of the 1000 estimates from the simulation scenario in Table 5.11. The median estimated treatment effect, for each method, is very close to the true value (shown in red), indicating there is no bias. A similar pattern was seen in the boxplots of the treatment effect estimates for the other NICS scenario (see Appendix B.3 Figure B.25).

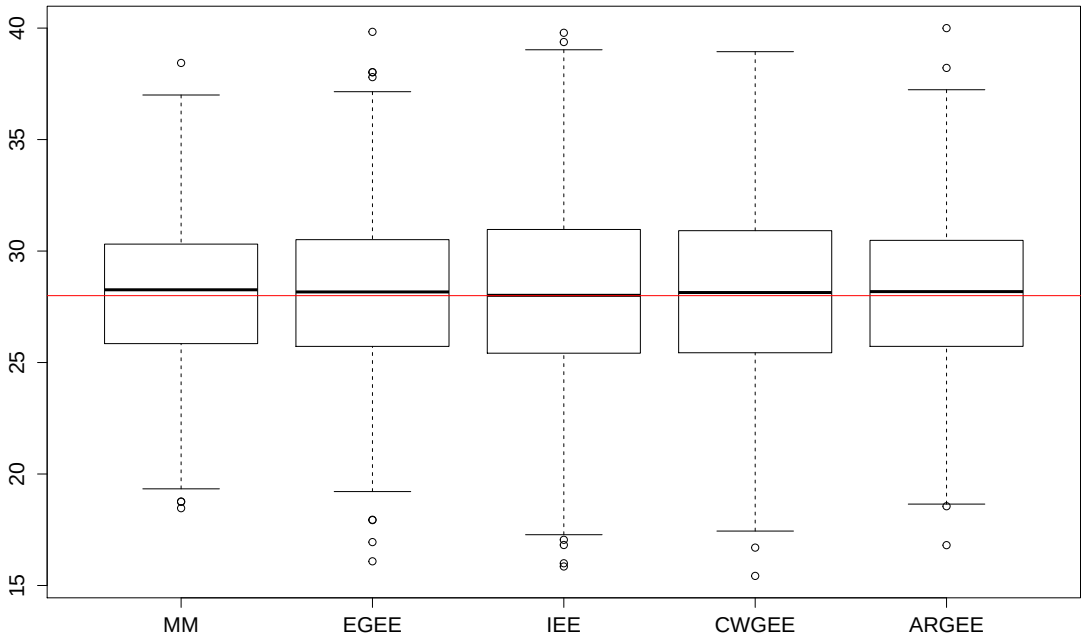

Figure 5.3: Extension 2: Boxplots of Interaction Effect estimates for Scenario 1 in Table 5.10

Standard Errors of Interaction Term

For the interaction effect in Tables 5.10 and 5.11, the IEE and CWGEE have the largest SE. In this extension, all methods have very small differences between SE and SD. Consider Figure 5.4, which gives a boxplot of the ratio of the standard error for each of the

1000 simulated datasets divided by the standard deviation for each method for the simulation scenario in Table 5.11. The median ratio is now always greater than 1, indicating a tendency for the SE to be slightly larger than the SD. Overall, the ratio of SE divided by SD for each method is very close to 1 with an average ratio of around 1.025. A similar pattern was seen in the boxplots of the SE/SD for the other NICS scenario, however this plot showed the median ratio was slightly less than 1 (see Appendix B.3 Figure B.37).

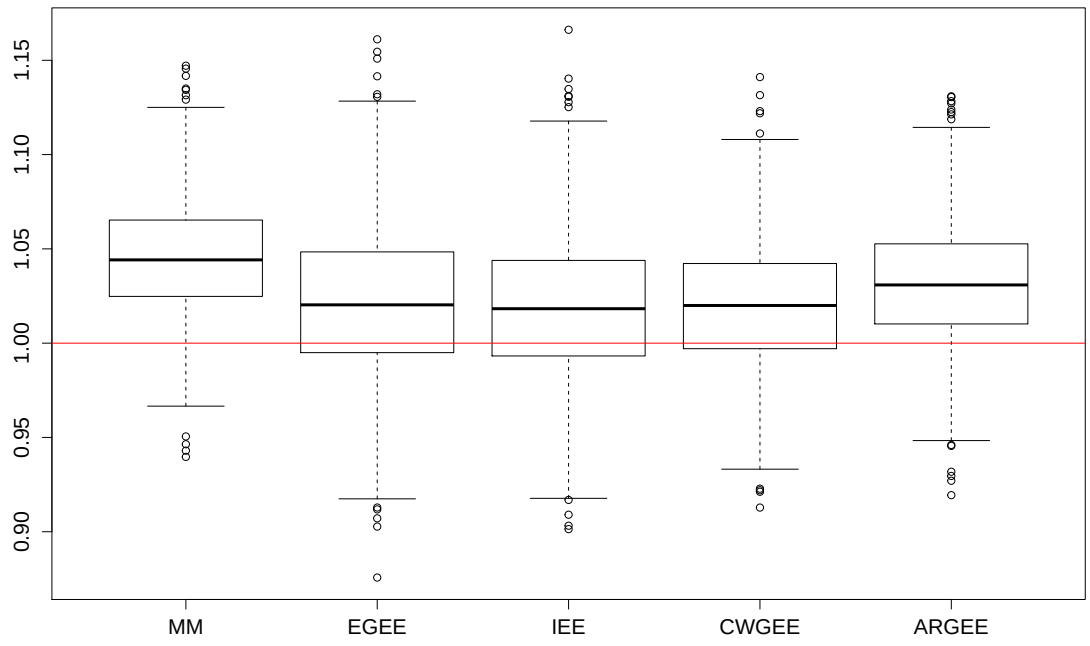

Figure 5.4: Extension 2: Boxplots of Interaction Effect standard error / standard deviation for Scenario 1 in Table 5.10

Coverage Probabilities

For the NICS simulation scenarios, the coverage probabilities of the Wald type confidence interval for the interaction effect are given in Table 5.12. The coverage probabilities are all very close to the expected value of 0.95 and are slightly higher than the coverage probabilities in the original simulation study.

Simulation Scenario	1	2
MM	0.96	0.939
EGEE	0.95	0.941
IEE	0.95	0.94
CWGEE	0.96	0.94
ARGEE	0.96	0.94

Table 5.12: Extension 2: Coverage probabilities for NICS scenarios

Summary

In summary, for the NICS scenarios, there is no bias in the estimated treatment effect, the SE is roughly equal to the SD and the coverage probabilities are all close to 0.95. This is what is expected when the methods are performing well and agrees with the findings in Section 4.4, when the sample size was only 60.

5.2.2 ICS Scenario Results for Extension 2

The results for the ICS simulation scenarios with a sample size of 600 and a treatment effect of 28 grams a week are given in Tables 5.13 and 5.22.

Scenarios with Smaller SD of Cluster Size using $\gamma_0 = 3.069142$

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.09	3.37	3.48	42.05	3.74	3.69	0.72	0.01	0.01	245.01	2.39	2.39	-1044.95	100.56	99.36
EGEE	28.27	5.98	6.29	42.13	5.85	5.80	0.72	0.02	0.02	244.40	4.15	4.52	-1046.01	157.23	154.76
IEE	28.29	5.99	6.27	42.12	6.91	7.00	0.72	0.02	0.03	223.79	5.75	6.52	-1011.64	185.95	187.04
CWGEE	28.18	4.50	4.53	42.12	5.44	5.43	0.72	0.02	0.02	222.33	4.15	4.27	-983.38	146.33	145.11
ARGEE	28.07	3.92	4.06	42.20	5.89	5.94	0.72	0.02	0.02	245.75	2.68	2.92	-1044.93	158.40	158.51

Table 5.13: Extension 2 Scenario 3 (γ_1=-0.50, γ_2=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.95	3.34	3.37	42.04	3.75	3.80	0.72	0.01	0.01	243.59	2.36	2.41	-1042.84	100.75	102.17
EGEE	27.99	8.72	10.06	42.21	5.44	5.62	0.72	0.02	0.02	202.33	5.92	6.61	-964.82	146.53	151.26
IEE	28.00	6.85	7.89	42.41	6.54	6.95	0.72	0.02	0.02	172.88	7.15	8.29	-910.73	176.66	186.75
CWGEE	28.01	4.86	5.03	42.18	5.12	5.31	0.72	0.02	0.02	188.85	5.02	4.75	-926.27	138.12	143.10
ARGEE	27.82	5.40	5.95	42.10	5.93	6.25	0.72	0.02	0.02	239.69	3.32	3.94	-1041.89	159.69	169.38

Table 5.14: Extension 2 Scenario 4 (γ_2=-0.50, γ_1=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.07	3.36	3.29	42.19	3.74	3.87	0.72	0.01	0.01	244.10	2.38	2.32	-1049.06	100.70	104.34
EGEE	28.12	7.62	8.57	42.19	5.39	5.53	0.72	0.02	0.02	214.50	5.18	5.54	-990.24	145.34	149.48
IEE	28.17	6.27	7.11	42.24	6.13	6.45	0.72	0.02	0.02	178.78	6.38	7.28	-924.84	165.59	175.46
CWGEE	28.17	4.39	4.51	42.14	4.59	4.69	0.72	0.02	0.02	189.25	4.49	4.33	-917.94	123.84	127.18
ARGEE	28.02	6.30	6.87	41.99	5.76	5.78	0.72	0.02	0.02	243.74	3.72	3.89	-1025.98	155.21	156.13

Table 5.15: Extension 2 Scenario 5 (γ_1=-0.35, γ_2=-0.35, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.92	3.38	3.51	41.98	3.74	3.86	0.72	0.01	0.01	244.71	2.39	2.34	-1042.36	100.66	104.18
EGEE	27.65	6.59	7.11	41.68	5.62	5.75	0.72	0.02	0.02	228.39	4.59	4.74	-1001.59	151.28	155.31
IEE	27.70	6.04	6.56	41.65	6.42	6.76	0.72	0.02	0.02	197.77	5.99	6.59	-945.59	173.31	183.72
CWGEE	27.76	4.36	4.49	41.72	4.93	5.02	0.72	0.02	0.02	202.70	4.23	4.06	-930.64	132.75	135.49
ARGEE	27.91	5.06	5.35	41.64	5.78	5.78	0.72	0.02	0.02	245.33	3.19	3.23	-1021.88	155.66	155.47

Table 5.16: Extension 2 Scenario 6 (γ_1=-0.46, γ_2=-0.19, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.15	3.35	3.43	42.01	3.74	3.87	0.72	0.01	0.01	243.86	2.37	2.37	-1042.32	100.67	103.49
EGEE	28.53	8.28	9.23	42.00	5.33	5.42	0.72	0.02	0.02	205.66	5.67	6.11	-966.70	143.55	146.33
IEE	28.41	6.49	7.25	41.85	6.17	6.55	0.72	0.02	0.02	170.71	6.79	7.84	-895.52	166.36	178.40
CWGEE	28.31	4.55	4.55	41.99	4.69	4.79	0.72	0.02	0.02	184.76	4.77	4.46	-907.41	126.47	129.37
ARGEE	28.16	6.23	6.62	42.13	5.82	5.83	0.72	0.02	0.02	242.03	3.72	3.97	-1034.80	156.53	156.79

Table 5.17: Extension 2 Scenario 7 (γ_2=-0.46, γ_1=-0.19, treatment effect = 28 g/week)

Scenarios With Larger SD of Cluster Size, $\gamma_0 = 2.787813$

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.90	3.42	3.39	41.87	3.75	3.79	0.72	0.01	0.01	245.46	2.42	2.35	-1041.21	100.96	101.48
EGEE	27.31	10.65	11.32	41.85	6.21	6.25	0.72	0.02	0.02	245.07	7.32	7.88	-1039.97	167.52	167.45
IEE	27.31	10.29	11.11	41.77	11.39	12.30	0.72	0.04	0.04	228.27	9.04	10.00	-1026.20	307.35	331.70
CWGEE	27.65	6.44	6.55	41.87	5.64	5.66	0.72	0.02	0.02	222.87	6.02	6.27	-979.58	152.16	151.75
ARGEE	27.85	4.01	3.95	41.94	5.69	5.74	0.72	0.02	0.02	247.30	2.74	2.73	-1035.77	152.94	153.56

Table 5.18: Extension 2 Scenario 36 (γ_1=-0.90, γ_2=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.19	3.33	3.32	42.27	3.75	3.82	0.72	0.01	0.01	242.33	2.36	2.39	-1047.42	100.89	102.80
EGEE	27.82	9.96	10.99	42.19	4.77	4.74	0.72	0.02	0.02	174.13	6.87	7.55	-911.12	128.78	128.79
IEE	27.98	6.89	7.92	41.93	7.30	7.91	0.72	0.03	0.03	149.58	6.09	7.22	-854.49	197.26	214.61
CWGEE	28.02	5.38	5.70	42.17	4.53	4.52	0.72	0.02	0.02	166.84	5.21	5.32	-890.88	122.48	122.67
ARGEE	28.20	5.29	5.44	42.60	5.48	5.51	0.72	0.02	0.02	234.11	3.37	3.30	-1049.25	147.63	148.76

Table 5.19: Extension 2 Scenario 37 (γ_2=-0.90, γ_1=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.98	3.38	3.46	41.99	3.76	3.85	0.72	0.01	0.01	243.83	2.39	2.45	-1042.38	101.11	103.88
EGEE	26.87	10.39	10.99	42.02	4.71	4.86	0.72	0.02	0.02	195.22	7.05	7.59	-947.00	127.27	131.87
IEE	27.12	7.67	8.38	42.00	7.54	8.41	0.72	0.03	0.03	165.77	6.69	7.71	-901.19	204.44	227.74
CWGEE	27.54	5.00	5.17	42.06	3.67	3.79	0.72	0.01	0.01	174.35	4.78	5.09	-891.30	99.11	102.95
ARGEE	27.73	5.59	5.69	42.22	5.25	5.26	0.72	0.02	0.02	239.47	3.50	3.49	-1031.68	141.34	140.99

Table 5.20: Extension 2 Scenario 38 (γ_1=-0.64, γ_2=-0.64, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.01	3.41	3.34	41.88	3.75	3.83	0.72	0.01	0.01	244.71	2.41	2.37	-1040.37	100.96	103.59
EGEE	27.63	10.48	11.26	41.96	5.49	5.62	0.72	0.02	0.02	217.83	7.16	7.91	-989.75	148.39	151.73
IEE	27.64	9.09	9.96	41.81	9.69	10.64	0.72	0.03	0.04	192.75	7.98	9.14	-954.24	262.30	285.22
CWGEE	27.89	5.65	5.76	41.96	4.59	4.69	0.72	0.02	0.02	194.70	5.33	5.67	-926.42	123.90	126.68
ARGEE	27.94	4.72	4.62	41.95	5.34	5.32	0.72	0.02	0.02	243.46	3.07	3.07	-1028.44	143.68	143.16

Table 5.21: Extension 2 Scenario 39 (γ_1=-0.83, γ_2=-0.34, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.95	3.35	3.46	42.02	3.76	3.80	0.72	0.01	0.01	243.08	2.37	2.52	-1041.38	101.11	102.20
EGEE	28.08	10.07	11.57	41.97	4.41	4.46	0.72	0.02	0.02	179.38	7.02	7.78	-916.18	119.18	119.76
IEE	28.11	6.54	7.80	42.04	6.44	7.39	0.72	0.02	0.03	150.23	5.86	7.16	-867.08	174.33	198.72
CWGEE	28.00	4.85	5.28	41.97	3.66	3.72	0.72	0.01	0.01	164.38	4.80	5.14	-874.61	98.94	100.06
ARGEE	27.88	5.85	5.85	41.86	5.32	5.29	0.72	0.02	0.02	236.20	3.67	3.54	-1022.83	143.24	141.64

Table 5.22: Extension 2 Scenario 40 (γ_2=-0.83, γ_1=-0.34, treatment effect = 28 g/week)

Bias

The simulation scenarios with ICS shows no significant bias in the estimated treatment effect for any method. This can be seen in Figures 5.5 and 5.6, which are boxplots by method of the 1000 estimates from the simulation scenarios in Tables 5.13 and 5.18 with a small and large standard deviation of cluster size, respectively. These boxplots show that under both situations there is no bias in the estimated treatment effect, with the EGEE and IEE having the largest spread. This agrees with the findings presented in Sections 4.4 and 5.1 with a smaller sample size. A similar pattern can be seen in the boxplots of the treatment effect estimates for the other ICS scenarios (see Appendix B.3 Figures B.27 : B.36).

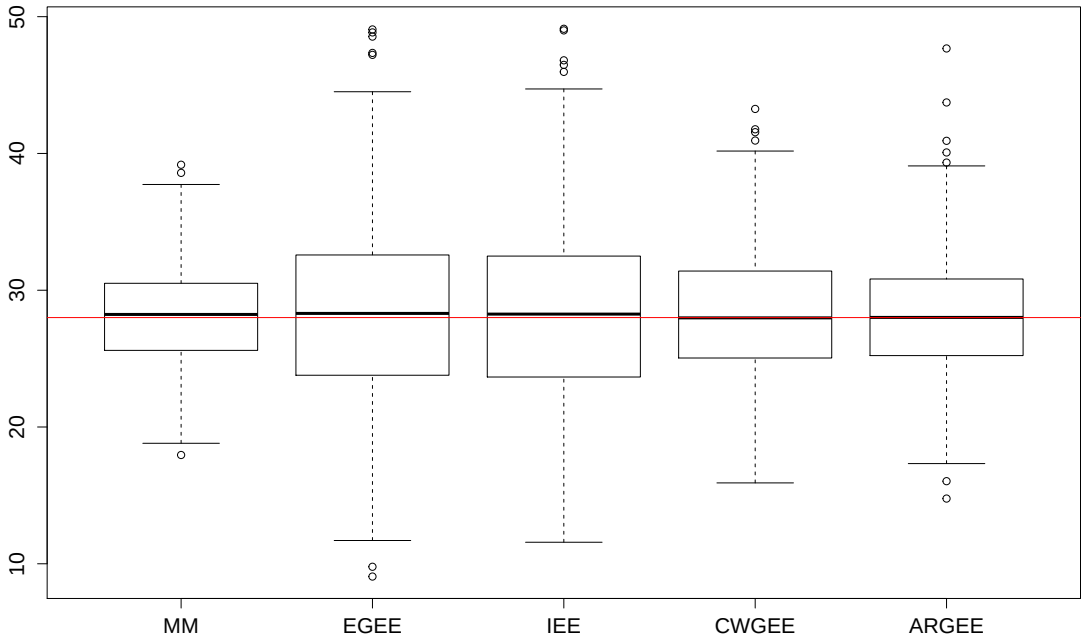

Figure 5.5: Extension 2: Boxplots of Interaction Effect estimates for Scenario 3 in Table 5.13 using smaller SD of cluster size

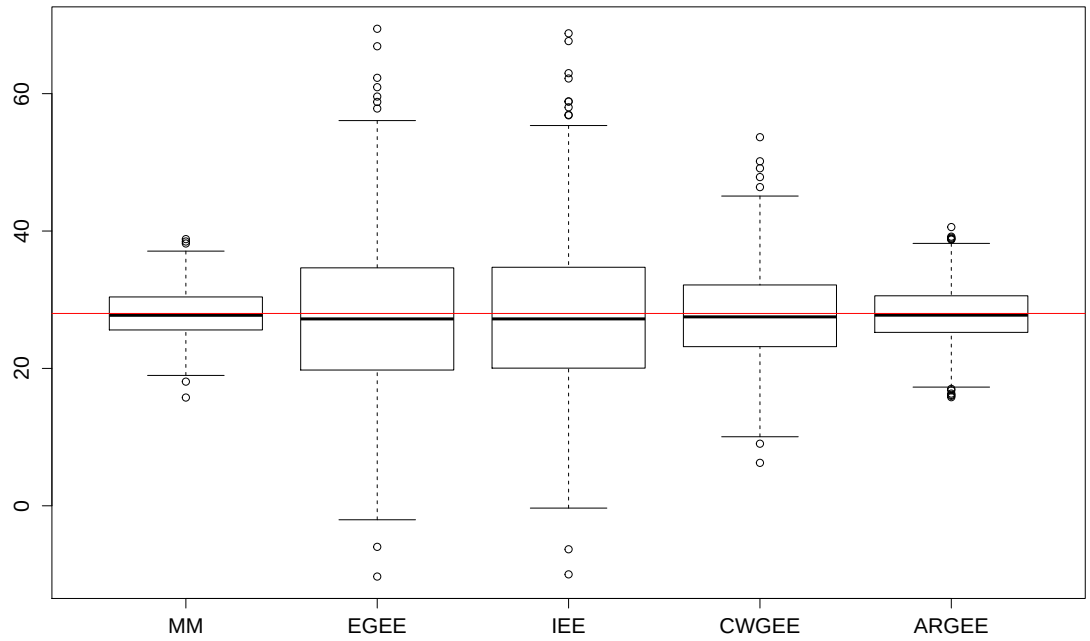

Figure 5.6: Extension 2: Boxplots of Interaction Effect estimates for Scenario 3 in Table 5.18 using larger SD of cluster size

Standard Errors of Interaction Term

The major difference seen in Extension 2 compared to the original simulation study and Extension 1, is that for the interaction effect, the model based standard error is now much closer to the standard deviation of the estimates.

This smaller difference can be seen in Figures 5.7 and 5.8. These plots show the combined data from Extension 2 for the ICS simulation scenarios using the smaller and larger standard SD of cluster size respectively. Each plot gives the ratio of SE for each of the 1000 simulated datasets divided by SD for each method. Both plots show that the MM has a ratio closest to 1 and the EGEE and IEE have ratios furthest from 1, consistent with the original simulation study and Extension 1 results. The ARGEE varies between the two, with a lower ratio in the smaller SD scenarios. Overall, the median ratios are much closer to 1 for all methods than those seen in Figures 4.8 and 5.2, which showed the equivalent results for the n=60 trial. Notice the variability appears to be larger in the smaller SD scenarios (Figure 5.7) compared with the larger SD scenarios (Figure 5.8). This is because of a few more outliers present mainly for the ARGEE. These large outliers for the ARGEE occured in a small number of seemingly random scenarios and this was not a pattern present across all scenarios. For example, the SE divided by SD ratios for a treatment effect of 14g/week resulted in similar outliers (See Appendix B.3 Figures B.49 and B.50). Overall, the variability for the ARGEE is smaller than in the original simulation study and extension 1. Separate plots for each simulation scenario showed similar results to the combined plot (see Appendix B.3 Figures B.39 : B.48)

The improvement in the SE can also be seen in Tables 5.23 and 5.24, which give the ratio of mean SE divided by SD, for the five smaller SD scenarios and the five larger SD scenarios respectively. Comparing Tables 5.23 and 5.24 to Tables 4.13 and 5.8 from the Chapter 4 simulation study and Extension 1, they shows that the results for the MM and ARGEE are similar to those seen previously. In contrast, there is significant improvement with ratios much closer to 1 for the EGEE, IEE and CWGEE methods. Similar to the Chapter 4 simulation study, the CWGEE performs the best out of these three methods.

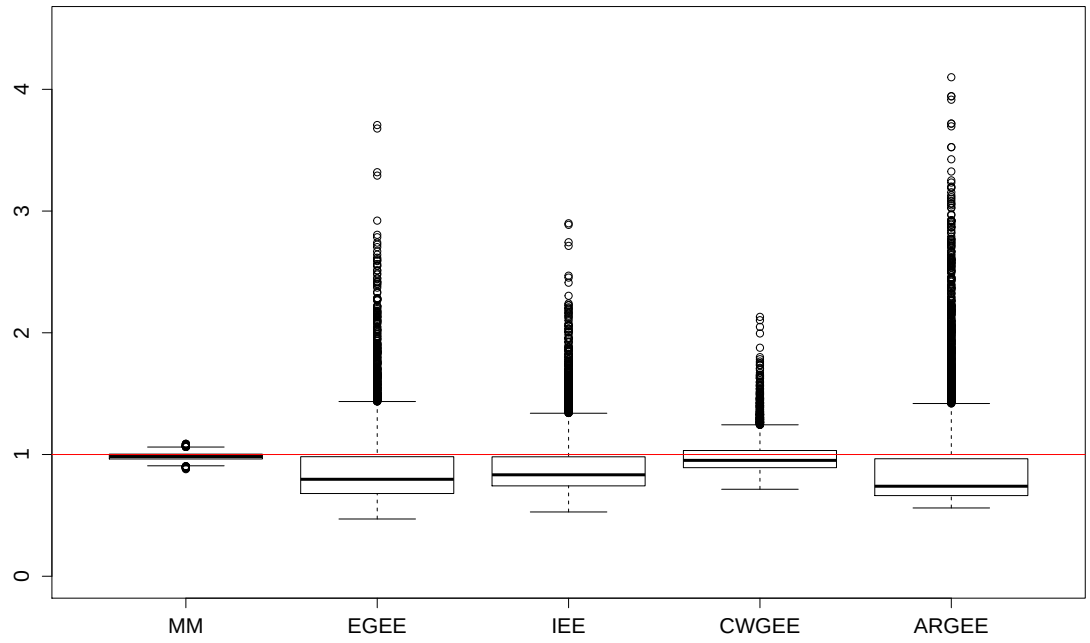

Figure 5.7: Extension 2: Boxplots of Interaction Effect standard error / standard deviation for all ICS data with smaller SD of cluster size

The EGEE and IEE methods generally performed best when the random intercept was highly correlated with duration (Scenarios 3 and 6). Notably, the MM still produces ratios closest to 1 in most scenarios. It is also worth noting that in Tables 5.13 to 5.22 the EGEE and IEE do result in larger SE and SD, whereas the CWGEE method gives a more efficient estimate in most scenarios.

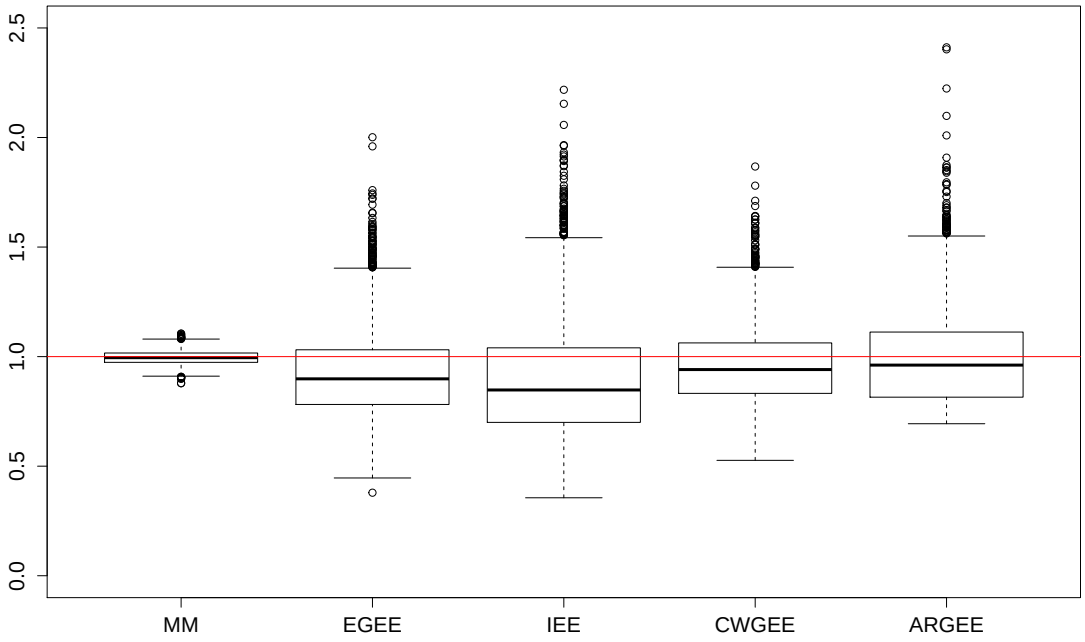

Figure 5.8: Extension 2: Boxplots of Interaction Effect standard error / standard deviation for all ICS data with larger SD of cluster size

Simulation Scenario	3	4	5	6	7
MM	0.97	0.99	1.02	0.96	0.98
EGEE	0.95	0.87	0.89	0.93	0.90
IEE	0.96	0.87	0.88	0.92	0.90
CWGEE	0.99	0.97	0.97	0.97	1
ARGEE	0.97	0.91	0.92	0.95	0.94

Table 5.23: Extension 2: Ratios of mean SE divided by SD using smaller SD of cluster size

Simulation Scenario	36	37	38	39	40
MM	1.01	1.00	0.98	1.02	0.98
EGEE	0.94	0.91	0.95	0.93	0.87
IEE	0.93	0.87	0.92	0.91	0.84
CWGEE	0.98	0.94	0.97	0.98	0.92
ARGEE	1.02	0.97	0.98	1.02	1

Table 5.24: Extension 2: Ratios of mean SE divided by SD using larger SD of cluster size

Coverage Probabilities

The coverage probabilities of the Wald type confidence interval for the interaction effect are given in Tables 5.25 and 5.26 for the original scenarios and larger standard deviation scenarios respectively. For all methods, the coverage probabilities are now closer to 0.95 than in the Chapter 4 simulation study and Extension 1, where the EGEE, IEE and CWGEE had coverage probabilities that were too low. Table 5.26 shows coverage probabilities that are much larger than those found in Extension 1, where undercoverage was present. For example, for the EGEE in scenario 37 the coverage probability increased from 0.594 to 0.908, which is a vast improvement. For the EGEE and IEE, the pattern of the coverage probabilities being higher when duration was highly correlated with the random intercept (Scenarios 3 and 6) can still be seen here, however the effect is not as great as in the Chapter 4 simulation study or Extension 1. It is worth noting that, although the coverage probabilities are quite high in this case and the ratio of SE to SD is close to one, there is substantial variability in the SE's for the EGEE, IEE, CWGEE and ARGEE methods.

Simulation Scenario	3	4	5	6	7
MM	0.949	0.95	0.958	0.942	0.95
EGEE	0.937	0.926	0.939	0.938	0.922
IEE	0.941	0.924	0.919	0.93	0.921
CWGEE	0.955	0.954	0.94	0.929	0.931
ARGEE	0.944	0.966	0.954	0.95	0.958

Table 5.25: Extension 2: Coverage probabilities for smaller SD of cluster size

Simulation Scenario	36	37	38	39	40
MM	0.96	0.943	0.943	0.954	0.932
EGEE	0.927	0.908	0.912	0.918	0.899
IEE	0.925	0.893	0.908	0.914	0.879
CWGEE	0.944	0.933	0.922	0.945	0.919
ARGEE	0.962	0.948	0.944	0.962	0.946

Table 5.26: Extension 2: Coverage probabilities for larger SD of cluster size

Summary

In summary, when the sample size for the trial was increased from 60 to 600, the MM and ARGEE performed well, as they did in the Chapter 4 simulation study. The EGEE, IEE and CWGEE also performed well with small differences between SE and SD and coverage probabilities close to the nominal level. It could be concluded that the size of the study was insufficient in the Chapter 4 simulation study and a larger study size is needed for the EGEE, IEE and CWGEE to perform well.

The results for the simulation scenarios using the other treatment effects of 0, 7, 14 and 21 g/week gave similar results to those discussed above. For the ICS scenarios, the value of the treatment effect estimate changed accordingly but the analysis concerning bias, estimates of SEs and coverage probabilities were unaffected. For the NICS scenarios the value of the treatment effect estimate changed and the SE and SD also changed slightly. Depending on the scenario, SE was sometimes larger than SD across all methods, although this was only by a small amount (less than 0.05%). In all cases, SE was very close to SD, as expected. Full results tables for these scenarios can be found in Appendix C Figure C.11 : C.60.

5.3 Extension 3: Simulating from a GEE Framework

In the Chapter 4 simulation study the MM performed well in all situations, including when ICS was present. Part of the explanation may be that the MM was used to simulate

the data, hence it was able to effectively analyse it. That is, the poor relative performance of the GEE methods may be due, in part, to the model misspecification that is not present for MM. To investigate this possible bias, Extension 3 was conducted, where data were simulated using a GEE framework. This extension was conducted on both the original sample size and the larger sample size in Extension 2. The method of simulation is described Section 5.3.1.

5.3.1 Generating Simulated Data from a GEE Framework

To simulate the infant weights using a GEE framework, a multivariate normal distribution was used, where the dependence in the data was taken into account via the covariance structure. A set of weights was generated, using `mvrnorm` in R, separately for each infant. To simulate the data, it is necessary to specify the mean vector and the covariance matrix.

The mean vectors are given by model 5.1. This is the same as the fixed effects specified in the MM.

$$\begin{aligned} Weight =& \beta_0 + \beta_1 Treatment : Time + \beta_2 Birthweight + \beta_3 GestationalAge \\ & \beta_4 Time \end{aligned} \tag{5.1}$$

The covariance matrix is the same for each infant and is given by the formula:

$$\text{Covariance} = (\text{diag}[V(1), ..., V(N])^{1/2} R (\text{diag}[V(1), ..., V(N])^{1/2},$$

where R is the working correlation structure and $V(i)$ is the variance for time i.

It is expected that as the time between outcomes increases, the correlation between those outcomes should decrease, hence an autoregressive structure was chosen to generate the data. The corresponding correlation matrix has the form:

$$R_i = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \dots & & \rho^{n-1} \\ \rho & 1 & & & & & \rho^{n-2} \\ \rho^2 & & \ddots & & & & \rho^{n-3} \\ \vdots & & & & & & \vdots \\ & & & & & 1 & \rho \\ \rho^{n-1} & \dots & & & & \rho & 1 \end{bmatrix}$$

Model 5.1 was fit to the singleton infants from the POPPET trial, using an autoregressive GEE, and the estimates of ρ and V were obtained. Only singleton infants were considered as only one level of clustering (longitudinal) will be considered in the simulation study. The value for ρ was set to 0.984 and the value for all V was set to the constant value of 24224.

A total of 180 weights were generated for each infant in the first instance. The cluster size was then simulated for each infant, and the superfluous weights discarded. As the aim here is to investigate methods when informative cluster size is present, a method of simulating cluster size in such a way that it induces informative cluster size is needed.

In the Chapter 4 simulation study, recall that the cluster size was determined from the random effects Poisson distribution Poisson$(e^{\gamma_0+\gamma_1 b_0+\gamma_2 b_1}) + N_{min}$, where b_0 is the random intercept, b_1 is the random slope and N_{min} is the minimum cluster size (Neuhaus and Mcculloch, 2011). It is desirable to use a comparable method here, so that the two data generation approaches are as similar as possible.

For the MM the assumption of non informative cluster size means that the random effects are independent of the cluster size. Hence, Neuhaus and Mcculloch's method induces informative cluster size by making the cluster size depend on the random effects. As there are no random effects in a GEE, to induce ICS we need the cluster size to depend on the outcome directly. To do this using Neuhaus and Mcculloch's method, the random effects

need to be replaced by terms based on the infants weights.

To match the original model for cluster size, an intercept term and a slope term are needed. The random intercept was replaced with the infant's weight on first day of the study (denoted W_1). The random slope was replaced with the infant's weight on day 14 minus the infants weight on the first day (denoted D). This term was used as the difference gives a measure of growth, which makes sense to replace a random slope term. Day 14 was used as this was the minimum duration in POPPET. W_1 and D are marginally normally distributed and are also jointly normal, as required for the method to work. The two weight terms were standardised for simplicity of calculation.

The cluster size was generated from the Poisson distribution
Poisson$(e^{\gamma_0+\gamma_1 W_1+\gamma_2 D}) + N_{min}$. As in the original simulation study, we want the simulated cluster size distribution to be the same as that for the POPPET trial. Hence, we will find the parameters $(\gamma_0, \gamma_1, \gamma_2)$ needed to give the same expected value and variance as the observed duration distribution for POPPET. This is done by repeating the calculation from Section 4.2.2.

The expected value of the cluster size is given by

$$E[N] = e^{\gamma_0+\frac{1}{2}(\gamma_1^2\sigma_{W_1}^2+\gamma_2^2\sigma_D^2+2\gamma_1\gamma_2\sigma_{W_1,D})} + N_{\min}. \tag{5.2}$$

Now assume,

$$\begin{pmatrix} W_1 \\ D \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.439 \\ 0.439 & 1 \end{pmatrix}\right).$$

Substituting these values into equation 5.2 and setting the expectation equal to the sample mean duration from POPPET gives

$$
\begin{aligned}
38.36667 &= e^{\gamma_0+\frac{1}{2}(\gamma_1^2+\gamma_2^2+2*0.439\gamma_1\gamma_2)} + 14 \\
\ln(24.36667) &= \gamma_0 + \frac{1}{2}(\gamma_1^2+\gamma_2^2+2*0.439\gamma_1\gamma_2) \\
2(\ln(24.36667)-\gamma_0) &= \gamma_1^2+\gamma_2^2+2*0.439\gamma_1\gamma_2 \qquad (5.3)
\end{aligned}
$$

and for the variance:

$$
\begin{aligned}
191.5921 &= e^{\gamma_0+\frac{1}{2}(\gamma_1^2\sigma_{0,0}^2+\gamma_2^2\sigma_{1,1}^2+2\gamma_1\gamma_2\sigma_{0,1})} + e^{2\gamma_0+2\gamma_1^2\sigma_{0,0}^2+2\gamma_2^2\sigma_{1,1}^2+4\gamma_1\gamma_2\sigma_{0,1}} \\
&\quad - e^{2\gamma_0+\gamma_1^2\sigma_{0,0}^2+\gamma_2^2\sigma_{1,1}^2+2\gamma_1\gamma_2\sigma_{0,1}} \\
191.5921 &= e^{\gamma_0+\frac{1}{2}(\gamma_1^2+\gamma_2^2+2*0.439\gamma_1\gamma_2)} + e^{2\gamma_0+2\gamma_1^2+2\gamma_2^2+4*0.439\gamma_1\gamma_2)} - e^{2\gamma_0+\gamma_1^2+\gamma_2^2+2*0.439\gamma_1\gamma_2)}.
\end{aligned}
$$

Now substituting in Equation 5.3 gives

$$
\begin{aligned}
191.59217 &= e^{\gamma_0+\ln(24.36667)-\gamma_0} + e^{2\gamma_0+4\ln(24.36667)-4\gamma_0} - e^{2\gamma_0+2\ln(24.36667)-2\gamma_0} \\
191.5921 &= 24.36667 + e^{-2\gamma_0+4\ln(24.36667)} - 24.36667^2 \\
\ln(760.96) &= -2\gamma_0 + 4\ln(24.36667) \\
\gamma_0 &= \frac{4\ln(24.36667)-\ln(760.96)}{2} \\
&= 3.069142.
\end{aligned}
$$

Substituting this value for γ_0 into Equation 5.3 gives

$$
\begin{aligned}
2(\ln(24.36667)-3.069142) &= \gamma_1^2+\gamma_2^2+2*0.439\gamma_1\gamma_2 \\
0.248 &= \gamma_1^2+\gamma_2^2+0.878\gamma_1\gamma_2.
\end{aligned}
$$

This is an ellipse with equation $0.248 = \gamma_1^2 + \gamma_2^2 + 0.878\gamma_1\gamma_2$, as opposed to the equation of a circle obtained when simulating under a MM. Notice the value of γ_0 is the same as in the original method.

To investigate whether ICS has been induced and how changing γ_1 and γ_2 affects the strength of the correlation, a MM was fit to the generated data and the correlation between the random effects and cluster size was found for different values for γ_1 and γ_2. Through experimentation it was determined that choosing the same values for γ_1 and γ_2 as in the Chapter 4 simulation study did not result in the same strength of correlation. Nonetheless, the same general pattern was present, in that increasing the coefficient of the intercept term (γ_1) increased the correlation between the random intercept and duration. Likewise, increasing the coefficient of the slope term (γ_2) increased the correlation between the random slope and duration. However, it did not occur to the same extent as in the original study. Setting the coefficient of one term to zero did not result in the maximum correlation strength between duration and the other term. The magnitude of the correlations from the original study could also not be reached. Through trial and error it was found that it was possible to obtain nearly zero correlation between one random effect and duration and achieve a maximum correlation between duration and the other random effect by selecting different combinations of values for γ_1 and γ_2.

To keep the extension as similar as possible to the Chapter 4 simulation study it was determined that the γ values would be chosen to give the same 5 correlation patterns as considered in the original study (i.e. correlation between duration and intercept only, correlation between duration and slope only, equal correlation between duration and slope/intercept, more correlation between duration and intercept than duration and slope, and more correlation between duration and slope than duration and intercept). Through trial and error, the γ values were chose to give correlations as close to the original cases as could be achieved, as listed in Table 5.27.

These five ICS cases were considered for the five treatment effects (0, 7, 14, 21 and 28 g/week), resulting in 25 simulation scenarios. There were also 2 NICS simulation sce-

Cases	γ_1	γ_2	$cor(b_0, N)$	$cor(b_1, N)$
1	-0.36	-0.22	-0.05	-0.30
2	-0.039	-0.48	-0.32	-0.003
3	-0.24	-0.35	-0.172	-0.206
4	-0.29	-0.29	-0.23	-0.143
5	-0.17	-0.40	-0.107	-0.262

Table 5.27: Extension 3: Values of γ_1 and γ_2 and resulting correlations

narios considered for the five treatment effects, a fixed trial length and a non fixed trial length that was unrelated to the outcome, resulting in an additional 10 simulation scenarios. These NICS simulation scenarios used the same parameters as the NICS examples from the original study. For each scenario, data were generated for 60 infants to match the original simulation study with 10000 datasets generated per scenario. In addition, Extension 2 was replicated to investigate the effect of increasing the sample size to 600 infants. Again only 1000 simulations were performed for each of these scenarios due to the time it takes to run a simulation scenario with a larger sample size.

5.3.2 NICS Scenario Results for Extension 3

Consider the results of the Extension 3, using the sample sizes of 60 and 600. The results for the NICS simulation scenarios with a treatment effect of 28 grams a week are given in Tables 5.28 - 5.31.

Scenarios for smaller sample size, n=60

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.91	6.72	6.83	42.16	17.06	17.66	0.72	0.06	0.06	244.46	5.19	5.28	-1049.11	458.36	473.81
EGEE	27.93	7.63	7.80	42.17	16.54	17.52	0.72	0.06	0.06	244.44	5.43	5.59	-1049.28	445.40	470.03
IEE	27.85	9.82	10.13	42.16	16.38	17.54	0.72	0.06	0.06	244.49	6.31	6.50	-1049.21	441.13	470.31
CWGEE	27.85	9.82	10.13	42.16	16.38	17.54	0.72	0.06	0.06	244.49	6.31	6.50	-1049.21	441.13	470.31
ARGEE	27.90	6.09	6.27	42.07	15.89	17.00	0.72	0.06	0.06	244.49	4.66	4.83	-1046.82	427.98	455.81

Table 5.28: Extension 3 Scenario 1 (fixed trial length=38, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.01	6.75	6.95	42.10	17.08	17.66	0.72	0.06	0.06	244.42	5.21	5.33	-1045.18	458.87	474.37
EGEE	27.92	7.70	7.96	42.05	16.52	17.49	0.72	0.06	0.06	244.49	5.46	5.68	-1043.97	445.18	469.52
IEE	28.08	9.59	10.11	42.06	16.38	17.59	0.72	0.06	0.06	244.45	6.48	6.77	-1044.28	441.33	472.32
CWGEE	28.10	9.59	10.05	42.05	16.34	17.50	0.72	0.06	0.06	244.44	6.48	6.73	-1044.07	440.31	469.71
ARGEE	28.03	6.03	6.25	42.06	15.86	16.88	0.72	0.06	0.06	244.44	4.59	4.76	-1043.86	427.35	452.62

Table 5.29: Extension 3 Scenario 2 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 28 g/week)

Scenarios for larger sample size, n=600

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.95	2.14	2.08	41.61	5.32	5.36	0.72	0.02	0.02	244.50	1.65	1.66	-1033.02	143.16	143.73
EGEE	27.94	2.46	2.40	41.61	5.30	5.36	0.72	0.02	0.02	244.51	1.76	1.76	-1033.13	142.63	143.54
IEE	27.98	3.19	3.13	41.61	5.29	5.36	0.72	0.02	0.02	244.48	2.03	2.03	-1033.01	142.50	143.57
CWGEE	27.98	3.19	3.13	41.61	5.29	5.36	0.72	0.02	0.02	244.48	2.03	2.03	-1033.01	142.50	143.57
ARGEE	27.97	1.97	1.93	41.61	5.12	5.21	0.72	0.02	0.02	244.48	1.51	1.52	-1032.77	137.82	139.44

Table 5.30: Extension 3 (Larger Trial) Scenario 1 (fixed trial length=38, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.87	2.16	2.14	42.09	5.33	5.41	0.72	0.02	0.02	244.60	1.66	1.70	-1046.82	143.46	146.04
EGEE	27.94	2.50	2.50	42.08	5.29	5.37	0.72	0.02	0.02	244.56	1.78	1.82	-1046.76	142.51	145.12
IEE	27.86	3.14	3.17	42.09	5.31	5.32	0.72	0.02	0.02	244.59	2.12	2.15	-1047.04	142.96	143.90
CWGEE	27.85	3.12	3.15	42.08	5.29	5.37	0.72	0.02	0.02	244.60	2.10	2.13	-1046.71	142.38	145.13
ARGEE	27.92	1.95	1.91	42.09	5.11	5.20	0.72	0.02	0.02	244.55	1.49	1.51	-1046.76	137.58	140.48

Table 5.31: Extension 3 (Larger Trial) Scenario 2 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 28 g/week)

Bias

For the four simulation scenarios with NICS, the results are similar to the Chapter 4 simulation study. There is very little bias in the estimated treatment effect, which is shown in boxplots by method of the 10000 estimates from the simulation scenario in Table 5.29 (see Figure 5.9). The plot shows the median treatment effect estimates are very close to the true value shown in red. There is some variation between methods in the spread of the data, with the IEE and CWGEE having the largest spread. Overall,

the spread is larger than what was found for the NICS cases in the MM based simulation scenarios. A similar pattern was seen in the boxplot of the other NICS scenario, with the only difference being that the plots using the larger sample size had a smaller spread of the data using all methods. This is expected when increasing the size of the sample (see Appendix B.4 Figures B.51 : B.53).

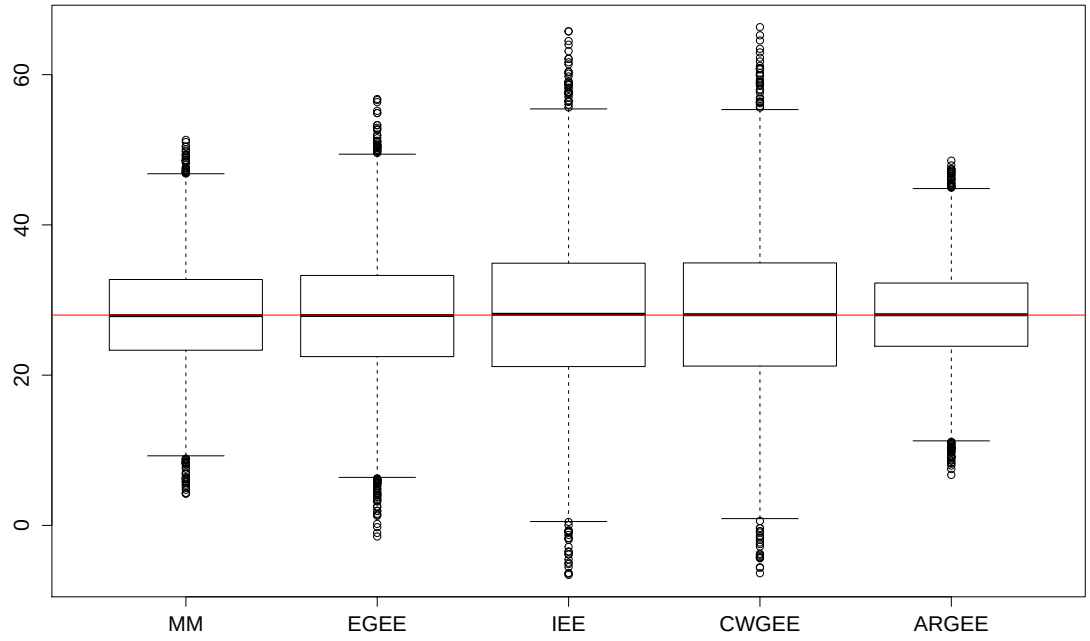

Figure 5.9: Extension 3: Boxplots of Interaction Effect estimates for Scenario 1 in Table 5.28 for n=60

Standard Errors of Interaction Term

For the interaction effect, the standard error of the estimates and the model based standard error were similar for the NICS scenarios. Considering n=60 first, Figure 5.10 gives the ratio of the standard error for each of the 10000 simulated datasets divided by the standard deviation for the simulation scenario in Table 5.29. This shows that for all the methods the standard error is generally being underestimated, but not by much. A similar pattern is seen in the other NICS scenario (see Appendix B.4 Figure B.66). Now considering n=600 results, Figure 5.11 gives the ratio of standard error for each of the

1000 simulated datasets divided by standard deviation for the simulation scenario in Table 5.31. This shows that the standard error is being overestimated by some methods and very slightly underestimated by others, which is different to the smaller trial where all methods were underestimating the SE. Overestimating the standard error is better than underestimating it because this is more conservative, so this would be preferable. Overall, the standard error is similar to the standard deviation as the median ratio is close to 1. A similar pattern was seen in the plot for the other NICS scenario except here all the methods were overestimating the standard error (see Appendix B.4 Figure B.72).

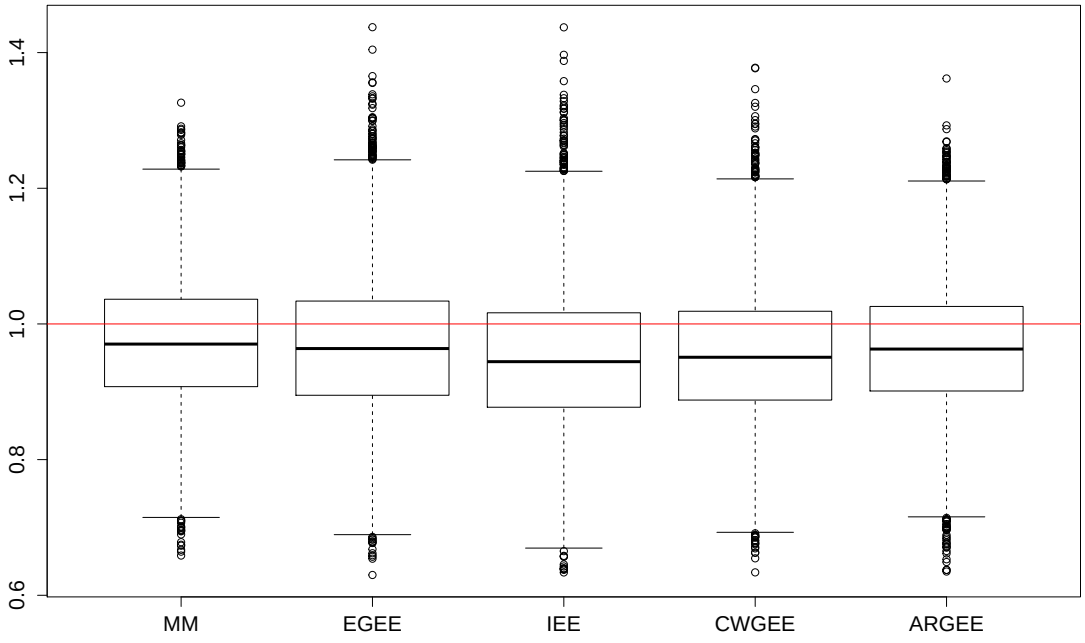

Figure 5.10: Extension 3: Boxplots of Interaction Effect standard error / standard deviation for Scenario 1 in Table 5.28 for n=60

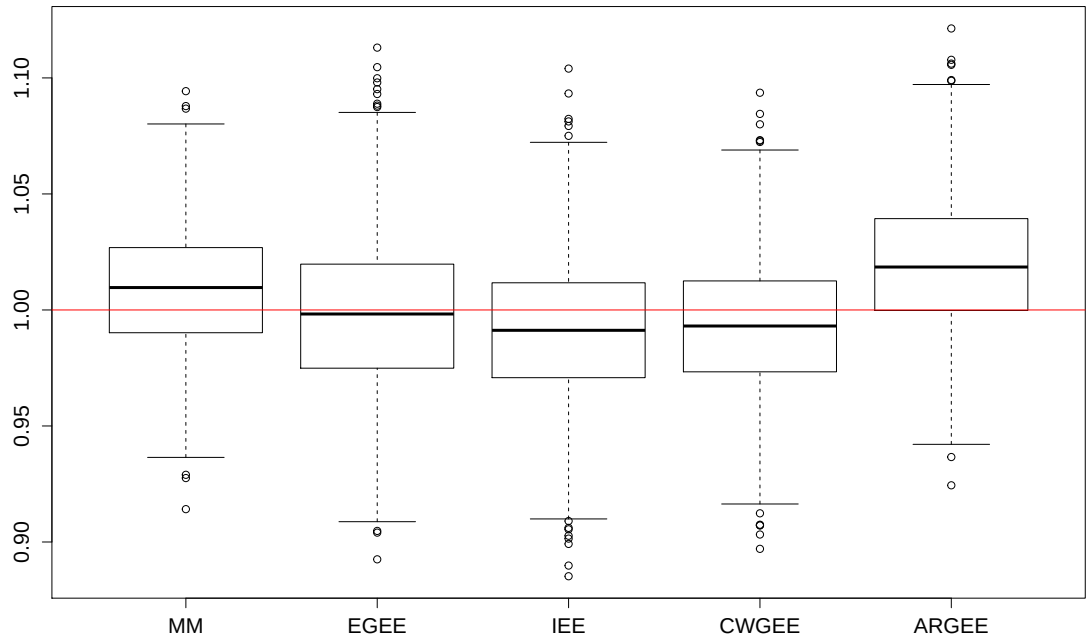

Figure 5.11: Extension 3: Boxplots of Interaction Effect standard error / standard deviation for Scenario 1 in Table 5.30for n=600

Coverage Probabilities

The coverage probabilities of the Wald type 95% confidence interval for the interaction effect in NICS scenarios are given in Tables 5.32 and 5.33 for the smaller and larger sample size respectively. For n=60, the coverage probabilities are all quite close to the expected value of 0.95. For n=600, the coverage probabilities are all extremely close to 0.95 and for the scenarios with equal follow up are slightly larger than 0.95.

Simulation Scenario	1	2
MM	0.937	0.945
EGEE	0.938	0.944
IEE	0.933	0.941
CWGEE	0.935	0.941
ARGEE	0.936	0.940

Table 5.32: Extension 3: Coverage probabilities for NICS scenarios for n=60

Simulation Scenario	1	2
MM	0.949	0.952
EGEE	0.944	0.955
IEE	0.949	0.963
CWGEE	0.944	0.963
ARGEE	0.947	0.953

Table 5.33: Extension 3: Coverage probabilities for NICS scenarios for n=600

Summary

In summary, for the GEE simulation scenarios with NICS there is very little bias in the estimated treatment effect, very little difference between the SE and SD, and correct coverage probabilities for the 95% confidence interval.

5.3.3 ICS Scenario Results for Extension 3

The results for the ICS simulation scenarios with a treatment effect of 28 grams a week are given in Tables 5.34 - 5.43.

Scenarios for smaller sample size, n=60

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.97	7.22	7.35	42.30	17.31	17.72	0.72	0.06	0.06	247.43	5.51	5.43	-1059.86	465.39	476.42
EGEE	27.73	8.16	8.74	42.15	16.57	17.39	0.72	0.06	0.06	247.08	5.63	6.07	-1052.48	447.18	467.80
IEE	27.06	9.21	9.91	41.46	16.11	17.31	0.71	0.06	0.06	235.66	6.42	6.89	-992.90	435.05	465.98
CWGEE	26.95	9.27	9.62	41.18	15.96	16.98	0.71	0.06	0.06	234.59	6.46	6.73	-966.19	431.06	458.11
ARGEE	27.85	6.24	6.58	42.09	15.84	16.76	0.72	0.06	0.06	244.83	4.58	4.82	-1044.01	427.18	450.76

Table 5.34: Extension 3 Scenario 3 (γ_1=-0.36, γ_2=-0.22, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.41	7.38	7.46	41.56	17.36	17.76	0.72	0.06	0.06	251.27	5.61	5.35	-1044.47	466.43	476.92
EGEE	26.88	7.93	8.71	41.49	16.47	17.35	0.72	0.06	0.06	244.40	5.29	5.97	-1029.41	443.75	465.95
IEE	26.85	8.93	9.77	41.42	16.35	17.80	0.72	0.06	0.06	239.48	5.92	6.84	-1017.96	440.61	478.07
CWGEE	27.01	9.11	9.52	41.48	16.19	17.29	0.72	0.06	0.06	240.46	6.22	6.64	-1017.16	436.24	464.67
ARGEE	27.75	6.02	6.31	41.53	15.77	16.90	0.72	0.06	0.06	244.78	4.26	4.51	-1030.76	424.85	453.82

Table 5.35: Extension 3 Scenario 4 (γ_1=-0.039, γ_2= -0.48, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.35	7.31	7.43	42.14	17.34	17.62	0.72	0.06	0.06	249.35	5.57	5.45	-1058.68	465.70	473.99
EGEE	27.32	8.12	8.80	41.91	16.52	17.15	0.72	0.06	0.06	245.76	5.52	5.99	-1040.85	445.36	461.39
IEE	26.92	9.28	10.03	41.65	16.20	17.27	0.71	0.06	0.06	236.53	6.28	6.85	-1006.76	436.83	465.11
CWGEE	26.91	9.37	9.80	41.48	16.07	16.92	0.71	0.06	0.06	236.63	6.43	6.75	-991.19	433.41	456.25
ARGEE	27.85	6.22	6.51	41.92	15.85	16.63	0.72	0.06	0.06	244.81	4.49	4.70	-1039.25	427.16	447.40

Table 5.36: Extension 3 Scenario 5 (γ_1=-0.24, γ_2=-0.35, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.35	7.26	7.42	42.15	17.36	18.13	0.72	0.06	0.07	248.52	5.54	5.46	-1059.10	466.36	485.34
EGEE	27.58	8.14	8.67	41.95	16.58	17.66	0.72	0.06	0.06	246.31	5.55	6.04	-1044.51	447.12	473.39
IEE	27.05	9.29	10.07	41.50	16.18	17.70	0.71	0.06	0.06	235.99	6.33	6.91	-999.05	436.72	474.84
CWGEE	27.02	9.36	9.85	41.30	16.07	17.34	0.71	0.06	0.06	235.62	6.44	6.81	-979.48	433.61	466.07
ARGEE	27.98	6.35	6.73	42.03	16.36	17.91	0.72	0.06	0.07	244.78	4.61	4.93	-1043.12	438.62	473.22

Table 5.37: Extension 3 Scenario 6 (γ_1=-0.29, γ_2=-0.29, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.49	7.33	7.60	41.91	17.40	18.19	0.72	0.06	0.06	250.20	5.58	5.42	-1054.49	467.29	488.43
EGEE	27.13	8.05	8.66	41.78	16.55	17.74	0.72	0.06	0.06	245.30	5.44	5.95	-1037.33	446.12	476.39
IEE	26.86	9.19	10.02	41.51	16.28	17.92	0.72	0.06	0.06	237.54	6.16	6.87	-1009.83	439.02	481.29
CWGEE	26.95	9.31	9.84	41.54	16.17	17.60	0.72	0.06	0.06	238.02	6.37	6.74	-1003.32	436.17	473.10
ARGEE	27.88	6.15	6.51	41.85	15.85	17.21	0.72	0.06	0.06	244.83	4.40	4.67	-1038.71	427.26	461.93

Table 5.38: Extension 3 Scenario 7 (γ_1=-0.17, γ_2=-0.40, treatment effect = 28 g/week)

Scenarios for larger sample size, n=600

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.28	2.30	2.39	42.34	5.41	5.34	0.72	0.02	0.02	247.37	1.75	1.78	-1061.01	145.66	144.16
EGEE	27.86	2.76	2.77	42.20	5.32	5.27	0.72	0.02	0.02	247.20	1.92	1.93	-1054.28	143.30	142.14
IEE	27.25	3.09	3.15	41.45	5.29	5.29	0.71	0.02	0.02	234.78	2.18	2.21	-990.46	142.48	143.47
CWGEE	27.21	3.02	3.08	41.21	5.18	5.14	0.71	0.02	0.02	233.85	2.11	2.16	-964.12	139.60	139.17
ARGEE	28.03	1.98	2.06	42.15	5.13	5.15	0.72	0.02	0.02	244.74	1.46	1.56	-1047.13	138.03	138.61

Table 5.39: Extension 3 (Larger Trial) Scenario 3 (γ_1=-0.36, γ_2=-0.22, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.63	2.36	2.42	42.12	5.43	5.23	0.72	0.02	0.02	251.33	1.79	1.70	-1059.33	146.17	140.33
EGEE	27.00	2.82	2.95	42.12	5.29	5.20	0.72	0.02	0.02	244.38	1.91	1.99	-1045.74	142.50	139.50
IEE	26.72	3.06	3.01	42.14	5.42	5.32	0.72	0.02	0.02	239.11	2.16	2.26	-1035.55	146.04	142.86
CWGEE	26.97	2.97	2.92	42.09	5.27	5.18	0.72	0.02	0.02	240.12	2.06	2.08	-1031.95	141.77	138.93
ARGEE	28.04	1.93	1.92	42.19	5.12	5.07	0.72	0.02	0.02	244.63	1.37	1.40	-1047.84	137.71	135.68

Table 5.40: Extension 3 (Larger Trial) Scenario 4 (γ_1=-0.039, γ_2= -0.48, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.38	2.33	2.34	42.11	5.42	5.39	0.72	0.02	0.02	249.39	1.77	1.73	-1059.20	145.94	145.26
EGEE	27.34	2.77	2.81	41.85	5.30	5.29	0.72	0.02	0.02	245.87	1.91	1.91	-1040.42	142.68	142.61
IEE	26.79	3.12	3.17	41.43	5.32	5.37	0.71	0.02	0.02	235.91	2.17	2.25	-1000.23	143.38	144.75
CWGEE	26.81	3.05	3.10	41.39	5.21	5.22	0.71	0.02	0.02	236.13	2.11	2.17	-987.87	140.37	141.11
ARGEE	27.99	1.98	1.95	41.95	5.13	5.12	0.72	0.02	0.02	244.68	1.44	1.47	-1042.33	138.00	137.53

Table 5.41: Extension 3 (Larger Trial) Scenario 5 (γ_1=-024, γ_2=-0.35, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.37	2.32	2.24	42.01	5.43	5.41	0.72	0.02	0.02	248.58	1.76	1.73	-1055.85	145.91	146.91
EGEE	27.55	2.77	2.70	41.80	5.32	5.33	0.72	0.02	0.02	246.45	1.92	1.92	-1041.24	143.18	144.82
IEE	27.03	3.12	3.12	41.23	5.32	5.35	0.71	0.02	0.02	235.10	2.17	2.13	-990.10	143.19	145.40
CWGEE	27.05	3.04	3.04	41.12	5.20	5.21	0.71	0.02	0.02	234.89	2.11	2.07	-972.48	140.20	142.02
ARGEE	28.07	1.98	1.92	41.83	5.14	5.17	0.72	0.02	0.02	244.73	1.45	1.50	-1039.43	138.23	139.98

Table 5.42: Extension 3 (Larger Trial) Scenario 6 (γ_1=-0.29, γ_2=-0.29, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.34	2.35	2.32	42.16	5.43	5.35	0.72	0.02	0.02	250.32	1.78	1.67	-1062.12	146.10	143.75
EGEE	26.89	2.78	2.71	41.96	5.31	5.26	0.72	0.02	0.02	245.40	1.90	1.87	-1042.96	143.02	141.35
IEE	26.78	3.11	3.09	41.70	5.36	5.30	0.72	0.02	0.02	236.89	2.17	2.23	-1014.62	144.35	142.43
CWGEE	26.85	3.04	2.98	41.70	5.25	5.21	0.72	0.02	0.02	237.52	2.10	2.09	-1006.69	141.45	139.79
ARGEE	27.96	1.96	1.99	42.03	5.14	5.06	0.72	0.02	0.02	244.71	1.42	1.41	-1044.76	138.35	135.79

Table 5.43: Extension 3 (Larger Trial) Scenario 7 (γ_1=-0.17, γ_2=-0.40, treatment effect = 28 g/week)

Bias

Examining the scenarios with ICS, for both sample sizes there appears to be a small amount of bias present in the estimated treatment effects. This is different to all of the results considered thus far, where very little bias was ever present. This bias can be seen in Figures 5.12 and 5.13, which are boxplots by method of the 10000 and 1000 estimates from the simulation scenarios in Table 5.34 and 5.39 respectively. It should be noted that the bias appears relatively large in Figure 5.13 because of the reduced SD arising for n=600. In all cases, the bias is small relative to the variability of the estimates. In both plots, there is a small amount of bias present (about 1 gram difference between the median estimate and the true value for some methods), with varying levels depending on the method. The ARGEE has virtually no bias in either case, the EGEE, IEE and CWGEE have the most bias and the MM has minor bias which is larger for n=600. A similar pattern was present in the other simulation scenarios for both trial sizes, where the larger trial size has consistently smaller spread in all cases (see Appendix B.4 Figure B.55 : B.64).

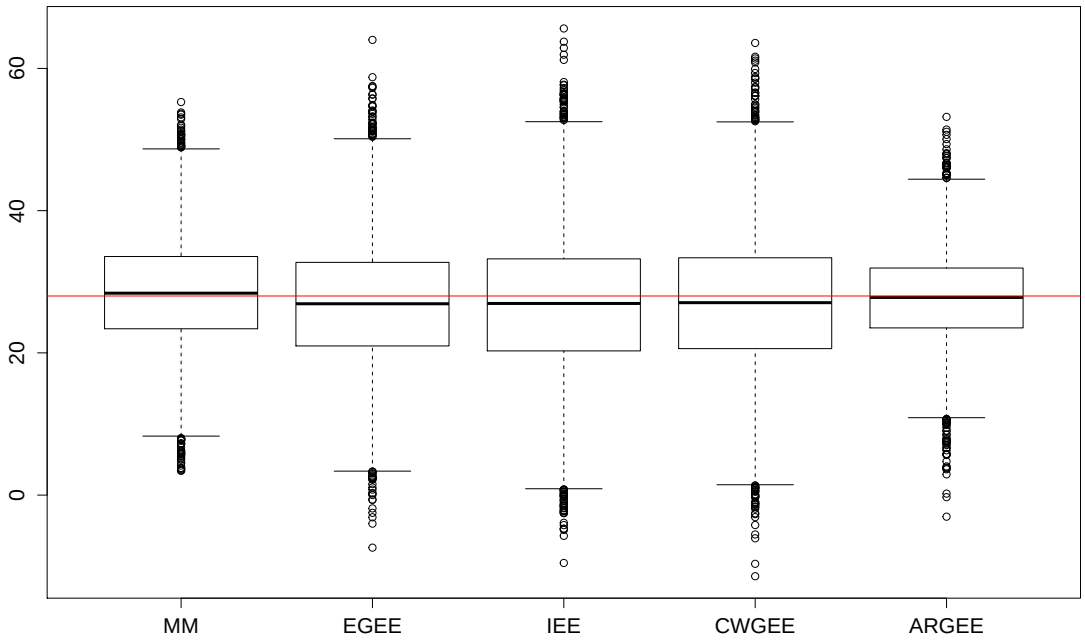

Figure 5.12: Extension 3: Boxplots of Interaction Effect estimates for Scenario 3 in Table 5.34 for n=60

The amount of bias the methods produced differed between simulation scenarios. This relationship can be seen in Tables 5.44 and 5.45, which give the difference between the mean estimated treatment value and the true value for each method, for n=60 and n=600 respectively. In both tables for the MM, EGEE and IEE, the largest bias is generally found in simulation scenarios 4 and 7, when the random slope was highly correlated with duration. It should be noted that compared to the size of the true parameter (28) this bias is not large for any method.

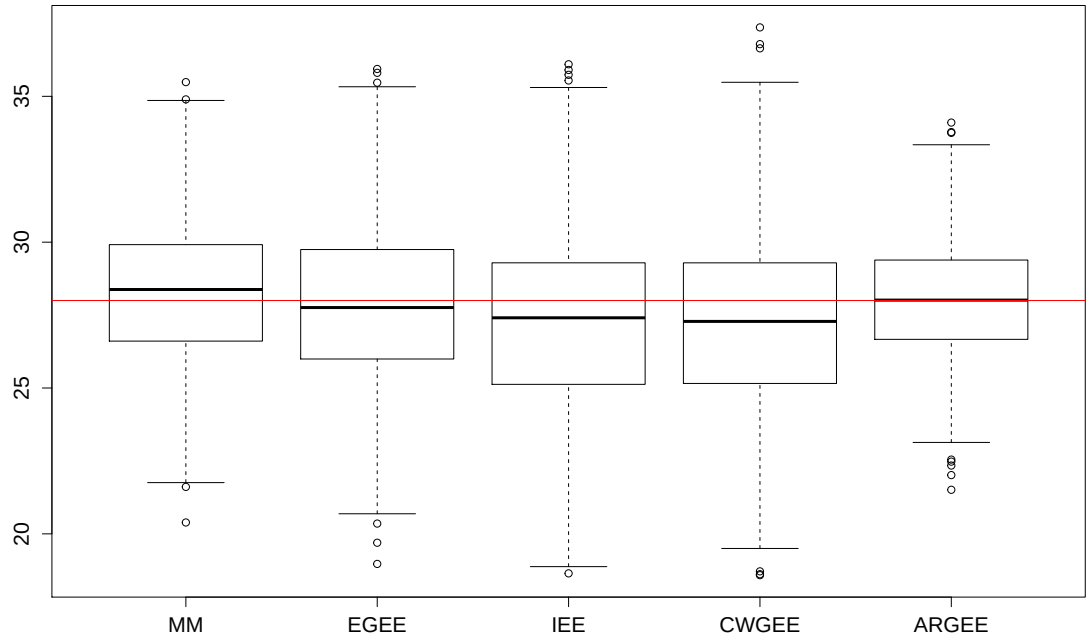

Figure 5.13: Extension 3: Boxplots of Interaction Effect estimates for Scenario 3 in Table 5.39 for n=600

Simulation Scenario	3	4	5	6	7
MM	0.03	-0.41	-0.35	-0.35	-0.49
EGEE	0.28	1.12	0.68	0.42	0.87
IEE	0.94	1.15	1.08	0.95	1.14
CWGEE	1.05	0.98	1.09	0.92	1.05
ARGEE	0.15	0.25	0.15	0.02	0.12

Table 5.44: Extension 3: Bias amount for n=60

Simulation Scenario	3	4	5	6	7
MM	-0.28	-0.63	-0.38	-0.37	-0.34
EGEE	0.14	1.00	0.66	0.45	1.11
IEE	0.75	1.28	1.21	0.97	1.22
CWGEE	0.79	1.03	1.19	0.95	1.15
ARGEE	-0.03	-0.04	0.01	-0.07	0.04

Table 5.45: Extension 3: Bias amount for n=600

Standard Errors of Interaction Term

For the interaction effect, there is very little difference between the model based standard error and the standard deviation of the estimates. There is slightly more difference when n=60 than when n=600, which can be seen in Tables 5.46 and 5.47. These tables give the mean ratio of the standard error divided by the standard deviation for n=60 and n=600 respectively. Table 5.46 shows that all the ratios are greater than 0.90, with the smallest ratios occurring for the EGEE and IEE and the closest to 1 for the MM. In all cases, the SE is being slightly underestimated. In comparison, Table 5.47 shows that the ratios for all methods are all close to 1. Here SE is larger sometimes and smaller sometimes compared to SD.

Simulation Scenario	3	4	5	6	7
MM	0.98	0.99	0.98	0.98	0.96
EGEE	0.93	0.91	0.93	0.94	0.93
IEE	0.93	0.91	0.93	0.92	0.92
CWGEE	0.96	0.96	0.96	0.95	0.95
ARGEE	0.95	0.95	0.96	0.94	0.95

Table 5.46: Extension 3: Ratios of mean SE divided by SD for n=60

Simulation Scenario	3	4	5	6	7
MM	0.96	0.98	1	1.04	1.01
EGEE	1	0.96	0.99	1.03	1.03
IEE	0.98	1.02	1.01	1	1.01
CWGEE	0.98	1.02	0.98	1	1.02
ARGEE	0.96	1.01	1.02	1.03	0.98

Table 5.47: Extension 3: Ratios of mean SE divided by SD for n=600

The difference between SE and SD for different methods can be seen more clearly in Figure 5.14 and 5.15 for n=60 and n=600 respectively. These give the ratio of the SE

for each of the 10000 and 1000 simulated datasets divided by the SD for each method when the data from all of the ICS simulation scenarios are combined for n=60 and n=600 respectively. Figure 5.14 shows that all methods tend to underestimate the SE, with the EGEE and the IEE having the smallest ratios. The ARGEE has the largest spread with a lot of positive outliers. Individual plots for these 5 scenarios showed a similar pattern (See Appendix B.4 Figures B.67 : B.71). For n=600, the ratio is very close to 1 and has no obvious pattern, as seen in Figure 5.15. Individual plots varied, however all showed ratios very close to 1 (See Appendix B.4 Figures B.74 : B.78).

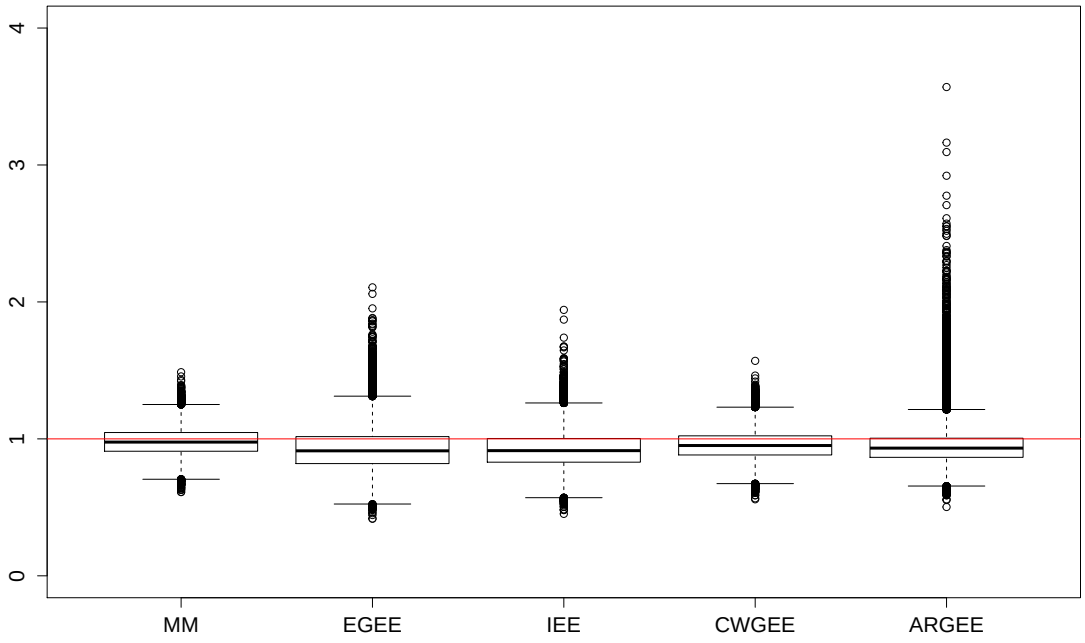

Figure 5.14: Extension 3: Boxplots of Interaction Effect standard error / standard deviation for all ICS data for n=60

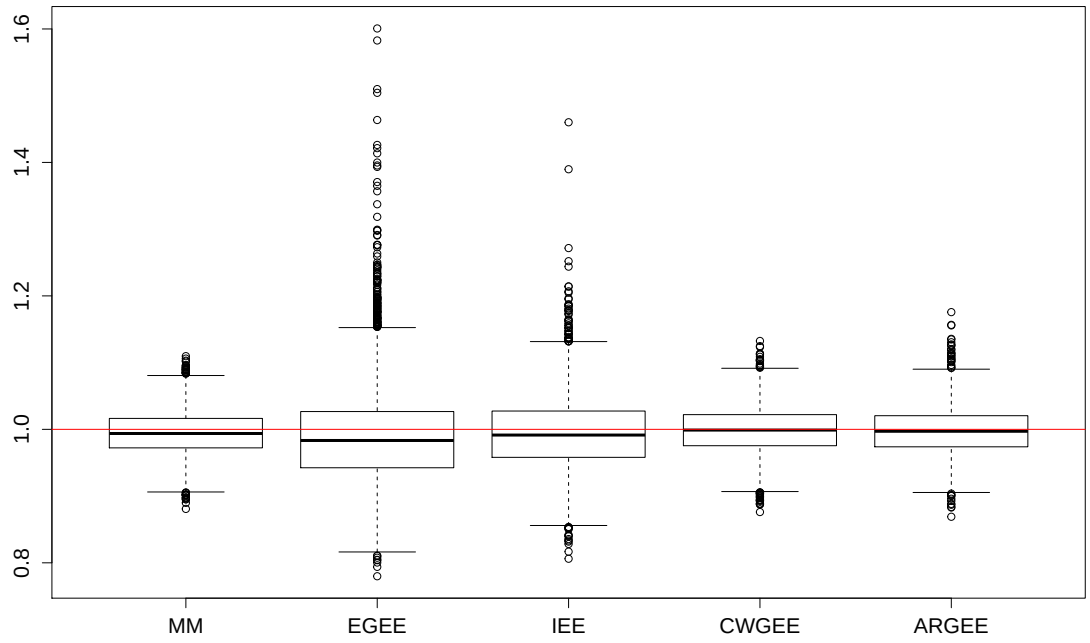

Figure 5.15: Extension 3: Boxplots of Interaction Effect standard error / standard deviation for all ICS data for n=600

Coverage Probabilities

The coverage probabilities of the Wald type confidence interval for the interaction effect by method, for both n=60 and n=600 scenarios, are given in Tables 5.48 and 5.49 respectively. Table 5.48 shows that all methods had coverage probabilities that were above 0.9. The MM had the highest coverage probabilities, which were consistently around 0.94, and the ARGEE was a close second with probabilities around 0.935. The EGEE and IEE performed worst overall, with the CWGEE being marginally better. The EGEE and IEE performed the best when the random intercept term was strongly correlated with duration (Scenarios 3 and 6). For the larger sample size, Table 5.49 shows coverage probabilities that are higher than Table 5.48 overall and are all higher than 0.92. Here the ARGEE and MM performed the best with coverage probabilities which were generally close to 0.95. The EGEE did marginally worse and the IEE and CWGEE performed the worst overall.

Simulation Scenario	3	4	5	6	7
MM	0.942	0.946	0.941	0.944	0.936
EGEE	0.926	0.913	0.925	0.928	0.92
IEE	0.923	0.914	0.922	0.924	0.917
CWGEE	0.936	0.933	0.934	0.933	0.928
ARGEE	0.935	0.934	0.933	0.936	0.933

Table 5.48: Extension 3: Coverage probabilities for n=60

Simulation Scenario	3	4	5	6	7
MM	0.944	0.934	0.949	0.951	0.949
EGEE	0.948	0.920	0.946	0.949	0.944
IEE	0.931	0.924	0.924	0.935	0.943
CWGEE	0.931	0.940	0.927	0.937	0.934
ARGEE	0.942	0.959	0.960	0.963	0.940

Table 5.49: Extension 3: Coverage probabilities for n=600

Summary

In summary, for the ICS scenarios, there was some bias in the treatment effect estimate present for all of the methods except the ARGEE, but it was fairly small. This is in contrast to the original simulation studies based on the MM, that did not show discernible bias. For both sample sizes, the SE was similar to SD, with SE being underestimated slightly in the original study size scenarios and no constant pattern in the larger study size. In addition, the difference between the SE and SD was much smaller than what was found in the MM simulations. For n=60 the ratio of mean SE divided by SD had a lot of variability for the ARGEE, similar to the original simulation study. For n=60, the coverage probabilities were quite low for the EGEE and IEE and slightly low for the CWGEE and ARGEE. In comparison, for n=600 all coverage probabilities were high, with the ARGEE and MM performing the best. For n=60, the EGEE and IEE methods generally performed better when the random slope term was not high correlated with duration. Overall, it appears that the MM performed the best when ICS was present.

The results for the simulation scenarios using the other treatment effects of 0, 7, 14, and 21 g/week gave similar results to those discussed above for the coverage probabilities and difference between SE and SD. The amount of bias did change depending on the treatment effect, in that as the treatment effect decreased, the amount of bias present decreased proportionally. Full results tables for these scenarios can be found in Appendix D.

5.4 Extensions Conclusions

In Chapter 5, 3 extensions to the original simulation study presented in Chapter 4 were considered. Extension 1 changed the distribution of the cluster size to have a larger spread, in order to investigate whether changing the distribution had an effect on the method's performance. The results from this extension showed that when the spread is larger, the EGEE, IEE and CWGEE all performed worse, however the MM and ARGEE performed similarly. These results show that changing the distribution of the cluster size can have an effect on how some of the methods perform, but the conclusions for the MM and ARGEE are unchanged.

Extension 2 investigated the effects of increasing the size of the study, that is, increasing the number of infants included in the trial. This was done to investigate whether the poor performance of the EGEE, IEE and CWGEE occurred because of the small trial size. The results from Extension 2 showed that for the larger trial size, the performance of the EGEE, IEE and CWGEE methods improved. In comparison, for both sample sizes, the MM and ARGEE performed well. This indicates that the poor performance of some methods in the initial simulation study may be due to the size of the study being too small.

Extension 3 changed the way the data were simulated, from a MM to a GEE. This was done to test whether the MM was performing relatively well because it had the additional advantage of being the correct model for the simulations in Chapter 4. The results from this extension showed that the MM still performed well when a GEE framework

was used to simulate the data, with little bias and high coverage probabilities. Thus, the MM performs well, regardless of whether or not it is used to simulate the data. The EGEE, IEE and CWGEE methods performed better when the trial size was large, which is consistent with the finding of the simulations under the MM. Hence, it can be concluded that the good performance of the MM was not solely due to its use in generating the data.

In summary, the results of all simulations conducted indicate that the MM and ARGEE are the most appropriate methods for analysing longitudinal data when the length of followup is informative. The EGEE, IEE and CWGEE methods do not always perform well and can result in large differences between the SE and the SD and low coverage probabilities. These methods performed the best when the trial size was large and when only the random intercept was correlated with duration.

5.5 Implications for the POPPET data

The objective of the POPPET trial analysis was to determine the effect of the intervention on the growth of preterm infants. The results of the simulation studies can now be used to address this objective. It is concluded that a MM or a ARGEE should be used to analyse the POPPET data, as these methods were determined to provide valid inference in the presence of ICS.

The POPPET results for the MM and ARGEE are re-listed in Tables 5.50 and 5.51 respectively. When comparing the parameter estimates between to two tables, some of them appear quite different. For example, for the MM the estimates for sex and plurality are 13.2 and -8.63 respectively, whereas for the ARGEE they are -106.13 and 40.99 respectively. However, when these are considered in conjunction with their standard errors, both methods appear to be reaching similar conclusions. For the MM, the confidence intervals for sex and plurality are (-31.21,57.62) and (-66.42, 49.16), and for the ARGEE they are (-269,56.7) and (-85.3,167). These are quite similar when considering that the outcome is weight in grams which is in the thousands, and these are factor variables. The estimates for gestational age and time squared are also quite different between the

two methods. Again, the standard errors are quite large for the ARGEE method and only slightly smaller for the MM, so when considering how large the outcomes are, the conclusions are not very different. Notably, both methods do result in similar estimates for time and birthweight.

Recall that interest lies in the interaction terms involving time and group, as they describe the treatment effect. The two sets of interaction term estimates are quite different between the two models, but when combined they lead to similar conclusions. For example, the predicted difference in weight between treatment groups, on day 45 is 132g for the MM and 148g for the ARGEE with standard errors of 98g and 108g respectively. This indicates the combined treatment effect is similar between the two models, with the ARGEE estimating only a slightly larger effect. In Chapter 3, a test for whether the two interaction terms are jointly significantly different from zero was performed for each method. For the MM, the null hypothesis of no difference was retained (p=0.2721). For the ARGEE, the null hypothesis was rejected as the p value was 0.048. Note that there may be an issue due to multiple testing present here. The p-value for the ARGEE is only marginally less than 0.05 and hence there was only limited evidence to suggest a difference from zero. This implies the two methods don't attain the same effect of the intervention, with the ARGEE producing a slightly larger effect than the MM. This agrees with the differences on day 45, where the ARGEE resulted in a slightly larger difference. As the MM results indicated no treatment effect and the ARGEE did not show strong evidence of a treatment effect, it is concluded that there is not enough evidence to show that the treatment affects the growth of preterm infants. It is possible that the high protein milk fortifier improves the growth of infants, however a larger trial is needed before any definitive conclusions can be drawn.

The slightly differing conclusions for the MM and ARGEE could be happening for a number of reasons. Firstly, the POPPET data are more complicated than the simulated data (discussed in Chapter 4 and 5) as they contain two levels of clustering, both longitudinal and maternal. Further research is needed to understand how the MM and ARGEE perform when both types of clustering are present. Secondly, it could be due to real world

data problems. It was briefly discussed in Chapter 1 that misspecification of the correlation matrix is a potential problem for the POPPET trial data. Recall, this is because the correlation between twins will be stronger if they are monozygotic (one egg fertilised by one sperm that splits) vs disygotic (two eggs fertilised by two sperm) and also monochorionic (shared placenta) vs dichorionic (separate placenta). This problem could be affecting the results of the two methods differently, since GEEs are robust to misspecification of the working correlation structure.

Fixed Effect	estimate	s.e	t-value	p-value	ci lower	ci upper
Intercept	-549.86	341.74	-1.61	0.11	-1219.67	119.94
Time	208.10	12.82	16.23	< 0.0001	182.95	233.21
Sex(Male)	13.20	22.66	0.58	0.57	-31.21	57.62
Plurality(2)	-8.63	29.49	-0.29	0.77	-66.42	49.16
Gestational Age	28.13	12.09	2.33	0.025	4.44	51.83
Birthweight	0.70	0.037	18.80	< 0.0001	0.63	0.78
Time Squared	6.60	2.09	3.15	0.0016	2.50	10.71
Group*Time	-5.02	18.00	-0.28	0.78	-40.32	30.27
Group*Time Squared	3.98	2.93	1.36	0.17	-1.76	9.73

Table 5.50: Results for the POPPET data for a MM

	estimate	s.e	Wald	p-value	ci lower	ci upper
Intercept	532.38	1304.29	0.17	0.683	-2024	3089
Time	275.63	32.69	71.10	< 0.0001	212	340
Sex(Male)	-106.13	83.06	1.63	0.201	-269	56.7
Plurality(2)	40.99	64.4	0.40	0.525	-85.3	167
Gestational Age	-32.92	54.21	0.37	0.544	-139	73.3
Birthweight	1.20	0.27	19.90	< 0.0001	0.675	1.73
Time Squared	-5.89	4.15	2.01	0.156	-14	2.24
Group*Time	-31.34	33.84	0.86	0.354	-97.7	35
Group*Time Squared	8.46	4.38	3.73	0.054	-0.125	17

Table 5.51: Results for the POPPET data for a ARGEE

Chapter 6

Conclusion

This thesis investigated the issue of informative cluster size in longitudinal data. The research objective was to determine suitable methods for the analysis of longitudinal data when cluster size (length of follow up) is informative. Chapter 1 provided the motivation for the research, and included a discussion of the POPPET trial. Chapter 2 explored different strategies for analysing clustered data, including methods that can be used when ICS is present. Following from this, Chapters 3 and 4 compared analysis methods using real and simulated data respectively, with extensions to the simulation study considered in Chapter 5. The main findings of this research will now be summarised and final recommendations made regarding the analysis of longitudinal data when cluster size is informative.

6.1 Main Findings

In Chapter 3, different analysis methods were applied to real data from the POPPET trial. Informative cluster size was suspected, which was confirmed thorough investigation. This analysis showed, with the exception of the ARGEE, that there was no significant effect of treatment on weight. Specifically, the MM, EGEE, IEE and CWGEE all showed similar results regarding the intervention's effect. Through the simulation studies conducted in Chapter 4 and 5, it was found that the MM and ARGEE were the most suitable analysis methods for analysing longitudinal data with informative cluster size. The results of these methods did not show strong evidence of a significant effect of the intervention on the

growth of preterm infants.

To further explore the research aim, a simulation study was conducted to compare methods when the true parameter values were known. Three measures were used to assess the performance of methods in the simulation study: the amount of bias in the estimated treatment effect; the difference between the model based standard error and the standard error of the estimates; and the coverage probability.

In the Chapter 4 simulation study, it was found that all analysis methods were valid when there was non informative cluster size. Importantly, there was no significant bias when ICS was present. However, when the EGEE, IEE and CWGEE methods were used, they did not have the expected coverage probabilities and their mean model based standard error and standard deviation of the estimates were not similar. These methods performed best when cluster size was correlated with the random intercept but not the random slope. The ARGEE method performed well with high coverage probabilities and the ratio of the mean model based standard error divided by the standard deviation of the estimates was close to 1, but there was a large variability in this ratio. Although the mean of the standard error is close to the standard deviation, there are lots of really large values which means this method could produce misleading estimates. Overall, the best performing method was the MM.

Three main questions arose from the Chapter 4 simulation results: would changing the distribution of cluster size affect the results, would increasing the sample size improve the results for the EGEE, IEE and CWGEE methods, and would generating data from a different model change how the methods perform. These three questions became three extension simulation studies, which were discussed in Chapter 5.

Extension 1 considered changing the distribution of the cluster size to be more variable in order to determine if this would alter the performance of the different analysis approaches. This analysis resulted in similar findings to the original simulation study in Chapter 4 with one notable exception; the performance of EGEE, IEE and CWGEE were all sub-

stantially worse. In particular, both the difference between SE and SD increased and the coverage probabilities decreased. Notably, optimal results still occurred when cluster size was correlated with the random intercept but not the random slope. Accordingly, it was concluded that the spread of the cluster size distribution negatively affected the performance of the EGEE, IEE and CWGEE methods.

In Extension 2, the sample size was increased to see if this improved the performance of the analysis approaches considered. Similar to all previous simulation studies, all methods performed well in NICS cases. Importantly, all methods also performed well in the ICS cases - this was in contrast to the poor performance of the EGEE, IEE and CWGEE methods in previous simulations (Chapter 4 and Extension 1 Chapter 5). In this extension, there were no large differences between SE and SD and the methods displayed coverage probabilities that were close to the expected value. The ratio of mean SE divided by SD was also less variable for the ARGEE in this study. Overall in this second extension, the MM and ARGEE performed the best, but the EGEE, IEE and CWGEE were only slightly worse. Hence, it was concluded that a larger sample size did improve the performance of some methods.

To determine if the MM still performed optimally, Extension 3 considered simulated data which was generated from a GEE framework rather than a MM. Two sample sizes were considered for this purpose: the small trial size in Chapter 4 and the large trial size from Extension 2 in Chapter 5. For the NICS cases, all methods performed well. In both sample sizes for ICS cases, some bias was present for the EGEE, IEE and CWGEE. For the EGEE and IEE, when cluster size was correlated with the random slope, the bias was worse. For the small sample size, there was a noticeable difference between the SE and SD and the coverage probabilities were quite low for the EGEE, IEE and CWGEE. These methods improved for the larger sample size, similar to Extension 2. The ARGEE performed well for the large sample size, but there was a large amount of variability in the ratio of the mean SE divided by SD for the small sample size. The MM performed well for both sample sizes. Hence, it is reasonable to conclude that the strong performance of MM seen in the original simulation study was not solely due to its use in generating the

data alone.

Overall, based on the results of all simulation studies it was concluded that all methods performed well when cluster size was non informative. Notably, the CWGEE and IEE were the most inefficient methods. However, when cluster size was informative,]the MM performed well across all scenarios considered and the ARGEE performed quite well, but there was a large variability in the ratio of the SE divided by SD in scenarios with a small cluster size. This variability means that sometimes the estimates are ok and other they are very wrong. The performance of the other three methods (EGEE, IEE and CWGEE) was sub-optimal in many scenarios. Of note, these methods did not perform well for the small sample sizes, resulting in large differences between SE and SD and low coverage probabilities. Regardless of sample size, these methods generally performed better when cluster size was correlated with the random intercept but not with the random slope, that is, when there is no correlation between individual growth trajectories and cluster size.

6.2 Significance and Innovation

The real data set considered in this thesis was the POPPET data set, which is from a recently completed Australian clinical trial. In Chapter 3, this data set was analysed to provide an estimate of the effect of treatment on the growth of preterm infants, which may be helpful for informing clinical practice regarding feeding practices for future preterm infants.

To my knowledge this thesis provides the most comprehensive comparison of the performance of analysis methods for longitudinal data with informative cluster size. Five methods were compared for several differing levels of informative cluster size, that is, different degrees of correlation between the cluster size and the outcome. Also considered were different sample sizes, distributions of cluster size and methods for simulating the data. Overall, this thesis provided a clear recommendation regarding which methods are suitable for analysing longitudinal data with ICS: some are suitable consistently and other

only under certain circumstances. These results will guide future researchers to make informed decisions regarding which methods of analysis should be used for longitudinal data with informative cluster size.

6.3 Limitations and Future Work

Within this thesis a large number of simulations were conducted across a broad range of scenarios relevant to longitudinal data in perinatal trials, but the simulation studies are necessarily limited. The data were simulated with one level of clustering; that is longitudinal data. In addition, only a linear growth relationship was considered and a gaussian distribution was assumed for the outcome. The results of these analyses cannot be assumed to apply to more general settings and hence are only applicable to data akin to that discussed throughout this thesis.

Additional simulation studies are needed to explore the performance of analysis methods for more complex data sets with ICS. This could include data with multiple levels of clustering, such as non longitudinal clustering, or data with more complex relationships, such as nonlinear growth. A broader range of sample sizes could also be considered, such as a smaller number of clusters than 60.

6.4 Final Recommendations

Based on the results of this thesis, it is recommended that the mixed model is used for the analysis of longitudinal data when the cluster size is informative. The ARGEE could also be used but the mixed model is preferred. The IEE, EGEE and CWGEE should only be used if there is a very large sample size or when there is no dependence between individual growth trajectories and cluster size.

Appendix A

Results

In Section 4.4 the results for the simulation study were discussed and the results for simulation scenarios with a treatment effect of 28 were provided. This appendix, A.1, gives the results for the full simulation study, including all treatment effects. That is, all 35 simulation scenarios listed in table 4.4.

In section 5.1 the results for Extension 1, which looked at doubling the standard deviation of the cluster size were discussed. The results for the simulation scenarios with a treatment effect of 28 were provided. This appendix, A.2, gives the results for the full extension to the simulation study, including all treatment effects. That is, all 25 simulation scenarios listed in table 5.2.

A.1 Chapter 4 Simulation Results

A.1.1 Scenarios with Non Informative Cluster Size

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.93	10.47	10.73	42.04	11.95	12.23	0.72	0.04	0.04	244.50	7.44	7.47	-1043.64	320.97	328.90
EGEE	27.91	10.38	10.65	41.93	17.31	18.32	0.72	0.06	0.07	244.52	7.32	7.43	-1040.82	466.64	491.60
IEE	27.87	12.19	12.74	41.94	17.34	18.42	0.72	0.06	0.07	244.54	8.00	8.23	-1040.99	466.84	493.70
CWGEE	27.87	12.19	12.74	41.94	17.34	18.42	0.72	0.06	0.07	244.54	8.00	8.23	-1040.99	466.84	493.70
ARGEE	27.93	10.81	11.26	41.96	17.56	18.66	0.72	0.06	0.07	244.51	7.63	7.82	-1041.84	473.26	500.66

Table A.1: Scenario 1 (fixed trial length of 38, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.91	10.48	10.82	41.80	11.93	12.42	0.72	0.04	0.04	244.51	7.45	7.63	-1036.97	320.62	333.65
EGEE	27.94	10.96	11.44	41.58	17.43	18.69	0.72	0.06	0.07	244.50	7.71	8.10	-1030.93	469.45	500.80
IEE	27.88	12.59	13.31	41.56	17.79	19.22	0.72	0.06	0.07	244.53	8.94	9.56	-1030.42	478.66	514.59
CWGEE	27.88	12.34	12.95	41.57	17.45	18.78	0.72	0.06	0.07	244.53	8.67	9.17	-1030.59	469.60	502.78
ARGEE	27.93	10.93	11.43	41.52	17.77	19.10	0.72	0.06	0.07	244.49	7.71	8.08	-1029.09	478.31	511.79

Table A.2: Scenario 2 ($\gamma_0 = log(24.3667), \gamma_1 = 0, \gamma_2 = 0$, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.94	10.47	10.68	41.73	11.95	12.23	0.72	0.04	0.04	244.52	7.43	7.53	-1035.06	320.85	328.50
EGEE	20.92	10.37	10.60	41.72	17.28	18.13	0.72	0.06	0.06	244.53	7.32	7.51	-1034.02	465.46	488.40
IEE	20.97	12.19	12.72	41.72	17.31	18.24	0.72	0.06	0.07	244.50	8.01	8.29	-1034.12	465.80	490.69
CWGEE	20.97	12.19	12.72	41.72	17.31	18.24	0.72	0.06	0.07	244.50	8.01	8.29	-1034.12	465.80	490.69
ARGEE	20.99	10.81	11.16	41.72	17.52	18.45	0.72	0.06	0.07	244.51	7.64	7.88	-1033.92	471.82	496.66

Table A.3: Scenario 8 (fixed trial length of 38, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.00	10.48	10.68	41.74	11.94	12.37	0.72	0.04	0.04	244.53	7.44	7.51	-1036.43	320.71	332.73
EGEE	20.98	10.97	11.31	41.75	17.46	18.82	0.72	0.06	0.07	244.49	7.71	8.01	-1036.58	470.18	505.69
IEE	20.99	12.57	13.15	41.72	17.82	19.38	0.72	0.06	0.07	244.42	8.94	9.49	-1035.69	479.47	520.46
CWGEE	21.01	12.32	12.80	41.75	17.48	18.91	0.72	0.06	0.07	244.44	8.67	9.10	-1036.58	470.27	507.76
ARGEE	21.02	10.92	11.32	41.71	17.80	19.24	0.72	0.06	0.07	244.51	7.72	8.04	-1035.66	479.16	516.81

Table A.4: Scenario 9 ($\gamma_0 = log(24.3667), \gamma_1 = 0, \gamma_2 = 0$, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.06	10.48	10.73	41.93	11.93	12.27	0.72	0.04	0.04	244.39	7.45	7.60	-1041.25	320.56	330.26
EGEE	14.04	10.39	10.64	41.96	17.33	18.47	0.72	0.06	0.07	244.40	7.32	7.57	-1042.01	466.65	496.05
IEE	13.97	12.20	12.72	41.95	17.35	18.56	0.72	0.06	0.07	244.44	8.01	8.32	-1041.89	466.84	497.97
CWGEE	13.97	12.20	12.72	41.95	17.35	18.56	0.72	0.06	0.07	244.44	8.01	8.32	-1041.89	466.84	497.97
ARGEE	14.00	10.82	11.20	41.95	17.57	18.74	0.72	0.06	0.07	244.42	7.64	7.92	-1041.80	473.05	502.86

Table A.5: Scenario 15 (fixed trial length of 38, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.06	10.49	10.70	42.01	11.98	12.22	0.72	0.04	0.04	244.37	7.45	7.51	-1042.90	321.74	328.08
EGEE	14.03	10.98	11.25	41.77	17.50	18.45	0.72	0.06	0.07	244.38	7.72	7.98	-1035.59	471.43	496.44
IEE	14.10	12.58	13.29	41.76	17.85	18.99	0.72	0.06	0.07	244.32	8.95	9.51	-1035.37	480.56	510.41
CWGEE	14.11	12.33	12.96	41.77	17.51	18.55	0.72	0.06	0.07	244.32	8.68	9.11	-1035.57	471.37	498.64
ARGEE	14.05	10.93	11.32	41.81	17.84	18.92	0.72	0.06	0.07	244.34	7.73	8.01	-1036.47	480.35	508.66

Table A.6: Scenario 16 ($\gamma_0 = log(24.3667), \gamma_1 = 0, \gamma_2 = 0$, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.98	10.49	10.59	42.13	11.94	12.36	0.72	0.04	0.04	244.57	7.45	7.51	-1047.10	320.71	330.63
EGEE	6.99	10.39	10.51	42.13	17.35	18.24	0.72	0.06	0.07	244.57	7.32	7.49	-1046.62	467.34	489.26
IEE	7.12	12.20	12.54	42.15	17.37	18.32	0.72	0.06	0.07	244.50	8.01	8.26	-1046.95	467.47	490.96
CWGEE	7.12	12.20	12.54	42.15	17.37	18.32	0.72	0.06	0.07	244.50	8.01	8.26	-1046.95	467.47	490.96
ARGEE	7.07	10.82	11.06	42.14	17.61	18.53	0.72	0.06	0.07	244.56	7.64	7.86	-1047.07	474.08	496.63

Table A.7: Scenario 22 (fixed trial length of 38, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.03	10.48	10.76	42.16	11.95	12.27	0.72	0.04	0.04	244.53	7.45	7.62	-1047.20	321.27	329.55
EGEE	6.97	10.97	11.41	41.89	17.45	18.38	0.72	0.06	0.07	244.59	7.73	8.11	-1039.56	470.48	493.81
IEE	6.98	12.61	13.18	41.92	17.82	18.91	0.72	0.06	0.07	244.62	8.98	9.53	-1040.26	479.93	507.81
CWGEE	6.99	12.35	12.83	41.89	17.47	18.48	0.72	0.06	0.07	244.62	8.70	9.14	-1039.72	470.63	496.26
ARGEE	7.02	10.93	11.36	41.87	17.79	18.78	0.72	0.06	0.07	244.56	7.74	8.05	-1039.11	479.46	504.19

Table A.8: Scenario 23 ($\gamma_0 = log(24.3667), \gamma_1 = 0, \gamma_2 = 0$, treatment effect = 7 g/week)

Interaction =0

	Int	SE(Int)	SD(Int)	GA	SE(GA)	SD(GA)	BW	SE(BW)	SD(BW)	T	SE(T)	SD(T)	Intercept	SE	SD
MM	-0.03	10.48	10.70	42.11	11.94	12.21	0.72	0.04	0.04	244.48	7.45	7.60	-1045.35	320.81	327.30
EGEE	-0.03	10.39	10.62	42.20	17.32	18.33	0.72	0.06	0.07	244.48	7.30	7.57	-1047.62	466.44	492.04
IEE	-0.03	12.18	12.72	42.21	17.34	18.43	0.72	0.06	0.07	244.48	7.98	8.30	-1047.87	466.55	494.08
CWGEE	-0.03	12.18	12.72	42.21	17.34	18.43	0.72	0.06	0.07	244.48	7.98	8.30	-1047.87	466.55	494.08
ARGEE	-0.07	10.82	11.14	42.20	17.57	18.63	0.72	0.06	0.07	244.52	7.62	7.90	-1047.51	473.06	499.88

Table A.9: Scenario 29 (fixed trial length of 38, treatment effect = 0 g/week)

	Int	SE(Int)	SD(Int)	GA	SE(GA)	SD(GA)	BW	SE(BW)	SD(BW)	T	SE(T)	SD(T)	Intercept	SE	SD
MM	-0.02	10.50	10.75	41.94	11.92	12.17	0.72	0.04	0.04	244.53	7.46	7.68	-1041.81	320.26	327.21
EGEE	0.02	10.98	11.32	42.19	17.44	18.45	0.72	0.06	0.07	244.49	7.74	8.13	-1048.89	469.95	495.83
IEE	0.01	12.59	13.30	42.22	17.78	18.99	0.72	0.06	0.07	244.49	8.96	9.64	-1049.66	478.90	510.05
CWGEE	0.00	12.34	12.97	42.18	17.45	18.56	0.72	0.06	0.07	244.50	8.69	9.24	-1048.66	469.88	498.35
ARGEE	-0.01	10.95	11.39	42.19	17.76	18.89	0.72	0.06	0.07	244.49	7.75	8.15	-1048.54	478.54	507.67

Table A.10: Scenario 30 ($\gamma_0 = log(24.3667), \gamma_1 = 0, \gamma_2 = 0$, treatment effect = 0 g/week)

A.1.2 Scenarios with Informative Cluster Size using $\gamma_0 = 3.069142$

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.99	10.56	10.76	42.12	11.95	12.38	0.72	0.04	0.04	245.03	7.50	7.50	-1047.11	321.14	332.70
EGEE	27.63	14.40	17.61	42.18	17.73	18.94	0.72	0.06	0.07	244.78	9.97	12.43	-1047.55	477.78	507.73
IEE	27.73	14.20	17.72	42.30	18.95	21.57	0.72	0.07	0.08	226.08	12.13	16.51	-1020.56	511.03	578.20
CWGEE	27.89	12.79	14.21	42.13	16.52	17.79	0.72	0.06	0.06	223.83	11.00	12.73	-987.43	444.97	476.84
ARGEE	27.96	11.81	12.53	42.02	18.02	19.47	0.72	0.06	0.07	244.77	8.19	8.60	-1040.56	485.32	522.08

Table A.11: Scenario 3 ($\gamma_1 = -0.50$, $\gamma_2 = 0$, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.99	10.44	10.70	41.84	12.00	12.25	0.72	0.04	0.04	243.64	7.42	7.61	-1037.92	322.26	329.33
EGEE	27.78	16.14	23.33	41.83	16.51	17.51	0.72	0.06	0.06	206.65	10.94	14.93	-966.41	444.69	471.39
IEE	27.77	13.22	18.91	42.02	17.12	20.13	0.72	0.06	0.07	181.66	11.55	17.04	-923.75	461.43	542.24
CWGEE	27.85	12.58	14.86	41.84	15.50	16.70	0.72	0.06	0.06	192.99	11.50	12.87	-928.36	417.67	449.53
ARGEE	28.06	12.57	13.31	41.75	18.17	19.52	0.72	0.06	0.07	234.93	8.46	9.81	-1033.89	489.19	524.94

Table A.12: Scenario 4 ($\gamma_1 = 0, \gamma_2 = -0.50$, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.09	10.51	10.75	42.02	11.94	12.40	0.72	0.04	0.04	244.24	7.46	7.61	-1043.79	320.95	331.96
EGEE	27.99	15.26	21.01	41.85	16.33	17.73	0.72	0.06	0.06	217.60	10.43	13.89	-986.85	440.03	474.24
IEE	27.93	12.45	17.69	41.95	15.91	19.28	0.72	0.06	0.07	186.73	10.87	16.47	-933.98	429.23	516.92
CWGEE	27.98	11.58	13.63	41.87	13.96	15.37	0.72	0.05	0.06	193.29	10.64	12.68	-918.87	376.24	411.28
ARGEE	27.97	13.46	14.22	41.92	17.67	19.14	0.72	0.06	0.07	240.88	8.77	9.68	-1033.65	475.64	511.67

Table A.13: Scenario 5 ($\gamma_1 = -0.35, \gamma_2 = -0.35$, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.82	10.53	10.74	42.08	11.97	12.38	0.72	0.04	0.04	244.87	7.48	7.56	-1045.29	321.65	332.29
EGEE	27.74	14.63	18.41	41.95	16.99	18.13	0.72	0.06	0.06	230.15	10.04	12.82	-1013.01	457.59	487.61
IEE	27.74	13.18	17.06	42.05	17.20	20.10	0.72	0.06	0.07	203.26	11.32	16.24	-968.88	463.70	541.05
CWGEE	27.78	11.98	13.37	41.93	14.97	16.15	0.72	0.05	0.06	205.68	10.61	12.54	-943.64	403.01	433.95
ARGEE	27.80	12.70	13.33	41.91	17.78	19.16	0.72	0.06	0.07	242.63	8.49	9.19	-1034.20	478.37	514.49

Table A.14: Scenario 6 ($\gamma_1 = -0.46, \gamma_2 = -0.19$, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.79	10.48	10.73	41.93	11.98	12.34	0.72	0.04	0.04	244.03	7.45	7.51	-1040.58	321.70	332.50
EGEE	27.53	15.91	22.85	42.15	16.10	17.35	0.72	0.06	0.06	209.71	10.79	14.63	-979.75	433.65	465.57
IEE	27.63	12.49	18.31	42.29	15.83	19.14	0.72	0.06	0.07	179.41	10.98	16.75	-927.32	426.74	514.37
CWGEE	27.74	11.84	14.15	42.11	14.19	15.47	0.72	0.05	0.06	189.06	11.01	12.78	-921.01	382.28	415.99
ARGEE	27.70	13.34	14.09	42.04	17.93	19.46	0.72	0.06	0.07	238.50	8.70	9.84	-1040.07	482.14	520.97

Table A.15: Scenario 7 ($\gamma_1 = -0.19, \gamma_2 = -0.46$, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.74	10.56	10.84	41.97	11.95	12.22	0.72	0.04	0.04	245.08	7.50	7.63	-1043.18	321.05	328.75
EGEE	20.82	14.42	17.39	41.91	17.74	18.89	0.72	0.06	0.07	244.42	9.94	12.33	-1041.11	477.70	507.65
IEE	20.79	14.20	17.62	41.97	18.94	21.75	0.72	0.07	0.08	225.70	12.10	16.37	-1012.34	510.32	585.40
CWGEE	20.78	12.80	14.17	41.90	16.52	17.77	0.72	0.06	0.06	223.58	11.00	12.69	-981.90	444.72	477.33
ARGEE	20.78	11.82	12.46	41.91	18.03	19.36	0.72	0.06	0.07	244.76	8.20	8.68	-1039.35	485.41	520.05

Table A.16: Scenario 10 ($\gamma_1 = -0.50$, $\gamma_2 = 0$, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.97	10.43	10.66	41.98	11.98	12.30	0.72	0.04	0.04	243.66	7.41	7.57	-1040.68	321.79	330.91
EGEE	21.05	16.18	23.61	42.07	16.53	17.84	0.72	0.06	0.06	206.51	11.04	15.23	-970.59	445.24	479.99
IEE	21.01	13.21	19.01	42.26	17.14	20.33	0.72	0.06	0.07	181.47	11.63	17.52	-927.64	462.01	547.25
CWGEE	20.99	12.61	14.86	42.02	15.51	16.95	0.72	0.06	0.06	192.89	11.60	13.08	-931.49	417.82	456.50
ARGEE	21.02	12.65	13.49	41.91	18.17	19.72	0.72	0.06	0.07	234.97	8.51	9.92	-1037.13	489.26	530.21

Table A.17: Scenario 11 ($\gamma_1 = 0, \gamma_2 = -0.50$, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.94	10.52	10.76	41.94	11.96	12.25	0.72	0.04	0.04	244.26	7.47	7.52	-1040.84	321.31	327.64
EGEE	21.45	15.18	20.39	42.02	16.29	17.51	0.72	0.06	0.06	217.55	10.44	13.85	-992.33	438.77	470.41
IEE	21.34	12.42	17.23	42.09	15.81	18.96	0.72	0.06	0.07	186.85	10.85	16.49	-939.55	426.66	510.63
CWGEE	21.21	11.53	13.38	42.04	13.93	15.18	0.72	0.05	0.05	193.35	10.63	12.61	-924.63	375.45	408.06
ARGEE	21.16	13.35	14.03	42.09	17.66	19.16	0.72	0.06	0.07	240.67	8.77	9.63	-1039.00	475.26	513.01

Table A.18: Scenario 12 ($\gamma_1 = -0.35, \gamma_2 = -0.35$, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.10	10.55	10.67	42.09	11.96	12.41	0.72	0.04	0.04	244.60	7.50	7.58	-1045.37	321.41	332.78
EGEE	21.08	14.62	18.39	42.21	16.97	18.45	0.72	0.06	0.07	229.97	10.00	12.72	-1019.39	457.18	493.78
IEE	21.16	13.19	17.15	42.23	17.19	20.27	0.72	0.06	0.07	203.14	11.28	16.12	-973.97	463.66	543.84
CWGEE	21.16	11.98	13.45	42.16	14.95	16.40	0.72	0.05	0.06	205.49	10.56	12.50	-949.45	402.93	438.87
ARGEE	21.04	12.66	13.16	42.00	17.70	19.31	0.72	0.06	0.07	242.38	8.44	9.17	-1035.67	476.85	515.84

Table A.19: Scenario 13 ($\gamma_1 = -0.46, \gamma_2 = -0.19$, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.00	10.48	10.63	41.86	11.97	12.33	0.72	0.04	0.04	243.79	7.44	7.55	-1038.54	321.57	330.60
EGEE	21.06	15.83	22.96	41.55	16.12	17.19	0.72	0.06	0.06	209.33	10.81	14.91	-963.91	434.16	461.59
IEE	21.06	12.47	18.42	41.52	15.85	18.88	0.72	0.06	0.07	179.10	10.98	17.14	-907.17	427.61	508.34
CWGEE	21.07	11.82	14.27	41.59	14.19	15.35	0.72	0.05	0.05	188.75	11.03	13.05	-907.06	382.51	412.62
ARGEE	20.96	13.25	14.04	41.57	17.85	19.09	0.72	0.06	0.07	238.23	8.71	9.71	-1027.62	480.44	511.22

Table A.20: Scenario 14 ($\gamma_1 = -0.19, \gamma_2 = -0.46$, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.90	10.54	10.74	41.93	11.94	12.36	0.72	0.04	0.04	245.09	7.49	7.61	-1042.03	320.90	330.91
EGEE	14.06	14.45	17.49	41.81	17.71	18.86	0.72	0.06	0.07	244.39	10.00	12.39	-1038.55	477.31	506.54
IEE	14.11	14.22	17.65	41.99	18.94	21.61	0.72	0.07	0.08	225.64	12.16	16.55	-1013.55	510.48	580.86
CWGEE	14.07	12.82	14.17	41.80	16.50	17.72	0.72	0.06	0.06	223.53	11.05	12.78	-979.45	444.49	476.06
ARGEE	13.93	11.84	12.42	41.82	18.02	19.26	0.72	0.06	0.07	244.78	8.21	8.67	-1036.44	485.58	516.60

Table A.21: Scenario 17 ($\gamma_1 = -0.50$, $\gamma_2 = 0$, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.91	10.42	10.64	41.80	11.98	12.27	0.72	0.04	0.04	243.61	7.40	7.62	-1036.40	321.82	328.89
EGEE	14.04	16.10	23.36	41.91	16.47	17.55	0.72	0.06	0.06	206.53	10.95	15.27	-967.54	443.65	470.47
IEE	14.01	13.20	18.86	41.96	17.07	20.06	0.72	0.06	0.07	181.63	11.56	17.40	-922.02	460.14	539.10
CWGEE	13.98	12.58	14.73	41.89	15.47	16.68	0.72	0.06	0.06	193.00	11.54	13.14	-929.38	416.75	447.38
ARGEE	13.96	12.61	13.42	41.98	18.16	19.78	0.72	0.06	0.07	234.93	8.49	9.90	-1038.57	488.77	529.14

Table A.22: Scenario 18 ($\gamma_1 = 0, \gamma_2 = -0.50$, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.91	10.51	10.66	41.88	11.96	12.24	0.72	0.04	0.04	244.21	7.47	7.53	-1039.68	321.25	328.44
EGEE	13.98	15.24	20.74	42.07	16.32	17.53	0.72	0.06	0.06	217.60	10.39	13.75	-992.65	439.72	470.34
IEE	13.99	12.44	17.49	42.14	15.88	19.06	0.72	0.06	0.07	186.76	10.84	16.26	-939.25	428.40	511.72
CWGEE	13.96	11.55	13.42	42.06	13.96	15.22	0.72	0.05	0.05	193.28	10.62	12.48	-924.00	376.31	408.41
ARGEE	13.95	13.48	14.16	42.07	17.68	19.18	0.72	0.06	0.07	240.77	8.78	9.69	-1037.60	475.62	513.14

Table A.23: Scenario 19 ($\gamma_1 = -0.35, \gamma_2 = -0.35$, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.01	10.56	10.81	41.97	11.97	12.30	0.72	0.04	0.04	244.63	7.50	7.57	-1042.24	321.68	330.13
EGEE	13.97	14.63	18.60	42.17	17.06	18.19	0.72	0.06	0.06	229.90	10.07	12.92	-1019.48	459.50	489.25
IEE	13.98	13.19	17.30	42.26	17.29	20.12	0.72	0.06	0.07	203.05	11.35	16.38	-975.31	466.34	541.29
CWGEE	13.94	11.98	13.60	42.17	15.02	16.20	0.72	0.05	0.06	205.44	10.62	12.65	-950.70	404.72	435.74
ARGEE	14.03	12.72	13.73	42.21	17.88	19.81	0.72	0.06	0.08	242.36	8.50	9.41	-1042.69	480.41	519.23

Table A.24: Scenario 20 ($\gamma_1 = -0.46, \gamma_2 = -0.19$, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.93	10.45	10.68	42.18	11.96	12.20	0.72	0.04	0.04	243.96	7.42	7.52	-1046.95	321.42	327.98
EGEE	13.39	15.84	22.78	42.18	16.09	17.16	0.72	0.06	0.06	209.91	10.72	14.51	-979.19	433.52	460.91
IEE	13.63	12.46	18.20	42.24	15.82	18.94	0.72	0.06	0.07	179.60	10.94	16.80	-924.79	426.82	509.68
CWGEE	13.78	11.82	14.03	42.18	14.16	15.33	0.72	0.05	0.06	189.18	10.97	12.82	-921.94	381.83	412.15
ARGEE	13.75	13.25	14.08	42.20	17.79	19.09	0.72	0.06	0.07	238.41	8.65	9.84	-1042.20	478.87	513.12

Table A.25: Scenario 21 ($\gamma_1 = -0.19, \gamma_2 = -0.46$, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.01	10.57	10.82	42.12	11.93	12.25	0.72	0.04	0.04	244.92	7.51	7.66	-1046.79	320.64	328.35
EGEE	7.26	14.42	17.39	41.94	17.75	19.00	0.72	0.06	0.07	244.17	9.98	12.32	-1041.07	478.33	510.62
IEE	7.17	14.21	17.54	41.85	18.96	21.71	0.72	0.07	0.08	225.47	12.17	16.43	-1008.68	511.28	583.83
CWGEE	7.06	12.82	14.12	41.89	16.53	17.91	0.72	0.06	0.06	223.41	11.06	12.72	-981.00	445.29	481.31
ARGEE	7.09	11.82	12.56	41.83	18.02	19.51	0.72	0.06	0.07	244.60	8.21	8.74	-1036.49	485.52	524.25

Table A.26: Scenario 24 ($\gamma_1 = -0.50$, $\gamma_2 = 0$, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.97	10.44	10.63	41.86	11.97	12.30	0.72	0.04	0.04	243.60	7.41	7.49	-1038.50	321.51	331.64
EGEE	7.17	16.14	23.81	41.74	16.53	17.61	0.72	0.06	0.06	206.37	11.03	15.24	-963.49	445.40	474.26
IEE	7.10	13.19	19.10	41.66	17.13	20.27	0.72	0.06	0.07	181.40	11.63	17.43	-913.47	461.95	547.46
CWGEE	7.04	12.58	14.98	41.72	15.50	16.74	0.72	0.06	0.06	192.84	11.59	13.14	-924.96	417.91	451.02
ARGEE	6.96	12.62	13.43	41.79	18.18	19.54	0.72	0.06	0.07	234.94	8.50	9.76	-1035.61	489.51	524.93

Table A.27: Scenario 25 ($\gamma_1 = 0, \gamma_2 = -0.50$, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.04	10.49	10.80	41.88	11.96	12.41	0.72	0.04	0.04	244.20	7.45	7.61	-1039.03	321.23	332.85
EGEE	7.15	15.16	20.84	41.60	16.29	17.74	0.72	0.06	0.06	217.71	10.37	13.94	-979.37	438.90	477.44
IEE	7.14	12.41	17.63	41.62	15.82	19.11	0.72	0.06	0.07	186.97	10.85	16.56	-924.84	426.88	515.72
CWGEE	7.09	11.53	13.62	41.64	13.94	15.37	0.72	0.05	0.05	193.46	10.61	12.74	-912.35	375.62	413.83
ARGEE	7.01	13.39	14.23	41.58	17.64	19.32	0.72	0.06	0.07	240.79	8.72	9.68	-1023.80	474.95	518.26

Table A.28: Scenario 26 ($\gamma_1 = -0.35, \gamma_2 = -0.35$, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.84	10.54	10.73	42.08	11.95	12.15	0.72	0.04	0.04	244.67	7.49	7.62	-1045.27	321.19	326.97
EGEE	6.85	14.67	18.54	42.25	16.95	18.09	0.72	0.06	0.07	229.88	10.11	12.86	-1021.71	456.80	486.44
IEE	6.89	13.23	17.41	42.39	17.20	20.03	0.72	0.06	0.07	202.95	11.39	16.24	-978.83	464.04	539.56
CWGEE	6.87	12.00	13.64	42.22	14.94	16.11	0.72	0.05	0.06	205.40	10.65	12.53	-952.09	402.55	433.57
ARGEE	6.85	12.68	13.28	42.25	17.68	19.04	0.72	0.06	0.07	242.42	8.49	9.22	-1043.16	476.17	512.02

Table A.29: Scenario 27 ($\gamma_1 = -0.46, \gamma_2 = -0.19$, treatment effect = 7 g/week) *one gross outlier was excluded from the ARGEE on the basis that it would be clear the value was wrong if it was obtained.*

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.98	10.46	10.72	41.75	11.99	12.31	0.72	0.04	0.04	243.81	7.43	7.55	-1035.20	322.14	330.15
EGEE	7.27	15.74	22.45	41.87	16.12	17.31	0.72	0.06	0.06	209.45	10.77	14.56	-972.47	434.30	462.82
IEE	7.23	12.44	18.10	42.04	15.86	19.02	0.72	0.06	0.07	179.41	10.98	16.90	-922.35	427.89	510.14
CWGEE	7.13	11.78	14.10	41.87	14.21	15.46	0.72	0.05	0.06	189.00	10.99	12.94	-915.17	383.02	414.03
ARGEE	7.13	13.14	13.95	41.70	17.84	19.21	0.72	0.06	0.07	238.11	8.64	9.78	-1030.50	480.42	513.41

Table A.30: Scenario 28 ($\gamma_1 = -0.19, \gamma_2 = -0.46$, treatment effect = 7 g/week)

Interaction =0

	Int	SE(Int)	SD(Int)	GA	SE(GA)	SD(GA)	BW	SE(BW)	SD(BW)	T	SE(T)	SD(T)	Intercept	SE	SD
MM	0.16	10.55	10.81	41.96	11.95	12.27	0.72	0.04	0.04	244.74	7.49	7.56	-1042.89	321.21	329.88
EGEE	0.25	14.39	17.45	41.65	17.68	19.02	0.72	0.06	0.07	244.18	9.94	12.13	-1034.01	476.50	510.80
IEE	0.23	14.17	17.59	41.69	18.84	21.65	0.72	0.07	0.08	225.50	12.12	16.16	-1004.71	508.31	582.21
CWGEE	0.19	12.77	14.19	41.64	16.46	17.94	0.72	0.06	0.06	223.33	11.01	12.52	-974.47	443.54	481.67
ARGEE	0.18	11.80	12.41	41.74	17.98	19.36	0.72	0.06	0.07	244.39	8.21	8.60	-1033.94	484.40	519.11

Table A.31: Scenario 31 ($\gamma_1 = -0.50$, $\gamma_2 = 0$, treatment effect = 0 g/week)

	Int	SE(Int)	SD(Int)	GA	SE(GA)	SD(GA)	BW	SE(BW)	SD(BW)	T	SE(T)	SD(T)	Intercept	SE	SD
MM	-0.00	10.43	10.73	41.74	11.94	12.27	0.72	0.04	0.04	243.65	7.41	7.50	-1033.99	320.82	328.94
EGEE	0.08	16.09	23.64	41.77	16.48	17.59	0.72	0.06	0.06	206.52	10.97	14.94	-963.99	443.89	472.67
IEE	0.12	13.17	19.13	41.87	17.07	20.22	0.72	0.06	0.07	181.50	11.58	17.30	-919.75	460.49	543.51
CWGEE	0.08	12.57	15.00	41.75	15.46	16.72	0.72	0.05	0.06	192.87	11.54	13.01	-925.48	416.57	449.61
ARGEE	0.02	12.57	13.40	41.64	18.14	19.60	0.72	0.06	0.07	234.97	8.49	9.71	-1030.63	488.71	527.09

Table A.32: Scenario 32 ($\gamma_1 = 0, \gamma_2 = -0.50$, treatment effect = 0 g/week)

	Int	SE(Int)	SD(Int)	GA	SE(GA)	SD(GA)	BW	SE(BW)	SD(BW)	T	SE(T)	SD(T)	Intercept	SE	SD
MM	0.03	10.51	10.75	41.84	11.96	12.31	0.72	0.04	0.04	244.24	7.47	7.51	-1038.58	321.30	330.32
EGEE	0.11	15.18	20.71	41.94	16.28	17.63	0.72	0.06	0.06	217.69	10.39	14.00	-989.41	438.46	473.32
IEE	0.02	12.43	17.46	42.04	15.84	19.14	0.72	0.06	0.07	186.95	10.82	16.60	-937.02	427.25	514.26
CWGEE	0.01	11.54	13.44	41.95	13.92	15.28	0.72	0.05	0.06	193.42	10.60	12.73	-921.51	375.00	410.53
ARGEE	0.16	13.36	14.20	41.95	17.60	18.95	0.72	0.06	0.07	240.73	8.73	9.72	-1034.74	473.86	508.47

Table A.33: Scenario 33 ($\gamma_1 = -0.35, \gamma_2 = -0.35$, treatment effect = 0 g/week)

	Int	SE(Int)	SD(Int)	GA	SE(GA)	SD(GA)	BW	SE(BW)	SD(BW)	T	SE(T)	SD(T)	Intercept	SE	SD
MM	0.19	10.55	10.81	41.79	11.97	12.23	0.72	0.04	0.04	244.58	7.49	7.61	-1038.31	321.54	329.62
EGEE	0.15	14.67	18.63	41.92	17.01	18.31	0.72	0.06	0.07	229.78	10.08	12.94	-1013.12	458.11	491.43
IEE	0.08	13.20	17.33	42.00	17.22	20.26	0.72	0.06	0.07	202.95	11.37	16.36	-968.98	464.16	543.79
CWGEE	0.07	11.99	13.60	41.94	14.99	16.34	0.72	0.05	0.06	205.36	10.63	12.68	-944.91	403.68	438.90
ARGEE	0.23	12.67	13.34	41.99	17.77	19.28	0.72	0.06	0.07	242.27	8.47	9.25	-1036.79	478.49	516.56

Table A.34: Scenario 34 ($\gamma_1 = -0.46, \gamma_2 = -0.19$, treatment effect = 0 g/week)

	Int	SE(Int)	SD(Int)	GA	SE(GA)	SD(GA)	BW	SE(BW)	SD(BW)	T	SE(T)	SD(T)	Intercept	SE	SD
MM	-0.12	10.47	10.66	42.06	11.99	12.28	0.72	0.04	0.04	243.89	7.44	7.49	-1043.92	322.14	330.43
EGEE	0.05	15.90	22.87	42.18	16.16	17.18	0.72	0.06	0.06	209.34	10.87	14.95	-980.80	435.01	461.95
IEE	-0.03	12.48	18.38	42.25	15.89	19.07	0.72	0.06	0.07	179.14	11.05	17.13	-927.40	428.71	513.41
CWGEE	-0.09	11.84	14.22	42.16	14.23	15.35	0.72	0.05	0.06	188.87	11.09	13.03	-922.55	383.45	412.88
ARGEE	-0.07	13.29	14.11	42.13	17.86	19.15	0.72	0.06	0.07	238.29	8.72	9.75	-1042.28	480.78	514.29

Table A.35: Scenario 35 ($\gamma_1 = -0.19, \gamma_2 = -0.46$, treatment effect = 0 g/week)

A.2 Extension 1: Simulation With Larger Standard Deviation of Cluster Size, $\gamma_0 = 2.787813$

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.06	10.69	10.87	41.95	12.02	12.31	0.72	0.04	0.04	245.44	7.60	7.66	-1043.43	322.78	331.19
EGEE	27.96	17.25	28.00	41.94	18.16	19.52	0.72	0.06	0.07	244.61	11.59	19.93	-1041.49	488.95	523.95
IEE	27.97	15.72	26.31	42.08	23.38	29.97	0.72	0.08	0.11	228.53	12.85	24.49	-1032.30	630.61	806.55
CWGEE	28.08	14.48	18.73	41.93	16.23	17.80	0.72	0.06	0.06	223.03	12.65	18.02	-981.20	437.11	478.01
ARGEE	28.09	12.56	13.27	41.95	17.78	19.17	0.72	0.06	0.07	246.22	8.65	9.14	-1037.23	478.70	514.17

Table A.36: Scenario 36 (γ_1=-0.90, γ_2=0, treatment effect =28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.08	10.42	10.69	42.01	11.99	12.39	0.72	0.04	0.04	242.49	7.40	7.64	-1040.62	322.14	332.88
EGEE	27.71	14.21	31.44	41.92	14.70	15.83	0.72	0.05	0.06	184.12	9.73	20.53	-928.41	395.96	423.92
IEE	27.88	9.53	21.32	41.97	15.87	20.92	0.72	0.06	0.07	162.45	7.57	19.05	-891.52	428.02	563.49
CWGEE	27.90	11.37	17.06	41.92	13.44	14.81	0.72	0.05	0.05	173.23	10.27	15.63	-899.67	362.11	398.23
ARGEE	28.04	13.33	14.07	41.81	17.50	18.82	0.72	0.06	0.07	232.64	8.85	9.59	-1034.62	471.11	505.41

Table A.37: Scenario 37 (γ_2=-0.90, γ_1=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.91	10.59	10.64	41.80	12.02	12.39	0.72	0.04	0.04	243.88	7.53	7.67	-1036.60	322.98	334.80
EGEE	27.93	15.85	29.77	41.87	14.38	15.81	0.72	0.05	0.06	202.11	10.69	19.98	-960.35	387.34	425.76
IEE	27.93	10.39	21.06	42.28	14.96	21.00	0.72	0.05	0.07	175.50	8.66	19.02	-935.78	404.19	566.27
CWGEE	27.89	10.37	15.09	41.91	10.75	12.20	0.72	0.04	0.04	179.31	9.49	14.57	-899.13	289.72	329.03
ARGEE	27.83	13.88	14.30	41.73	16.53	17.97	0.72	0.06	0.06	239.27	9.08	9.49	-1026.04	445.16	483.97

Table A.38: Scenario 38 (γ_1=-0.64, γ_2=-0.64, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.23	10.65	10.94	42.04	12.00	12.29	0.72	0.04	0.04	244.60	7.57	7.77	-1043.13	322.48	329.58
EGEE	27.81	16.89	29.22	41.76	16.30	17.26	0.72	0.06	0.06	221.90	11.33	20.13	-992.33	439.00	462.73
IEE	27.84	13.51	24.45	41.67	19.32	25.52	0.72	0.07	0.09	198.93	11.15	22.02	-964.23	521.62	686.85
CWGEE	27.99	12.46	17.01	41.74	13.23	14.41	0.72	0.05	0.05	197.89	11.07	16.16	-926.03	356.49	387.32
ARGEE	28.21	13.22	13.83	41.88	16.94	18.01	0.72	0.06	0.07	242.03	8.83	9.40	-1029.13	455.91	483.74

Table A.39: Scenario 39 (γ_1=-0.83, γ_2=-0.34, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.97	10.48	10.62	41.93	12.01	12.30	0.72	0.04	0.04	243.07	7.45	7.62	-1038.93	322.62	330.28
EGEE	28.24	14.76	31.03	41.70	13.66	14.87	0.72	0.05	0.05	188.50	10.04	20.56	-930.58	368.10	399.64
IEE	28.12	8.38	19.89	41.63	13.18	18.51	0.72	0.05	0.07	162.53	6.87	18.42	-888.39	355.42	499.40
CWGEE	28.11	9.87	15.33	41.75	10.85	12.17	0.72	0.04	0.04	170.49	9.16	15.02	-882.98	292.37	328.16
ARGEE	28.15	13.86	14.34	41.90	16.81	18.28	0.72	0.06	0.07	235.77	9.08	9.46	-1033.34	452.14	490.29

Table A.40: Scenario 40 (γ_2=-0.83, γ_1=-0.34, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.92	10.68	10.83	41.80	12.00	12.39	0.72	0.04	0.04	245.52	7.59	7.70	-1039.29	322.39	332.98
EGEE	20.71	17.30	28.33	41.53	18.14	19.54	0.72	0.06	0.07	244.74	11.62	20.14	-1030.86	488.85	525.25
IEE	20.71	15.79	26.57	41.40	23.25	29.98	0.72	0.08	0.11	228.76	12.89	24.71	-1016.00	627.75	807.38
CWGEE	20.75	14.51	18.74	41.53	16.20	17.81	0.72	0.06	0.06	223.25	12.67	18.21	-971.12	436.63	479.42
ARGEE	20.90	12.55	13.42	41.58	17.75	19.22	0.72	0.06	0.07	246.37	8.62	9.29	-1027.44	478.26	516.73

Table A.41: Scenario 41 (γ_1=-0.90, γ_2=0, treatment effect =21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.91	10.42	10.68	42.16	12.00	12.31	0.72	0.04	0.04	242.55	7.41	7.64	-1043.23	322.37	329.89
EGEE	20.89	14.25	31.50	42.13	14.65	15.70	0.72	0.05	0.06	183.84	9.78	20.83	-932.55	394.70	420.71
IEE	20.90	9.55	21.24	42.22	15.82	20.67	0.72	0.06	0.07	162.22	7.62	19.29	-895.97	426.76	556.78
CWGEE	20.90	11.42	17.03	42.13	13.39	14.72	0.72	0.05	0.05	173.07	10.33	15.92	-903.82	360.96	394.91
ARGEE	20.86	13.36	13.90	42.08	17.45	18.80	0.72	0.06	0.07	232.71	8.88	9.62	-1041.23	469.96	503.81

Table A.42: Scenario 42 (γ_2=-0.90, γ_1=0, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.00	10.57	10.74	42.16	12.02	12.34	0.72	0.04	0.04	243.85	7.52	7.64	-1046.09	322.87	330.84
EGEE	20.72	15.88	29.92	42.09	14.39	15.76	0.72	0.05	0.06	202.16	10.68	19.96	-967.08	387.58	422.47
IEE	20.81	10.39	21.12	42.14	14.93	21.15	0.72	0.05	0.07	175.46	8.65	18.93	-933.36	403.13	568.94
CWGEE	20.84	10.38	15.19	41.97	10.76	12.20	0.72	0.04	0.04	179.32	9.48	14.52	-901.90	290.18	327.87
ARGEE	20.99	13.92	14.41	41.95	16.56	17.86	0.72	0.06	0.06	239.26	9.07	9.40	-1032.84	445.71	479.89

Table A.43: Scenario 43 (γ_1=-0.64, γ_2=-0.64, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.93	10.65	10.79	41.93	12.01	12.58	0.72	0.04	0.04	244.67	7.57	7.70	-1041.80	322.64	337.13
EGEE	21.40	16.86	28.56	41.83	16.28	17.89	0.72	0.06	0.06	221.40	11.35	19.72	-997.02	438.43	479.59
IEE	21.31	13.48	23.82	41.88	19.28	25.84	0.72	0.07	0.09	198.36	11.10	21.48	-973.12	520.65	695.39
CWGEE	21.17	12.41	16.53	41.82	13.20	14.88	0.72	0.05	0.05	197.52	11.05	15.74	-930.56	355.65	399.63
ARGEE	21.01	13.24	13.61	41.84	16.94	18.44	0.72	0.06	0.07	242.13	8.86	9.31	-1029.74	455.86	493.99

Table A.44: Scenario 44 (γ_1=-0.83, γ_2=-0.34, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.05	10.50	10.60	41.88	12.01	12.27	0.72	0.04	0.04	242.99	7.46	7.50	-1037.47	322.58	329.39
EGEE	21.55	14.68	30.83	42.08	13.65	14.78	0.72	0.05	0.05	188.41	10.01	20.57	-941.99	367.58	395.52
IEE	21.37	8.35	19.74	42.36	13.15	18.35	0.72	0.05	0.07	162.46	6.86	18.35	-909.75	354.74	493.50
CWGEE	21.26	9.85	15.22	42.02	10.85	12.10	0.72	0.04	0.04	170.40	9.16	14.92	-891.16	292.21	324.56
ARGEE	21.17	13.85	14.31	41.89	16.77	18.15	0.72	0.06	0.07	235.74	9.09	9.32	-1033.80	451.33	485.88

Table A.45: Scenario 45 (γ_2=-0.83, γ_1=-0.34, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.91	10.67	10.87	41.85	11.98	12.20	0.72	0.04	0.04	245.69	7.58	7.65	-1040.60	321.96	328.30
EGEE	14.22	17.26	28.52	41.85	18.09	19.10	0.72	0.06	0.07	244.71	11.57	20.08	-1039.31	487.32	512.54
IEE	14.17	15.74	26.78	42.04	23.17	29.59	0.72	0.08	0.11	228.66	12.81	24.64	-1032.30	625.47	795.36
CWGEE	14.06	14.48	18.87	41.88	16.15	17.39	0.72	0.06	0.06	223.26	12.64	18.04	-980.44	435.10	467.34
ARGEE	14.01	12.56	13.27	41.81	17.70	18.90	0.72	0.06	0.07	246.50	8.63	9.07	-1034.42	476.72	507.11

Table A.46: Scenario 46 (γ_1=-0.90, γ_2=0, treatment effect =14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.15	10.41	10.62	41.95	11.98	12.26	0.72	0.04	0.04	242.28	7.40	7.58	-1039.04	321.90	329.81
EGEE	14.30	14.18	31.17	41.92	14.64	15.50	0.72	0.05	0.06	183.77	9.74	20.78	-929.70	394.28	416.42
IEE	14.17	9.55	21.07	41.78	15.80	20.51	0.72	0.06	0.07	162.21	7.62	19.41	-886.83	426.08	554.04
CWGEE	14.14	11.38	16.86	41.87	13.39	14.54	0.72	0.05	0.05	173.00	10.32	15.94	-899.52	360.74	391.99
ARGEE	14.15	13.30	13.99	42.07	17.45	18.62	0.72	0.06	0.07	232.40	8.86	9.58	-1043.80	469.63	498.84

Table A.47: Scenario 47 (γ_2=-0.90, γ_1=0, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.79	10.58	10.91	41.97	12.02	12.30	0.72	0.04	0.04	243.91	7.52	7.79	-1042.04	323.10	330.53
EGEE	14.03	15.83	30.01	42.02	14.36	15.61	0.72	0.05	0.06	201.76	10.70	20.34	-964.71	387.02	419.61
IEE	14.01	10.36	21.24	42.39	14.96	21.04	0.72	0.05	0.07	175.11	8.64	19.34	-938.49	404.00	566.20
CWGEE	13.98	10.38	15.25	42.01	10.72	12.07	0.72	0.04	0.04	178.97	9.51	14.85	-902.25	289.01	324.99
ARGEE	13.81	13.87	14.36	41.81	16.56	17.69	0.72	0.06	0.06	239.21	9.09	9.52	-1029.00	445.90	475.07

Table A.48: Scenario 48 (γ_1=-0.64, γ_2=-0.64, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.95	10.65	10.89	41.81	12.02	12.16	0.72	0.04	0.04	244.64	7.57	7.72	-1038.12	323.08	325.78
EGEE	13.61	16.75	28.61	41.83	16.27	17.42	0.72	0.06	0.06	221.88	11.22	19.95	-995.88	438.43	466.36
IEE	13.70	13.38	23.82	41.95	19.29	25.38	0.72	0.07	0.09	198.73	11.03	21.66	-974.07	520.57	682.17
CWGEE	13.80	12.33	16.72	41.88	13.19	14.55	0.72	0.05	0.05	197.67	10.95	16.01	-931.03	355.52	390.04
ARGEE	13.91	13.20	13.78	41.86	16.95	17.99	0.72	0.06	0.07	242.13	8.82	9.33	-1029.15	456.37	481.44

Table A.49: Scenario 49 (γ_1=-0.83, γ_2=-0.34, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.83	10.47	10.71	42.04	12.00	12.32	0.72	0.04	0.04	243.14	7.44	7.59	-1042.41	322.45	331.35
EGEE	13.94	14.72	30.76	42.03	13.64	14.78	0.72	0.05	0.05	188.68	10.02	20.61	-941.23	367.39	396.29
IEE	13.93	8.37	19.77	42.12	13.12	18.50	0.72	0.05	0.07	162.74	6.82	18.45	-903.71	353.87	499.57
CWGEE	13.95	9.87	15.21	42.02	10.82	12.09	0.72	0.04	0.04	170.64	9.12	15.04	-891.66	291.61	325.41
ARGEE	13.84	13.86	14.30	42.05	16.80	18.20	0.72	0.06	0.07	235.96	9.06	9.39	-1038.97	452.25	489.35

Table A.50: Scenario 50 (γ_2=-0.83, γ_1=-0.34, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.98	10.68	10.93	41.90	11.98	12.34	0.72	0.04	0.04	245.49	7.59	7.72	-1041.84	321.90	331.95
EGEE	6.99	17.30	28.64	41.85	18.10	19.77	0.72	0.06	0.07	244.63	11.59	20.06	-1040.14	487.85	531.60
IEE	7.09	15.83	26.96	42.13	23.15	30.37	0.72	0.08	0.11	228.51	12.85	24.43	-1037.45	625.69	817.99
CWGEE	7.07	14.55	19.05	41.88	16.18	18.07	0.72	0.06	0.06	223.10	12.66	17.98	-980.89	436.25	486.12
ARGEE	6.93	12.54	13.47	41.77	17.71	19.53	0.72	0.06	0.07	246.36	8.60	9.21	-1033.29	477.45	524.74

Table A.51: Scenario 51 (γ_1=-0.90, γ_2=0, treatment effect =7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.03	10.40	10.73	42.11	11.98	12.31	0.72	0.04	0.04	242.48	7.39	7.64	-1043.03	321.85	330.85
EGEE	6.24	14.15	31.34	42.00	14.66	15.71	0.72	0.05	0.06	184.55	9.68	20.65	-931.93	394.86	421.43
IEE	6.53	9.52	21.34	42.05	15.84	20.77	0.72	0.06	0.07	162.86	7.54	19.23	-895.36	427.47	559.82
CWGEE	6.69	11.35	17.07	42.00	13.39	14.67	0.72	0.05	0.05	173.49	10.20	15.88	-903.07	361.02	394.68
ARGEE	6.94	13.32	13.95	41.97	17.44	18.78	0.72	0.06	0.07	232.73	8.83	9.54	-1039.70	469.60	503.74

Table A.52: Scenario 52 (γ_2=-0.90, γ_1=0, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.80	10.58	10.78	41.89	12.05	12.25	0.72	0.04	0.04	243.84	7.52	7.75	-1039.13	323.86	328.32
EGEE	6.98	15.80	29.11	41.73	14.39	15.50	0.72	0.05	0.06	202.09	10.67	19.85	-957.88	387.92	416.90
IEE	7.01	10.37	20.60	41.73	14.94	20.65	0.72	0.05	0.07	175.46	8.64	18.94	-922.34	403.65	558.24
CWGEE	6.98	10.32	14.83	41.80	10.75	11.96	0.72	0.04	0.04	179.18	9.43	14.62	-897.09	290.24	322.74
ARGEE	6.84	13.78	14.07	42.00	16.58	17.79	0.72	0.06	0.06	239.22	9.03	9.41	-1034.15	446.55	477.82

Table A.53: Scenario 53 (γ_1=-0.64, γ_2=-0.64, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.89	10.67	10.89	41.92	12.00	12.36	0.72	0.04	0.04	244.72	7.58	7.70	-1040.84	322.58	332.02
EGEE	6.80	16.80	28.51	41.67	16.20	17.42	0.72	0.06	0.06	221.77	11.31	19.85	-991.66	436.65	468.54
IEE	6.78	13.45	23.80	41.59	19.20	25.44	0.72	0.07	0.09	198.72	11.10	21.63	-964.17	519.05	686.84
CWGEE	6.90	12.38	16.57	41.73	13.13	14.44	0.72	0.05	0.05	197.72	11.01	15.89	-927.80	354.12	388.97
ARGEE	6.83	13.18	13.67	41.81	16.83	17.99	0.72	0.06	0.07	242.24	8.81	9.28	-1029.37	453.39	482.19

Table A.54: Scenario 54 (γ_1=-0.83, γ_2=-0.34, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.03	10.49	10.64	42.12	12.03	12.38	0.72	0.04	0.04	243.03	7.45	7.55	-1044.52	323.16	332.01
EGEE	6.32	14.75	30.74	42.25	13.68	14.71	0.72	0.05	0.05	189.05	9.99	20.35	-947.04	368.68	395.80
IEE	6.55	8.39	19.75	42.39	13.19	18.28	0.72	0.05	0.06	162.94	6.82	18.22	-910.91	356.18	494.40
CWGEE	6.71	9.87	15.21	42.20	10.87	12.04	0.72	0.04	0.04	170.74	9.05	14.91	-896.59	293.19	324.95
ARGEE	6.82	13.82	14.22	42.32	16.79	18.05	0.72	0.06	0.06	235.94	9.02	9.45	-1045.37	452.22	485.16

Table A.55: Scenario 55 (γ_2=-0.83, γ_1=-0.34, treatment effect = 7 g/week)

Interaction =0

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.03	10.68	10.86	41.55	11.99	12.22	0.72	0.04	0.04	245.45	7.59	7.61	-1033.29	322.06	328.48
EGEE	7.12	17.23	29.07	41.43	18.10	19.38	0.72	0.06	0.07	244.33	11.51	20.48	-1029.35	487.50	518.92
IEE	7.09	15.75	27.24	41.55	23.14	29.92	0.72	0.08	0.11	228.20	12.77	24.98	-1021.01	624.87	804.07
CWGEE	7.02	14.50	19.18	41.53	16.14	17.71	0.72	0.06	0.06	222.82	12.63	18.29	-971.82	434.87	474.62
ARGEE	6.98	12.58	13.42	41.38	17.72	19.13	0.72	0.06	0.07	246.25	8.63	9.19	-1022.76	477.31	512.82

Table A.56: Scenario 56 (γ_1=-0.90, γ_2=0, treatment effect =0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.01	10.42	10.66	42.04	11.97	12.27	0.72	0.04	0.04	242.45	7.40	7.73	-1040.96	321.81	329.28
EGEE	0.22	14.17	31.15	41.91	14.67	15.78	0.72	0.05	0.06	183.73	9.70	20.74	-928.74	395.41	422.61
IEE	0.09	9.51	21.23	41.98	15.83	20.95	0.72	0.06	0.07	162.30	7.54	19.29	-891.59	426.96	563.34
CWGEE	0.13	11.37	16.95	41.91	13.42	14.79	0.72	0.05	0.05	173.01	10.24	15.91	-899.75	361.80	396.60
ARGEE	-0.02	13.33	13.98	41.90	17.50	19.00	0.72	0.06	0.07	232.57	8.88	9.71	-1038.70	471.20	509.36

Table A.57: Scenario 57 (γ_2=-0.90, γ_1=0, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.03	10.58	10.78	41.97	12.03	12.24	0.72	0.04	0.04	243.85	7.52	7.59	-1040.80	323.36	328.84
EGEE	0.33	15.80	29.91	42.00	14.36	15.62	0.72	0.05	0.06	201.91	10.69	20.19	-963.89	386.96	419.83
IEE	0.19	10.34	21.18	42.13	14.86	20.85	0.72	0.05	0.07	175.31	8.62	19.22	-932.36	401.74	561.97
CWGEE	0.14	10.34	15.17	41.99	10.71	12.10	0.72	0.04	0.04	179.14	9.48	14.70	-901.60	288.98	325.91
ARGEE	0.04	13.88	14.32	41.84	16.54	17.72	0.72	0.06	0.06	239.21	9.10	9.41	-1029.05	445.36	475.80

Table A.58: Scenario 58 (γ_1=-0.64, γ_2=-0.64, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.19	10.66	10.87	41.92	12.01	12.39	0.72	0.04	0.05	244.54	7.58	7.64	-1042.06	322.69	332.35
EGEE	0.40	16.84	28.82	42.14	16.27	17.65	0.72	0.06	0.06	221.46	11.25	19.67	-1004.55	438.36	472.69
IEE	0.34	13.44	24.04	42.45	19.33	26.08	0.72	0.07	0.09	198.40	11.06	21.40	-985.85	521.70	700.70
CWGEE	0.29	12.38	16.81	42.14	13.21	14.72	0.72	0.05	0.05	197.50	10.98	15.81	-938.26	356.00	394.93
ARGEE	0.19	13.26	13.73	42.16	16.93	18.14	0.72	0.06	0.07	242.06	8.84	9.33	-1038.58	455.81	486.08

Table A.59: Scenario 59 (γ_1=-0.83, γ_2=-0.34, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.04	10.46	10.61	42.06	12.03	12.28	0.72	0.04	0.04	242.99	7.43	7.58	-1043.18	323.28	330.20
EGEE	0.08	14.69	30.89	42.04	13.70	14.86	0.72	0.05	0.05	188.70	10.04	20.65	-941.21	368.90	397.69
IEE	-0.00	8.36	19.87	42.11	13.16	18.41	0.72	0.05	0.07	162.74	6.85	18.46	-903.08	355.01	495.43
CWGEE	0.01	9.87	15.24	42.04	10.87	12.11	0.72	0.04	0.04	170.62	9.15	14.95	-892.16	293.07	325.41
ARGEE	0.01	13.84	14.37	42.15	16.81	18.20	0.72	0.06	0.07	235.82	9.05	9.42	-1040.72	452.50	487.46

Table A.60: Scenario 60 (γ_2=-0.83, γ_1=-0.34, treatment effect = 0 g/week)

Appendix B

Results Plots

B.1 Plots for Chapter 4 Simulation Study Results

Boxplots of Bias in Treatment Effect

NICS Scenarios

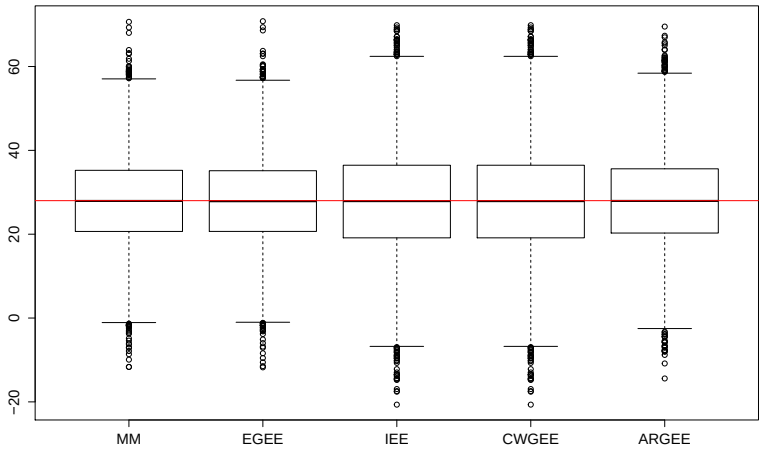

Figure B.1: Boxplots of Interaction Effect Estimates for scenario 1

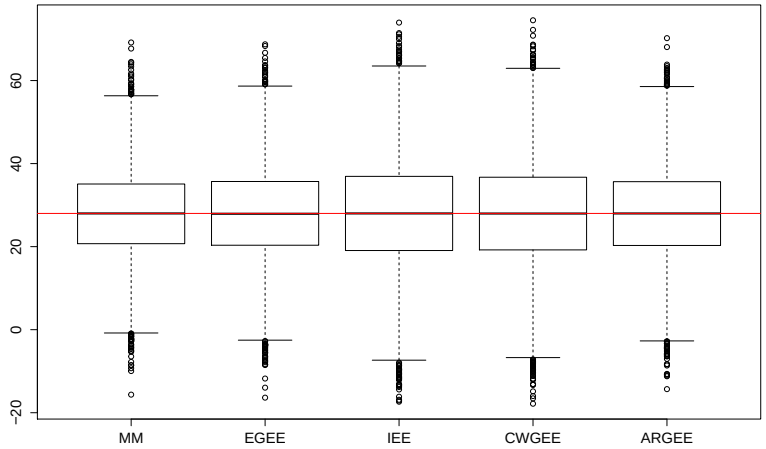

Figure B.2: Boxplots of Interaction Effect Estimates for scenario 2

ICS Scenarios

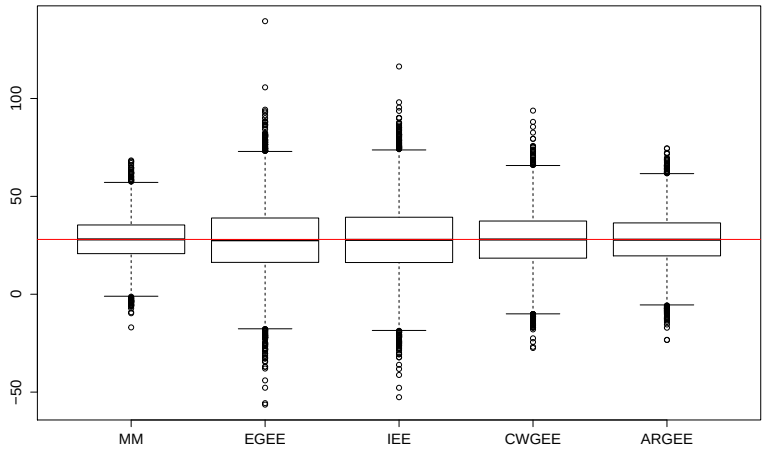

Figure B.3: Boxplots of Interaction Effect Estimates for scenario 3

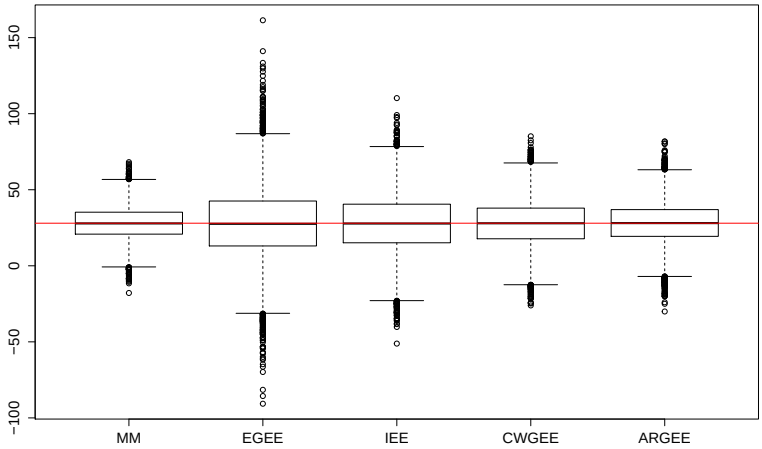

Figure B.4: Boxplots of Interaction Effect Estimates for scenario 4

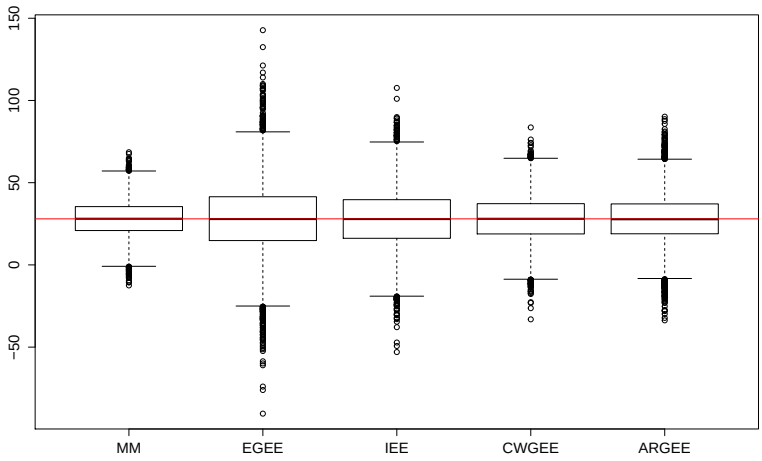

Figure B.5: Boxplots of Interaction Effect Estimates for scenario 5

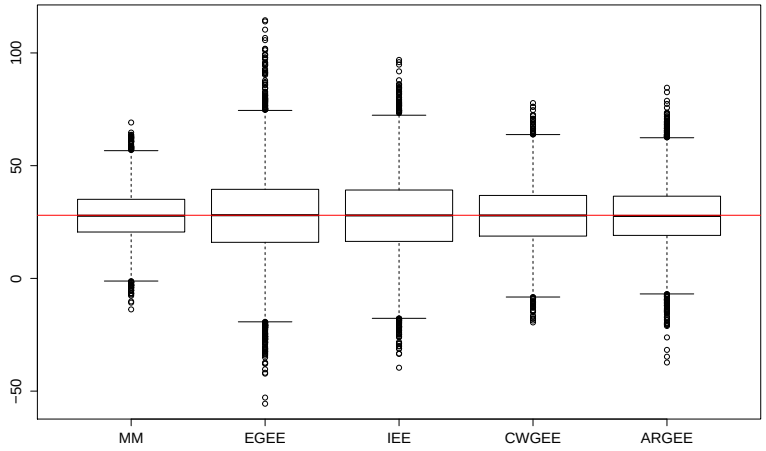

Figure B.6: Boxplots of Interaction Effect Estimates for scenario 6

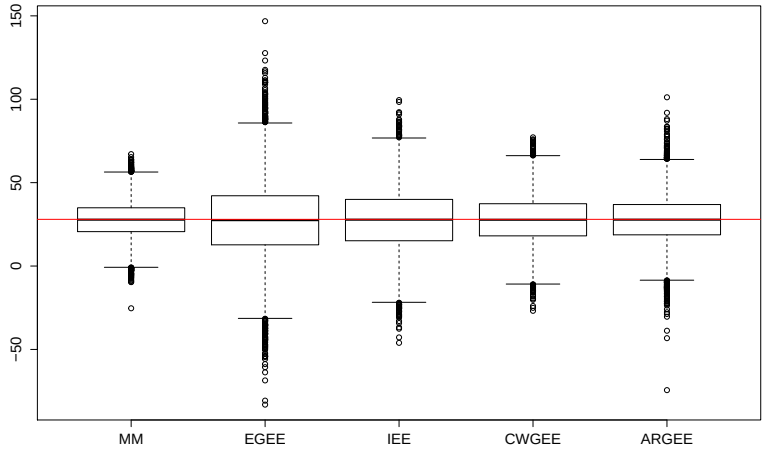

Figure B.7: Boxplots of Interaction Effect Estimates for scenario 7

Boxplots of Standard Error / Standard Deviation

NICS Scenarios

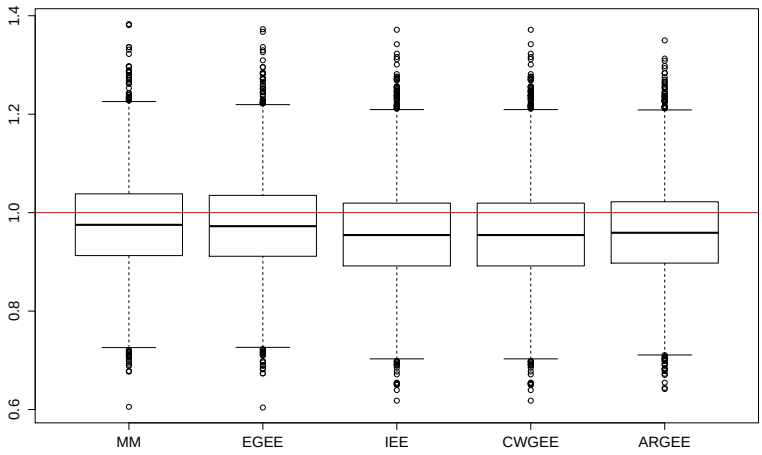

Figure B.8: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 1

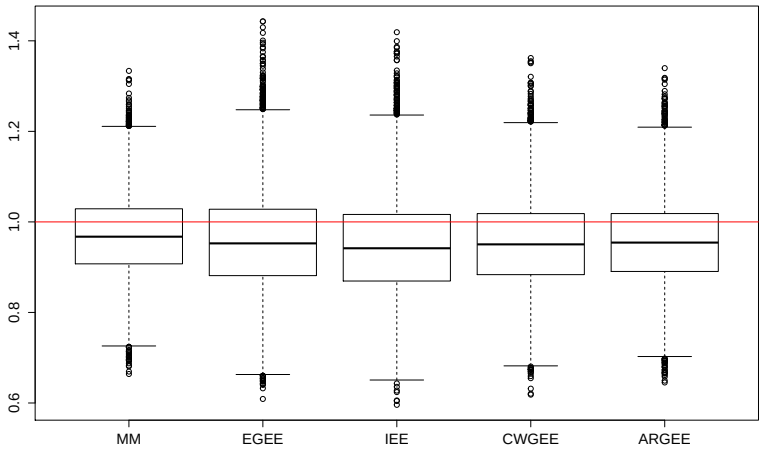

Figure B.9: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 2

ICS Scenrios

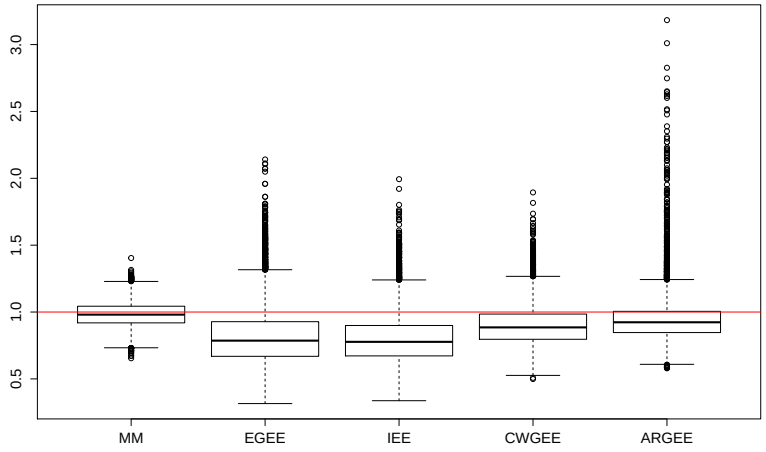

Figure B.10: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 3

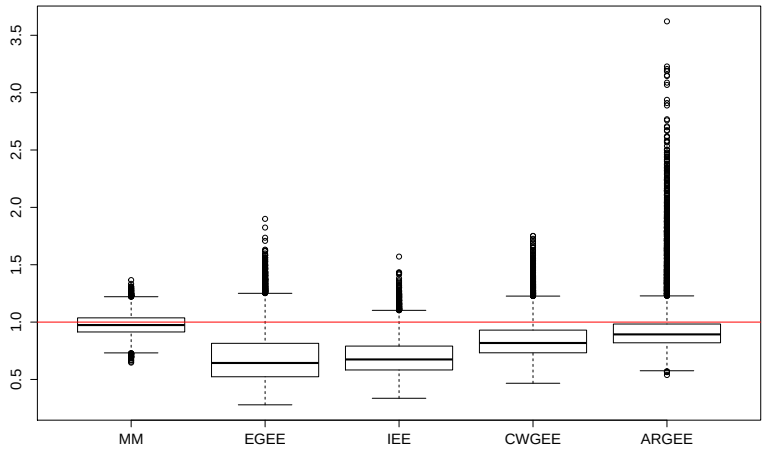

Figure B.11: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 4

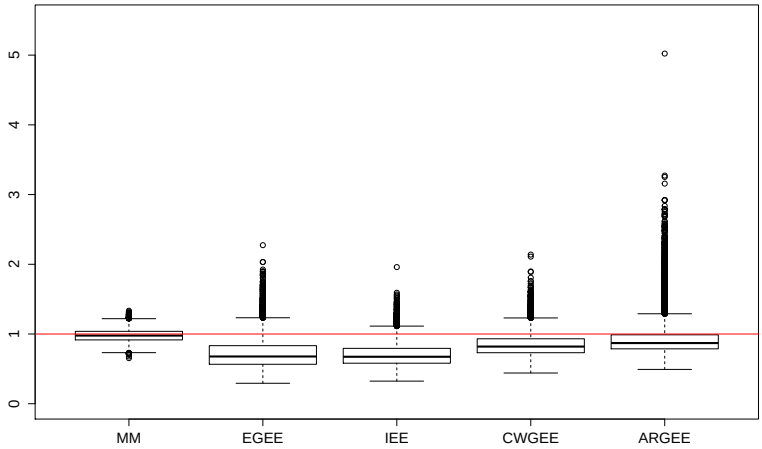

Figure B.12: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 5

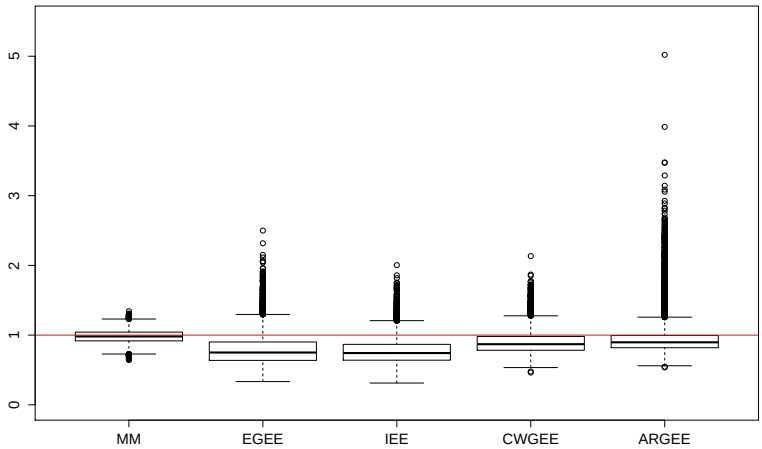

Figure B.13: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 6

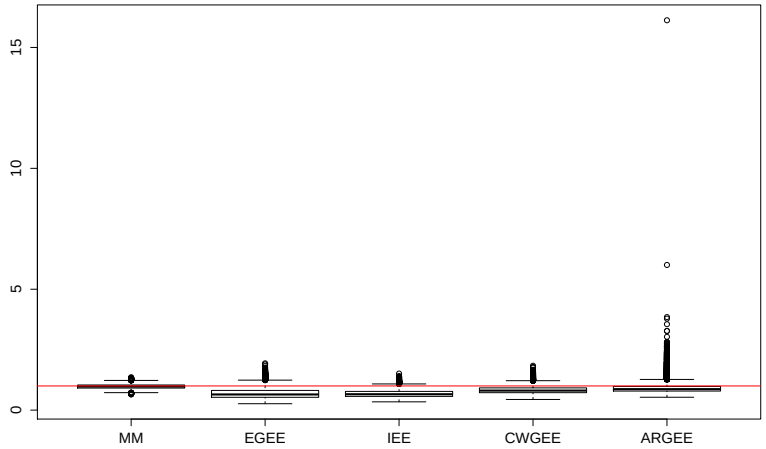

Figure B.14: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 7

B.2 Extension 1: Plots for Larger Standard Deviation of Cluster Size

Boxplots of Bias in Treatment Effect

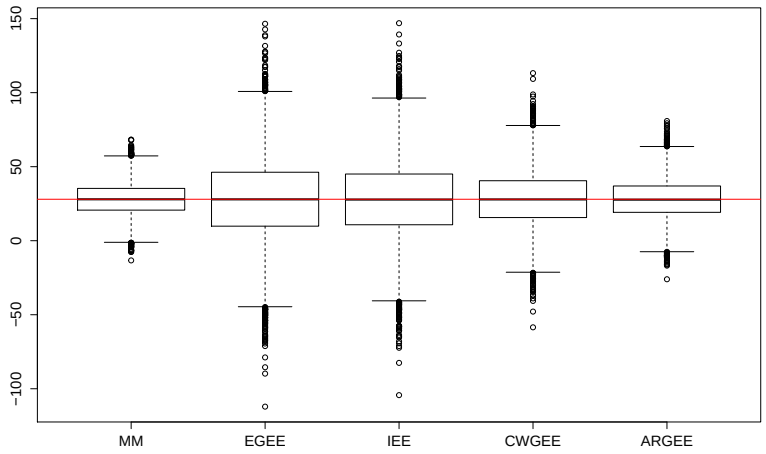

Figure B.15: Extension 1: Boxplots of Interaction Effect Estimates for scenario 3

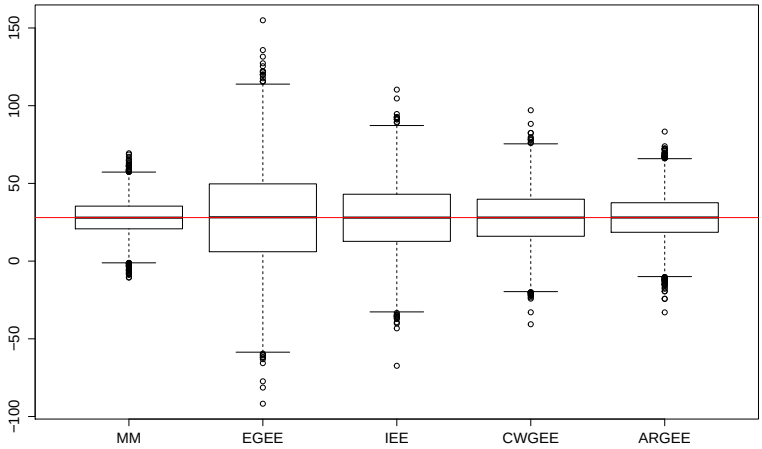

Figure B.16: Extension 1: Boxplots of Interaction Effect Estimates for scenario 4

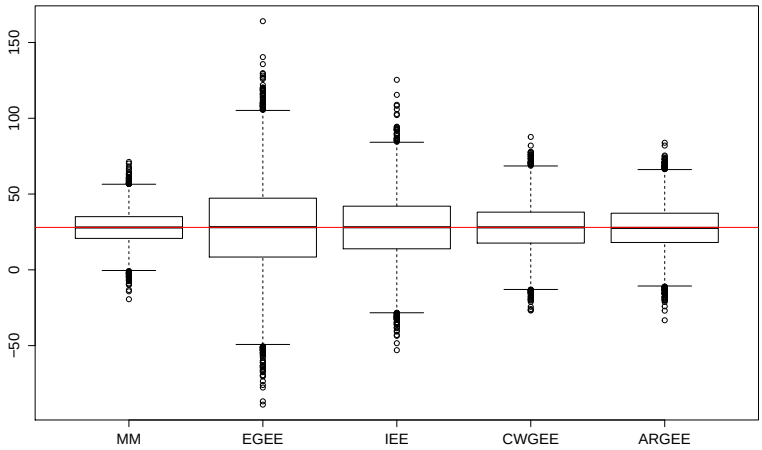

Figure B.17: Extension 1: Boxplots of Interaction Effect Estimates for scenario 5

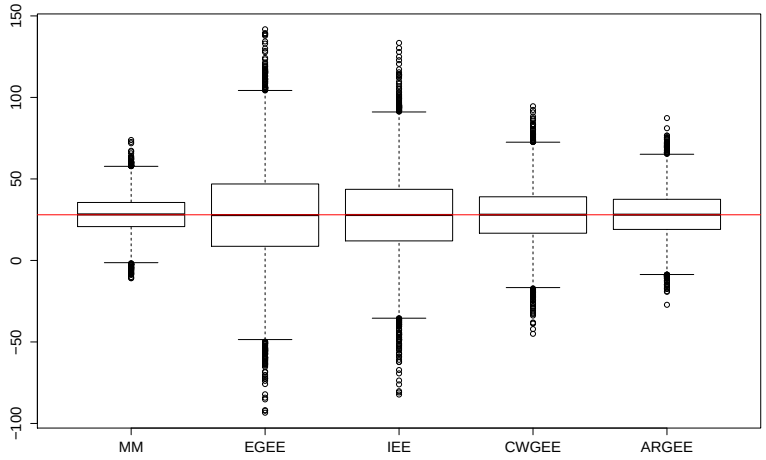

Figure B.18: Extension 1: Boxplots of Interaction Effect Estimates for scenario 6

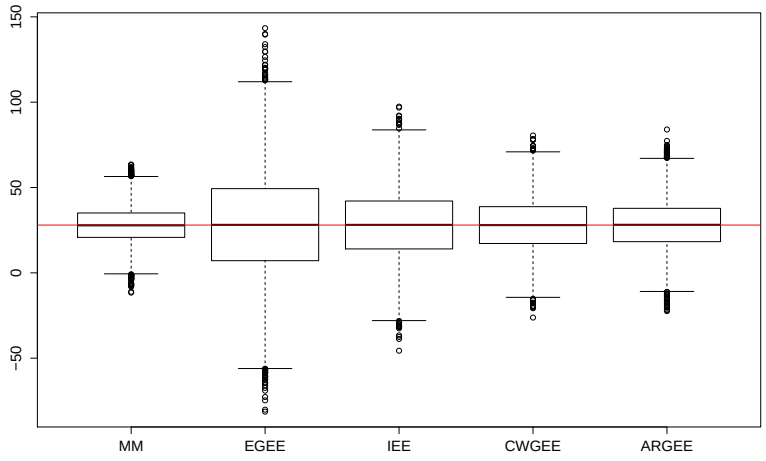

Figure B.19: Extension 1: Boxplots of Interaction Effect Estimates for scenario 7

Boxplots of Standard Error / Standard Deviation

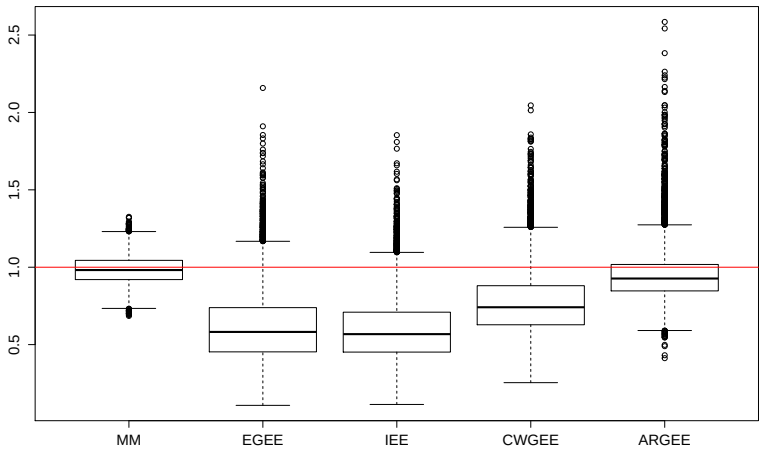

Figure B.20: Extension 1: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 3

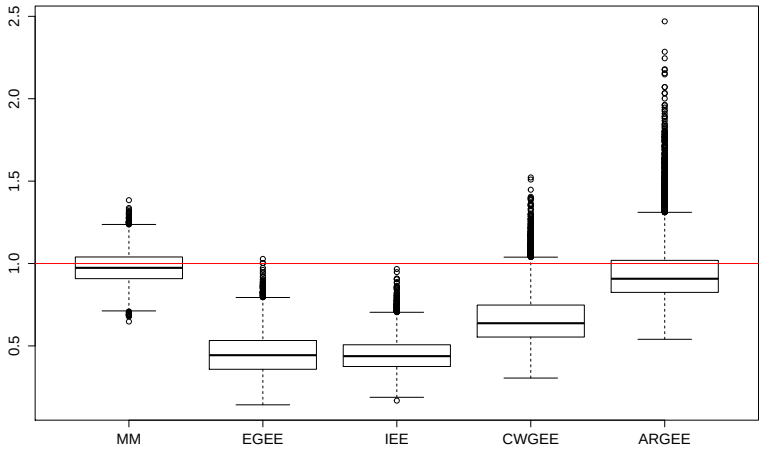

Figure B.21: Extension 1: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 4

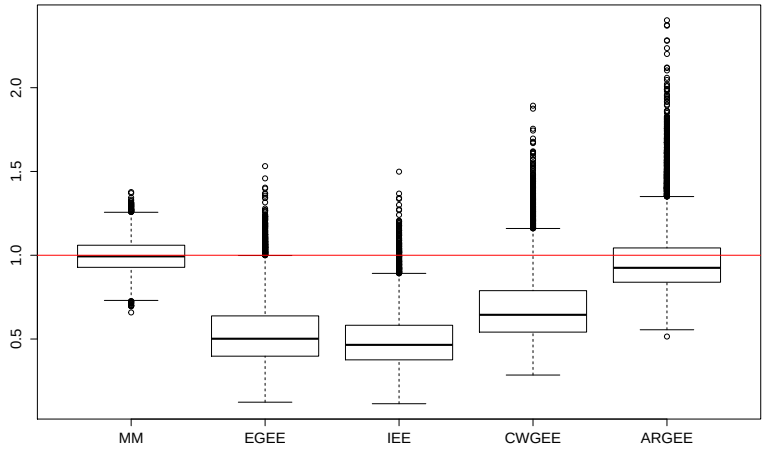

Figure B.22: Extension 1: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 5

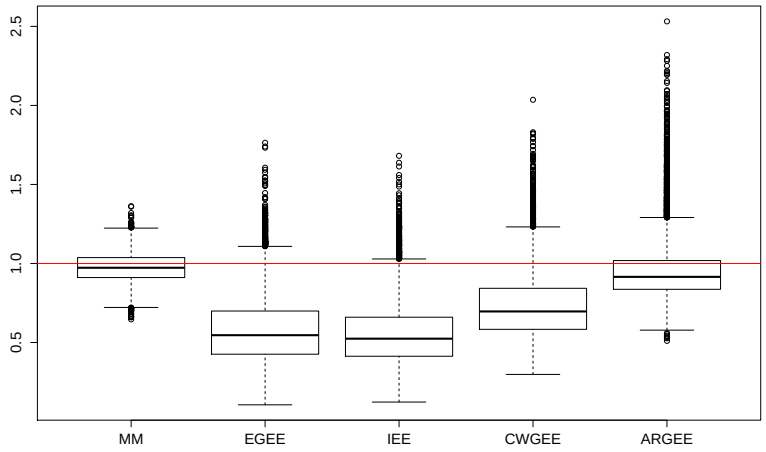

Figure B.23: Extension 1: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 6

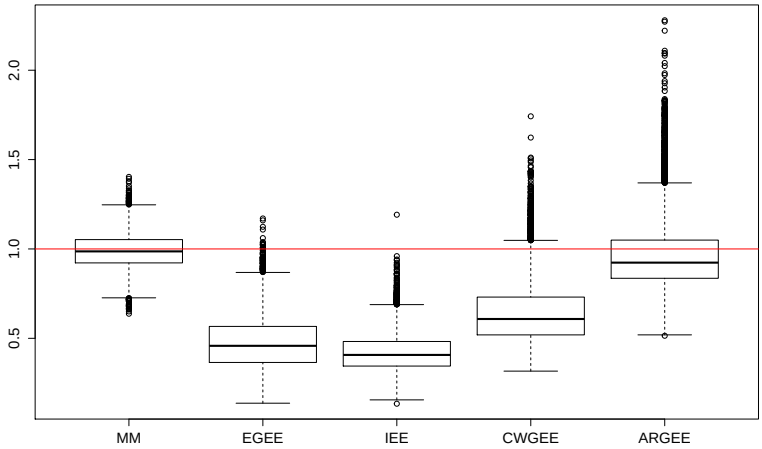

Figure B.24: Extension 1: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 7

B.3 Extension 2: Plots for Larger Sample Size Results

Boxplots of Bias in Treatment Effect

NICS Scenarios

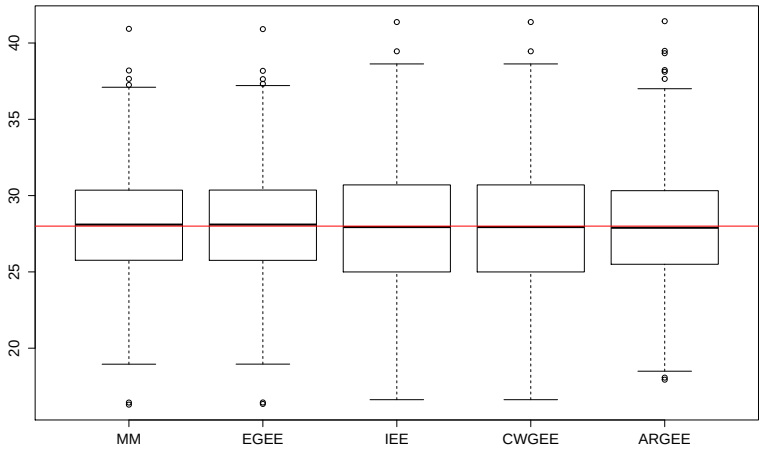

Figure B.25: Extension 2: Boxplots of Interaction Effect Estimates for scenario 1

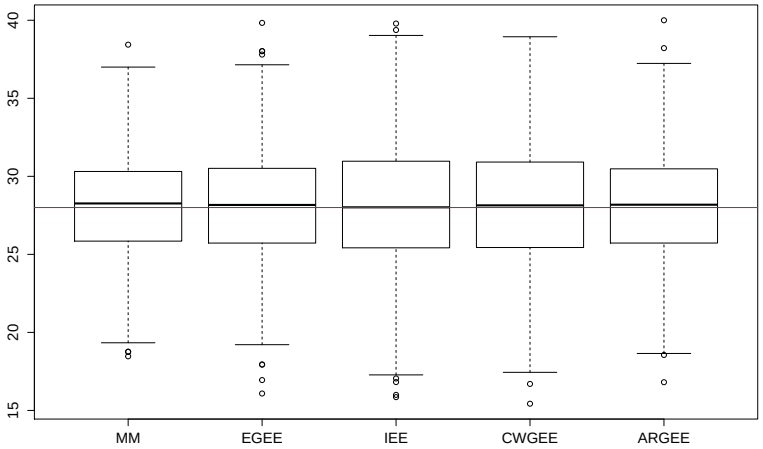

Figure B.26: Extension 2: Boxplots of Interaction Effect Estimates for scenario 2

ICS Scenrios Smaller SD of Cluster Size

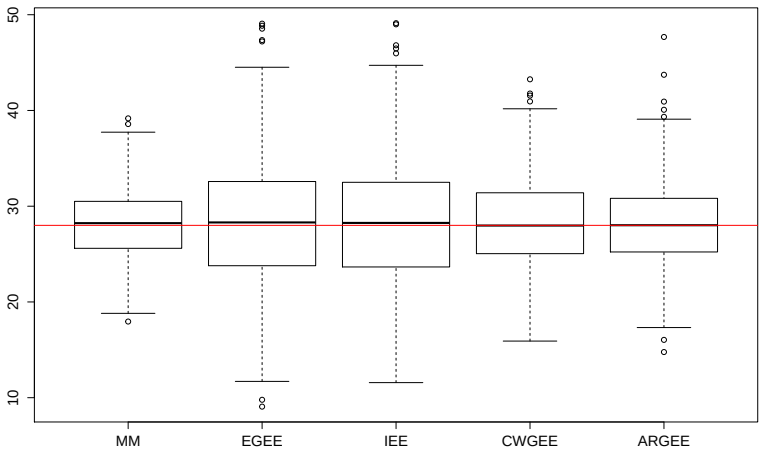

Figure B.27: Extension 2: Boxplots of Interaction Effect Estimates for scenario 3

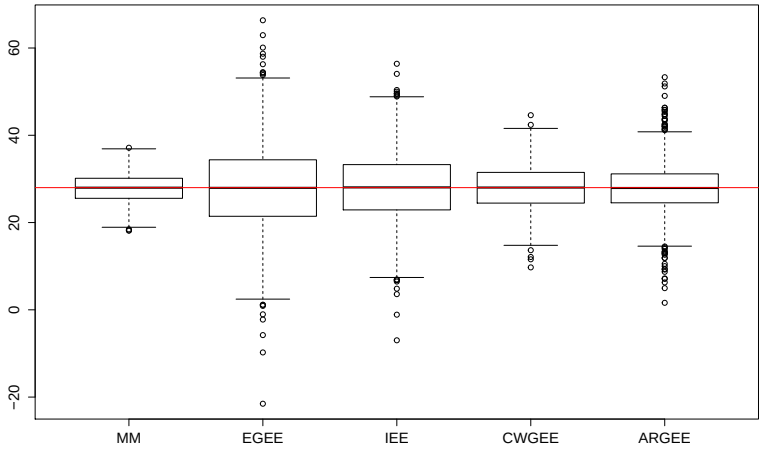

Figure B.28: Extension 2: Boxplots of Interaction Effect Estimates for scenario 4

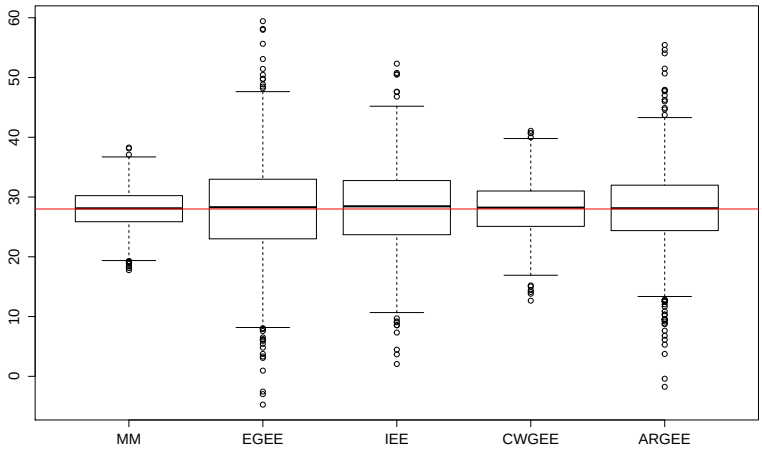

Figure B.29: Extension 2: Boxplots of Interaction Effect Estimates for scenario 5

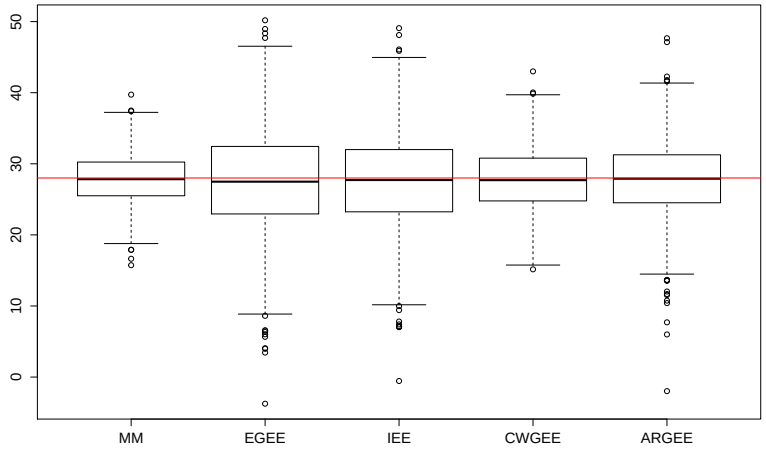

Figure B.30: Extension 2: Boxplots of Interaction Effect Estimates for scenario 6

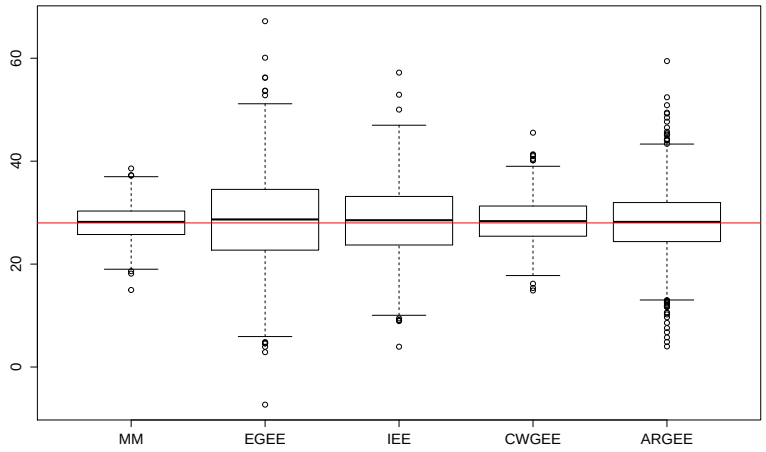

Figure B.31: Extension 2: Boxplots of Interaction Effect Estimates for scenario 7

ICS Scenrios Larger SD of Cluster Size

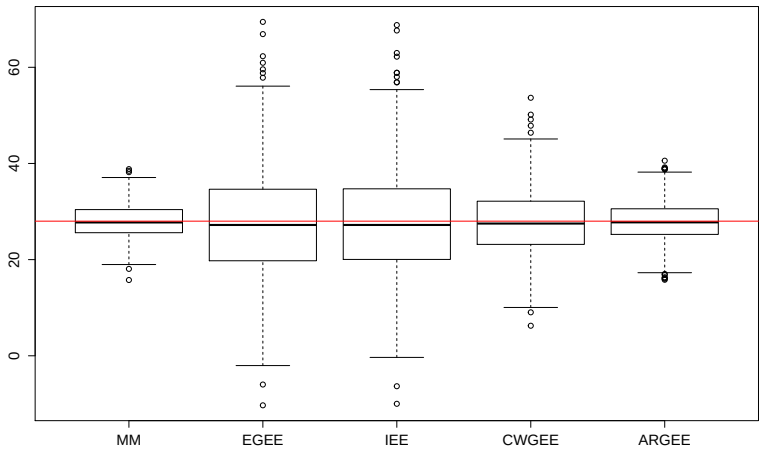

Figure B.32: Extension 2:(Larger Sd of Cluster Size) Boxplots of Interaction Effect Estimates for scenario 3

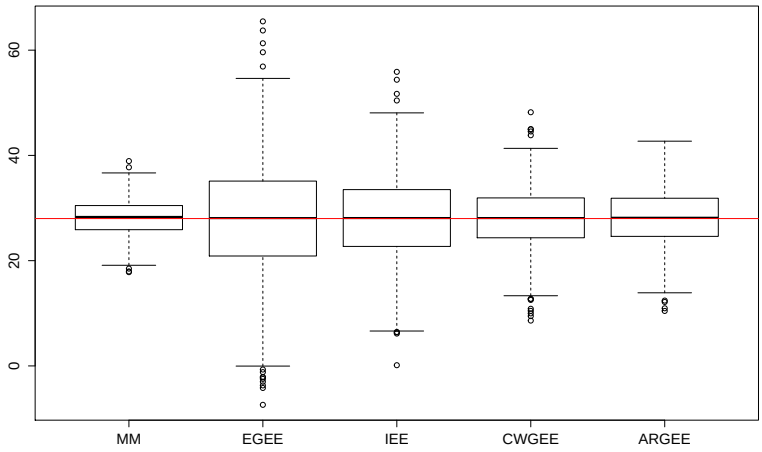

Figure B.33: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Estimates for scenario 4

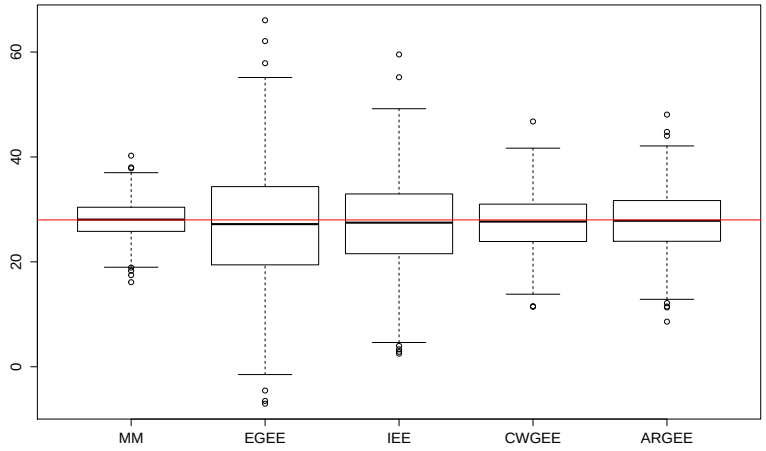

Figure B.34: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Estimates for scenario 5

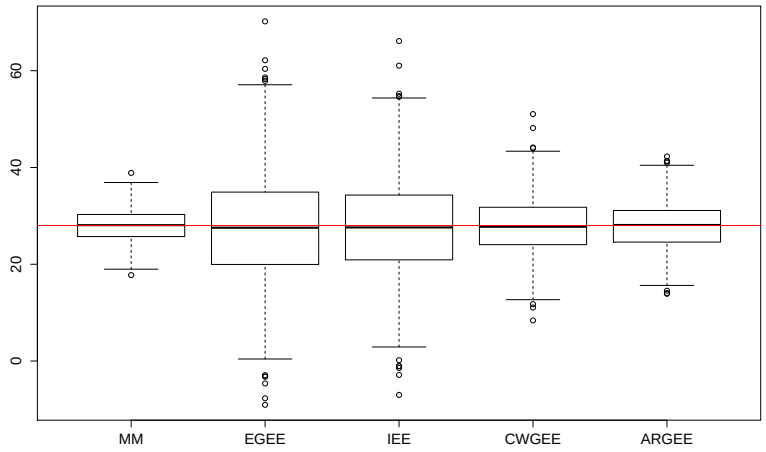

Figure B.35: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Estimates for scenario 6

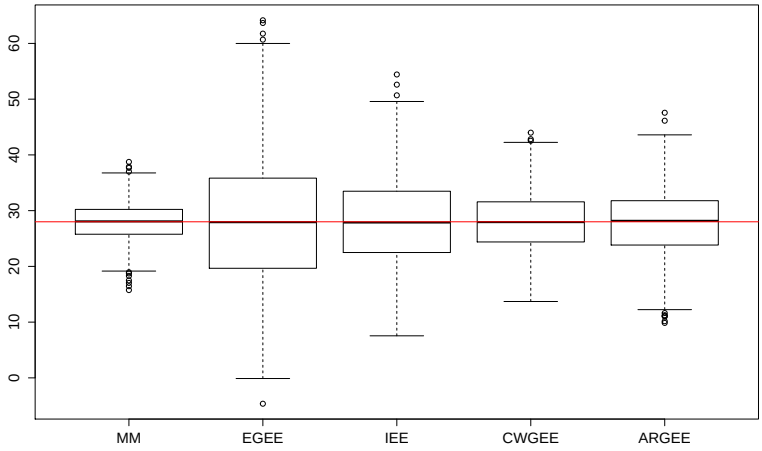

Figure B.36: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Estimates for scenario 7

Boxplots of Standard Error / Standard Deviation

NICS Scenarios

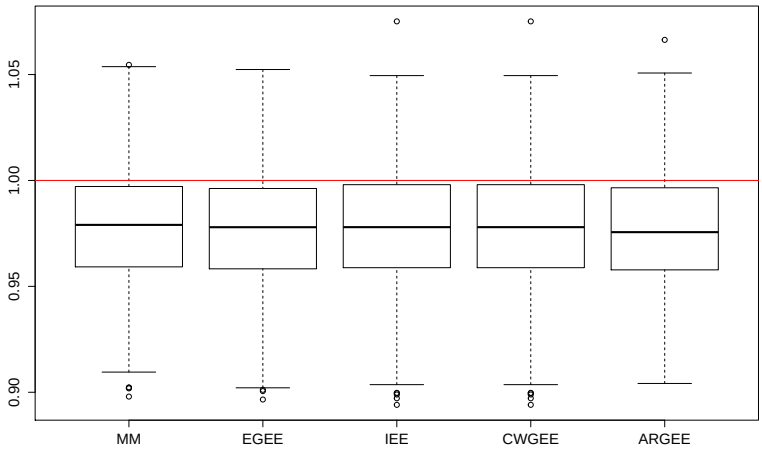

Figure B.37: Extension 2: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 1

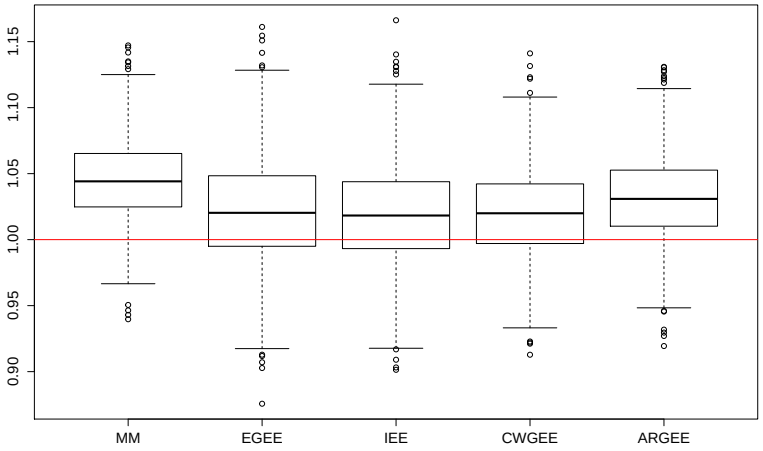

Figure B.38: Extension 2: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 2

ICS Scenrios Smaller SD of Cluster SIze

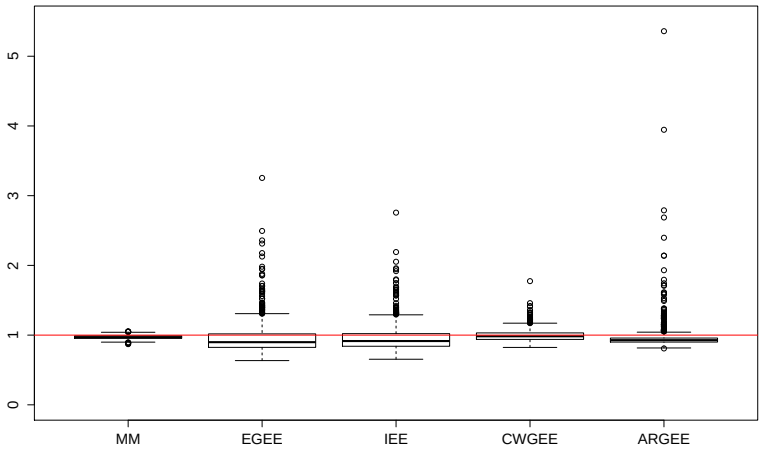

Figure B.39: Extension 2: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 3

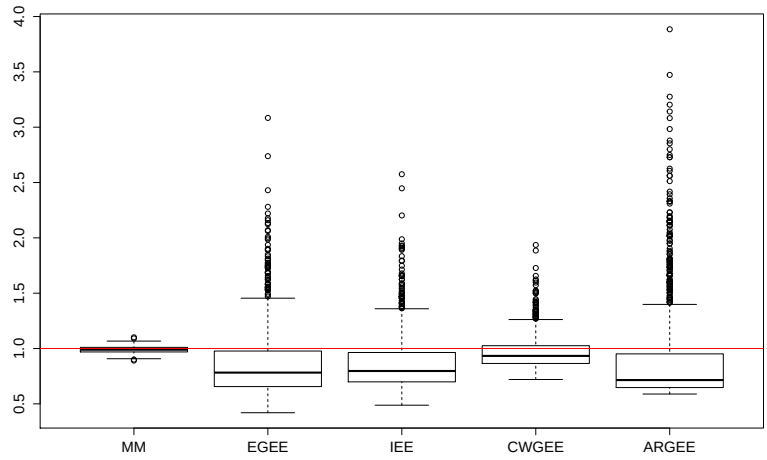

Figure B.40: Extension 2: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 4

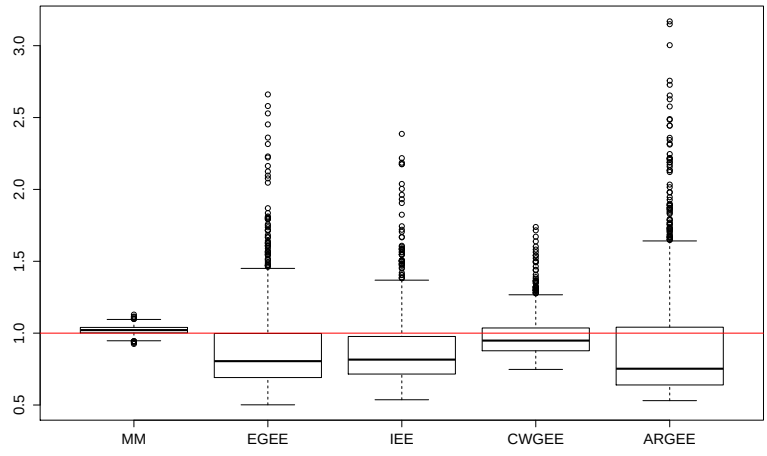

Figure B.41: Extension 2: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 5

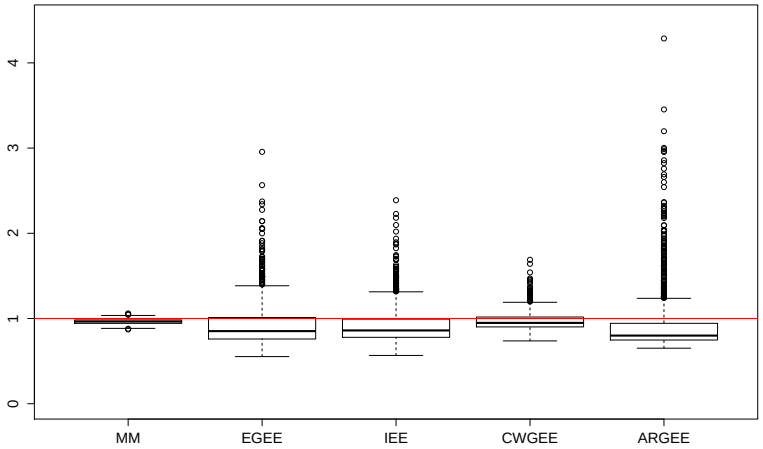

Figure B.42: Extension 2: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 6

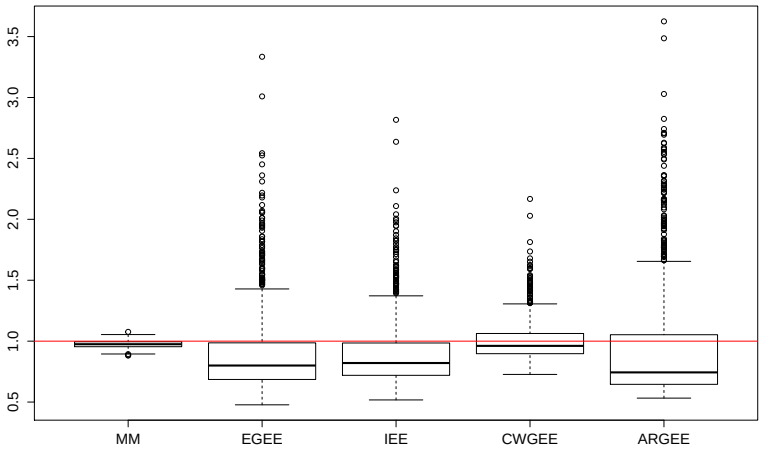

Figure B.43: Extension 2: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 7

ICS Scenrios Larger SD of Cluster Size

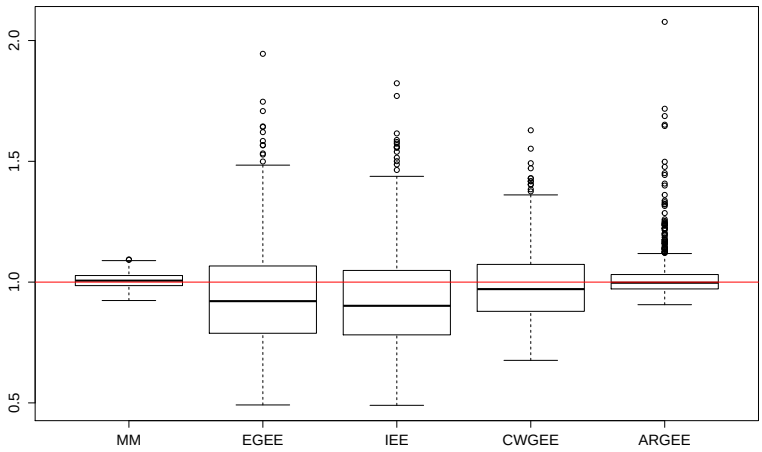

Figure B.44: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 3

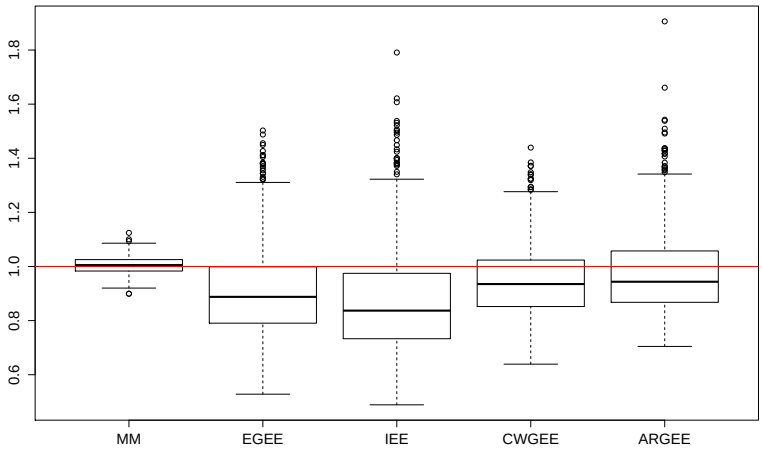

Figure B.45: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 4

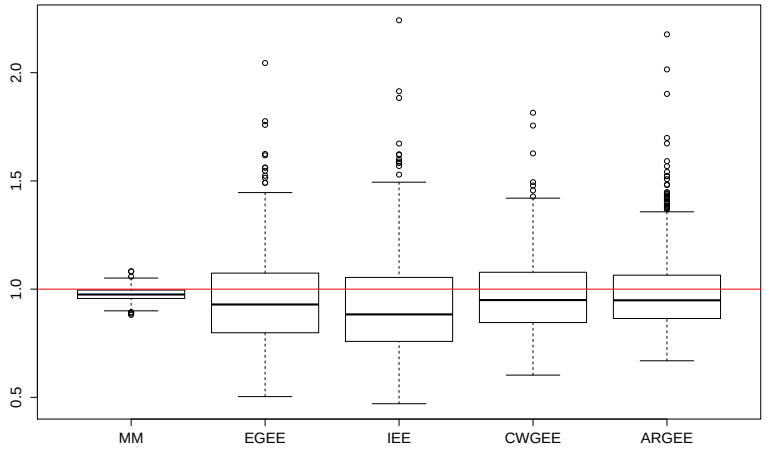

Figure B.46: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 5

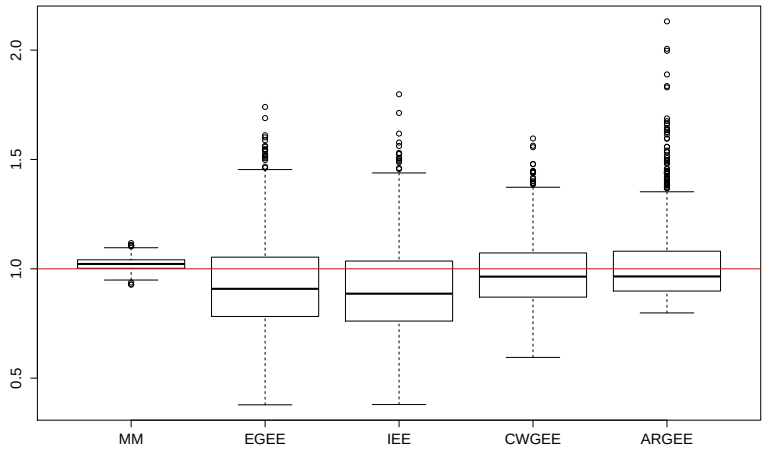

Figure B.47: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 6

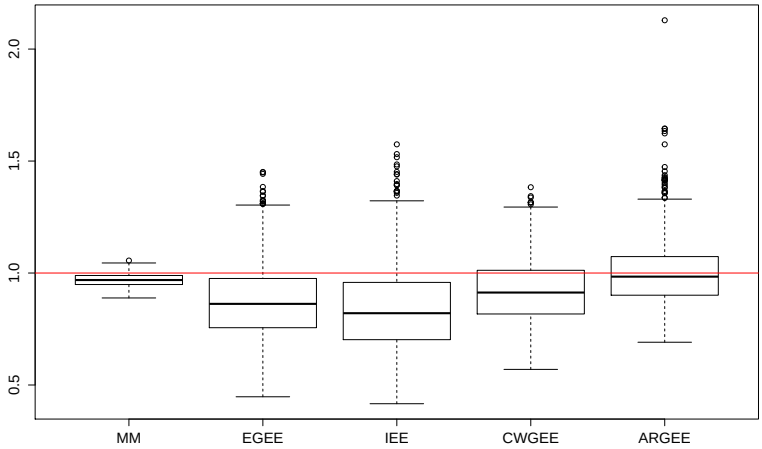

Figure B.48: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 7

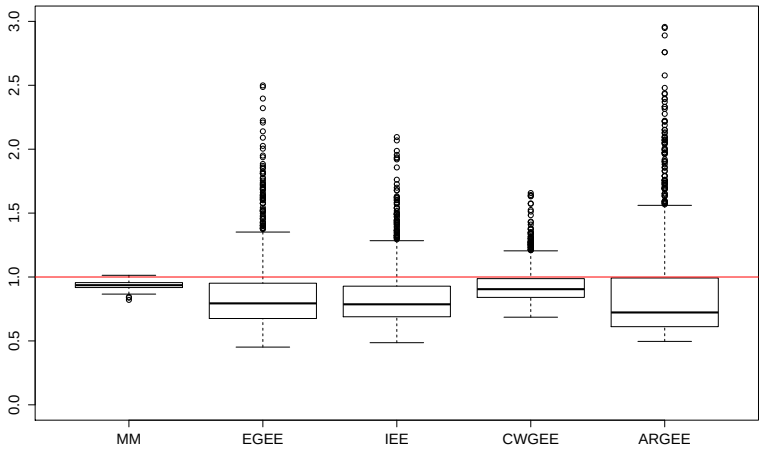

Figure B.49: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Standard Error / Standard Deviation for all ICS scenarios with treatment effect of 14 g/week

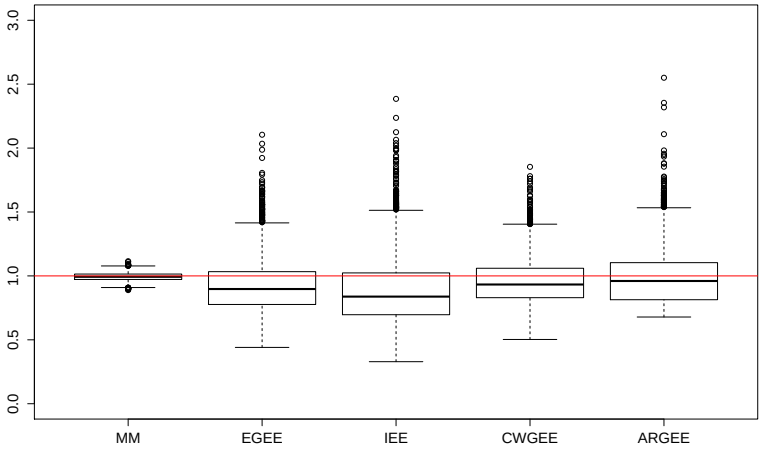

Figure B.50: Extension 2:(Larger SD of Cluster Size) Boxplots of Interaction Effect Standard Error / Standard Deviation for all ICS scenarios with treatment effect of 14 g/week

B.4 Extension 3: Plots for GEE Study Results

Boxplots of Bias in Treatment Effect

NICS Scenarios n=60

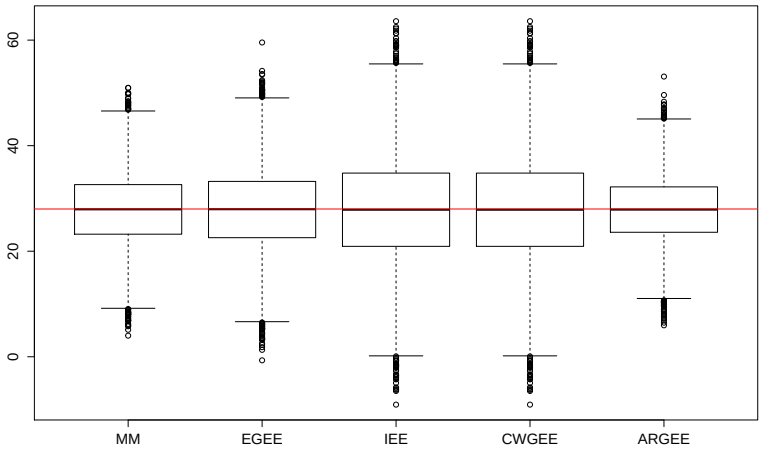

Figure B.51: Extension 3: Boxplots of Interaction Effect Estimates for scenario 1

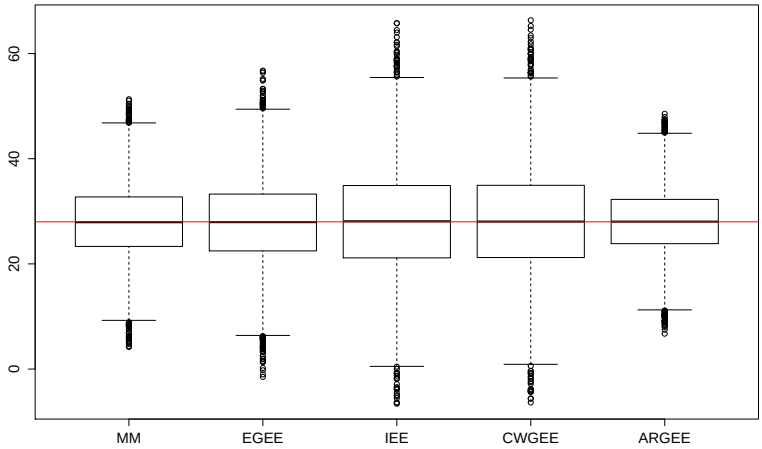

Figure B.52: Extension 3: Boxplots of Interaction Effect Estimates for scenario 2

NICS Scenarios n=600

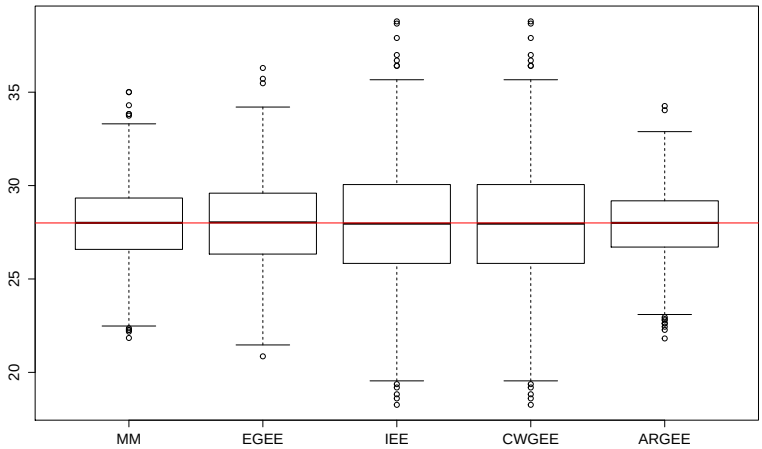

Figure B.53: Extension 3:(Larger Trial) Boxplots of Interaction Effect Estimates for scenario 1

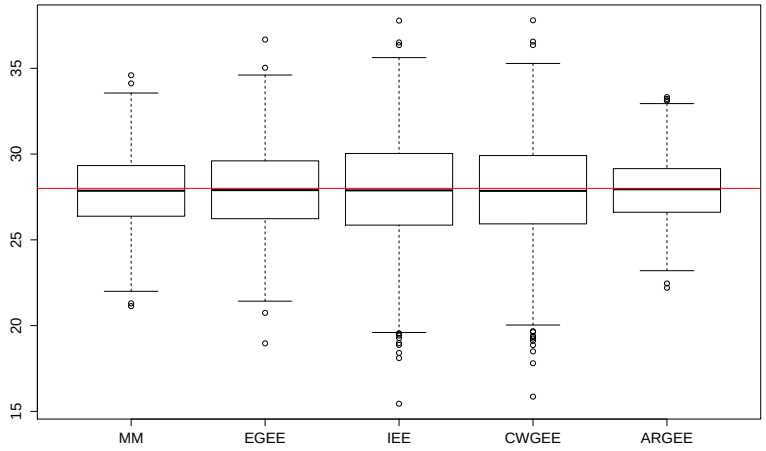

Figure B.54: Extension 3:(Larger Trial) Boxplots of Interaction Effect Estimates for scenario 2

ICS Scenrios n=60

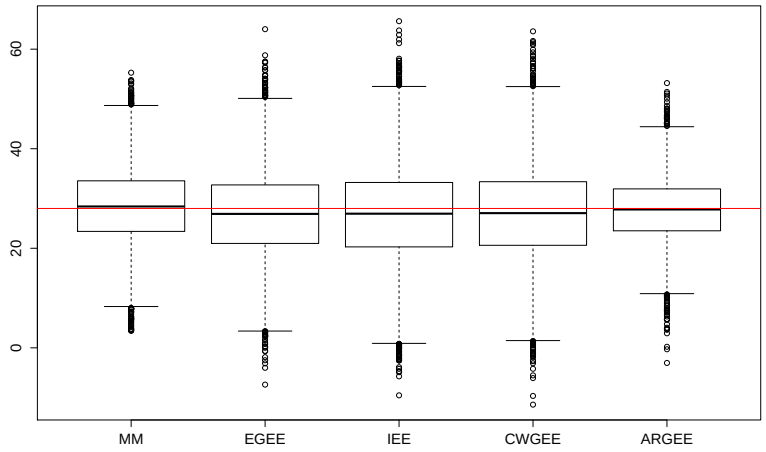

Figure B.55: Extension 3: Boxplots of Interaction Effect Estimates for scenario 3

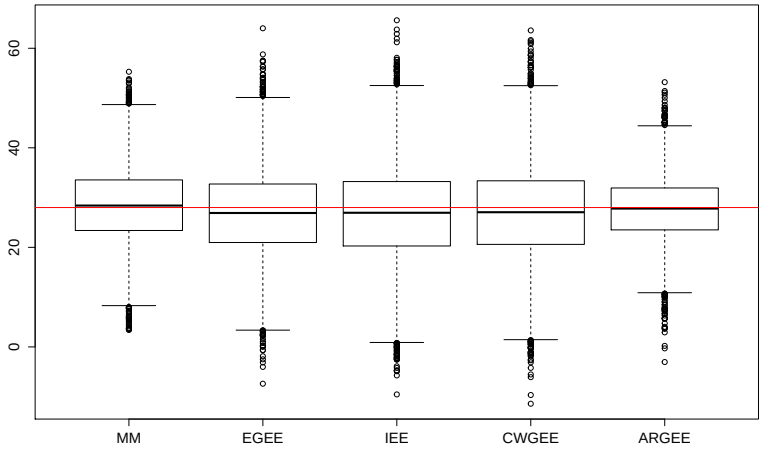

Figure B.56: Extension 3: Boxplots of Interaction Effect Estimates for scenario 4

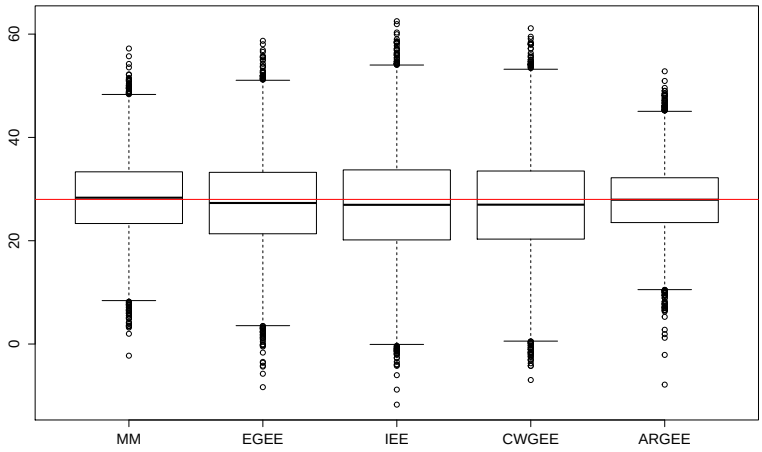

Figure B.57: Extension 3: Boxplots of Interaction Effect Estimates for scenario 5

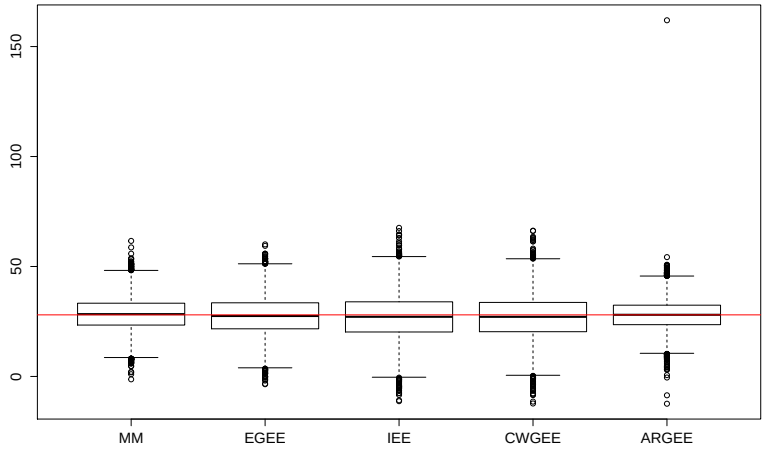

Figure B.58: Extension 3: Boxplots of Interaction Effect Estimates for scenario 6

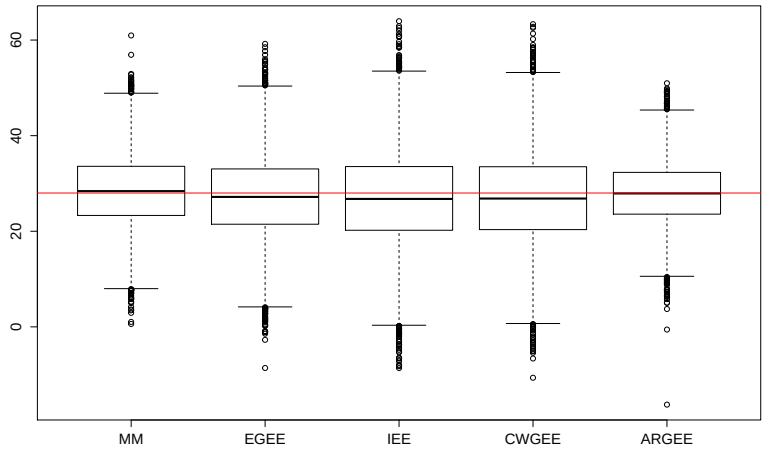

Figure B.59: Extension 3: Boxplots of Interaction Effect Estimates for scenario 7

ICS Scenrios n=600

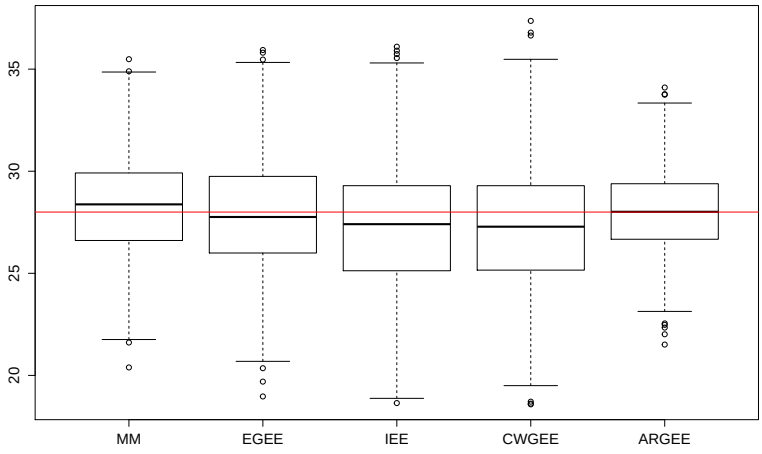

Figure B.60: Extension 3:(n=600) Boxplots of Interaction Effect Estimates for scenario 3

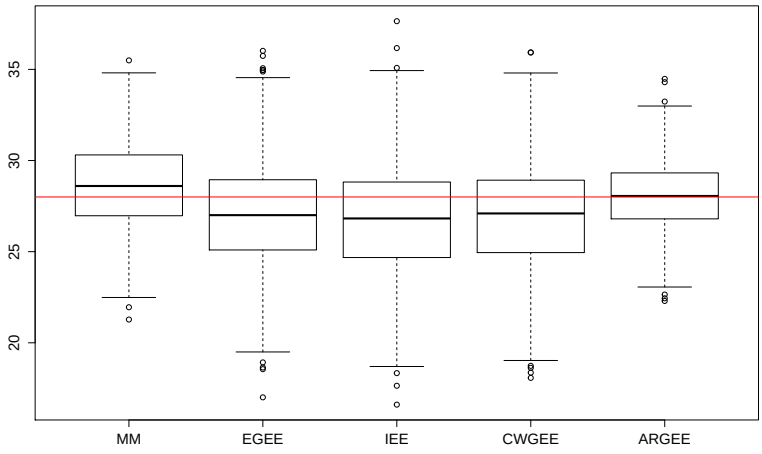

Figure B.61: Extension 3:(n=600) Boxplots of Interaction Effect Estimates for scenario 4

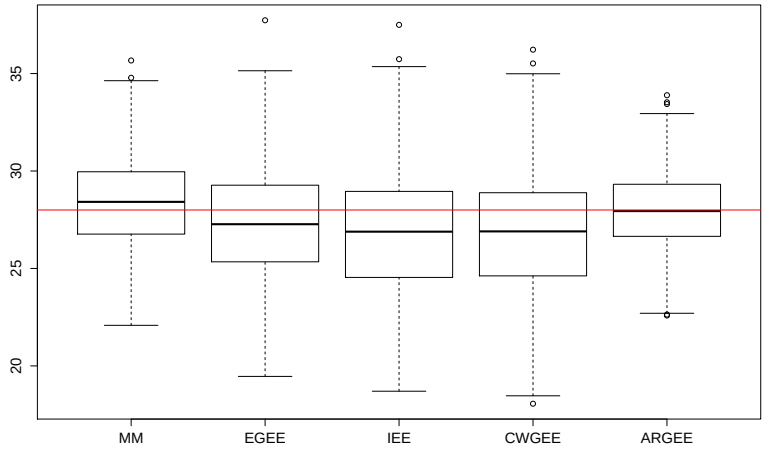

Figure B.62: Extension 3:(n=600) Boxplots of Interaction Effect Estimates for scenario 5

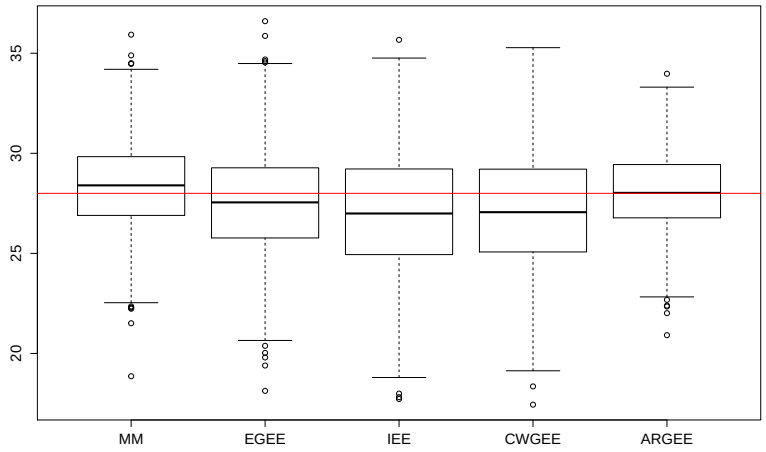

Figure B.63: Extension 3:(n=600) Boxplots of Interaction Effect Estimates for scenario 6

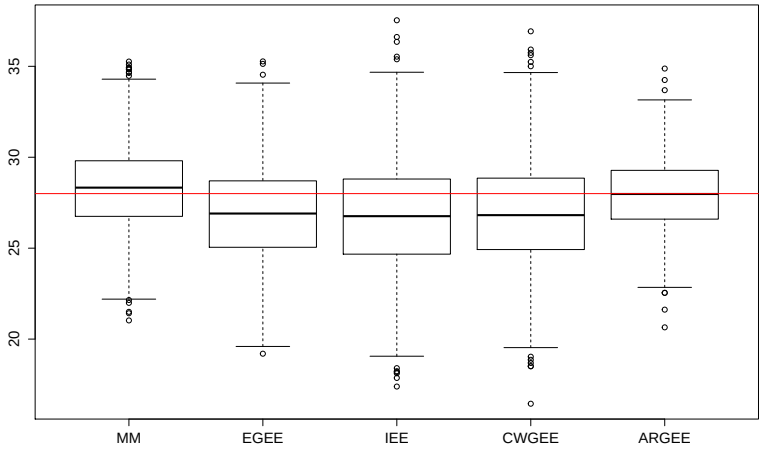

Figure B.64: Extension 3:(n=600) Boxplots of Interaction Effect Estimates for scenario 7

Boxplots of Standard Error / Standard Deviation

NICS Scenarios n=60

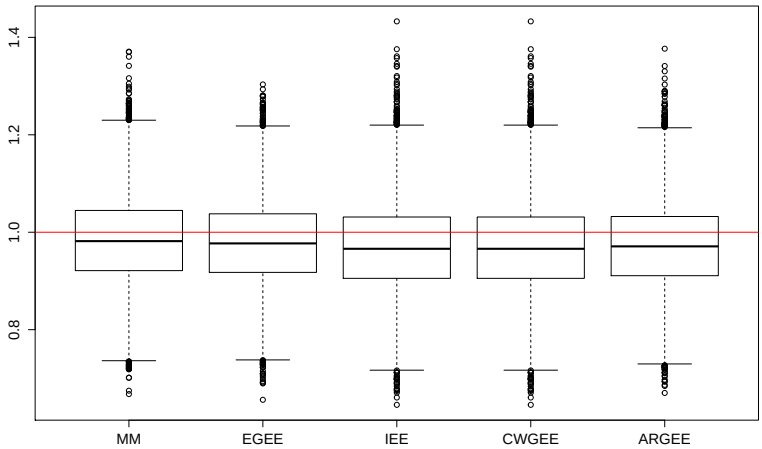

Figure B.65: Extension 3: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 1

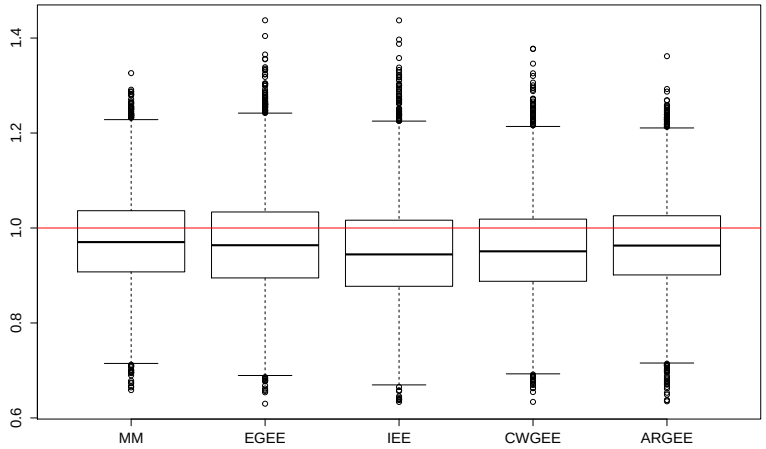

Figure B.66: Extension 3: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 2

ICS Scenrios n=60

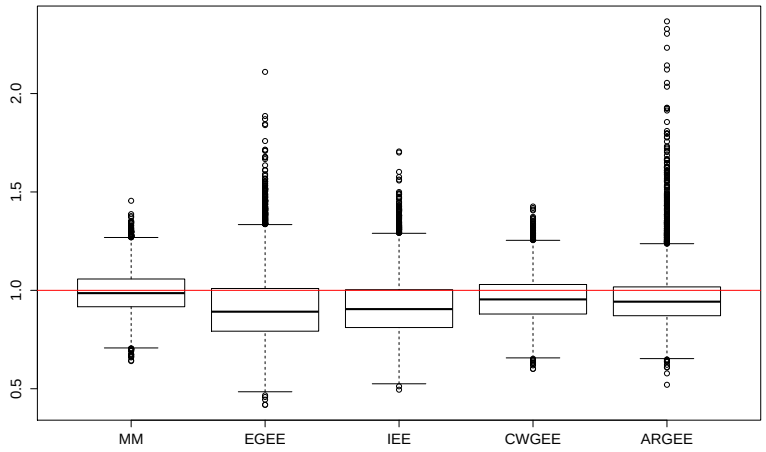

Figure B.67: Extension 3: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 3

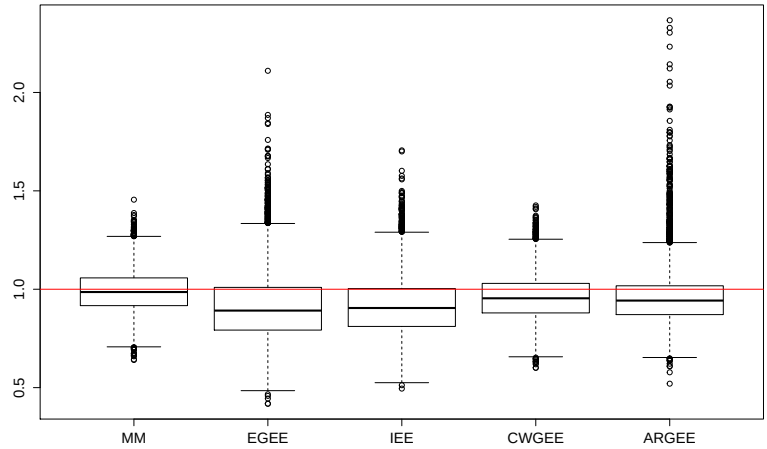

Figure B.68: Extension 3: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 4

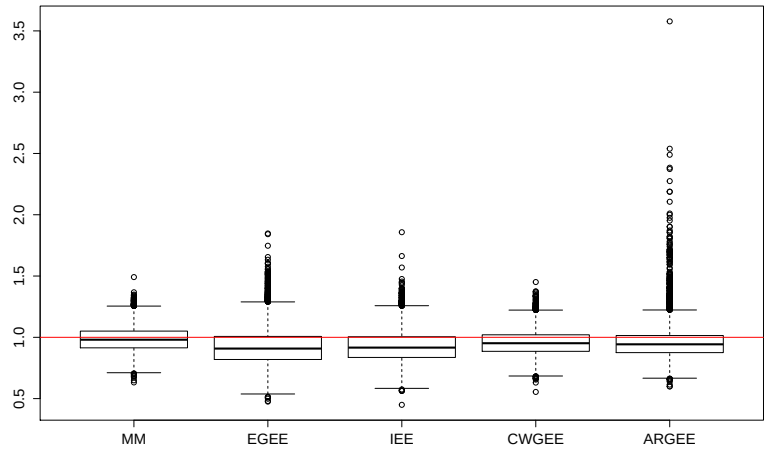

Figure B.69: Extension 3: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 5

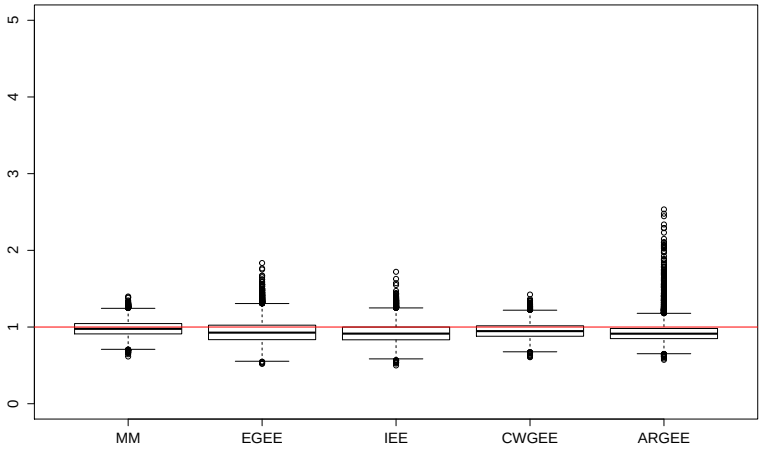

Figure B.70: Extension 3: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 6

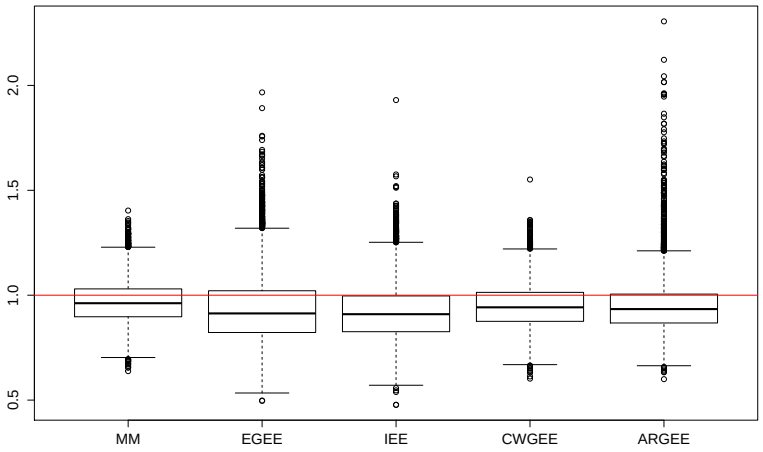

Figure B.71: Extension 3: Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 7

NICS Scenarios n=600

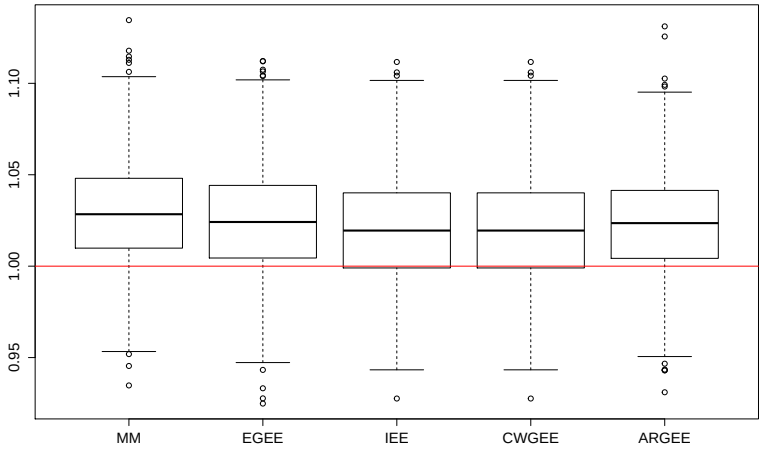

Figure B.72: Extension 3:(n=600) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 1

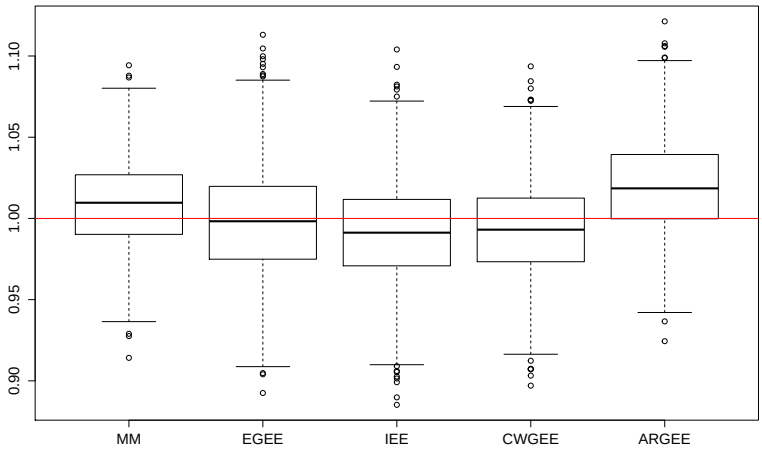

Figure B.73: Extension 3:(n=600) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 2

ICS Scenrios n=600

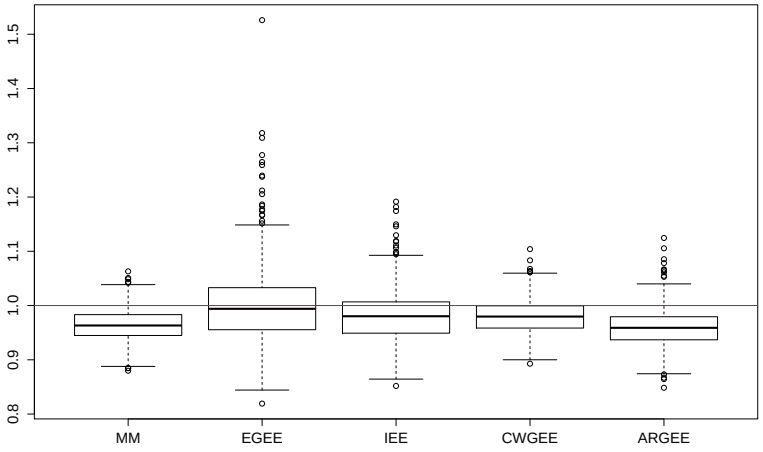

Figure B.74: Extension 3:(n=600) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 3

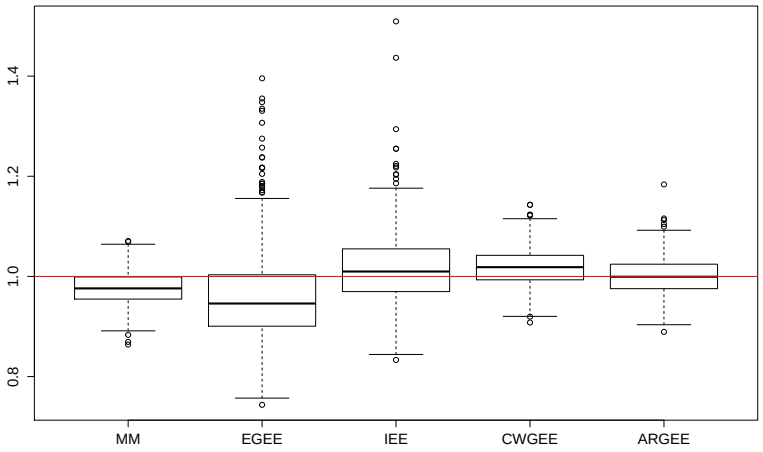

Figure B.75: Extension 3:(n=600) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 4

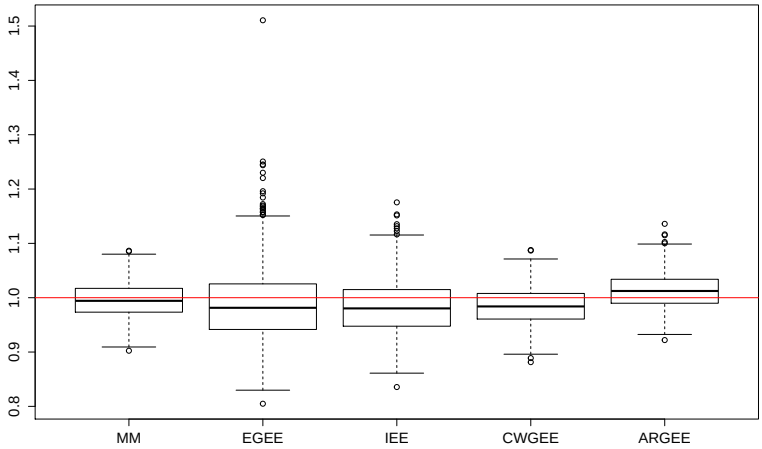

Figure B.76: Extension 3:(n=600) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 5

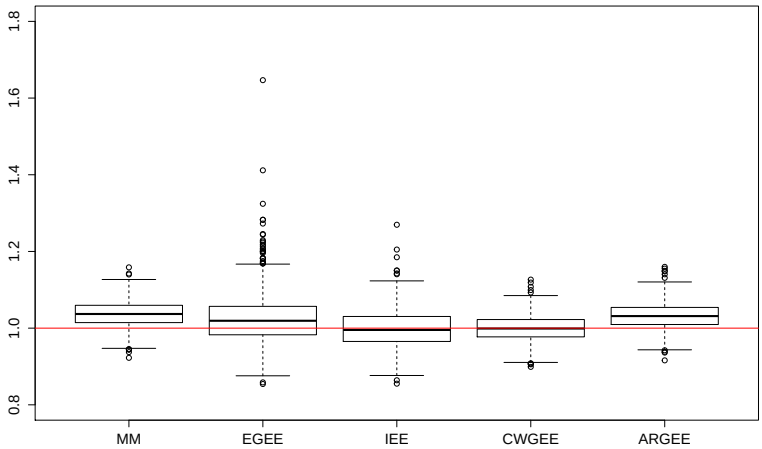

Figure B.77: Extension 3:(n=600) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 6

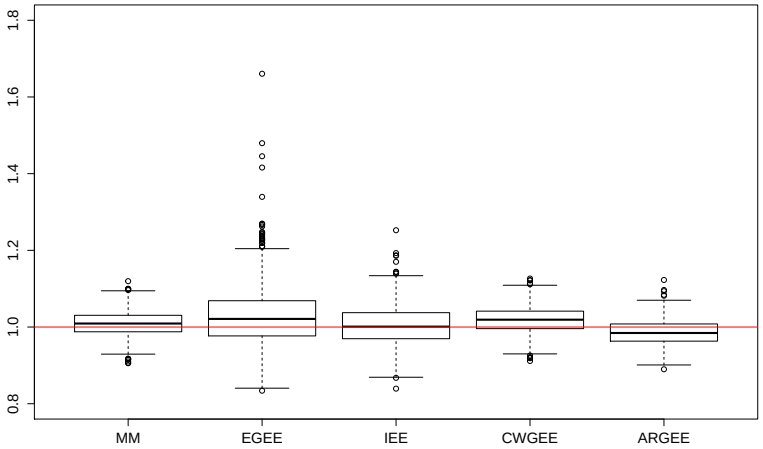

Figure B.78: Extension 3: (n=600) Boxplots of Interaction Effect Standard Error / Standard Deviation for scenario 7

Appendix C

Results for Larger Trial (600 individuals)

In Section 5.2, the results for the larger sample size trial were discussed, which looked at repeated the Chapter 4 simulation study and the Extension 1, using a larger trial size of 600 infants. The results for the simulation scenarios with a treatment effect of 28 were provided. This appendix, C, gives the results for the full larger simulation study, that is all 60 simulations discussed in section 5.2.

C.1 Scenarios with Non Informative Cluster Size

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.05	3.35	3.43	41.73	3.73	3.77	0.72	0.01	0.01	244.37	2.37	2.41	-1035.37	100.46	101.64
EGEE	28.05	3.35	3.43	41.91	5.62	5.68	0.72	0.02	0.02	244.37	2.37	2.41	-1040.00	151.23	152.66
IEE	27.90	3.97	4.06	41.91	5.62	5.69	0.72	0.02	0.02	244.44	2.60	2.66	-1040.03	151.24	152.75
CWGEE	27.90	3.97	4.06	41.91	5.62	5.69	0.72	0.02	0.02	244.44	2.60	2.66	-1040.03	151.24	152.75
ARGEE	28.00	3.52	3.60	41.93	5.70	5.80	0.72	0.02	0.02	244.40	2.48	2.54	-1040.65	153.33	155.94

Table C.1: Extension 2 Scenario 1 (fixed trial length=38, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.07	3.36	3.21	41.92	3.74	3.84	0.72	0.01	0.01	244.42	2.37	2.34	-1041.55	100.57	103.23
EGEE	28.08	3.58	3.51	42.07	5.69	5.80	0.72	0.02	0.02	244.42	2.54	2.55	-1045.55	153.14	155.58
IEE	28.07	4.15	4.08	42.10	5.84	5.97	0.72	0.02	0.02	244.44	3.00	3.05	-1046.70	157.15	160.31
CWGEE	28.07	4.04	3.96	42.07	5.69	5.80	0.72	0.02	0.02	244.44	2.87	2.92	-1045.63	153.14	155.74
ARGEE	28.07	3.57	3.46	42.08	5.81	5.91	0.72	0.02	0.02	244.43	2.53	2.54	-1046.11	156.18	158.79

Table C.2: Extension 2 Scenario 2 (γ_0=log(24.3667),γ_1=0, γ_2=0, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.06	3.36	3.48	42.03	3.73	3.76	0.72	0.01	0.01	244.48	2.37	2.37	-1043.06	100.38	101.32
EGEE	21.06	3.35	3.47	42.12	5.62	5.80	0.72	0.02	0.02	244.48	2.37	2.37	-1045.56	151.20	155.91
IEE	21.15	3.97	4.07	42.12	5.62	5.80	0.72	0.02	0.02	244.43	2.60	2.61	-1045.56	151.21	155.89
CWGEE	21.15	3.97	4.07	42.12	5.62	5.80	0.72	0.02	0.02	244.43	2.60	2.61	-1045.56	151.21	155.89
ARGEE	21.11	3.52	3.68	42.18	5.70	5.87	0.72	0.02	0.02	244.46	2.48	2.52	-1047.16	153.38	157.78

Table C.3: Extension 2 Scenario 8 (fixed trial length=38, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.02	3.35	3.26	41.93	3.73	3.77	0.72	0.01	0.01	244.46	2.37	2.30	-1041.41	100.47	101.78
EGEE	21.06	3.58	3.51	41.88	5.67	5.73	0.72	0.02	0.02	244.44	2.53	2.48	-1040.02	152.55	153.88
IEE	21.14	4.14	4.11	41.90	5.81	5.86	0.72	0.02	0.02	244.40	2.99	2.95	-1040.49	156.40	157.42
CWGEE	21.13	4.03	3.99	41.88	5.67	5.73	0.72	0.02	0.02	244.41	2.86	2.81	-1040.03	152.55	154.18
ARGEE	21.06	3.57	3.48	41.86	5.79	5.77	0.72	0.02	0.02	244.43	2.52	2.46	-1039.46	155.74	155.40

Table C.4: Extension 2 Scenario 9 (γ_0=log(24.3667),γ_1=0, γ_2=0, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.87	3.35	3.30	42.05	3.74	3.54	0.72	0.01	0.01	244.52	2.37	2.36	-1043.95	100.56	95.26
EGEE	13.87	3.35	3.30	42.32	5.64	5.76	0.72	0.02	0.02	244.52	2.37	2.35	-1051.33	151.67	153.94
IEE	13.76	3.97	3.92	42.32	5.64	5.77	0.72	0.02	0.02	244.58	2.59	2.59	-1051.24	151.68	154.09
CWGEE	13.76	3.97	3.92	42.32	5.64	5.77	0.72	0.02	0.02	244.58	2.59	2.59	-1051.24	151.68	154.09
ARGEE	13.79	3.52	3.45	42.30	5.72	5.79	0.72	0.02	0.02	244.57	2.48	2.45	-1050.41	153.85	154.84

Table C.5: Extension 2 Scenario 15 (fixed trial length=38, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.98	3.35	3.41	42.18	3.72	3.83	0.72	0.01	0.01	244.50	2.37	2.26	-1048.79	100.18	103.32
EGEE	13.93	3.58	3.69	42.35	5.66	5.76	0.72	0.02	0.02	244.55	2.53	2.46	-1053.25	152.45	155.53
IEE	13.95	4.14	4.29	42.32	5.81	5.95	0.72	0.02	0.02	244.57	3.00	3.01	-1052.58	156.48	160.34
CWGEE	13.96	4.03	4.14	42.34	5.66	5.76	0.72	0.02	0.02	244.55	2.86	2.85	-1053.26	152.45	155.55
ARGEE	13.99	3.56	3.64	42.36	5.78	5.90	0.72	0.02	0.02	244.51	2.52	2.45	-1053.64	155.50	159.10

Table C.6: Extension 2 Scenario 16 (γ_0=log(24.3667),γ_1=0, γ_2=0, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.19	3.35	3.40	42.04	3.74	3.76	0.72	0.01	0.01	244.40	2.37	2.49	-1043.88	100.60	100.92
EGEE	7.19	3.35	3.40	41.87	5.63	5.48	0.72	0.02	0.02	244.40	2.36	2.49	-1039.69	151.51	146.09
IEE	7.26	3.97	4.08	41.87	5.63	5.48	0.72	0.02	0.02	244.37	2.59	2.74	-1039.72	151.52	146.20
CWGEE	7.26	3.97	4.08	41.87	5.63	5.48	0.72	0.02	0.02	244.37	2.59	2.74	-1039.72	151.52	146.20
ARGEE	7.21	3.51	3.57	41.86	5.71	5.52	0.72	0.02	0.02	244.38	2.48	2.59	-1039.57	153.61	147.44

Table C.7: Extension 2 Scenario 22 (fixed trial length=38, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.00	3.36	3.44	42.12	3.74	3.74	0.72	0.01	0.01	244.50	2.38	2.41	-1046.68	100.68	100.24
EGEE	7.08	3.59	3.63	42.14	5.69	5.61	0.72	0.02	0.02	244.46	2.54	2.53	-1047.48	153.12	149.77
IEE	7.04	4.16	4.18	42.12	5.84	5.74	0.72	0.02	0.02	244.48	3.01	2.91	-1046.95	157.16	153.29
CWGEE	7.02	4.04	4.07	42.14	5.69	5.61	0.72	0.02	0.02	244.49	2.87	2.80	-1047.54	153.12	149.85
ARGEE	7.03	3.57	3.65	42.11	5.81	5.73	0.72	0.02	0.02	244.49	2.53	2.55	-1046.85	156.22	153.10

Table C.8: Extension 2 Scenario 23 (γ_0=log(24.3667),γ_1=0, γ_2=0, treatment effect = 7 g/week)

Interaction =0

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.03	3.36	3.36	41.97	3.72	3.88	0.72	0.01	0.01	244.51	2.37	2.39	-1041.49	100.22	104.97
EGEE	0.02	3.35	3.36	42.07	5.62	5.80	0.72	0.02	0.02	244.52	2.37	2.39	-1044.49	151.31	156.30
IEE	0.02	3.97	3.97	42.07	5.63	5.80	0.72	0.02	0.02	244.52	2.60	2.60	-1044.49	151.31	156.23
CWGEE	0.02	3.97	3.97	42.07	5.63	5.80	0.72	0.02	0.02	244.52	2.60	2.60	-1044.49	151.31	156.23
ARGEE	-0.00	3.52	3.49	42.05	5.71	5.83	0.72	0.02	0.02	244.51	2.48	2.49	-1044.09	153.50	157.12

Table C.9: Extension 2 Scenario 29 (fixed trial length=38, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.03	3.36	3.30	41.86	3.74	3.96	0.72	0.01	0.01	244.50	2.38	2.35	-1039.67	100.59	106.12
EGEE	-0.06	3.59	3.61	42.02	5.68	5.88	0.72	0.02	0.02	244.57	2.54	2.54	-1044.19	152.86	157.60
IEE	-0.13	4.16	4.18	42.03	5.84	6.05	0.72	0.02	0.02	244.64	3.00	2.98	-1044.57	157.06	162.25
CWGEE	-0.13	4.04	4.06	42.02	5.68	5.88	0.72	0.02	0.02	244.62	2.87	2.84	-1044.30	152.87	157.63
ARGEE	-0.06	3.57	3.55	42.06	5.79	5.98	0.72	0.02	0.02	244.54	2.52	2.52	-1045.33	155.94	160.09

Table C.10: Extension 2 Scenario 30 (γ_0=log(24.3667),γ_1=0, γ_2=0, treatment effect = 0 g/week)

C.2 Scenarios with smaller SD of Cluster Size, γ_0 = 3.069142

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.09	3.37	3.48	42.05	3.74	3.69	0.72	0.01	0.01	245.01	2.39	2.39	-1044.95	100.56	99.36
EGEE	28.27	5.98	6.29	42.13	5.85	5.80	0.72	0.02	0.02	244.40	4.15	4.52	-1046.01	157.23	154.76
IEE	28.29	5.99	6.27	42.12	6.91	7.00	0.72	0.02	0.03	223.79	5.75	6.52	-1011.64	185.95	187.04
CWGEE	28.18	4.50	4.53	42.12	5.44	5.43	0.72	0.02	0.02	222.33	4.15	4.27	-983.38	146.33	145.11
ARGEE	28.07	3.92	4.06	42.20	5.89	5.94	0.72	0.02	0.02	245.75	2.68	2.92	-1044.93	158.40	158.51

Table C.11: Extension 2 Scenario 3 (γ_1=-0.50, γ_2=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.95	3.34	3.37	42.04	3.75	3.80	0.72	0.01	0.01	243.59	2.36	2.41	-1042.84	100.75	102.17
EGEE	27.99	8.72	10.06	42.21	5.44	5.62	0.72	0.02	0.02	202.33	5.92	6.61	-964.82	146.53	151.26
IEE	28.00	6.85	7.89	42.41	6.54	6.95	0.72	0.02	0.02	172.88	7.15	8.29	-910.73	176.66	186.75
CWGEE	28.01	4.86	5.03	42.18	5.12	5.31	0.72	0.02	0.02	188.85	5.02	4.75	-926.27	138.12	143.10
ARGEE	27.82	5.40	5.95	42.10	5.93	6.25	0.72	0.02	0.02	239.69	3.32	3.94	-1041.89	159.69	169.38

Table C.12: Extension 2 Scenario 4 (γ_2=-0.50, γ_1=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.07	3.36	3.29	42.19	3.74	3.87	0.72	0.01	0.01	244.10	2.38	2.32	-1049.06	100.70	104.34
EGEE	28.12	7.62	8.57	42.19	5.39	5.53	0.72	0.02	0.02	214.50	5.18	5.54	-990.24	145.34	149.48
IEE	28.17	6.27	7.11	42.24	6.13	6.45	0.72	0.02	0.02	178.78	6.38	7.28	-924.84	165.59	175.46
CWGEE	28.17	4.39	4.51	42.14	4.59	4.69	0.72	0.02	0.02	189.25	4.49	4.33	-917.94	123.84	127.18
ARGEE	28.02	6.30	6.87	41.99	5.76	5.78	0.72	0.02	0.02	243.74	3.72	3.89	-1025.98	155.21	156.13

Table C.13: Extension 2 Scenario 5 (γ_1=-0.35, γ_2=-0.35, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.92	3.38	3.51	41.98	3.74	3.86	0.72	0.01	0.01	244.71	2.39	2.34	-1042.36	100.66	104.18
EGEE	27.65	6.59	7.11	41.68	5.62	5.75	0.72	0.02	0.02	228.39	4.59	4.74	-1001.59	151.28	155.31
IEE	27.70	6.04	6.56	41.65	6.42	6.76	0.72	0.02	0.02	197.77	5.99	6.59	-945.59	173.31	183.72
CWGEE	27.76	4.36	4.49	41.72	4.93	5.02	0.72	0.02	0.02	202.70	4.23	4.06	-930.64	132.75	135.49
ARGEE	27.91	5.06	5.35	41.64	5.78	5.78	0.72	0.02	0.02	245.33	3.19	3.23	-1021.88	155.66	155.47

Table C.14: Extension 2 Scenario 6 (γ_1=-0.46, γ_2=-0.19, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.15	3.35	3.43	42.01	3.74	3.87	0.72	0.01	0.01	243.86	2.37	2.37	-1042.32	100.67	103.49
EGEE	28.53	8.28	9.23	42.00	5.33	5.42	0.72	0.02	0.02	205.66	5.67	6.11	-966.70	143.55	146.33
IEE	28.41	6.49	7.25	41.85	6.17	6.55	0.72	0.02	0.02	170.71	6.79	7.84	-895.52	166.36	178.40
CWGEE	28.31	4.55	4.55	41.99	4.69	4.79	0.72	0.02	0.02	184.76	4.77	4.46	-907.41	126.47	129.37
ARGEE	28.16	6.23	6.62	42.13	5.82	5.83	0.72	0.02	0.02	242.03	3.72	3.97	-1034.80	156.53	156.79

Table C.15: Extension 2 Scenario 7 (γ_2=-0.46, γ_1=-0.19, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.06	3.38	3.48	41.88	3.74	3.67	0.72	0.01	0.01	245.04	2.39	2.42	-1040.60	100.58	98.75
EGEE	21.20	5.98	6.37	41.68	5.84	6.21	0.72	0.02	0.02	244.40	4.15	4.50	-1033.32	157.37	167.47
IEE	21.20	6.01	6.39	41.58	6.90	7.50	0.72	0.02	0.03	223.69	5.74	6.52	-995.31	186.37	202.85
CWGEE	21.13	4.50	4.60	41.71	5.44	5.80	0.72	0.02	0.02	222.32	4.15	4.29	-971.96	146.50	156.55
ARGEE	21.17	3.94	4.02	41.76	5.90	6.23	0.72	0.02	0.02	245.77	2.70	2.87	-1032.51	158.66	168.57

Table C.16: Extension 2 Scenario 10 (γ_1=-0.50, γ_2=0, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.95	3.34	3.31	41.70	3.73	3.83	0.72	0.01	0.01	243.59	2.36	2.36	-1034.17	100.35	103.36
EGEE	21.20	8.77	10.27	41.83	5.43	5.47	0.72	0.02	0.02	202.26	5.98	6.56	-955.51	146.42	147.70
IEE	21.27	6.89	7.95	41.76	6.56	7.00	0.72	0.02	0.03	172.74	7.26	8.13	-894.23	177.22	188.60
CWGEE	21.20	4.86	4.98	41.79	5.12	5.17	0.72	0.02	0.02	188.78	5.06	4.56	-916.79	137.97	139.56
ARGEE	20.91	5.49	6.26	41.89	5.94	6.01	0.72	0.02	0.02	239.77	3.39	3.98	-1037.04	159.93	162.40

Table C.17: Extension 2 Scenario 11 (γ_2=-0.50, γ_1=0, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.13	3.36	3.40	41.96	3.75	3.78	0.72	0.01	0.01	244.20	2.38	2.36	-1041.07	100.93	101.23
EGEE	21.50	7.33	8.21	42.02	5.39	5.25	0.72	0.02	0.02	214.74	5.11	5.68	-984.71	145.07	141.89
IEE	21.34	6.07	6.84	42.07	6.03	6.31	0.72	0.02	0.02	179.41	6.24	7.27	-920.00	162.96	171.46
CWGEE	21.18	4.32	4.49	42.01	4.58	4.48	0.72	0.02	0.02	189.58	4.43	4.47	-913.68	123.59	120.83
ARGEE	21.38	5.98	6.67	41.92	5.70	5.57	0.72	0.02	0.02	243.77	3.63	3.95	-1024.78	153.27	148.97

Table C.18: Extension 2 Scenario 12 (γ_1=-0.35, γ_2=-0.35, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.92	3.37	3.32	42.07	3.75	3.66	0.72	0.01	0.01	244.65	2.39	2.37	-1045.26	100.82	98.81
EGEE	20.94	6.51	7.11	42.41	5.63	5.44	0.72	0.02	0.02	228.37	4.50	4.97	-1022.25	151.61	146.72
IEE	20.97	5.97	6.54	42.53	6.43	6.45	0.72	0.02	0.02	197.91	5.86	6.84	-970.78	173.51	174.40
CWGEE	20.98	4.34	4.43	42.36	4.94	4.79	0.72	0.02	0.02	202.71	4.19	4.33	-948.98	133.08	129.26
ARGEE	20.96	4.96	5.14	42.22	5.79	5.62	0.72	0.02	0.02	245.19	3.13	3.33	-1037.95	155.83	152.14

Table C.19: Extension 2 Scenario 13 (γ_1=-0.46, γ_2=-0.19, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.07	3.35	3.28	41.98	3.75	3.78	0.72	0.01	0.01	243.84	2.37	2.38	-1041.88	100.80	101.88
EGEE	21.31	8.62	9.53	41.93	5.33	5.43	0.72	0.02	0.02	205.30	5.89	6.60	-963.43	143.86	146.89
IEE	21.30	6.72	7.49	41.79	6.18	6.64	0.72	0.02	0.02	169.89	7.06	8.39	-890.36	167.72	181.43
CWGEE	21.25	4.64	4.69	41.91	4.69	4.81	0.72	0.02	0.02	184.35	4.89	4.81	-904.42	126.76	129.88
ARGEE	21.07	6.52	7.02	42.22	5.85	5.80	0.72	0.02	0.02	242.01	3.88	4.18	-1036.22	157.25	157.69

Table C.20: Extension 2 Scenario 14 (γ_2=-0.46, γ_1=-0.19, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.02	3.38	3.35	42.05	3.74	3.82	0.72	0.01	0.01	245.03	2.39	2.43	-1045.22	100.51	102.86
EGEE	13.89	5.98	6.40	42.09	5.84	5.81	0.72	0.02	0.02	244.52	4.11	4.39	-1045.47	157.05	156.65
IEE	13.88	6.01	6.40	42.06	6.88	7.09	0.72	0.02	0.02	223.86	5.69	6.38	-1009.81	185.21	191.64
CWGEE	14.00	4.51	4.52	42.07	5.43	5.41	0.72	0.02	0.02	222.38	4.13	4.24	-982.85	146.24	145.80
ARGEE	13.93	3.95	3.99	41.94	5.90	5.84	0.72	0.02	0.02	245.83	2.69	2.90	-1038.56	158.75	157.76

Table C.21: Extension 2 Scenario 17 (γ_1=-0.50, γ_2=0, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.92	3.34	3.39	42.02	3.74	3.79	0.72	0.01	0.01	243.56	2.36	2.29	-1042.42	100.74	101.51
EGEE	13.74	8.72	10.61	41.86	5.44	5.58	0.72	0.02	0.02	202.47	5.89	6.77	-955.55	146.67	150.80
IEE	13.86	6.87	8.31	41.82	6.55	7.06	0.72	0.02	0.02	172.90	7.19	8.56	-895.25	177.26	190.52
CWGEE	13.91	4.87	5.29	41.88	5.12	5.28	0.72	0.02	0.02	188.84	5.03	4.93	-918.29	138.19	142.45
ARGEE	13.89	5.50	6.44	42.02	5.93	6.05	0.72	0.02	0.02	239.58	3.35	3.89	-1038.93	159.82	163.38

Table C.22: Extension 2 Scenario 18 (γ_2=-0.50, γ_1=0, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.10	3.36	3.43	42.05	3.74	3.69	0.72	0.01	0.01	244.11	2.38	2.45	-1044.37	100.74	98.40
EGEE	14.21	7.54	8.68	41.88	5.41	5.40	0.72	0.02	0.02	214.56	5.17	5.73	-981.18	145.71	143.74
IEE	14.16	6.20	7.18	41.85	6.13	6.41	0.72	0.02	0.02	179.02	6.36	7.45	-914.41	165.51	170.81
CWGEE	14.12	4.36	4.62	41.89	4.60	4.62	0.72	0.02	0.02	189.34	4.47	4.49	-910.77	124.12	122.90
ARGEE	14.32	6.18	6.93	41.84	5.77	5.84	0.72	0.02	0.02	243.65	3.68	4.02	-1021.78	155.41	157.22

Table C.23: Extension 2 Scenario 19 (γ_1=-0.35, γ_2=-0.35, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.03	3.38	3.56	42.16	3.75	3.71	0.72	0.01	0.01	244.72	2.39	2.50	-1047.14	100.77	98.97
EGEE	14.00	6.38	7.04	42.11	5.64	5.74	0.72	0.02	0.02	228.62	4.45	5.01	-1014.12	151.85	153.78
IEE	13.97	5.87	6.51	41.99	6.47	6.79	0.72	0.02	0.02	198.37	5.77	6.76	-956.59	174.29	182.43
CWGEE	13.99	4.31	4.49	42.08	4.95	5.03	0.72	0.02	0.02	202.93	4.15	4.40	-941.29	133.28	135.10
ARGEE	14.20	4.81	5.26	42.07	5.78	5.81	0.72	0.02	0.02	245.23	3.07	3.40	-1034.07	155.29	154.59

Table C.24: Extension 2 Scenario 20 (γ_1=-0.46, γ_2=-0.19, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.12	3.35	3.58	41.86	3.74	3.73	0.72	0.01	0.01	243.77	2.37	2.54	-1037.92	100.60	100.69
EGEE	13.94	8.46	9.76	41.84	5.35	5.29	0.72	0.02	0.02	205.51	5.73	6.18	-961.41	144.04	143.95
IEE	13.88	6.60	7.77	41.80	6.24	6.57	0.72	0.02	0.02	170.25	6.90	7.64	-892.78	168.39	178.59
CWGEE	13.92	4.60	4.93	41.83	4.71	4.68	0.72	0.02	0.02	184.55	4.82	4.55	-902.36	127.01	127.86
ARGEE	14.05	6.33	7.12	41.83	5.84	5.70	0.72	0.02	0.02	242.03	3.73	4.14	-1023.94	157.17	154.98

Table C.25: Extension 2 Scenario 21 (γ_2=-0.46, γ_1=-0.19, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.05	3.37	3.36	41.91	3.73	3.76	0.72	0.01	0.01	245.02	2.38	2.42	-1042.12	100.40	100.34
EGEE	7.01	6.06	6.49	42.07	5.85	5.67	0.72	0.02	0.02	244.44	4.28	4.67	-1045.20	157.41	153.27
IEE	7.11	6.07	6.53	42.02	6.90	6.83	0.72	0.02	0.02	223.72	5.95	6.74	-1008.90	186.06	185.61
CWGEE	7.03	4.52	4.53	42.11	5.44	5.27	0.72	0.02	0.02	222.33	4.22	4.32	-983.92	146.52	142.53
ARGEE	6.95	3.95	3.98	42.03	5.91	5.81	0.72	0.02	0.02	245.89	2.70	2.97	-1041.44	158.82	155.99

Table C.26: Extension 2 Scenario 24 (γ_1=-0.50, γ_2=0, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.12	3.34	3.29	41.89	3.74	3.80	0.72	0.01	0.01	243.62	2.36	2.32	-1039.24	100.70	101.93
EGEE	7.02	8.74	10.10	41.96	5.45	5.40	0.72	0.02	0.02	202.52	5.96	6.48	-959.11	146.80	145.50
IEE	7.01	6.89	8.08	41.86	6.58	6.73	0.72	0.02	0.02	173.04	7.27	8.12	-896.83	178.01	180.84
CWGEE	6.97	4.87	5.08	41.96	5.13	5.10	0.72	0.02	0.02	189.04	5.06	4.62	-921.30	138.35	137.41
ARGEE	7.18	5.50	6.22	42.05	5.96	6.08	0.72	0.02	0.02	239.69	3.38	3.95	-1041.62	160.22	164.35

Table C.27: Extension 2 Scenario 25 (γ_2=-0.50, γ_1=0, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.08	3.36	3.31	42.02	3.75	3.82	0.72	0.01	0.01	244.13	2.38	2.40	-1043.09	100.94	102.30
EGEE	6.82	7.46	8.45	42.00	5.43	5.62	0.72	0.02	0.02	214.92	5.02	5.54	-985.29	145.98	150.09
IEE	6.87	6.15	6.96	42.04	6.14	6.61	0.72	0.02	0.02	179.44	6.16	7.27	-920.81	165.33	175.97
CWGEE	7.05	4.35	4.45	42.00	4.62	4.80	0.72	0.02	0.02	189.53	4.40	4.33	-914.43	124.33	128.30
ARGEE	6.67	6.07	6.53	41.86	5.76	5.89	0.72	0.02	0.02	243.96	3.59	3.91	-1023.31	154.99	157.53

Table C.28: Extension 2 Scenario 26 (γ_1=-0.35, γ_2=-0.35, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.89	3.38	3.41	41.96	3.74	3.84	0.72	0.01	0.01	244.75	2.39	2.43	-1042.65	100.64	103.92
EGEE	6.92	6.52	7.11	41.95	5.62	5.57	0.72	0.02	0.02	228.48	4.54	4.76	-1010.53	151.45	151.28
IEE	6.84	5.98	6.60	42.01	6.43	6.48	0.72	0.02	0.02	198.07	5.93	6.42	-957.43	173.67	176.26
CWGEE	6.82	4.34	4.43	41.92	4.93	4.88	0.72	0.02	0.02	202.90	4.20	4.18	-937.67	132.92	132.59
ARGEE	6.99	5.00	5.40	41.88	5.76	5.86	0.72	0.02	0.02	245.31	3.14	3.45	-1029.53	154.90	159.30

Table C.29: Extension 2 Scenario 27 (γ_1=-0.46, γ_2=-0.19, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.03	3.34	3.35	41.87	3.75	3.92	0.72	0.01	0.01	243.86	2.36	2.38	-1039.51	100.88	105.05
EGEE	6.81	8.37	9.68	41.73	5.34	5.67	0.72	0.02	0.02	205.92	5.67	6.28	-960.78	143.90	151.93
IEE	6.93	6.55	7.52	41.66	6.19	6.70	0.72	0.02	0.02	170.77	6.85	7.97	-891.43	167.47	181.36
CWGEE	6.98	4.58	4.62	41.75	4.70	4.98	0.72	0.02	0.02	184.88	4.79	4.53	-902.50	126.74	133.69
ARGEE	7.02	6.33	7.15	41.86	5.81	6.00	0.72	0.02	0.02	242.03	3.74	4.15	-1027.02	156.53	161.30

Table C.30: Extension 2 Scenario 28 (γ_2=-0.46, γ_1=-0.19, treatment effect = 7 g/week)

Interaction =0

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.15	3.37	3.50	41.99	3.73	3.67	0.72	0.01	0.01	245.10	2.38	2.40	-1043.71	100.40	98.57
EGEE	-0.39	5.98	6.30	41.76	5.83	5.99	0.72	0.02	0.02	244.65	4.16	4.60	-1037.00	157.06	161.00
IEE	-0.33	6.00	6.32	41.69	6.86	7.10	0.72	0.02	0.03	224.03	5.78	6.54	-1000.31	185.38	191.33
CWGEE	-0.24	4.49	4.55	41.75	5.42	5.60	0.72	0.02	0.02	222.47	4.15	4.32	-974.50	146.09	150.87
ARGEE	-0.26	3.93	4.12	41.84	5.89	6.08	0.72	0.02	0.02	245.92	2.69	2.84	-1036.74	158.47	163.91

Table C.31: Extension 2 Scenario 31 (γ_1=-0.50, γ_2=0, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.01	3.33	3.29	41.94	3.74	3.75	0.72	0.01	0.01	243.50	2.36	2.31	-1039.45	100.73	101.30
EGEE	-0.04	8.71	9.86	42.01	5.44	5.40	0.72	0.02	0.02	202.32	5.86	6.52	-959.17	146.78	145.83
IEE	0.04	6.87	7.70	42.01	6.56	6.80	0.72	0.02	0.02	172.81	7.13	8.42	-900.46	177.26	183.11
CWGEE	0.09	4.86	4.89	42.01	5.13	5.10	0.72	0.02	0.02	188.77	5.02	4.75	-921.44	138.23	138.02
ARGEE	-0.03	5.46	6.31	42.16	5.94	5.97	0.72	0.02	0.02	239.55	3.34	3.93	-1041.95	159.60	161.00

Table C.32: Extension 2 Scenario 32 (γ_2=-0.50, γ_1=0, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.18	3.36	3.53	42.14	3.75	3.71	0.72	0.01	0.01	244.09	2.38	2.37	-1046.50	100.86	99.62
EGEE	0.51	7.48	8.50	42.20	5.39	5.23	0.72	0.02	0.02	214.54	5.16	5.85	-990.28	145.04	141.06
IEE	0.38	6.15	7.03	42.28	6.07	6.32	0.72	0.02	0.02	179.10	6.31	7.69	-925.68	163.73	169.98
CWGEE	0.26	4.35	4.55	42.18	4.59	4.48	0.72	0.02	0.02	189.43	4.46	4.47	-918.96	123.63	120.97
ARGEE	0.50	6.11	6.78	42.05	5.73	5.74	0.72	0.02	0.02	243.59	3.69	3.93	-1029.31	154.17	154.74

Table C.33: Extension 2 Scenario 33 (γ_1=-0.35, γ_2=-0.35, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.08	3.37	3.25	41.88	3.75	3.63	0.72	0.01	0.01	244.65	2.39	2.33	-1040.66	100.85	97.15
EGEE	0.23	6.44	6.98	42.12	5.63	5.64	0.72	0.02	0.02	228.29	4.52	4.81	-1015.05	151.65	151.68
IEE	0.24	5.91	6.40	42.12	6.45	6.61	0.72	0.02	0.02	197.89	5.88	6.60	-960.43	174.22	178.00
CWGEE	0.09	4.31	4.34	42.12	4.94	4.95	0.72	0.02	0.02	202.70	4.19	4.20	-942.83	133.17	133.33
ARGEE	0.03	4.94	5.35	42.17	5.79	5.89	0.72	0.02	0.02	245.22	3.14	3.36	-1036.85	155.83	159.76

Table C.34: Extension 2 Scenario 34 (γ_1=-0.46, γ_2=-0.19, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.32	3.35	3.39	41.80	3.74	3.62	0.72	0.01	0.01	243.77	2.37	2.34	-1037.34	100.60	97.73
EGEE	-0.06	8.43	9.80	41.73	5.32	5.18	0.72	0.02	0.02	205.75	5.67	6.27	-959.35	143.54	139.35
IEE	-0.05	6.60	7.65	41.69	6.16	6.15	0.72	0.02	0.02	170.49	6.85	8.04	-889.81	167.03	166.09
CWGEE	0.09	4.60	4.75	41.76	4.68	4.58	0.72	0.02	0.02	184.68	4.80	4.61	-901.22	126.42	123.52
ARGEE	0.19	6.41	7.53	41.85	5.81	5.94	0.72	0.02	0.02	241.97	3.76	4.33	-1026.83	156.44	159.63

Table C.35: Extension 2 Scenario 35 (γ_2=-0.46, γ_1=-0.19, treatment effect = 0 g/week)

C.3 Scenarios with Larger SD of Cluster Size, γ_0 = 2.787813

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.90	3.42	3.39	41.87	3.75	3.79	0.72	0.01	0.01	245.46	2.42	2.35	-1041.21	100.96	101.48
EGEE	27.31	10.65	11.32	41.85	6.21	6.25	0.72	0.02	0.02	245.07	7.32	7.88	-1039.97	167.52	167.45
IEE	27.31	10.29	11.11	41.77	11.39	12.30	0.72	0.04	0.04	228.27	9.04	10.00	-1026.20	307.35	331.70
CWGEE	27.65	6.44	6.55	41.87	5.64	5.66	0.72	0.02	0.02	222.87	6.02	6.27	-979.58	152.16	151.75
ARGEE	27.85	4.01	3.95	41.94	5.69	5.74	0.72	0.02	0.02	247.30	2.74	2.73	-1035.77	152.94	153.56

Table C.36: Extension 2 Scenario 36 (γ_1=-0.90, γ_2=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.19	3.33	3.32	42.27	3.75	3.82	0.72	0.01	0.01	242.33	2.36	2.39	-1047.42	100.89	102.80
EGEE	27.82	9.96	10.99	42.19	4.77	4.74	0.72	0.02	0.02	174.13	6.87	7.55	-911.12	128.78	128.79
IEE	27.98	6.89	7.92	41.93	7.30	7.91	0.72	0.03	0.03	149.58	6.09	7.22	-854.49	197.26	214.61
CWGEE	28.02	5.38	5.70	42.17	4.53	4.52	0.72	0.02	0.02	166.84	5.21	5.32	-890.88	122.48	122.67
ARGEE	28.20	5.29	5.44	42.60	5.48	5.51	0.72	0.02	0.02	234.11	3.37	3.30	-1049.25	147.63	148.76

Table C.37: Extension 2 Scenario 37 (γ_2=-0.90, γ_1=0, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.98	3.38	3.46	41.99	3.76	3.85	0.72	0.01	0.01	243.83	2.39	2.45	-1042.38	101.11	103.88
EGEE	26.87	10.39	10.99	42.02	4.71	4.86	0.72	0.02	0.02	195.22	7.05	7.59	-947.00	127.27	131.87
IEE	27.12	7.67	8.38	42.00	7.54	8.41	0.72	0.03	0.03	165.77	6.69	7.71	-901.19	204.44	227.74
CWGEE	27.54	5.00	5.17	42.06	3.67	3.79	0.72	0.01	0.01	174.35	4.78	5.09	-891.30	99.11	102.95
ARGEE	27.73	5.59	5.69	42.22	5.25	5.26	0.72	0.02	0.02	239.47	3.50	3.49	-1031.68	141.34	140.99

Table C.38: Extension 2 Scenario 38 (γ_1=-0.64, γ_2=-0.64, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.01	3.41	3.34	41.88	3.75	3.83	0.72	0.01	0.01	244.71	2.41	2.37	-1040.37	100.96	103.59
EGEE	27.63	10.48	11.26	41.96	5.49	5.62	0.72	0.02	0.02	217.83	7.16	7.91	-989.75	148.39	151.73
IEE	27.64	9.09	9.96	41.81	9.69	10.64	0.72	0.03	0.04	192.75	7.98	9.14	-954.24	262.30	285.22
CWGEE	27.89	5.65	5.76	41.96	4.59	4.69	0.72	0.02	0.02	194.70	5.33	5.67	-926.42	123.90	126.68
ARGEE	27.94	4.72	4.62	41.95	5.34	5.32	0.72	0.02	0.02	243.46	3.07	3.07	-1028.44	143.68	143.16

Table C.39: Extension 2 Scenario 39 (γ_1=-0.83, γ_2=-0.34, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.95	3.35	3.46	42.02	3.76	3.80	0.72	0.01	0.01	243.08	2.37	2.52	-1041.38	101.11	102.20
EGEE	28.08	10.07	11.57	41.97	4.41	4.46	0.72	0.02	0.02	179.38	7.02	7.78	-916.18	119.18	119.76
IEE	28.11	6.54	7.80	42.04	6.44	7.39	0.72	0.02	0.03	150.23	5.86	7.16	-867.08	174.33	198.72
CWGEE	28.00	4.85	5.28	41.97	3.66	3.72	0.72	0.01	0.01	164.38	4.80	5.14	-874.61	98.94	100.06
ARGEE	27.88	5.85	5.85	41.86	5.32	5.29	0.72	0.02	0.02	236.20	3.67	3.54	-1022.83	143.24	141.64

Table C.40: Extension 2 Scneario 40 (γ_2=-0.83, γ_1=-0.34, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.95	3.41	3.43	41.83	3.75	3.74	0.72	0.01	0.01	245.56	2.42	2.43	-1039.61	100.87	99.81
EGEE	20.65	10.60	11.85	41.79	6.19	6.30	0.72	0.02	0.02	244.61	7.30	8.13	-1036.55	166.98	170.56
IEE	20.67	10.26	11.63	42.16	11.30	11.98	0.72	0.04	0.04	227.57	9.02	10.32	-1032.39	305.04	323.20
CWGEE	20.76	6.42	6.82	41.88	5.62	5.74	0.72	0.02	0.02	222.54	6.00	6.40	-977.97	151.46	155.44
ARGEE	20.82	4.03	4.00	41.78	5.67	5.86	0.72	0.02	0.02	247.35	2.76	2.78	-1031.02	152.60	158.22

Table C.41: Extension 2 Scenario 41 (γ_1=-0.90, γ_2=0, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.88	3.33	3.28	42.27	3.75	3.77	0.72	0.01	0.01	242.54	2.36	2.42	-1046.54	100.81	100.70
EGEE	20.66	9.90	11.01	42.37	4.76	4.85	0.72	0.02	0.02	174.26	6.85	7.73	-915.50	128.48	130.02
IEE	20.71	6.82	7.85	42.54	7.24	8.19	0.72	0.03	0.03	149.91	6.00	7.45	-870.41	195.39	219.71
CWGEE	20.72	5.35	5.60	42.36	4.53	4.63	0.72	0.02	0.02	167.13	5.18	5.47	-895.97	122.19	124.12
ARGEE	20.94	5.26	5.24	42.35	5.48	5.63	0.72	0.02	0.02	234.26	3.35	3.32	-1042.86	147.29	151.93

Table C.42: Extension 2 Scenario 42 (γ_2=-0.90, γ_1=0, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.15	3.39	3.48	41.84	3.76	3.82	0.72	0.01	0.01	243.76	2.40	2.48	-1038.63	101.26	103.45
EGEE	21.40	10.49	11.50	41.90	4.77	4.82	0.72	0.02	0.02	194.45	7.21	7.71	-944.29	128.63	130.98
IEE	21.36	7.76	8.73	42.14	7.74	8.60	0.72	0.03	0.03	165.08	6.87	7.67	-906.37	208.46	231.95
CWGEE	21.22	5.04	5.30	41.92	3.71	3.76	0.72	0.01	0.01	173.91	4.89	5.10	-887.82	99.94	101.98
ARGEE	21.16	5.63	5.93	41.73	5.26	5.19	0.72	0.02	0.02	239.25	3.55	3.56	-1017.52	141.66	138.91

Table C.43: Extension 2 Scenario 43 (γ_1=-0.64, γ_2=-0.64, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.00	3.40	3.54	42.22	3.76	3.77	0.72	0.01	0.01	244.63	2.41	2.50	-1049.12	101.14	101.45
EGEE	21.25	10.52	11.44	42.02	5.51	5.57	0.72	0.02	0.02	217.63	7.21	8.17	-991.85	148.68	150.93
IEE	21.25	9.14	10.10	41.81	9.69	10.47	0.72	0.03	0.04	192.67	8.02	9.47	-955.38	261.79	282.40
CWGEE	21.17	5.67	5.94	41.96	4.60	4.68	0.72	0.02	0.02	194.60	5.36	5.83	-926.78	124.16	126.73
ARGEE	21.01	4.66	4.71	42.06	5.35	5.25	0.72	0.02	0.02	243.34	3.05	3.10	-1030.98	144.04	141.06

Table C.44: Extension 2 Scenario 44 (γ_1=-0.83, γ_2=-0.34, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.02	3.36	3.25	41.98	3.76	3.86	0.72	0.01	0.01	243.04	2.38	2.35	-1040.99	101.17	104.07
EGEE	21.21	10.14	10.69	41.75	4.40	4.47	0.72	0.02	0.02	178.98	7.04	7.35	-910.54	119.12	120.05
IEE	21.17	6.60	7.20	41.67	6.44	7.15	0.72	0.02	0.03	149.82	5.86	6.90	-857.08	174.61	194.16
CWGEE	21.05	4.88	4.91	41.80	3.66	3.75	0.72	0.01	0.01	164.09	4.81	5.05	-870.53	98.91	100.91
ARGEE	21.09	5.87	5.77	41.98	5.34	5.40	0.72	0.02	0.02	236.00	3.69	3.45	-1026.55	143.59	144.36

Table C.45: Extension 2 Scneario 45 (γ_2=-0.83, γ_1=-0.34, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.95	3.42	3.39	41.99	3.76	3.86	0.72	0.01	0.01	245.49	2.42	2.38	-1045.29	101.10	103.99
EGEE	13.79	10.61	11.82	41.94	6.21	6.36	0.72	0.02	0.02	245.02	7.32	8.10	-1042.35	167.38	171.59
IEE	13.80	10.25	11.62	41.84	11.37	12.02	0.72	0.04	0.04	228.23	9.03	10.20	-1026.65	306.95	323.87
CWGEE	13.89	6.43	6.78	41.89	5.64	5.73	0.72	0.02	0.02	223.00	6.02	6.32	-979.82	152.01	154.36
ARGEE	13.92	4.01	4.03	41.87	5.68	5.89	0.72	0.02	0.02	247.32	2.74	2.77	-1034.33	152.65	158.41

Table C.46: Extension 2 Scenario 46 (γ_1=-0.90, γ_2=0, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.07	3.32	3.30	42.04	3.75	3.85	0.72	0.01	0.01	242.46	2.35	2.31	-1040.80	100.97	103.06
EGEE	13.85	9.97	10.94	41.99	4.77	4.84	0.72	0.02	0.02	173.98	6.95	7.66	-905.52	128.84	131.07
IEE	14.03	6.90	7.78	41.89	7.35	8.02	0.72	0.03	0.03	149.53	6.13	7.44	-852.86	198.29	216.21
CWGEE	13.95	5.40	5.68	41.99	4.55	4.62	0.72	0.02	0.02	166.97	5.26	5.54	-886.33	122.65	124.70
ARGEE	14.01	5.29	5.22	42.12	5.48	5.51	0.72	0.02	0.02	234.23	3.38	3.22	-1036.79	147.36	148.34

Table C.47: Extension 2 Scenario 47 (γ_2=-0.90, γ_1=0, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.85	3.38	3.37	41.87	3.76	3.68	0.72	0.01	0.01	243.91	2.39	2.43	-1039.29	101.20	99.47
EGEE	13.49	10.44	11.40	42.06	4.75	4.69	0.72	0.02	0.02	194.91	7.14	7.87	-947.85	128.46	126.52
IEE	13.62	7.72	8.67	42.21	7.70	8.26	0.72	0.03	0.03	165.44	6.82	7.95	-907.17	209.07	223.05
CWGEE	13.82	5.03	5.21	42.06	3.69	3.67	0.72	0.01	0.01	174.15	4.86	5.17	-890.98	99.86	98.95
ARGEE	13.71	5.64	5.52	41.97	5.27	5.45	0.72	0.02	0.02	239.50	3.54	3.40	-1023.62	141.81	145.03

Table C.48: Extension 2 Scenario 48 (γ_1=-0.64, γ_2=-0.64, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.06	3.41	3.55	42.06	3.77	3.93	0.72	0.01	0.01	244.73	2.42	2.46	-1045.20	101.30	105.71
EGEE	13.51	10.66	11.46	41.95	5.51	5.58	0.72	0.02	0.02	218.04	7.36	8.03	-990.03	148.71	150.65
IEE	13.58	9.26	10.22	41.83	9.74	10.31	0.72	0.03	0.04	192.91	8.21	9.28	-956.62	262.92	277.48
CWGEE	13.84	5.72	5.93	41.93	4.60	4.67	0.72	0.02	0.02	194.82	5.43	5.69	-925.43	123.99	125.92
ARGEE	13.95	4.72	4.94	41.96	5.35	5.51	0.72	0.02	0.02	243.49	3.08	3.12	-1027.15	143.95	147.26

Table C.49: Extension 2 Scenario 49 (γ_1=-0.83, γ_2=-0.34, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.01	3.35	3.40	42.00	3.76	3.83	0.72	0.01	0.01	243.00	2.37	2.39	-1040.62	101.08	103.09
EGEE	14.35	10.20	10.99	41.73	4.39	4.55	0.72	0.02	0.02	179.16	7.08	7.62	-909.08	118.73	122.56
IEE	14.29	6.63	7.36	41.38	6.39	7.32	0.72	0.02	0.03	150.01	5.93	7.10	-849.64	173.44	198.25
CWGEE	14.11	4.89	5.02	41.78	3.65	3.78	0.72	0.01	0.01	164.32	4.85	5.09	-869.45	98.73	102.03
ARGEE	14.27	5.90	6.01	42.14	5.33	5.30	0.72	0.02	0.02	235.93	3.71	3.68	-1029.01	143.46	142.39

Table C.50: Extension 2 Scneario 50 (γ_2=-0.83, γ_1=-0.34, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.88	3.42	3.41	42.06	3.74	3.79	0.72	0.01	0.01	245.57	2.42	2.36	-1045.69	100.73	102.48
EGEE	7.33	10.80	11.48	42.31	6.20	6.28	0.72	0.02	0.02	244.48	7.46	8.16	-1050.80	166.95	169.67
IEE	7.30	10.42	11.24	42.59	11.40	12.55	0.72	0.04	0.05	227.64	9.20	10.51	-1046.96	307.61	334.12
CWGEE	7.05	6.50	6.49	42.26	5.63	5.73	0.72	0.02	0.02	222.66	6.09	6.35	-988.69	151.69	154.22
ARGEE	6.90	4.03	3.97	42.20	5.67	5.68	0.72	0.02	0.02	247.34	2.76	2.63	-1041.47	152.50	153.43

Table C.51: Extension 2 Scenario 51 (γ_1=-0.90, γ_2=0, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.08	3.34	3.41	42.04	3.74	3.81	0.72	0.01	0.01	242.44	2.36	2.42	-1042.63	100.73	102.73
EGEE	7.17	9.97	10.77	42.09	4.77	4.80	0.72	0.02	0.02	173.73	6.95	7.26	-909.44	128.90	130.48
IEE	7.07	6.88	7.74	42.17	7.33	7.82	0.72	0.03	0.03	149.30	6.10	7.06	-861.81	198.32	211.62
CWGEE	7.01	5.40	5.65	42.07	4.54	4.60	0.72	0.02	0.02	166.74	5.25	5.33	-889.62	122.60	124.73
ARGEE	7.18	5.28	5.21	41.96	5.47	5.57	0.72	0.02	0.02	234.05	3.38	3.26	-1033.85	147.26	150.30

Table C.52: Extension 2 Scenario 52 (γ_2=-0.90, γ_1=0, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.00	3.39	3.46	42.01	3.77	3.86	0.72	0.01	0.01	243.67	2.40	2.44	-1041.51	101.48	103.58
EGEE	6.50	10.31	11.07	42.03	4.75	4.93	0.72	0.02	0.02	194.75	7.05	7.89	-947.35	128.32	133.03
IEE	6.61	7.61	8.38	42.05	7.62	8.43	0.72	0.03	0.03	165.35	6.71	7.88	-905.69	206.07	226.55
CWGEE	6.81	4.98	5.17	42.00	3.69	3.84	0.72	0.01	0.01	173.93	4.80	5.11	-889.88	99.68	103.62
ARGEE	6.99	5.55	5.73	42.10	5.27	5.50	0.72	0.02	0.02	239.17	3.49	3.50	-1026.11	141.82	149.25

Table C.53: Extension 2 Scenario 53 (γ_1=-0.64, γ_2=-0.64, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.17	3.41	3.36	42.05	3.75	3.68	0.72	0.01	0.01	244.53	2.42	2.43	-1045.34	101.02	99.30
EGEE	7.21	10.50	11.45	42.10	5.48	5.39	0.72	0.02	0.02	217.57	7.30	8.02	-994.31	147.92	145.38
IEE	7.11	9.13	10.18	42.02	9.60	10.39	0.72	0.03	0.04	192.61	8.13	9.23	-960.25	259.44	282.04
CWGEE	7.08	5.65	5.83	42.07	4.56	4.55	0.72	0.02	0.02	194.58	5.40	5.57	-929.87	123.19	122.64
ARGEE	7.33	4.69	4.55	42.16	5.35	5.27	0.72	0.02	0.02	243.13	3.08	2.96	-1034.71	143.90	140.69

Table C.54: Extension 2 Scenario 54 (γ_1=-0.83, γ_2=-0.34, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.12	3.35	3.46	42.07	3.76	3.80	0.72	0.01	0.01	242.94	2.37	2.45	-1042.76	101.04	101.65
EGEE	7.24	10.11	11.02	42.21	4.39	4.53	0.72	0.02	0.02	179.22	6.97	7.54	-922.52	118.54	122.07
IEE	7.10	6.55	7.43	42.35	6.41	7.15	0.72	0.02	0.03	150.17	5.80	6.89	-875.19	172.89	192.58
CWGEE	7.12	4.85	5.14	42.18	3.64	3.75	0.72	0.01	0.01	164.29	4.77	5.07	-880.26	98.38	101.18
ARGEE	7.22	5.85	5.94	42.01	5.31	5.67	0.72	0.02	0.02	235.95	3.67	3.61	-1027.43	142.86	152.52

Table C.55: Extension 2 Scneario 55 (γ_2=-0.83, γ_1=-0.34, treatment effect = 7 g/week)

Interaction =0

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.05	3.41	3.47	41.78	3.75	3.73	0.72	0.01	0.01	245.55	2.41	2.46	-1038.58	100.91	101.63
EGEE	-0.24	10.60	12.04	41.74	6.17	6.49	0.72	0.02	0.02	244.61	7.35	8.40	-1034.55	166.41	176.05
IEE	-0.26	10.24	11.83	41.45	11.20	12.77	0.72	0.04	0.04	227.61	9.06	10.60	-1012.97	302.94	343.08
CWGEE	-0.13	6.41	6.95	41.82	5.59	5.94	0.72	0.02	0.02	222.58	6.01	6.58	-975.92	150.80	161.07
ARGEE	-0.05	4.02	4.11	41.84	5.65	5.92	0.72	0.02	0.02	247.34	2.76	2.83	-1031.73	152.08	159.98

Table C.56: Extension 2 Scenario 56 (γ_1=-0.90, γ_2=0, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.12	3.33	3.44	42.04	3.76	3.74	0.72	0.01	0.01	242.54	2.36	2.42	-1041.66	101.07	100.70
EGEE	-0.08	10.05	10.86	42.11	4.78	4.77	0.72	0.02	0.02	173.89	6.92	7.36	-909.25	129.10	128.34
IEE	-0.01	6.93	7.80	42.26	7.32	7.79	0.72	0.03	0.03	149.41	6.09	7.08	-862.77	198.17	209.11
CWGEE	-0.00	5.41	5.62	42.11	4.55	4.55	0.72	0.02	0.02	166.85	5.25	5.22	-889.96	122.89	122.44
ARGEE	-0.13	5.29	5.37	41.77	5.48	5.60	0.72	0.02	0.02	234.26	3.37	3.28	-1027.57	147.59	150.89

Table C.57: Extension 2 Scenario 57 (γ_2=-0.90, γ_1=0, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.15	3.39	3.45	41.93	3.77	3.63	0.72	0.01	0.01	243.76	2.40	2.40	-1040.71	101.35	97.54
EGEE	0.50	10.31	11.85	41.93	4.74	4.67	0.72	0.02	0.02	194.36	7.08	7.81	-944.77	128.20	125.29
IEE	0.39	7.60	8.99	41.78	7.59	8.54	0.72	0.03	0.03	165.03	6.71	7.79	-896.40	206.11	230.26
CWGEE	0.33	4.98	5.50	41.90	3.68	3.67	0.72	0.01	0.01	173.83	4.82	5.15	-887.06	99.57	98.47
ARGEE	0.17	5.58	5.84	41.80	5.25	5.27	0.72	0.02	0.02	239.17	3.53	3.48	-1019.27	141.32	141.12

Table C.58: Extension 2 Scenario 58 (γ_1=-0.64, γ_2=-0.64, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.11	3.41	3.46	41.98	3.75	3.76	0.72	0.01	0.01	244.68	2.41	2.52	-1043.04	100.93	101.27
EGEE	0.51	10.67	11.73	41.82	5.52	5.65	0.72	0.02	0.02	217.31	7.31	7.92	-986.45	148.85	151.00
IEE	0.49	9.26	10.45	41.73	9.76	10.57	0.72	0.03	0.04	192.23	8.12	9.02	-953.05	263.90	284.11
CWGEE	0.31	5.73	6.04	41.83	4.60	4.74	0.72	0.02	0.02	194.45	5.41	5.61	-923.18	124.28	126.77
ARGEE	0.08	4.71	4.74	41.80	5.35	5.47	0.72	0.02	0.02	243.44	3.08	3.14	-1024.30	144.07	146.17

Table C.59: Extension 2 Scenario 59 (γ_1=-0.83, γ_2=-0.34, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.13	3.35	3.40	42.14	3.76	3.73	0.72	0.01	0.01	243.11	2.37	2.34	-1044.89	101.26	99.38
EGEE	-0.00	10.25	11.10	42.07	4.43	4.32	0.72	0.02	0.02	179.18	7.13	7.75	-918.07	119.82	117.29
IEE	-0.07	6.68	7.55	41.94	6.60	7.17	0.72	0.02	0.03	150.04	5.94	7.25	-863.02	178.12	194.04
CWGEE	-0.14	4.91	5.20	42.05	3.69	3.60	0.72	0.01	0.01	164.32	4.84	5.15	-876.68	99.69	97.72
ARGEE	-0.18	5.85	5.96	42.26	5.34	5.39	0.72	0.02	0.02	236.18	3.66	3.55	-1033.57	143.71	144.30

Table C.60: Extension 2 Scneario 60 (γ_2=-0.83, γ_1=-0.34, treatment effect = 0 g/week)

Appendix D

GEE Results

In Section 5.3 the results for the GEE simulation study were discussed, which looked at repeated the Chapter 4 simulation study and Extension 2 using a GEE to simulate the data from. The results for the simulation scenarios with a treatment effect of 28 were provided. This appendix, D, gives the results for the full GEE simulation study, that is all 70 simulations discussed in section 5.3.

D.1 GEE Results for n=60

D.1.1 Scenarios with Non Informative Cluster Size

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.91	6.72	6.83	42.16	17.06	17.66	0.72	0.06	0.06	244.46	5.19	5.28	-1049.11	458.36	473.81
EGEE	27.93	7.63	7.80	42.17	16.54	17.52	0.72	0.06	0.06	244.44	5.43	5.59	-1049.28	445.40	470.03
IEE	27.85	9.82	10.13	42.16	16.38	17.54	0.72	0.06	0.06	244.49	6.31	6.50	-1049.21	441.13	470.31
CWGEE	27.85	9.82	10.13	42.16	16.38	17.54	0.72	0.06	0.06	244.49	6.31	6.50	-1049.21	441.13	470.31
ARGEE	27.90	6.09	6.27	42.07	15.89	17.00	0.72	0.06	0.06	244.49	4.66	4.83	-1046.82	427.98	455.81

Table D.1: Extension 3 Scenario 1 (fixed trial length=38, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.01	6.75	6.95	42.10	17.08	17.66	0.72	0.06	0.06	244.42	5.21	5.33	-1045.18	458.87	474.37
EGEE	27.92	7.70	7.96	42.05	16.52	17.49	0.72	0.06	0.06	244.49	5.46	5.68	-1043.97	445.18	469.52
IEE	28.08	9.59	10.11	42.06	16.38	17.59	0.72	0.06	0.06	244.45	6.48	6.77	-1044.28	441.33	472.32
CWGEE	28.10	9.59	10.05	42.05	16.34	17.50	0.72	0.06	0.06	244.44	6.48	6.73	-1044.07	440.31	469.71
ARGEE	28.03	6.03	6.25	42.06	15.86	16.88	0.72	0.06	0.06	244.44	4.59	4.76	-1043.86	427.35	452.62

Table D.2: Extension 3 Scenario 2 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.96	6.72	6.88	42.14	16.99	17.60	0.72	0.06	0.06	244.55	5.20	5.22	-1046.55	456.59	472.71
EGEE	20.98	7.64	7.81	42.12	16.43	17.41	0.72	0.06	0.06	244.54	5.42	5.55	-1046.08	442.74	467.89
IEE	20.91	9.79	10.17	42.13	16.27	17.46	0.72	0.06	0.06	244.58	6.30	6.40	-1046.37	438.26	468.85
CWGEE	20.91	9.79	10.17	42.13	16.27	17.46	0.72	0.06	0.06	244.58	6.30	6.40	-1046.37	438.26	468.85
ARGEE	20.93	6.10	6.32	42.10	15.79	16.86	0.72	0.06	0.06	244.57	4.66	4.78	-1045.50	425.46	453.04

Table D.3: Extension 3 Scenario 8 (fixed trial length=38, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.07	6.76	6.99	41.76	17.08	17.71	0.72	0.06	0.06	244.46	5.22	5.27	-1036.01	458.76	474.32
EGEE	21.02	7.69	7.96	41.74	16.49	17.53	0.72	0.06	0.06	244.47	5.46	5.64	-1035.49	444.12	469.64
IEE	21.02	9.60	10.11	41.74	16.36	17.63	0.72	0.06	0.06	244.48	6.50	6.75	-1035.59	440.48	472.04
CWGEE	21.03	9.60	10.07	41.71	16.32	17.57	0.72	0.06	0.06	244.47	6.49	6.71	-1034.99	439.45	470.59
ARGEE	21.04	6.03	6.31	41.74	15.84	16.98	0.72	0.06	0.06	244.46	4.61	4.73	-1035.78	426.66	454.39

Table D.4: Extension 3 Scenario 9 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.08	6.72	6.89	41.85	17.05	17.39	0.72	0.06	0.06	244.50	5.19	5.26	-1038.81	457.92	467.46
EGEE	14.03	7.63	7.87	41.86	16.54	17.28	0.72	0.06	0.06	244.53	5.43	5.60	-1038.80	445.41	464.68
IEE	14.18	9.81	10.23	41.84	16.37	17.28	0.72	0.06	0.06	244.45	6.30	6.46	-1038.64	440.87	464.67
CWGEE	14.18	9.81	10.23	41.84	16.37	17.28	0.72	0.06	0.06	244.45	6.30	6.46	-1038.64	440.87	464.67
ARGEE	14.10	6.09	6.30	41.95	15.89	16.63	0.72	0.06	0.06	244.48	4.66	4.78	-1041.41	428.12	447.30

Table D.5: Extension 3 Scenario 15 (fixed trial length=38, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	13.94	6.75	6.95	41.85	17.09	17.65	0.72	0.06	0.06	244.47	5.21	5.28	-1038.44	458.93	473.56
EGEE	13.94	7.69	7.97	41.88	16.52	17.47	0.72	0.06	0.06	244.48	5.45	5.67	-1039.02	445.07	468.86
IEE	14.02	9.60	10.08	41.88	16.38	17.60	0.72	0.06	0.06	244.44	6.48	6.72	-1039.16	441.21	472.31
CWGEE	14.02	9.60	10.05	41.87	16.34	17.50	0.72	0.06	0.06	244.43	6.47	6.67	-1038.95	440.20	469.63
ARGEE	13.94	6.02	6.26	41.87	15.87	16.87	0.72	0.06	0.06	244.49	4.59	4.75	-1039.06	427.64	452.88

Table D.6: Extension 3 Scenario 16 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.05	6.71	6.86	41.99	17.03	17.61	0.72	0.06	0.06	244.50	5.19	5.25	-1042.39	457.52	473.71
EGEE	7.01	7.63	7.81	42.00	16.50	17.44	0.72	0.06	0.06	244.52	5.42	5.57	-1042.67	444.35	469.25
IEE	7.07	9.79	10.18	42.00	16.34	17.46	0.72	0.06	0.06	244.49	6.30	6.47	-1042.53	439.89	469.76
CWGEE	7.07	9.79	10.18	42.00	16.34	17.46	0.72	0.06	0.06	244.49	6.30	6.47	-1042.53	439.89	469.76
ARGEE	7.03	6.08	6.30	41.95	15.86	16.88	0.72	0.06	0.06	244.51	4.65	4.78	-1041.10	427.15	454.74

Table D.7: Extension 3 Scenario 22 (fixed trial length=38, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.99	6.75	6.89	42.05	17.06	17.51	0.72	0.06	0.06	244.51	5.22	5.28	-1045.21	458.42	469.47
EGEE	6.98	7.70	7.96	42.07	16.51	17.31	0.72	0.06	0.06	244.50	5.46	5.65	-1045.81	444.77	463.97
IEE	7.02	9.58	10.07	42.09	16.36	17.37	0.72	0.06	0.06	244.51	6.49	6.77	-1046.67	440.86	465.61
CWGEE	7.02	9.59	10.02	42.07	16.33	17.32	0.72	0.06	0.06	244.53	6.48	6.73	-1046.03	439.87	464.23
ARGEE	7.02	6.02	6.26	42.12	15.85	16.77	0.72	0.06	0.06	244.49	4.60	4.78	-1047.11	427.03	449.24

Table D.8: Extension 3 Scenario 23 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 7 g/week)

Interaction =0

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.22	6.71	6.88	41.88	17.03	17.61	0.72	0.06	0.06	244.34	5.19	5.30	-1039.98	457.74	472.82
EGEE	0.18	7.63	7.83	41.85	16.50	17.47	0.72	0.06	0.06	244.35	5.42	5.61	-1039.22	444.58	469.17
IEE	0.32	9.79	10.23	41.88	16.34	17.50	0.72	0.06	0.06	244.28	6.30	6.50	-1039.85	440.06	469.92
CWGEE	0.32	9.79	10.23	41.88	16.34	17.50	0.72	0.06	0.06	244.28	6.30	6.50	-1039.85	440.06	469.92
ARGEE	0.20	6.09	6.35	41.80	15.86	16.83	0.72	0.06	0.06	244.35	4.65	4.85	-1037.78	427.34	451.57

Table D.9: Extension 3 Scenario 29 (fixed trial length=38, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.06	6.75	6.91	41.96	17.06	17.64	0.72	0.06	0.06	244.60	5.21	5.28	-1042.94	458.54	473.53
EGEE	-0.02	7.70	7.90	41.98	16.51	17.44	0.72	0.06	0.06	244.58	5.47	5.65	-1043.26	445.10	468.11
IEE	-0.13	9.59	10.03	41.97	16.38	17.53	0.72	0.06	0.06	244.62	6.50	6.76	-1043.07	441.48	470.72
CWGEE	-0.13	9.59	10.00	41.99	16.34	17.45	0.72	0.06	0.06	244.62	6.49	6.71	-1043.34	440.45	468.47
ARGEE	-0.08	6.02	6.26	41.93	15.85	16.90	0.72	0.06	0.06	244.60	4.61	4.81	-1041.70	427.28	453.50

Table D.10: Extension 3 Scenario 30 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 0 g/week)

D.1.2 Scenarios with Informative Cluster Size using $\gamma_0 = 3.069142$

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.97	7.22	7.35	42.30	17.31	17.72	0.72	0.06	0.06	247.43	5.51	5.43	-1059.86	465.39	476.42
EGEE	27.73	8.16	8.74	42.15	16.57	17.39	0.72	0.06	0.06	247.08	5.63	6.07	-1052.48	447.18	467.80
IEE	27.06	9.21	9.91	41.46	16.11	17.31	0.71	0.06	0.06	235.66	6.42	6.89	-992.90	435.05	465.98
CWGEE	26.95	9.27	9.62	41.18	15.96	16.98	0.71	0.06	0.06	234.59	6.46	6.73	-966.19	431.06	458.11
ARGEE	27.85	6.24	6.58	42.09	15.84	16.76	0.72	0.06	0.06	244.83	4.58	4.82	-1044.01	427.18	450.76

Table D.11: Extension 3 Scenario 3 (γ_1=-0.36, γ_2=-0.22, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.41	7.38	7.46	41.56	17.36	17.76	0.72	0.06	0.06	251.27	5.61	5.35	-1044.47	466.43	476.92
EGEE	26.88	7.93	8.71	41.49	16.47	17.35	0.72	0.06	0.06	244.40	5.29	5.97	-1029.41	443.75	465.95
IEE	26.85	8.93	9.77	41.42	16.35	17.80	0.72	0.06	0.06	239.48	5.92	6.84	-1017.96	440.61	478.07
CWGEE	27.01	9.11	9.52	41.48	16.19	17.29	0.72	0.06	0.06	240.46	6.22	6.64	-1017.16	436.24	464.67
ARGEE	27.75	6.02	6.31	41.53	15.77	16.90	0.72	0.06	0.06	244.78	4.26	4.51	-1030.76	424.85	453.82

Table D.12: Extension 3 Scenario 4 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.35	7.31	7.43	42.14	17.34	17.62	0.72	0.06	0.06	249.35	5.57	5.45	-1058.68	465.70	473.99
EGEE	27.32	8.12	8.80	41.91	16.52	17.15	0.72	0.06	0.06	245.76	5.52	5.99	-1040.85	445.36	461.39
IEE	26.92	9.28	10.03	41.65	16.20	17.27	0.71	0.06	0.06	236.53	6.28	6.85	-1006.76	436.83	465.11
CWGEE	26.91	9.37	9.80	41.48	16.07	16.92	0.71	0.06	0.06	236.63	6.43	6.75	-991.19	433.41	456.25
ARGEE	27.85	6.22	6.51	41.92	15.85	16.63	0.72	0.06	0.06	244.81	4.49	4.70	-1039.25	427.16	447.40

Table D.13: Extension 3 Scenario 5 (γ_1=-0.24, γ_2=-0.35, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.35	7.26	7.42	42.15	17.36	18.13	0.72	0.06	0.07	248.52	5.54	5.46	-1059.10	466.36	485.34
EGEE	27.58	8.14	8.67	41.95	16.58	17.66	0.72	0.06	0.06	246.31	5.55	6.04	-1044.51	447.12	473.39
IEE	27.05	9.29	10.07	41.50	16.18	17.70	0.71	0.06	0.06	235.99	6.33	6.91	-999.05	436.72	474.84
CWGEE	27.02	9.36	9.85	41.30	16.07	17.34	0.71	0.06	0.06	235.62	6.44	6.81	-979.48	433.61	466.07
ARGEE	27.98	6.35	6.73	42.03	16.36	17.91	0.72	0.06	0.07	244.78	4.61	4.93	-1043.12	438.62	473.22

Table D.14: Extension 3 Scenario 6 (γ_1=-0.29, γ_2=-0.29, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.49	7.33	7.60	41.91	17.40	18.19	0.72	0.06	0.06	250.20	5.58	5.42	-1054.49	467.29	488.43
EGEE	27.13	8.05	8.66	41.78	16.55	17.74	0.72	0.06	0.06	245.30	5.44	5.95	-1037.33	446.12	476.39
IEE	26.86	9.19	10.02	41.51	16.28	17.92	0.72	0.06	0.06	237.54	6.16	6.87	-1009.83	439.02	481.29
CWGEE	26.95	9.31	9.84	41.54	16.17	17.60	0.72	0.06	0.06	238.02	6.37	6.74	-1003.32	436.17	473.10
ARGEE	27.88	6.15	6.51	41.85	15.85	17.21	0.72	0.06	0.06	244.83	4.40	4.67	-1038.71	427.26	461.93

Table D.15: Extension 3 Scenario 7 (γ_1=-0.17, γ_2=-0.40, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.97	7.22	7.39	42.36	17.30	17.84	0.72	0.06	0.06	247.50	5.51	5.45	-1061.89	464.99	480.31
EGEE	20.69	8.15	8.74	42.18	16.55	17.47	0.72	0.06	0.06	247.13	5.64	6.06	-1053.62	446.58	470.39
IEE	20.41	9.19	9.92	41.58	16.09	17.43	0.71	0.06	0.06	235.60	6.48	6.98	-996.56	434.66	470.30
CWGEE	20.35	9.25	9.68	41.26	15.95	17.06	0.71	0.06	0.06	234.50	6.51	6.81	-968.57	430.93	460.97
ARGEE	20.81	6.24	6.57	42.12	15.83	16.81	0.72	0.06	0.06	244.89	4.61	4.83	-1045.25	427.07	452.68

Table D.16: Extension 3 Scenario 10 (γ_1=-0.36, γ_2=-0.22, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.44	7.38	7.50	42.11	17.35	17.77	0.72	0.06	0.06	251.40	5.62	5.44	-1060.45	466.23	478.83
EGEE	20.39	7.94	8.76	42.01	16.48	17.31	0.72	0.06	0.06	244.20	5.34	6.07	-1044.52	444.21	466.48
IEE	20.02	8.84	9.76	41.97	16.34	17.77	0.72	0.06	0.06	239.37	6.05	7.01	-1033.79	440.64	478.36
CWGEE	20.21	9.05	9.55	41.98	16.19	17.27	0.72	0.06	0.06	240.33	6.33	6.78	-1031.45	436.44	465.60
ARGEE	20.99	6.00	6.30	42.00	15.78	16.76	0.72	0.06	0.06	244.71	4.31	4.56	-1044.68	425.21	451.81

Table D.17: Extension 3 Scenario 11 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.15	7.32	7.52	42.17	17.35	17.94	0.72	0.06	0.06	249.57	5.58	5.50	-1060.60	466.23	480.66
EGEE	20.37	8.12	8.75	41.96	16.55	17.50	0.72	0.06	0.06	245.81	5.57	6.02	-1042.70	446.38	469.07
IEE	20.15	9.26	10.01	41.58	16.22	17.66	0.71	0.06	0.06	236.42	6.37	6.99	-1004.93	437.82	473.84
CWGEE	20.18	9.35	9.77	41.49	16.10	17.29	0.71	0.06	0.06	236.51	6.51	6.82	-992.14	434.46	464.13
ARGEE	20.80	6.21	6.59	42.05	15.88	16.92	0.72	0.06	0.06	244.91	4.53	4.77	-1043.55	428.14	453.13

Table D.18: Extension 3 Scenario 12 (γ_1=-0.24, γ_2=-0.35, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.09	7.25	7.43	41.79	17.31	17.64	0.73	0.06	0.06	248.79	5.53	5.50	-1050.70	465.26	473.40
EGEE	20.70	8.13	8.78	41.55	16.54	17.19	0.72	0.06	0.06	246.34	5.59	6.02	-1034.89	446.37	461.41
IEE	20.10	9.23	9.95	41.10	16.14	17.26	0.71	0.06	0.06	236.10	6.42	7.04	-989.71	435.91	463.69
CWGEE	20.06	9.32	9.73	40.89	16.02	16.90	0.71	0.06	0.06	235.75	6.51	6.90	-969.99	432.62	454.90
ARGEE	20.83	6.22	6.66	41.60	15.86	16.71	0.72	0.06	0.06	244.94	4.57	4.86	-1033.00	427.65	448.25

Table D.19: Extension 3 Scenario 13 (γ_1=-0.29, γ_2=-0.29, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.39	7.35	7.57	41.96	17.36	17.80	0.72	0.06	0.06	250.33	5.60	5.46	-1056.26	466.57	477.51
EGEE	20.38	8.06	8.70	41.78	16.53	17.30	0.72	0.06	0.06	245.14	5.50	6.06	-1037.31	445.65	464.49
IEE	20.20	9.16	9.84	41.50	16.25	17.51	0.72	0.06	0.06	237.28	6.29	7.02	-1009.73	438.29	470.39
CWGEE	20.25	9.29	9.63	41.49	16.15	17.13	0.72	0.06	0.06	237.80	6.48	6.82	-1002.04	435.69	460.56
ARGEE	21.00	6.15	6.54	41.86	15.84	16.67	0.72	0.06	0.06	244.78	4.46	4.69	-1039.25	427.09	447.12

Table D.20: Extension 3 Scenario 14 (γ_1=-0.17, γ_2=-0.40, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.01	7.21	7.42	42.03	17.32	17.91	0.72	0.06	0.06	247.55	5.51	5.51	-1052.37	465.43	481.39
EGEE	13.80	8.15	8.90	41.89	16.58	17.53	0.72	0.06	0.06	247.10	5.67	6.20	-1045.15	446.83	471.71
IEE	13.50	9.18	10.04	41.19	16.09	17.50	0.71	0.06	0.06	235.58	6.56	7.13	-985.43	434.49	472.03
CWGEE	13.46	9.25	9.81	40.94	15.96	17.13	0.71	0.06	0.06	234.49	6.57	6.94	-959.56	430.76	462.71
ARGEE	13.89	6.24	6.65	41.85	15.84	16.81	0.72	0.06	0.06	244.84	4.63	4.89	-1037.21	426.97	452.73

Table D.21: Extension 3 Scenario 17 (γ_1=-0.36, γ_2=-0.22, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.36	7.39	7.56	41.88	17.35	17.68	0.72	0.06	0.06	251.48	5.63	5.54	-1052.67	466.15	474.53
EGEE	13.57	7.90	8.81	41.90	16.50	17.20	0.72	0.06	0.06	244.11	5.41	6.14	-1039.41	444.68	461.44
IEE	13.33	8.77	9.62	41.86	16.33	17.63	0.72	0.06	0.06	239.30	6.18	7.20	-1029.13	440.43	473.27
CWGEE	13.45	9.00	9.46	41.89	16.22	17.17	0.72	0.06	0.06	240.28	6.45	6.92	-1027.29	437.10	460.74
ARGEE	14.00	5.98	6.37	41.89	15.79	16.61	0.72	0.06	0.06	244.68	4.37	4.66	-1039.30	425.70	445.26

Table D.22: Extension 3 Scenario 18 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.13	7.31	7.56	42.11	17.34	17.99	0.72	0.06	0.07	249.55	5.58	5.54	-1058.22	465.64	481.43
EGEE	13.68	8.10	8.75	41.87	16.49	17.54	0.72	0.06	0.06	245.57	5.60	6.14	-1039.35	444.39	469.37
IEE	13.29	9.17	10.05	41.52	16.15	17.69	0.71	0.06	0.06	236.33	6.49	7.20	-1002.80	435.50	473.81
CWGEE	13.33	9.29	9.82	41.43	16.04	17.34	0.71	0.06	0.06	236.42	6.60	7.01	-989.63	432.51	464.46
ARGEE	13.89	6.22	6.56	41.98	15.83	16.98	0.72	0.06	0.06	244.76	4.58	4.84	-1040.72	426.30	454.10

Table D.23: Extension 3 Scenario 19 (γ_1=-0.24, γ_2=-0.35, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.10	7.26	7.42	42.23	17.31	17.76	0.72	0.06	0.06	248.78	5.55	5.49	-1061.72	465.33	477.56
EGEE	13.76	8.12	8.87	41.98	16.55	17.41	0.72	0.06	0.06	246.23	5.62	6.14	-1045.53	446.14	468.13
IEE	13.51	9.20	10.03	41.43	16.13	17.43	0.71	0.06	0.06	235.81	6.53	7.19	-997.86	435.42	468.65
CWGEE	13.48	9.29	9.76	41.33	16.02	17.08	0.71	0.06	0.06	235.45	6.59	6.98	-980.50	432.32	460.37
ARGEE	13.88	6.22	6.62	41.97	15.86	16.87	0.72	0.06	0.06	244.85	4.60	4.85	-1041.14	427.53	453.35

Table D.24: Extension 3 Scenario 20 (γ_1=-0.29, γ_2=-0.29, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.13	7.36	7.62	42.19	17.35	17.83	0.72	0.06	0.06	250.54	5.61	5.53	-1061.11	466.12	477.18
EGEE	13.60	8.07	8.75	41.92	16.55	17.37	0.72	0.06	0.06	245.08	5.56	6.12	-1039.95	445.83	465.14
IEE	13.29	9.11	9.93	41.71	16.28	17.61	0.72	0.06	0.06	237.30	6.45	7.18	-1013.93	438.93	471.31
CWGEE	13.29	9.25	9.75	41.68	16.17	17.20	0.72	0.06	0.06	237.84	6.60	6.95	-1006.14	435.83	461.15
ARGEE	13.88	6.15	6.57	41.96	15.86	16.85	0.72	0.06	0.06	244.84	4.52	4.80	-1040.90	427.33	450.68

Table D.25: Extension 3 Scenario 21 (γ_1=-0.17, γ_2=-0.40, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.07	7.23	7.38	42.00	17.36	17.91	0.72	0.06	0.06	247.45	5.52	5.59	-1051.16	466.33	480.79
EGEE	6.92	8.16	8.70	41.85	16.61	17.58	0.72	0.06	0.06	246.97	5.73	6.18	-1043.58	447.84	471.91
IEE	6.81	9.19	9.90	41.11	16.13	17.54	0.71	0.06	0.06	235.29	6.64	7.14	-982.79	435.55	471.49
CWGEE	6.81	9.25	9.65	40.90	16.00	17.18	0.71	0.06	0.06	234.17	6.63	6.95	-957.83	431.65	462.55
ARGEE	7.01	6.24	6.59	41.76	15.91	17.00	0.72	0.06	0.06	244.75	4.67	4.91	-1034.43	428.61	456.49

Table D.26: Extension 3 Scenario 24 (γ_1=-0.36, γ_2=-0.22, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.11	7.40	7.60	41.80	17.38	17.64	0.72	0.06	0.06	251.72	5.64	5.57	-1052.33	466.86	473.34
EGEE	6.78	7.90	8.83	41.71	16.55	17.10	0.72	0.06	0.06	244.05	5.46	6.25	-1036.01	445.54	458.87
IEE	6.66	8.72	9.83	41.71	16.39	17.59	0.72	0.06	0.06	239.16	6.35	7.52	-1027.15	441.55	472.03
CWGEE	6.71	8.98	9.59	41.67	16.25	17.07	0.72	0.06	0.06	240.17	6.59	7.19	-1022.94	437.67	458.07
ARGEE	6.94	5.97	6.28	41.79	15.83	16.60	0.72	0.06	0.06	244.73	4.43	4.69	-1038.15	426.23	445.49

Table D.27: Extension 3 Scenario 25 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 7 g/week)

3	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.03	7.29	7.44	41.98	17.32	17.45	0.72	0.06	0.06	249.65	5.57	5.53	-1054.01	465.41	468.35
EGEE	6.76	8.09	8.73	41.73	16.50	17.05	0.72	0.06	0.06	245.63	5.64	6.15	-1035.28	444.80	458.12
IEE	6.59	9.19	9.94	41.40	16.15	17.19	0.71	0.06	0.06	236.26	6.64	7.29	-999.36	435.84	462.57
CWGEE	6.60	9.30	9.69	41.30	16.04	16.79	0.71	0.06	0.06	236.36	6.71	7.09	-985.72	432.62	452.55
ARGEE	6.92	6.19	6.47	41.85	15.81	16.63	0.72	0.06	0.06	244.80	4.61	4.83	-1036.36	426.21	446.96

Table D.28: Extension 3 Scenario 26 (γ_1=-0.24, γ_2=-0.35, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.15	7.26	7.46	42.05	17.34	17.50	0.72	0.06	0.06	248.76	5.55	5.50	-1056.15	465.82	470.12
EGEE	6.78	8.13	8.67	41.83	16.55	17.07	0.72	0.06	0.06	246.22	5.68	6.10	-1040.83	446.05	458.84
IEE	6.68	9.21	10.03	41.37	16.15	17.08	0.71	0.06	0.06	235.65	6.63	7.22	-995.09	435.83	459.93
CWGEE	6.70	9.30	9.85	41.19	16.02	16.77	0.71	0.06	0.06	235.26	6.67	6.99	-976.05	432.24	451.73
ARGEE	6.96	6.21	6.54	41.91	15.86	16.52	0.72	0.06	0.06	244.82	4.63	4.84	-1039.42	427.22	443.88

Table D.29: Extension 3 Scenario 27 (γ_1=-0.29, γ_2=-0.29, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.11	7.36	7.61	42.21	17.38	17.79	0.72	0.06	0.06	250.49	5.62	5.57	-1061.94	466.97	477.33
EGEE	6.74	8.06	8.87	41.99	16.53	17.31	0.72	0.06	0.06	244.93	5.61	6.22	-1041.94	445.44	464.63
IEE	6.79	9.07	10.00	41.68	16.23	17.48	0.72	0.06	0.06	236.99	6.56	7.40	-1013.64	437.76	469.63
CWGEE	6.79	9.23	9.82	41.73	16.13	17.17	0.72	0.06	0.06	237.52	6.70	7.14	-1007.59	435.05	461.68
ARGEE	6.95	6.13	6.54	41.99	15.84	16.75	0.72	0.06	0.06	244.72	4.57	4.84	-1041.75	426.83	449.28

Table D.30: Extension 3 Scenario 28 (γ_1=-0.17, γ_2=-0.40, treatment effect = 7 g/week)

Interaction =0

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.04	7.23	7.22	42.14	17.37	17.94	0.72	0.06	0.06	247.54	5.53	5.50	-1056.67	466.58	482.13
EGEE	-0.11	8.17	8.68	41.91	16.64	17.51	0.72	0.06	0.06	247.05	5.77	6.24	-1046.92	448.72	470.42
IEE	0.05	9.16	9.85	41.28	16.15	17.46	0.71	0.06	0.06	235.18	6.71	7.32	-988.83	436.20	470.27
CWGEE	0.04	9.23	9.61	40.95	16.03	17.11	0.71	0.06	0.06	234.05	6.68	7.09	-960.36	432.80	461.61
ARGEE	0.01	6.24	6.55	41.94	15.90	16.88	0.72	0.06	0.06	244.78	4.70	4.97	-1040.56	428.56	453.92

Table D.31: Extension 3 Scenario 31 (γ_1=-0.36, γ_2=-0.22, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.03	7.40	7.55	42.08	17.36	17.84	0.72	0.06	0.06	251.67	5.65	5.55	-1059.76	466.71	479.07
EGEE	-0.04	7.90	8.80	41.99	16.50	17.41	0.72	0.06	0.06	243.88	5.54	6.29	-1043.38	444.59	467.49
IEE	-0.05	8.72	9.65	41.96	16.33	17.79	0.72	0.06	0.06	238.96	6.54	7.55	-1033.44	440.45	477.92
CWGEE	-0.07	8.98	9.46	41.94	16.21	17.37	0.72	0.06	0.06	240.00	6.74	7.21	-1030.21	437.01	466.19
ARGEE	-0.02	5.97	6.28	42.00	15.77	16.89	0.72	0.06	0.06	244.65	4.50	4.71	-1043.37	425.27	452.69

Table D.32: Extension 3 Scenario 32 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.03	7.32	7.39	41.79	17.41	17.80	0.72	0.06	0.06	249.69	5.59	5.61	-1050.03	467.74	477.62
EGEE	-0.11	8.10	8.77	41.60	16.58	17.40	0.72	0.06	0.06	245.50	5.71	6.34	-1032.49	446.94	467.05
IEE	0.06	9.18	9.88	41.25	16.23	17.51	0.71	0.06	0.06	235.95	6.76	7.47	-995.69	437.83	470.06
CWGEE	0.09	9.30	9.67	41.14	16.12	17.16	0.71	0.06	0.06	236.01	6.81	7.19	-982.01	435.02	461.40
ARGEE	-0.05	6.21	6.54	41.63	15.88	16.85	0.72	0.06	0.06	244.80	4.68	4.96	-1031.50	427.99	452.22

Table D.33: Extension 3 Scenario 33 (γ_1=-0.24, γ_2=-0.35, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.01	7.28	7.52	42.23	17.37	17.78	0.72	0.06	0.06	248.87	5.56	5.63	-1060.90	466.68	478.18
EGEE	-0.03	8.12	8.63	41.98	16.57	17.44	0.72	0.06	0.06	246.11	5.73	6.11	-1044.81	446.88	468.71
IEE	0.00	9.20	10.04	41.52	16.17	17.44	0.71	0.06	0.06	235.49	6.74	7.40	-998.96	436.49	469.40
CWGEE	0.03	9.30	9.82	41.35	16.05	17.13	0.71	0.06	0.06	235.10	6.76	7.20	-980.05	433.16	461.48
ARGEE	-0.02	6.21	6.55	41.95	15.87	16.84	0.72	0.06	0.06	244.82	4.69	4.90	-1040.33	427.87	452.58

Table D.34: Extension 3 Scenario 34 (γ_1=-0.29, γ_2=-0.29, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.10	7.35	7.55	41.94	17.38	18.06	0.72	0.06	0.06	250.68	5.61	5.64	-1056.17	466.97	484.08
EGEE	-0.03	8.03	8.62	41.71	16.54	17.59	0.72	0.06	0.06	244.91	5.65	6.24	-1035.67	445.76	471.54
IEE	-0.15	9.09	9.83	41.52	16.27	17.72	0.72	0.06	0.06	237.07	6.74	7.51	-1010.66	438.80	475.44
CWGEE	-0.16	9.25	9.67	41.45	16.15	17.41	0.72	0.06	0.06	237.58	6.83	7.25	-1001.26	435.55	467.77
ARGEE	-0.07	6.13	6.47	41.73	15.85	17.06	0.72	0.06	0.06	244.81	4.62	4.89	-1035.94	427.22	457.49

Table D.35: Extension 3 Scenario 35 (γ_1=-0.17, γ_2=-0.40, treatment effect = 0 g/week)

D.2 GEE Results for n=600

D.2.1 Scenarios with Non Informative Cluster Size

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.95	2.14	2.08	41.61	5.32	5.36	0.72	0.02	0.02	244.50	1.65	1.66	-1033.02	143.16	143.73
EGEE	27.94	2.46	2.40	41.61	5.30	5.36	0.72	0.02	0.02	244.51	1.76	1.76	-1033.13	142.63	143.54
IEE	27.98	3.19	3.13	41.61	5.29	5.36	0.72	0.02	0.02	244.48	2.03	2.03	-1033.01	142.50	143.57
CWGEE	27.98	3.19	3.13	41.61	5.29	5.36	0.72	0.02	0.02	244.48	2.03	2.03	-1033.01	142.50	143.57
ARGEE	27.97	1.97	1.93	41.61	5.12	5.21	0.72	0.02	0.02	244.48	1.51	1.52	-1032.77	137.82	139.44

Table D.36: Extension 3 (Larger Sample Size) Scenario 1 (fixed trial length=38, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	27.87	2.16	2.14	42.09	5.33	5.41	0.72	0.02	0.02	244.60	1.66	1.70	-1046.82	143.46	146.04
EGEE	27.94	2.50	2.50	42.08	5.29	5.37	0.72	0.02	0.02	244.56	1.78	1.82	-1046.76	142.51	145.12
IEE	27.86	3.14	3.17	42.09	5.31	5.32	0.72	0.02	0.02	244.59	2.12	2.15	-1047.04	142.96	143.90
CWGEE	27.85	3.12	3.15	42.08	5.29	5.37	0.72	0.02	0.02	244.60	2.10	2.13	-1046.71	142.38	145.13
ARGEE	27.92	1.95	1.91	42.09	5.11	5.20	0.72	0.02	0.02	244.55	1.49	1.51	-1046.76	137.58	140.48

Table D.37: Extension 3 (Larger Sample Size) Scenario 2 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.97	2.14	2.20	41.84	5.32	5.19	0.72	0.02	0.02	244.49	1.65	1.65	-1038.63	143.08	140.98
EGEE	20.99	2.46	2.48	41.84	5.30	5.18	0.72	0.02	0.02	244.48	1.75	1.74	-1038.66	142.77	140.89
IEE	20.94	3.19	3.23	41.85	5.30	5.19	0.72	0.02	0.02	244.51	2.03	2.05	-1038.80	142.64	141.09
CWGEE	20.94	3.19	3.23	41.85	5.30	5.19	0.72	0.02	0.02	244.51	2.03	2.05	-1038.80	142.64	141.09
ARGEE	21.02	1.97	2.05	41.89	5.12	5.01	0.72	0.02	0.02	244.46	1.50	1.52	-1040.06	137.95	136.24

Table D.38: Extension 3 (Larger Sample Size) Scenario 8 (fixed trial length=38, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.03	2.15	2.15	42.40	5.34	5.28	0.72	0.02	0.02	244.46	1.66	1.67	-1054.29	143.84	141.77
EGEE	21.08	2.50	2.45	42.38	5.31	5.27	0.72	0.02	0.02	244.44	1.78	1.78	-1053.73	142.98	141.38
IEE	20.97	3.15	3.11	42.38	5.33	5.29	0.72	0.02	0.02	244.52	2.13	2.06	-1053.84	143.49	141.92
CWGEE	20.97	3.13	3.08	42.38	5.30	5.27	0.72	0.02	0.02	244.52	2.11	2.03	-1053.80	142.84	141.34
ARGEE	21.03	1.95	1.99	42.30	5.13	5.08	0.72	0.02	0.02	244.47	1.49	1.54	-1051.74	138.04	136.39

Table D.39: Extension 3 (Larger Sample Size) Scenario 9 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.03	2.14	2.20	41.88	5.32	5.15	0.72	0.02	0.02	244.44	1.65	1.74	-1039.29	143.26	139.20
EGEE	14.08	2.46	2.59	41.88	5.30	5.14	0.72	0.02	0.02	244.42	1.75	1.88	-1039.39	142.78	138.89
IEE	13.95	3.19	3.07	41.87	5.30	5.14	0.72	0.02	0.02	244.48	2.03	2.01	-1039.22	142.62	139.09
CWGEE	13.95	3.19	3.07	41.87	5.30	5.14	0.72	0.02	0.02	244.48	2.03	2.01	-1039.22	142.62	139.09
ARGEE	14.05	1.98	2.02	41.88	5.13	4.93	0.72	0.02	0.02	244.44	1.50	1.59	-1039.56	138.03	133.51

Table D.40: Extension 3 (Larger Sample Size) Scenario 15 (fixed trial length=38, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.05	2.15	2.10	41.84	5.33	5.38	0.72	0.02	0.02	244.44	1.66	1.65	-1039.63	143.28	146.13
EGEE	14.07	2.50	2.44	41.84	5.29	5.36	0.72	0.02	0.02	244.41	1.78	1.76	-1039.40	142.40	145.46
IEE	13.92	3.14	3.10	41.85	5.31	5.40	0.72	0.02	0.02	244.44	2.12	2.12	-1039.63	142.92	146.33
CWGEE	13.93	3.12	3.09	41.84	5.29	5.36	0.72	0.02	0.02	244.44	2.10	2.10	-1039.43	142.26	145.30
ARGEE	14.02	1.95	1.91	41.86	5.11	5.16	0.72	0.02	0.02	244.43	1.49	1.47	-1039.80	137.45	140.16

Table D.41: Extension 3 (Larger Sample Size) Scenario 16 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.99	2.14	2.13	42.12	5.31	5.73	0.72	0.02	0.02	244.48	1.65	1.62	-1046.46	142.93	155.39
EGEE	7.00	2.46	2.52	42.12	5.29	5.71	0.72	0.02	0.02	244.48	1.76	1.75	-1046.42	142.52	154.75
IEE	6.99	3.19	3.08	42.11	5.29	5.72	0.72	0.02	0.02	244.48	2.03	1.97	-1046.19	142.39	154.91
CWGEE	6.99	3.19	3.08	42.11	5.29	5.72	0.72	0.02	0.02	244.48	2.03	1.97	-1046.19	142.39	154.91
ARGEE	6.99	1.97	2.01	42.16	5.12	5.48	0.72	0.02	0.02	244.48	1.51	1.51	-1047.71	137.76	148.53

Table D.42: Extension 3 (Larger Sample Size) Scenario 22 (fixed trial length=38, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.00	2.15	2.20	42.03	5.34	5.36	0.72	0.02	0.02	244.43	1.65	1.62	-1043.17	143.63	143.52
EGEE	6.95	2.49	2.51	42.04	5.31	5.36	0.72	0.02	0.02	244.48	1.78	1.74	-1043.39	142.84	143.56
IEE	6.91	3.14	3.29	42.05	5.32	5.38	0.72	0.02	0.02	244.53	2.12	2.15	-1043.74	143.34	144.31
CWGEE	6.94	3.13	3.27	42.03	5.30	5.36	0.72	0.02	0.02	244.51	2.11	2.13	-1043.38	142.69	143.70
ARGEE	6.95	1.95	2.02	42.00	5.13	5.15	0.72	0.02	0.02	244.47	1.49	1.48	-1042.97	137.99	138.07

Table D.43: Extension 3 (Larger Sample Size) Scenario 23 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 7 g/week)

Interaction =0

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.09	2.15	2.08	41.63	5.32	5.10	0.72	0.02	0.02	244.46	1.65	1.64	-1033.77	143.20	137.45
EGEE	0.09	2.46	2.43	41.63	5.31	5.11	0.72	0.02	0.02	244.46	1.75	1.76	-1033.81	142.74	137.45
IEE	0.11	3.19	3.18	41.63	5.30	5.11	0.72	0.02	0.02	244.45	2.03	2.03	-1033.91	142.60	137.53
CWGEE	0.11	3.19	3.18	41.63	5.30	5.11	0.72	0.02	0.02	244.45	2.03	2.03	-1033.91	142.60	137.53
ARGEE	0.11	1.98	1.95	41.66	5.13	4.94	0.72	0.02	0.02	244.46	1.51	1.48	-1034.58	137.95	133.19

Table D.44: Extension 3 (Larger Sample Size) Scenario 29 (fixed trial length=38, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.00	2.16	2.21	42.01	5.34	5.42	0.72	0.02	0.02	244.48	1.66	1.72	-1042.33	143.71	145.58
EGEE	-0.03	2.50	2.67	42.02	5.30	5.36	0.72	0.02	0.02	244.45	1.79	1.90	-1042.52	142.70	144.11
IEE	-0.05	3.14	3.11	42.02	5.32	5.39	0.72	0.02	0.02	244.46	2.12	2.13	-1042.53	143.19	144.81
CWGEE	-0.04	3.13	3.11	42.02	5.29	5.37	0.72	0.02	0.02	244.47	2.11	2.13	-1042.59	142.56	144.02
ARGEE	-0.02	1.95	2.00	42.02	5.12	5.24	0.72	0.02	0.02	244.48	1.49	1.56	-1042.67	137.91	140.82

Table D.45: Extension 3 (Larger Sample Size) Scenario 30 (γ_0=log(24.3667),γ_1=0,γ_2=0, treatment effect = 0 g/week)

D.2.2 Scenarios with Informative Cluster Size using $\gamma_0 = 3.069142$

Interaction =28

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.28	2.30	2.39	42.34	5.41	5.34	0.72	0.02	0.02	247.37	1.75	1.78	-1061.01	145.66	144.16
EGEE	27.86	2.76	2.77	42.20	5.32	5.27	0.72	0.02	0.02	247.20	1.92	1.93	-1054.28	143.30	142.14
IEE	27.25	3.09	3.15	41.45	5.29	5.29	0.71	0.02	0.02	234.78	2.18	2.21	-990.46	142.48	143.47
CWGEE	27.21	3.02	3.08	41.21	5.18	5.14	0.71	0.02	0.02	233.85	2.11	2.16	-964.12	139.60	139.17
ARGEE	28.03	1.98	2.06	42.15	5.13	5.15	0.72	0.02	0.02	244.74	1.46	1.56	-1047.13	138.03	138.61

Table D.46: Extension 3 (Larger Sample Size) Scenario 3 (γ_1=-0.36, γ_2=-0.22, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.63	2.36	2.42	42.12	5.43	5.23	0.72	0.02	0.02	251.33	1.79	1.70	-1059.33	146.17	140.33
EGEE	27.00	2.82	2.95	42.12	5.29	5.20	0.72	0.02	0.02	244.38	1.91	1.99	-1045.74	142.50	139.50
IEE	26.72	3.06	3.01	42.14	5.42	5.32	0.72	0.02	0.02	239.11	2.16	2.26	-1035.55	146.04	142.86
CWGEE	26.97	2.97	2.92	42.09	5.27	5.18	0.72	0.02	0.02	240.12	2.06	2.08	-1031.95	141.77	138.93
ARGEE	28.04	1.93	1.92	42.19	5.12	5.07	0.72	0.02	0.02	244.63	1.37	1.40	-1047.84	137.71	135.68

Table D.47: Extension 3 (Larger Sample Size) Scenario 4 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.38	2.33	2.34	42.11	5.42	5.39	0.72	0.02	0.02	249.39	1.77	1.73	-1059.20	145.94	145.26
EGEE	27.34	2.77	2.81	41.85	5.30	5.29	0.72	0.02	0.02	245.87	1.91	1.91	-1040.42	142.68	142.61
IEE	26.79	3.12	3.17	41.43	5.32	5.37	0.71	0.02	0.02	235.91	2.17	2.25	-1000.23	143.38	144.75
CWGEE	26.81	3.05	3.10	41.39	5.21	5.22	0.71	0.02	0.02	236.13	2.11	2.17	-987.87	140.37	141.11
ARGEE	27.99	1.98	1.95	41.95	5.13	5.12	0.72	0.02	0.02	244.68	1.44	1.47	-1042.33	138.00	137.53

Table D.48: Extension 3 (Larger Sample Size) Scenario 5 (γ_1=-024, γ_2=-0.35, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.37	2.32	2.24	42.01	5.43	5.41	0.72	0.02	0.02	248.58	1.76	1.73	-1055.85	145.91	146.91
EGEE	27.55	2.77	2.70	41.80	5.32	5.33	0.72	0.02	0.02	246.45	1.92	1.92	-1041.24	143.18	144.82
IEE	27.03	3.12	3.12	41.23	5.32	5.35	0.71	0.02	0.02	235.10	2.17	2.13	-990.10	143.19	145.40
CWGEE	27.05	3.04	3.04	41.12	5.20	5.21	0.71	0.02	0.02	234.89	2.11	2.07	-972.48	140.20	142.02
ARGEE	28.07	1.98	1.92	41.83	5.14	5.17	0.72	0.02	0.02	244.73	1.45	1.50	-1039.43	138.23	139.98

Table D.49: Extension 3 (Larger Sample Size) Scenario 6 (γ_1=-0.29, γ_2=-0.29, treatment effect = 28 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	28.34	2.35	2.32	42.16	5.43	5.35	0.72	0.02	0.02	250.32	1.78	1.67	-1062.12	146.10	143.75
EGEE	26.89	2.78	2.71	41.96	5.31	5.26	0.72	0.02	0.02	245.40	1.90	1.87	-1042.96	143.02	141.35
IEE	26.78	3.11	3.09	41.70	5.36	5.30	0.72	0.02	0.02	236.89	2.17	2.23	-1014.62	144.35	142.43
CWGEE	26.85	3.04	2.98	41.70	5.25	5.21	0.72	0.02	0.02	237.52	2.10	2.09	-1006.69	141.45	139.79
ARGEE	27.96	1.96	1.99	42.03	5.14	5.06	0.72	0.02	0.02	244.71	1.42	1.41	-1044.76	138.35	135.79

Table D.50: Extension 3 (Larger Sample Size) Scenario 7 (γ_1=-0.17, γ_2=-0.40, treatment effect = 28 g/week)

Interaction =21

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	20.99	2.31	2.26	42.21	5.43	5.44	0.72	0.02	0.02	247.62	1.76	1.68	-1058.50	146.10	147.83
EGEE	20.57	2.76	2.83	42.05	5.33	5.37	0.72	0.02	0.02	247.42	1.93	1.92	-1051.06	143.59	146.08
IEE	20.24	3.09	3.07	41.27	5.29	5.31	0.71	0.02	0.02	234.83	2.20	2.22	-986.24	142.74	145.00
CWGEE	20.24	3.02	2.97	41.06	5.18	5.24	0.71	0.02	0.02	233.89	2.13	2.16	-960.82	139.88	143.37
ARGEE	20.95	1.98	1.98	42.04	5.14	5.19	0.72	0.02	0.02	244.90	1.47	1.45	-1044.89	138.46	141.42

Table D.51: Extension 3 (Larger Sample Size) Scenario 10 (γ_1=-0.36, γ_2=-0.22, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.44	2.36	2.41	41.96	5.43	5.30	0.72	0.02	0.02	251.41	1.79	1.77	-1054.02	146.24	143.12
EGEE	20.17	2.84	2.90	41.97	5.29	5.15	0.72	0.02	0.02	244.24	1.94	2.00	-1040.55	142.59	138.98
IEE	20.05	3.02	2.97	41.86	5.42	5.21	0.72	0.02	0.02	238.92	2.20	2.28	-1027.03	145.90	140.57
CWGEE	20.23	2.95	2.89	41.94	5.27	5.13	0.72	0.02	0.02	239.93	2.10	2.10	-1026.82	141.91	138.60
ARGEE	20.97	1.92	1.93	42.05	5.12	4.95	0.72	0.02	0.02	244.57	1.39	1.43	-1043.15	137.80	133.37

Table D.52: Extension 3 (Larger Sample Size) Scenario 11 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.32	2.34	2.27	41.95	5.43	5.63	0.72	0.02	0.02	249.53	1.77	1.69	-1054.62	146.13	150.04
EGEE	20.39	2.77	2.75	41.71	5.30	5.52	0.72	0.02	0.02	245.85	1.93	1.90	-1036.05	142.87	147.10
IEE	20.13	3.11	2.99	41.27	5.32	5.54	0.71	0.02	0.02	235.71	2.21	2.11	-994.88	143.37	148.33
CWGEE	20.13	3.04	2.93	41.25	5.21	5.44	0.71	0.02	0.02	235.96	2.14	2.06	-983.58	140.54	145.87
ARGEE	20.97	1.98	1.93	41.83	5.13	5.30	0.72	0.02	0.02	244.72	1.45	1.47	-1038.26	138.13	141.58

Table D.53: Extension 3 (Larger Sample Size) Scenario 12 (γ_1=-024, γ_2=-0.35, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.24	2.32	2.33	41.99	5.41	5.55	0.72	0.02	0.02	248.76	1.77	1.69	-1055.39	145.73	150.36
EGEE	20.60	2.78	2.73	41.77	5.29	5.43	0.72	0.02	0.02	246.44	1.94	1.93	-1040.17	142.69	147.15
IEE	20.23	3.09	2.95	41.18	5.28	5.41	0.71	0.02	0.02	235.12	2.20	2.13	-988.72	142.54	146.62
CWGEE	20.24	3.03	2.91	41.09	5.18	5.33	0.71	0.02	0.02	234.91	2.14	2.08	-971.78	139.79	144.30
ARGEE	21.01	1.98	1.96	41.77	5.12	5.23	0.72	0.02	0.02	244.77	1.47	1.47	-1037.97	137.83	141.47

Table D.54: Extension 3 (Larger Sample Size) Scenario 13 (γ_1=-0.29, γ_2=-0.29, treatment effect = 21 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	21.28	2.34	2.40	42.28	5.42	5.60	0.72	0.02	0.02	250.40	1.78	1.72	-1064.93	145.84	150.90
EGEE	20.19	2.78	2.84	42.12	5.30	5.49	0.72	0.02	0.02	245.30	1.92	1.99	-1046.54	142.62	148.09
IEE	19.95	3.09	3.26	41.90	5.35	5.55	0.72	0.02	0.02	236.87	2.22	2.28	-1018.70	144.12	150.20
CWGEE	20.06	3.02	3.18	41.87	5.23	5.44	0.72	0.02	0.02	237.47	2.13	2.20	-1010.80	140.97	147.27
ARGEE	20.92	1.96	2.00	42.23	5.13	5.30	0.72	0.02	0.02	244.69	1.43	1.44	-1049.54	138.08	142.86

Table D.55: Extension 3 (Larger Sample Size) Scenario 14 (γ_1=-0.17, γ_2=-0.40, treatment effect = 21 g/week)

Interaction =14

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.18	2.30	2.25	42.22	5.41	5.46	0.72	0.02	0.02	247.53	1.75	1.70	-1057.95	145.49	148.63
EGEE	13.92	2.78	2.81	42.09	5.31	5.39	0.72	0.02	0.02	247.20	1.95	2.02	-1050.78	143.03	146.89
IEE	13.74	3.08	3.16	41.40	5.28	5.27	0.71	0.02	0.02	234.53	2.23	2.27	-988.29	142.49	144.21
CWGEE	13.70	3.01	3.07	41.10	5.16	5.27	0.71	0.02	0.02	233.61	2.15	2.18	-960.63	139.26	144.38
ARGEE	14.08	1.98	2.04	42.10	5.12	5.32	0.72	0.02	0.02	244.77	1.48	1.52	-1045.43	137.93	144.66

Table D.56: Extension 3 (Larger Sample Size) Scenario 17 (γ_1=-0.36, γ_2=-0.22, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.34	2.36	2.35	42.12	5.44	5.51	0.72	0.02	0.02	251.59	1.79	1.78	-1059.31	146.28	148.69
EGEE	13.51	2.83	2.77	42.03	5.30	5.36	0.72	0.02	0.02	244.18	1.96	1.99	-1043.43	142.73	144.59
IEE	13.47	3.01	3.01	41.90	5.42	5.42	0.72	0.02	0.02	238.92	2.26	2.25	-1029.85	145.98	145.85
CWGEE	13.59	2.94	2.92	42.00	5.28	5.34	0.72	0.02	0.02	239.93	2.14	2.08	-1029.70	142.05	144.33
ARGEE	14.05	1.91	1.86	42.00	5.13	5.18	0.72	0.02	0.02	244.59	1.41	1.44	-1042.78	138.04	139.80

Table D.57: Extension 3 (Larger Sample Size) Scenario 18 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.19	2.33	2.24	42.30	5.44	5.39	0.72	0.02	0.02	249.65	1.77	1.69	-1063.99	146.34	144.00
EGEE	13.58	2.76	2.84	42.07	5.32	5.28	0.72	0.02	0.02	245.72	1.92	1.95	-1045.41	143.32	140.87
IEE	13.51	3.10	3.05	41.72	5.34	5.28	0.71	0.02	0.02	235.41	2.25	2.26	-1006.80	143.88	141.54
CWGEE	13.54	3.04	2.95	41.59	5.23	5.19	0.71	0.02	0.02	235.65	2.17	2.17	-992.05	141.00	139.22
ARGEE	14.04	1.97	2.01	42.18	5.15	5.14	0.72	0.02	0.02	244.68	1.47	1.49	-1047.59	138.56	137.18

Table D.58: Extension 3 (Larger Sample Size) Scenario 19 (γ_1=-024, γ_2=-0.35, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.11	2.32	2.36	41.73	5.42	5.45	0.73	0.02	0.02	248.87	1.76	1.75	-1048.97	145.89	145.51
EGEE	13.70	2.78	2.76	41.46	5.32	5.40	0.72	0.02	0.02	246.35	1.95	1.98	-1032.11	143.34	144.35
IEE	13.43	3.09	3.11	40.90	5.31	5.36	0.71	0.02	0.02	234.98	2.24	2.24	-981.05	143.05	143.10
CWGEE	13.43	3.03	3.05	40.78	5.21	5.28	0.71	0.02	0.02	234.80	2.16	2.19	-963.81	140.40	141.48
ARGEE	13.97	1.98	2.00	41.55	5.14	5.11	0.72	0.02	0.02	244.76	1.48	1.51	-1032.11	138.50	136.56

Table D.59: Extension 3 (Larger Sample Size) Scenario 20 (γ_1=-0.29, γ_2=-0.29, treatment effect = 14 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	14.13	2.35	2.36	42.07	5.43	5.54	0.72	0.02	0.02	250.50	1.78	1.73	-1058.20	146.09	148.08
EGEE	13.41	2.78	2.90	41.89	5.29	5.37	0.72	0.02	0.02	245.08	1.94	2.05	-1039.21	142.56	143.97
IEE	13.33	3.08	3.07	41.69	5.35	5.43	0.72	0.02	0.02	236.56	2.26	2.37	-1011.80	144.12	146.12
CWGEE	13.38	3.02	3.02	41.61	5.23	5.31	0.72	0.02	0.02	237.22	2.17	2.27	-1002.88	140.93	142.37
ARGEE	13.94	1.96	2.02	42.03	5.12	5.17	0.72	0.02	0.02	244.66	1.45	1.51	-1043.08	137.97	138.06

Table D.60: Extension 3 (Larger Sample Size) Scenario 21 (γ_1=-0.17, γ_2=-0.40, treatment effect = 14 g/week)

Interaction =7

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	6.86	2.31	2.30	42.22	5.42	5.59	0.72	0.02	0.02	247.64	1.75	1.69	-1058.25	145.91	151.93
EGEE	6.77	2.77	2.85	42.07	5.32	5.53	0.72	0.02	0.02	247.21	1.95	1.93	-1050.79	143.33	150.25
IEE	6.60	3.08	3.16	41.29	5.29	5.47	0.71	0.02	0.02	234.53	2.25	2.29	-986.17	142.65	149.56
CWGEE	6.60	3.01	3.09	41.07	5.17	5.38	0.71	0.02	0.02	233.58	2.17	2.17	-960.26	139.60	146.89
ARGEE	6.87	1.97	2.00	42.02	5.13	5.36	0.72	0.02	0.02	244.83	1.49	1.49	-1043.61	138.21	145.95

Table D.61: Extension 3 (Larger Sample Size) Scenario 24 (γ_1=-0.36, γ_2=-0.22, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.08	2.37	2.29	42.19	5.44	5.14	0.72	0.02	0.02	251.78	1.80	1.75	-1061.07	146.25	138.13
EGEE	6.64	2.84	2.90	42.15	5.30	5.05	0.72	0.02	0.02	244.18	1.99	2.09	-1046.12	142.59	135.78
IEE	6.55	2.98	2.99	42.15	5.42	5.24	0.72	0.02	0.02	238.94	2.29	2.31	-1035.95	145.87	141.16
CWGEE	6.58	2.92	2.93	42.13	5.27	5.03	0.72	0.02	0.02	239.95	2.18	2.16	-1032.68	141.84	135.31
ARGEE	6.91	1.90	1.85	42.23	5.12	4.90	0.72	0.02	0.02	244.66	1.42	1.45	-1048.34	137.90	131.59

Table D.62: Extension 3 (Larger Sample Size) Scenario 25 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 7 g/week)

		β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
3	MM	7.11	2.33	2.28	42.31	5.43	5.48	0.72	0.02	0.02	249.75	1.77	1.74	-1064.52	146.08	147.20
	EGEE	6.87	2.78	2.70	42.09	5.31	5.40	0.72	0.02	0.02	245.68	1.96	1.97	-1046.22	143.06	144.94
	IEE	6.73	3.09	3.19	41.65	5.33	5.43	0.71	0.02	0.02	235.43	2.30	2.43	-1005.11	143.59	145.70
	CWGEE	6.73	3.03	3.14	41.61	5.22	5.31	0.71	0.02	0.02	235.66	2.21	2.34	-992.90	140.73	142.92
	ARGEE	7.04	1.97	1.90	42.20	5.14	5.17	0.72	0.02	0.02	244.73	1.48	1.51	-1047.94	138.25	138.67

Table D.63: Extension 3 (Larger Sample Size) Scenario 26 (γ_1=-024, γ_2=-0.35, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.02	2.32	2.35	41.88	5.43	5.35	0.72	0.02	0.02	248.91	1.77	1.70	-1053.40	146.01	143.87
EGEE	6.85	2.79	2.85	41.67	5.31	5.28	0.72	0.02	0.02	246.31	1.97	1.95	-1038.40	143.08	142.01
IEE	6.64	3.09	3.13	41.09	5.31	5.30	0.71	0.02	0.02	234.80	2.28	2.32	-987.23	143.05	143.04
CWGEE	6.65	3.03	3.05	41.00	5.20	5.18	0.71	0.02	0.02	234.61	2.20	2.24	-969.95	140.12	139.84
ARGEE	6.97	1.98	2.02	41.73	5.13	5.12	0.72	0.02	0.02	244.79	1.49	1.48	-1037.36	138.20	138.05

Table D.64: Extension 3 (Larger Sample Size) Scenario 27 (γ_1=-0.29, γ_2=-0.29, treatment effect = 7 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	7.12	2.35	2.28	41.99	5.43	5.52	0.72	0.02	0.02	250.60	1.79	1.71	-1055.59	146.18	149.94
EGEE	6.91	2.79	2.77	41.84	5.30	5.44	0.72	0.02	0.02	244.93	1.97	2.00	-1037.02	142.86	147.37
IEE	6.62	3.07	3.13	41.54	5.36	5.61	0.72	0.02	0.02	236.51	2.31	2.40	-1007.59	144.28	152.57
CWGEE	6.65	3.01	3.01	41.58	5.24	5.39	0.71	0.02	0.02	237.13	2.21	2.21	-1001.08	141.24	146.34
ARGEE	7.01	1.96	1.96	41.96	5.14	5.30	0.72	0.02	0.02	244.65	1.47	1.46	-1040.52	138.30	143.42

Table D.65: Extension 3 (Larger Sample Size) Scenario 28 (γ_1=-0.17, γ_2=-0.40, treatment effect = 7 g/week)

Interaction =0

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.07	2.31	2.35	42.00	5.42	5.38	0.72	0.02	0.02	247.63	1.76	1.80	-1052.98	145.79	143.57
EGEE	-0.18	2.76	2.90	41.82	5.32	5.34	0.72	0.02	0.02	247.18	1.96	2.02	-1044.43	143.47	142.64
IEE	-0.10	3.08	3.05	41.13	5.28	5.34	0.71	0.02	0.02	234.41	2.29	2.32	-982.59	142.39	142.41
CWGEE	-0.06	3.01	3.00	40.83	5.18	5.20	0.71	0.02	0.02	233.43	2.20	2.24	-954.23	139.78	139.33
ARGEE	-0.12	1.98	2.01	41.85	5.13	5.18	0.72	0.02	0.02	244.83	1.50	1.55	-1039.87	138.32	138.15

Table D.66: Extension 3 (Larger Sample Size) Scenario 31 (γ_1=-0.36, γ_2=-0.22, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.04	2.36	2.33	42.19	5.44	5.40	0.72	0.02	0.02	251.75	1.80	1.72	-1061.31	146.34	145.64
EGEE	-0.03	2.84	2.72	42.15	5.31	5.31	0.72	0.02	0.02	243.93	2.01	2.04	-1046.48	143.05	143.07
IEE	-0.06	3.00	3.08	42.13	5.44	5.48	0.72	0.02	0.02	238.53	2.37	2.43	-1035.59	146.44	147.98
CWGEE	-0.06	2.93	2.96	42.11	5.28	5.28	0.72	0.02	0.02	239.62	2.24	2.24	-1032.44	142.33	142.61
ARGEE	-0.04	1.91	1.84	42.16	5.14	5.12	0.72	0.02	0.02	244.55	1.45	1.46	-1046.98	138.38	138.37

Table D.67: Extension 3 (Larger Sample Size) Scenario 32 (γ_1=-0.039 , γ_2= -0.48, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	-0.19	2.34	2.32	42.28	5.43	5.50	0.72	0.02	0.02	249.89	1.78	1.66	-1063.51	146.05	147.77
EGEE	-0.00	2.77	2.82	42.01	5.31	5.41	0.72	0.02	0.02	245.55	1.97	1.98	-1043.76	142.98	145.36
IEE	-0.07	3.09	3.20	41.57	5.33	5.44	0.71	0.02	0.02	235.30	2.35	2.41	-1002.85	143.56	146.05
CWGEE	-0.09	3.03	3.11	41.53	5.22	5.33	0.71	0.02	0.02	235.58	2.25	2.23	-990.60	140.69	143.88
ARGEE	-0.11	1.97	1.95	42.11	5.14	5.17	0.72	0.02	0.02	244.77	1.50	1.46	-1045.31	138.26	138.89

Table D.68: Extension 3 (Larger Sample Size) Scenario 33 (γ_1=-024, γ_2=-0.35, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.05	2.32	2.32	42.21	5.43	5.43	0.72	0.02	0.02	248.91	1.77	1.70	-1060.63	146.10	145.33
EGEE	0.05	2.77	2.76	41.95	5.32	5.32	0.72	0.02	0.02	246.21	1.96	1.97	-1043.96	143.29	142.38
IEE	0.07	3.08	3.00	41.36	5.31	5.36	0.71	0.02	0.02	234.68	2.30	2.26	-992.81	142.90	143.89
CWGEE	0.08	3.02	2.96	41.26	5.20	5.21	0.71	0.02	0.02	234.44	2.21	2.18	-975.12	140.31	140.04
ARGEE	0.08	1.97	1.94	42.05	5.14	5.14	0.72	0.02	0.02	244.68	1.50	1.50	-1044.09	138.28	138.01

Table D.69: Extension 3 (Larger Sample Size) Scenario 34 (γ_1=-0.29, γ_2=-0.29, treatment effect = 0 g/week)

	β_1	SE(β_1)	SD(β_1)	β_2	SE(β_2)	SD(β_2)	β_3	SE(β_3)	SD(β_3)	β_4	SE(β_4)	SD(β_4)	β_0	SE(β_0)	SD(β_0)
MM	0.04	2.35	2.26	42.31	5.43	5.47	0.72	0.02	0.02	250.65	1.79	1.78	-1064.65	146.21	147.94
EGEE	0.01	2.78	2.72	42.10	5.30	5.36	0.72	0.02	0.02	244.94	1.98	2.01	-1045.03	142.85	144.90
IEE	0.12	3.07	3.23	41.90	5.36	5.41	0.71	0.02	0.02	236.20	2.36	2.47	-1017.36	144.51	146.32
CWGEE	0.12	3.01	3.16	41.83	5.24	5.30	0.71	0.02	0.02	236.85	2.25	2.34	-1008.15	141.18	143.53
ARGEE	0.04	1.96	1.88	42.22	5.13	5.15	0.72	0.02	0.02	244.67	1.49	1.54	-1048.57	138.26	139.17

Table D.70: Extension 3 (Larger Sample Size) Scenario 35 (γ_1=-0.17, γ_2=-0.40, treatment effect = 0 g/week)

Bibliography

Benhin, E., Rao, J., and Scott, A. (2005). Mean estimating equation approach to analysing cluster-correlated data with nonignorable cluster sizes. *Biometrika*, 92(2):435–540.

Chen, Z., Zhang, B., and Albert, P. S. (2011). A joint modeling approach to data with informative cluster size: Robustness to the cluster size model. *Statistics in Medicine*, 30:1825–1836.

Cummings, P. (2013). Missing data and multiple imputations. *JAMA Pediatrics*, 167(7):656–661.

Dunlop, D. D. (1994). Regression for longitudinal data: A bridge from least squares regression. *The American Statistician*, 48(4):299–303.

Dunson, D. B., Chen, Z., and Harry, J. (2003). A bayesian approach for joint modeling of cluster size and subunit-specific outcomes. *Biometrics*, 59:521–530.

Gueorguieva, R. V. (2005). Comments about joint modeling of cluster size and binary and continuous subunit-specific outcomes. *Biometrics*, 61:862–867.

Harville, D. (1974). Bayesian inference for variance components using only error contrasts. *Biometrika*, 61(2):383:385.

Harville, D. (1976). Extension of the gauss-markov theorem to include the estimation of random effects. *The Annals of Statistics*, 4(2):384:395.

Hoffman, E. B., Sen, P. K., and Weinberg, C. R. (2001). WIthin Cluster Resampling. *Biometrika*, 88(4):1121–1134.

Hu, F. B., Goldberg, J., Hedeker, D., Flay, B. R., and Pentz, M. A. (1998). Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *American Journal of Epidemiology*, 147(7):694–703.

J. M. Neuhaus, J. D. Kalbfleisch, W. H. (1991). A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. *International Statistical Review*, 59(1):25–35.

Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38:963–974.

Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal Data Analysis using Generalized Linear Models. *Biometrics*, 73:13–22.

Mancl, L. A. and DeRouen, T. A. (2001). A covariance estimator for gee with improved small-sampleproperties. *Biometrics*, 57:126–134.

McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*. Chapman and Hall, second edition.

Miglioretti, D. L. and Heagerty, P. J. (2006). Marginal modeling of nonnested multilevel data using standard software. *American Journal of Epidemiology*, 165(4):453–463.

Neuhaus, J. M. and Mcculloch, C. E. (2011). Estimation of covariate effects in generalized linear mixed models with informative cluster sizes. *Biometrika*, 98(1):147–162.

Ritz, J. and Spiegelman, D. (2004). Equivalence of conditional and marginal regression models for clustered and longitudinal data. *Statistical Methods in Medical Research*, 13:309–323.

Roberts, C. L. and Lancaster, P. A. (1999). Australian national birthweight percentiles by gestational age. *Medical Journal of Australia*, 170:114–118.

Schafer, J. L. and Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2):147–177.

Seaman, S., Pavlou, M., and Copas, A. (2014a). Review of methods for handling confounding by cluster and informative cluster size in clustered data. *Statistics in Medicine*, 33:5371–5387.

Seaman, S., Pavlou, M., and Copas, A. J. (2014b). Methods for observed-cluster inference when cluster size is informative: A review and clarification. *Biometrics*, 70:449–456.

the Committee for Proprietary Medicinal Products (2004). Points to consider on adjusment for baseline covariates. *Statistics in Medicine*, 23:701–709.

Williamson, J. M., Datta, S., and Satten, G. A. (2003). Marginal analysis of clustered data when cluster size is informative. *Biometrics*, 59:36–42.

Yelland, L., Sullivan, T., Pavlou, M., and Seaman, S. (2015). Analysis of randomised trials including multiple births when birth size is informative. *Paediatric and Perinatal Epidemiology*, 29:567–575.

Zeger, S. L. and Liang, K.-Y. (1986). Longitudinal Data Analysis for Discrete and Continuous Outcomes. *Biometrics*, 42:121–130.