COMPARING REPEATED-MEASURES ANOVA AND MULTILEVEL GROWTH CURVE MODELING: A SIMULATION STUDY

A Thesis in
Human Development and Family Studies

by

Siwei Liu

© 2008 Siwei Liu

Submitted in Partial Fulfillment of the Requirements of the Degree of

Master of Science

December 2008
This thesis of Siwei Liu was reviewed and approved* by the following:

Michael J. Rovine  
Professor of Human Development & Family Studies  
Thesis Advisor

Peter C. M. Molenaar  
Professor of Human Development & Family Studies

Douglas M. Teti  
Professor of Human Development & Family Studies and Psychology  
Graduate Professor-in-Charge, Human Development & Family Studies

*Signatures are on file in the Graduate School
ABSTRACT

Currently in developmental research, there seems to be a trend for multilevel growth curve modeling to take the place of the classic repeated-measures ANOVA as the first choice for analyzing longitudinal data. The preference for growth curve modeling is so strong that it is not uncommon to see researchers fit growth curve models without considering its appropriateness. This is problematic because the repeated-measures ANOVA model and the growth curve model are sub-models in the general linear mixed model family and are complementary to each other. This paper compares the two methods in analyzing simulated data that is assumed to come from a repeated-measures study with five equally spaced occasions and show a linear increase pattern. The simulation involves varying effect size, intraclass correlation size, and sample size. It is shown that repeated-measures ANOVA models generally have better fit when the error structure is selected properly. The growth curve model fits the data better only when sample size is not small and the variance of slopes is not small, even when the error structure is simulated to show a “growth curve” pattern. The results cast doubt on the popular practice of using growth curve modeling for longitudinal data without comparing the fit of different models. A general procedure of model selection is discussed.
# TABLE OF CONTENTS

LIST OF TABLES...........................................................................................................v
LIST OF FIGURES.........................................................................................................vi
Chapter 1: Introduction...............................................................................................1
Chapter 2: Literature Review......................................................................................4
  The General Linear Model (GLM) and the General Linear Mixed Model (GLMM)……4
  The Repeated-Measures ANOVA Model.................................................................5
  The Multilevel Growth Curve Model.......................................................................9
  Repeated-measures ANOVA or growth curve modeling?........................................13
Chapter 3: Method.....................................................................................................14
  Simulation...............................................................................................................14
  Analysis..................................................................................................................16
Chapter 4: Results.....................................................................................................17
  Estimates of fixed effects parameters.................................................................17
  Standard errors of fixed effects parameters......................................................18
  Estimates of covariance parameters..................................................................19
  Fit statistics............................................................................................................20
Chapter 5: Discussion...............................................................................................23
References................................................................................................................31
Appendix A: Simulation of data with compound symmetry error structure.........49
Appendix B: Estimating parameters in SAS PROC MIXED.....................................53
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table #</th>
<th>Page #</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35</td>
<td>Values of parameters used in simulation of data with compound symmetry error structure</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>Values of parameters used in simulation of data with random-effects error structure</td>
</tr>
<tr>
<td>3</td>
<td>37</td>
<td>Parameter estimates and fit statistics of CS, AR1, and GC models for data with CS error structure and sample size of 20</td>
</tr>
<tr>
<td>4</td>
<td>38</td>
<td>Parameter estimates and fit statistics of CS, AR1, and GC models for data with CS error structure and sample size of 100</td>
</tr>
<tr>
<td>5</td>
<td>39</td>
<td>Parameter estimates and fit statistics of CS, AR1, and GC models for data with CS error structure and sample size of 200</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>Parameter estimates and fit statistics of CS, AR1, and GC models for data with AR1 error structure and sample size of 20</td>
</tr>
<tr>
<td>7</td>
<td>41</td>
<td>Parameter estimates and fit statistics of CS, AR1, and GC models for data with AR1 error structure and sample size of 100</td>
</tr>
<tr>
<td>8</td>
<td>42</td>
<td>Parameter estimates and fit statistics of CS, AR1, and GC models for data with AR1 error structure and sample size of 200</td>
</tr>
<tr>
<td>9</td>
<td>43</td>
<td>Parameter estimates and fit statistics of CS, AR1, and GC models for data with random-effects error structure, sample size of 20, and effect size of .5</td>
</tr>
<tr>
<td>10</td>
<td>44</td>
<td>Parameter estimates and fit statistics of CS, AR1, and GC models for data with random-effects error structure, sample size of 100, and effect size of .5</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure #</th>
<th>Description</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Percentage correct of AIC and BIC for data with CS error structure, sample size=20</td>
<td>45</td>
</tr>
<tr>
<td>2</td>
<td>Percentage correct of AIC and BIC for data with AR1 error structure, sample size=20</td>
<td>46</td>
</tr>
<tr>
<td>3</td>
<td>Percentage correct of AIC and BIC for data with random-effects error structure, sample size=20</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>Percentage correct of AIC and BIC for data with random-effects error structure, sample size=100</td>
<td>48</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Developmental researchers often conduct longitudinal studies to examine stability and change, in which individuals are measured on multiple occasions. Repeated-measures analysis of variance (ANOVA; Box, 1954; Myers, 1979; Scheffé, 1959) and multilevel growth curve modeling are the most widely used methods for analyzing longitudinal data. As an extension to the traditional ANOVA which assumes independent observations, repeated-measures ANOVA accounts for the correlation of within person measures by allowing the covariance matrix of errors to show a variety of patterns. On the other hand, the multilevel growth curve model is based on the conceptualization that development is a function of time. Individual is assumed to follow a “growth curve” with a particular “level” and “shape”, which are estimated as latent variables. The means of these latent variables form the average “growth curve” for the sample. Since repeated-measures ANOVA and growth curve models are perceived to have little similarity, and are often presented using different languages, many researchers consider them as two very distinct methods.

It seems that relatively few researchers are aware that the repeated-measures ANOVA model and the growth curve model actually come from the same general family, the general linear mixed model (GLMM; Laird & Ware, 1982). The mixed model distinguishes between fixed effects, which are equal across individuals, and random effects, which vary among individuals. Typically, we estimate the value of fixed effects and model the covariance matrix of random effects, the *error structure*. 
Repeated-measures ANOVA and the growth curve model are sub-models from the mixed model family. They differ in two important ways. One is the way time is treated. Repeated-measures ANOVA treats time as a categorical variable, whereas growth curve modeling treats time as a continuous variable. The other difference is the way residuals are modeled. In repeated-measures ANOVA, residuals are defined as the deviation from the group means. In growth curve modeling, residuals are divided into two parts: the deviation of individual curves from the group curves, and the deviation around the individual curves. In other words, in repeated-measures ANOVA, residuals are modeled in one covariance matrix. In growth curve modeling, this covariance matrix is divided into a “random effects” part and a “true error” part.

In a typical developmental design where all participants are measured on the same relatively few occasions, the difference between repeated-measures ANOVA and growth curve modeling remains only in the types of error structure that they can model. In most cases, researchers are not interested in the error structure per se. However, it is important to use the most appropriate error structure in order to obtain the proper inference representing the optimal test of the fixed effects. Since we typically do not know the “true” error structure of the data in reality, it is recommended that researchers try out different error structures, and select the proper one based on certain goodness-of-fit indices. Yet, at present there seems to be a common misunderstanding that growth curve models are somewhat superior to repeated-measures ANOVA models with some researchers explicitly stating that “fact” (Llabre, Spitzer, Siegel, Saab, & Schneiderman, 2004; van der Leeden, 1998).
It is often the case that researchers fit growth curve models without any consideration of the appropriateness of doing so. This oversimplified approach to analyze longitudinal data reflects a lack of clear understanding of what these models really do, and can be dangerous to scientific studies.

The work presented here aims at helping researchers better understand and choose between repeated-measures ANOVA and growth curve modeling in analyzing longitudinal data. In the literature review section, I present the two models using the general linear mixed model framework. From the comparison, readers can see the similarity and difference between these methods, and also get a general picture of what a “true” error structure looks like under the assumption of these models. Next, I simulate a series of data using different types of error structure, and compare the results of the two methods. Specifically, I examine whether simple indices of goodness-of-fit, the AIC and BIC, are able to pick out the correct model for the data. For a comprehensive comparison, I include various sample size, effect size, and intraclass correlation size in the simulation. Implications of the results are discussed.
Chapter 2

Literature Review

The General Linear Model (GLM) and the General Linear Mixed Model (GLMM)

Most researchers are familiar with the general linear model (GLM), which is usually used to represent regression models. In the simplest terms, if $y$ is an $n \times 1$ vector of scores on the dependent variable, and $X$ is an $n \times k$ matrix (called the design matrix) with one column representing a constant and $k-1$ columns representing the $k-1$ independent variables, then the general linear model of regression is:

$$y = X\beta + \epsilon$$  \hspace{1cm} (1)

where $\beta$ is a $k \times 1$ vector of regression coefficients, and $\epsilon$ is an $n \times 1$ vector of errors with a distribution of $N(0, \sigma^2_\epsilon)$. The design matrix can include dummy coded variables to represent groups, which then turn the model into an ANOVA model. In either case, $\beta$ can be described as fixed across individuals, and $\epsilon$ can be described as random.

The general linear mixed model (GLMM) was first proposed by Laird and Ware (1982) based on the work of Harville (1977) as an extension of the GLM in the presence of correlated regression residuals. A random-effects part is added to the GLM, hence, the mixed model is also called the “random-effects model”. The mathematical representation of the mixed model is:

$$y_i = X_i \beta + Z_i \gamma_i + \epsilon_i$$  \hspace{1cm} (2)

where the subscript $i$ represents individual’s ID number, $y_i$ is an $n_i \times 1$ vector of response values, $X_i$ is an $n_i \times b$ design matrix for the fixed effects, $\beta$ is a $b \times 1$ vector of
fixed effect parameters, $Z_i$ is an $n_i \times g$ design matrix for the random effects, and $\gamma_i$ is a $g \times 1$ vector of random effect scores. As hinted by the subscript $i$, the fixed effects coefficients $\mathbf{b} = (\beta_1, ..., \beta_g)^T$ are common among all individuals, and the random effects $\gamma_i = (\gamma_{i1}, ..., \gamma_{ig})^T$ can vary across individuals. They are assumed to be independent, with a distribution $\gamma_i \sim N(0, \sigma^2 \mathbf{D})$. The $\epsilon_i$ are within subject errors with a distribution $\epsilon_i \sim N(0, \sigma_{\epsilon}^2 \mathbf{W}_i)$.

The Repeated-Measures ANOVA Model

Repeated-measures ANOVA is the classical approach for longitudinal data analysis. Fisher (Scheffe, 1959) developed the ANOVA model to model mean differences based on experimental design. His original formulation of the repeated-measures ANOVA model was designed for the analysis of a randomized experiment in which the repeated measures factor represented a randomized ordering of a repeatedly administered treatment factor. Since his repeated measures factor resulted in correlated residuals, Fisher recognized that the validity of the F-test rested on certain assumptions. It remained for Box (1954) to demonstrate the assumptions required to have a valid F-test. When first developed, the model assumed a strict error structure called compound symmetry, which requires homogeneous variances across time and covariances between each pair of occasions. Until Box it was generally assumed that the F-test was robust to violations of this assumption. Box (1954) showed that this assumption is often invalid in longitudinal studies, resulting in a positive F-test bias. When violated the F-test had a distribution in which the degrees of freedom were $\varepsilon$-adjusted ($\varepsilon \times df$) with $\varepsilon$ having an upper bound of 1 under no
violation of compound symmetry. Accordingly, a large body of literature emerged on alternatives to the test. Some researchers suggest ways to adjust the F-test with a correction factor. Greenhouse and Geisser (1959) developed an alternative $\varepsilon$-adjusted test; Huynh and Feldt (1970) and Rouanet and Lepine (1970) independently developed an alternative $\varepsilon$ adjustment based on the alternative assumption of sphericity in which the covariance matrix of a set of orthogonally transformed variables formed the basis of the assumption. McCall and Appelbaum (1973) recommended replacing the univariate ANOVA model with a multivariate ANOVA (MANOVA) model based on the work of Mauchly (1940). The MANOVA model requires no assumption on error structure, but it is also less powerful than the univariate omnibus test. Hence, routine use of MANOVA may be too conservative for longitudinal data analysis (Hertzog & Rovine, 1985).

A major contribution to the development of repeated-measures ANOVA was made by Jennrich and Schluchter (1986) when they introduced methods for modeling different types of error structure related to their work on the development of a general procedure for estimating mixed model (SAS PROC MIXED). Rather than adjust F-tests for violations of assumptions (either compound symmetry or sphericity), they included the proper error covariance structure as part of the model. They extended the model’s repertoire to include covariance patterns such as first-order autoregressive (AR1), general autoregressive (Toeplitz), and unstructured (UN), which greatly enhanced its flexibility. These alternative patterns are extremely important for repeated measures models in which time, not the randomized order of treatments, is
the repeated measures factor. When the levels of the repeated measures factor are ordered according to time, it is reasonable that the distance between occasions would affect the size of residual correlations. AR(1) and Toeplitz patterns take this into consideration.

In the GLMM framework, a one-way repeated measures ANOVA model is:

$$y_i = X_i \beta + \epsilon_i$$  \hspace{1cm} (3)

(Rovine & Molenaar, 2000). For example, for a repeated-measures design with 5 occasions, the model can be written as:

$$\begin{bmatrix}
y_{i1} \\
y_{i2} \\
y_{i3} \\
y_{i4} \\
y_{i5}
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0
\end{bmatrix} \begin{bmatrix}
\beta_1 \\
\beta_2 \\
\beta_3 \\
\beta_4 \\
\beta_5
\end{bmatrix} + \begin{bmatrix}
\epsilon_{i1} \\
\epsilon_{i2} \\
\epsilon_{i3} \\
\epsilon_{i4} \\
\epsilon_{i5}
\end{bmatrix}$$  \hspace{1cm} (4)

where $y_{i1}$ to $y_{i5}$ are the individual’s scores at the 5 occasions. The first column of the design matrix represents an intercept, and the next four columns represent dummy codes for the time effect. The fixed-effects coefficients $\beta_1$, $\beta_2$, $\beta_3$, $\beta_4$, and $\beta_5$ yield the expected values for $y_{i5}$, $y_{i1} - y_{i5}$, $y_{i2} - y_{i5}$, $y_{i3} - y_{i5}$, $y_{i4} - y_{i5}$, respectively. The vector of within person errors $\epsilon_i$ represents the random effects. They form a $5 \times 5$ covariance matrix, $\sigma^2 \mathbf{W}_i$, which is usually assumed to be identical across subjects, and thus can be written as $\sigma^2 \mathbf{W}$.

The error structure is modeled when estimating the fixed-effects parameters. One possibility is $\sigma^2 \mathbf{W} = \sigma^2 \mathbf{I}$, where $\mathbf{I}$ is an identity matrix with 1s on the diagonal and 0s off the diagonal. This pattern corresponds to the independent observation assumption of traditional ANOVA, which is unlikely in a repeated measures design.
Another possibility is the compound symmetry pattern:

\[
\begin{bmatrix}
\sigma_e^2 + \sigma & \sigma & \sigma & \sigma & \sigma \\
\sigma & \sigma^2 + \sigma & \sigma & \sigma & \sigma \\
\sigma & \sigma & \sigma^2 + \sigma & \sigma & \sigma \\
\sigma & \sigma & \sigma & \sigma^2 + \sigma & \sigma \\
\sigma & \sigma & \sigma & \sigma & \sigma^2 + \sigma \\
\end{bmatrix}
\]

where the variances of errors are the same across occasions, and the covariance are the same between any two occasions. In theory, compound symmetry is most appropriate for a repeated measures experimental design in which the order of treatment conditions is randomly assigned.

For a longitudinal study in which the effect of time is important, a reasonable error structure is AR(1):

\[
\begin{bmatrix}
1 & \rho & \rho^2 & \rho^3 & \rho^4 \\
\rho & 1 & \rho & \rho^2 & \rho^3 \\
\rho^2 & \rho & 1 & \rho & \rho^2 \\
\rho^3 & \rho^2 & \rho & 1 & \rho \\
\rho^4 & \rho^3 & \rho^2 & \rho & 1 \\
\end{bmatrix}
\]

where \(\sigma_e^2\) is the variance of errors and \(\rho\) is the autocorrelation. Conceptually, this pattern results from the first-order autoregressive process:

\[
e_{it} = \rho e_{i(t-1)} + v_{it}
\]

(5)

where \(v_{it} \sim N(0, \sigma^2)\). It assumes that the correlation of errors between any two consecutive occasions is identical; thus, the correlation decreases at a constant rate as two measurements get farther away in time.

A similar but less strict structure is general autoregressive, banded, or Toeplitz:
This pattern indicates that the correlation of errors depends only on the interval of time between two measurements.

A repeated-measures ANOVA model is estimated based on the presumed error structure. Notably, error structures do not influence the fixed effects estimates $\mathbf{B}$, which are simply mean differences. They do, however, influence the standard error estimates for the fixed effects, and thus the inferences of the significance tests. They also affect the goodness-of-fit of the model. Two frequently used goodness-of-fit indices are the Akaike information criterion (AIC; Akaike, 1974) and the Bayesian information criterion (BIC; McQuarrie & Tsai, 1998; Schwarz, 1978). They both penalize the number of parameters estimated, but by different standards. Generally, the AIC prefers more complex models and the BIC favors simpler models (Cudeck & Browne, 1983). However, these two indices are measures of relative fit and do not provide information about the absolute fit. One can only use them to compare different models on the same data and select the most appropriate one.

The Multilevel Growth Curve Model

The multilevel growth curve model is an outgrowth of the general linear mixed model (Laird & Ware, 1982) which followed two complementary traditions: growth models first described independently by Tucker (1958) and Rao (1958); and linear mixed models described by Henderson (1953) and Hartley and Rao (1967). Statistical
methods for analyzing multilevel data from this point of view have been developed and described as random coefficients models (Rosenberg, 1973), random effects models (Laird & Ware, 1982), mixed models (Littell, Milliken, Stroup, & Wolfinger, 1996), and variance component models (Longford, 1993). The three most popular approaches for estimating these models are the multilevel modeling approach (Goldstein, 1995), the hierarchical linear modeling approach (Bryk & Raudenbush, 1992), and the mixed model approach (Littell, Milliken, Stroup, & Wolfinger, 1996). Estimation approaches for these methods include empirical Bayes (Bryk & Raudenbush, 1992) and the estimator resulting from the Henderson mixed model equations (Henderson, 1990). It has been shown that the estimators resulting from these two methods are equivalent (Littell, Milliken, Stroup, & Wolfinger, 1996; Robinson, 1991).

A variant of the growth curve model is the latent growth curve model proposed by McArdle and Epstein (1987). It represented a structural modeling approach that estimated some of the effects of a multilevel model presented as a combination of repeated-measures ANOVA and longitudinal factor analysis. Rovine and Molenaar (2000; 2001) showed how to add a complete set of fixed effects to the model, demonstrating the close relationship between the latent growth curve model and the multilevel growth curve model. An alternative model was proposed by Muthen (1994). In the latent growth curve model, it is conceptualized that individual development follows a common trajectory that can be described by latent variables representing “level” and “shape”. Each individual has his or her own scores on these latent
variables. For example, a linear growth model assumes that individual development follows a straight line, but each person has his or her intercept and slope.

The growth curve model is often presented in the form of a multilevel model, which is an alternative way of representing a single mixed model equation. For example, a simplest multilevel linear growth model is:

Level 1: \[ y_{it} = \pi_{0i} + \pi_{1i} \text{time} + \epsilon_{it} \]
Level 2: \[ \pi_{0i} = \beta_{00} + \nu_{0i} \]
\[ \pi_{1i} = \beta_{10} + \nu_{1i} \] (6)

where \( \pi_{0i} \) and \( \pi_{1i} \) are individual intercepts and slopes, \( \epsilon_{it} \) are errors around individual lines that are assumed to be normally distributed and have a mean of zero, \( \beta_{00} \) and \( \beta_{10} \) are the average intercept and slope, \( \nu_{0i} \) are differences between the individual and average intercepts, and \( \nu_{1i} \) are differences between the individual and average slopes.

Combining the level 1 and level 2 equations we will get:

\[ y_{it} = [\beta_{00} + \beta_{10} \text{time}] + [\nu_{0i} + \nu_{1i} \text{time}] + \epsilon_{it} \] (7)

Suppose the data comes from a repeated-measures design with 5 equally spaced occasions, Equation (7) can be written as:

\[
\begin{bmatrix}
  y_{i1} \\
  y_{i2} \\
  y_{i3} \\
  y_{i4} \\
  y_{i5}
\end{bmatrix} =
\begin{bmatrix}
  1 & 1 \\
  1 & 2 \\
  1 & 3 \\
  1 & 4 \\
  1 & 5
\end{bmatrix}
\begin{bmatrix}
  \beta_{00} \\
  \beta_{10}
\end{bmatrix} +
\begin{bmatrix}
  1 & 1 \\
  1 & 2 \\
  1 & 3 \\
  1 & 4 \\
  1 & 5
\end{bmatrix}
\begin{bmatrix}
  \nu_{0i} \\
  \nu_{1i}
\end{bmatrix} +
\begin{bmatrix}
  \epsilon_{i1} \\
  \epsilon_{i2} \\
  \epsilon_{i3} \\
  \epsilon_{i4} \\
  \epsilon_{i5}
\end{bmatrix}
\] (8)

This is in the form of the GLMM which is:

\[ y_{i} = X_i \beta + Z_i \gamma_i + \epsilon_i \] (2)

where \( \gamma_i \sim N (0, \sigma^2 \mathbf{D}) \) and \( \epsilon_i \sim N (0, \sigma^2 \mathbf{I}) \). In the case of the linear growth model,
$\sigma^2D$ is the covariance matrix of the random intercept $\nu_{0i}$ and the random slope $\nu_{1i}$:

$$
\begin{bmatrix}
\sigma_i^2 & \sigma_{i\nu} \\
\sigma_{i\nu} & \sigma_{\nu}^2
\end{bmatrix}
$$

The error structure of growth curve model is called the random-effects structure (Jennrich & Schluchter, 1986; Laird & Ware, 1982). Using the denotations for GLMM, the total covariance matrix for the $i$th subject is:

$$
V_i = \sigma^2Z_iDZ_i^T + \sigma^2\epsilon I
$$

where $Z_i = Z$ is constant across subjects (Jennrich & Schluchter, 1986; Rovine & Molenaar, 2000).

As demonstrated by Rovine and Molenaar (1998), the magnitude of $\sigma_{i\nu}$ depends on the scaling of the shape basis vector, which is often an arbitrary decision made by researchers. For simplicity, therefore, I compute the random-effects structure for a five-occasion linear growth model with $\sigma_{i\nu}=\sigma_{i\nu}=0$:

$$
\begin{bmatrix}
\sigma_i^2 + \sigma_\epsilon^2 & \sigma_i^2 & \sigma_i^2 & \sigma_i^2 & \sigma_i^2 \\
\sigma_i^2 & \sigma_i^2 + \sigma_s^2 + \sigma_\epsilon^2 & \sigma_i^2 + 2\sigma_s^2 & \sigma_i^2 + 3\sigma_s^2 & \sigma_i^2 + 4\sigma_s^2 \\
\sigma_i^2 & \sigma_i^2 + 2\sigma_s^2 & \sigma_i^2 + 4\sigma_s^2 + \sigma_\epsilon^2 & \sigma_i^2 + 6\sigma_s^2 & \sigma_i^2 + 8\sigma_s^2 \\
\sigma_i^2 & \sigma_i^2 + 3\sigma_s^2 & \sigma_i^2 + 6\sigma_s^2 & \sigma_i^2 + 9\sigma_s^2 + \sigma_\epsilon^2 & \sigma_i^2 + 12\sigma_s^2 \\
\sigma_i^2 & \sigma_i^2 + 4\sigma_s^2 & \sigma_i^2 + 8\sigma_s^2 & \sigma_i^2 + 12\sigma_s^2 & \sigma_i^2 + 16\sigma_s^2 + \sigma_\epsilon^2
\end{bmatrix}
$$

This pattern assumes an increasing composite on the diagonal over time, which may be expected for some growth phenomenon. However, it is unrealistic in a number of ways. For example, it assumes larger covariances between equally spaced occasions over time, and larger covariances between occasions that are farther apart. In general, therefore, this is an error structure that is hardly seen in real life.
Repeated-measures ANOVA or growth curve modeling?

A comparison of Equation (8) and Equation (4) shows that the difference between the repeated-measures ANOVA model and the growth curve model is twofold. First, in the repeated-measures ANOVA model, time is dummy-coded and treated as a categorical variable, whereas in the growth curve model, time is continuous. Hence, growth curve models are more convenient when dealing with data where individuals are measured at varying occasions. Such data often comes from studies on actual growth phenomena, such as weight, height, or children’s vocabulary development. In many developmental studies, however, researchers are only interested in changes of a construct across several time points, and each participant is measured on the same occasions. In this case, treating time as categorical may be more appropriate.

Second, the types of error structure that the two can model are different. It is obvious that the selection of the correct model depends on the true error structure of the data. Yet, the question remains in whether there is a way to figure out what the true error structure is, and how the characteristics of the data influence the validity of the criterion.

In this study, I examine the validity of AIC and BIC in selecting the correct error structure, and how it is affected by effect size, intraclass correlation size, and sample size, etc. The study consists of two parts: simulation and analysis. Details of the study are described in Chapter 3.
Chapter 3

Method

Simulation

The simulation consists of two parts: setting the expected values, and simulating the errors. First, to make the comparison possible, the expected values are set as identical for all error structures. Specifically, all data are assumed to come from a repeated-measures study with five equally spaced occasions. The mean at each occasion is 5, 10, 15, 20, and 25, respectively. Next, errors are simulated based on three types of error structure:

*Compound symmetry.* For the compound symmetry structure, the simulation of errors follows a three dimensional design: 1) effect size; 2) intraclass correlation size; and 3) sample size. Effect size has two levels: medium (.5) and large (.8) (Cohen, 1988). Intraclass correlation size has three levels: small (\(\rho = .2\)), medium (\(\rho = .5\)), and large (\(\rho = .8\)). Sample size has three levels: small (20), medium (100), and large (200).

To construct the covariance matrix, I use the formula:

\[
d = \frac{M_1 - M_2}{\sigma'}
\]

(Cohen, 1988) where \(d\) is the effect size, \(M_1-M_2\) is the difference between two means, and \(\sigma'\) is the standard deviation of either group. In this case, \(M_1-M_2\) is 5, and

\[
\sigma' = \sqrt{\sigma^2 + \sigma}
\]

(11)

Combining the formula for intraclass correlation:

\[
\rho = \frac{\sigma}{\sigma^2 + \sigma}
\]

(12)
we can solve for $\sigma^2_e$ and $\sigma$. Table 1 lists the values of $\sigma^2_e$ and $\sigma$ using different combinations of parameters. All together, the three-dimension design yields 18 $(2 \times 3 \times 3)$ possible combinations of parameter values. For each combination, I simulate 100 sets of data using different seeds. The SAS program used for simulating data with compound symmetry error structure is given in Appendix A.

**First-order autoregressive (AR1).** The simulation of AR1 error structure follows the same design as before and yields 1800 sets of data. Note, however, that for the AR1 structure, $\sigma'$ is simply $\sqrt{\sigma^2_e}$. In other words, the compound symmetry structure and AR1 structure have the same error variances, but different $\sigma^2_e$. This is because $\sigma^2_e$ is defined in different ways in the two models.

**Random-effects structure.** Since $Z$ is constant in the simulation, the random-effects structure is determined by $\sigma^2_e$ and $\sigma^2D = \begin{bmatrix} \sigma^2_i & \sigma_{in} \\ \sigma_{in} & \sigma^2_v \end{bmatrix}$. To make the growth curve models comparable to the repeated-measures ANOVA models, I set the value of $\sigma^2_e$ and $\sigma_i^2$ as equivalent to the $\sigma^2_e$ and $\sigma$ in the CS structure, respectively. In other words, the total variance at the first occasion in the growth curve model is the same as the error variance in the CS and AR1 model. Moreover, I define:

$$r = \frac{\sigma^2_i}{\sigma^2_v} \tag{13}$$

which has two levels: small (0.1) and medium (0.25). Because $\sigma_{in}$ is arbitrary, it is set to 0. The simulation of random-effects structure thus follows a four dimensional design, with $r$ as the added dimension. For this structure, I only use a medium effect size (.5), and small (20) and medium (100) sample sizes. In total, there are 12
(1×2×3×2) combinations of parameter values, which yield 1200 datasets. Table 2 shows the values of the covariance parameters for the simulation.

Analysis

Each set of data is analyzed in SAS using three models: 1) repeated-measures ANOVA with compound symmetry error structure (i.e., the CS model); 2) repeated-measures ANOVA with AR1 error structure (i.e., the AR1 model); and 3) growth curve model (i.e., the GC model). Results are summarized and discussed in Chapter 4 and Chapter 5.
Chapter 4

Results

Results are summarized in Table 3 to Table 10, which include estimates of covariance parameter, fixed effects parameters, fixed effects standard errors, and fit statistics. Because 100 datasets are simulated for each combination of parameter values, it is also possible to compute the standard deviation of the estimates. The standard deviations are listed in the parentheses under the corresponding parameters.

Estimates of fixed effects parameters

In repeated-measures ANOVA models, $\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$, and $\beta_4$ is $\bar{y}_5$, $\bar{y}_1 - \bar{y}_5$, $\bar{y}_2 - \bar{y}_5$, $\bar{y}_3 - \bar{y}_5$, and $\bar{y}_4 - \bar{y}_5$, respectively. Since the means are 5, 10, 15, 20, and 25, we have:

$$
\begin{align*}
E(\beta_0) &= 25 \\
E(\beta_1) &= -20 \\
E(\beta_2) &= -15 \\
E(\beta_3) &= -10 \\
E(\beta_4) &= -5
\end{align*}
$$

In the growth curve model, $\beta_0$ and $\beta_1$ are the intercept and slope, therefore:

$$
\begin{align*}
E(\beta_0) &= 5 \\
E(\beta_1) &= 5
\end{align*}
$$

As shown in Table 3 to Table 10, all three models successfully recover the expected values. The CS model and the AR1 model always yield the same values, demonstrating that the error structure does not influence the estimates of fixed effects parameters. The standard deviations decrease as effect size and sample size get larger, which indicates that the simulation is more reliable for data with larger effect size and sample size.
Standard errors of fixed effects parameters

Comparison of models. In the repeated-measures ANOVA model, we are most interested in the standard errors of $\beta_1$ to $\beta_4$, because they are used to construct significance tests for contrasts between two occasions, i.e., the time effect. The results show that the CS model yields the same standard errors for $\beta_1$ to $\beta_4$, whereas in the AR1 model, the standard error decreases as $\beta$ goes from $\beta_1$ to $\beta_4$. The reason for that lies in the way standard errors are computed in these models:

$$ SE = \sqrt{L \hat{C} L'} $$

where $L$ is the vector representing the contrast and $\hat{C}$ is the estimated covariance matrix of $\beta$ (Littell, Milliken, Stroup, & Wolfinger, 1996). Let $\hat{R}$ denote the estimated covariance matrix of errors, and $\hat{C}$ can be computed as:

$$ \hat{C} = (X' \hat{R}^{-1} X) $$

where $X$ is the design matrix for the fixed effects. It can be demonstrated that when $\hat{R}$ has a compound symmetry structure, the covariance matrix of $\beta_1$ to $\beta_4$ also shows a compound symmetry pattern. This is equivalent to identifying compound symmetry as a one-factor model. On the other hand, when $\hat{R}$ has an AR1 structure, the covariance matrix of $\beta_1$ to $\beta_4$ has a similar pattern that results in decreasing standard errors.

The way that standard errors are calculated also leads to an unexpected result, that is, the AR1 model tends to yield larger standard errors than the CS model, even when the data under analysis is of an AR1 error structure.

The influence of intraclass correlation. In the CS and AR1 models, the standard
errors of $\beta_1$ to $\beta_4$ generally decline as intraclass correlation gets larger. In the growth curve models, the standard error of $\beta_0$ increases as intraclass correlation gets larger. The standard error of $\beta_1$ does not show an obvious trend.

*The influence of effect size.* For data with CS or AR1 error structure, standard errors decline as effect size increases, regardless of the analytic model (Table 3 to Table 8). This is consistent with expectation because larger effect size means smaller errors given that the fixed effects parameters are hold constant.

*The influence of r.* For data with a random-effects error structure, standard errors increase as $r$ increases, regardless of the analytic model (Table 9 and Table 10). This is not hard to understand because a larger $r$, i.e., a larger $\sigma_s^2$ compared to $\sigma_i^2$, means the data is more spread out and individual scores are farther away from the group means.

*The influence of sample size.* Standard errors always decline as sample size gets larger.

*Estimates of covariance parameters*

Generally, the “true” model successfully recovers the covariance parameters, and the other two models give estimates within a reasonable range. For example, for data with CS error structure, small intraclass correlation ($\rho=.2$), medium effect size (.5) and sample size of 100 (Table 4), the CS model gives an estimate of $\sigma_e^2=79.72$ and an estimate of $\sigma=19.97$, which are very close to the true values 80 and 20. In the AR1 model, $\sigma_e^2$ is estimated to be 99.72, which is very close to $\sigma_e^2+\sigma$ in the CS model. $\rho$ is estimated to be .20, which is the true value. The GC model gives estimates of $\sigma_e^2$ and
\(\sigma_i^2\) similar to the estimates of \(\sigma^2\) and \(\sigma\) in the CS model. It also gives estimates of \(\sigma_{s}^2\) and \(\sigma_{is}\) that are not significantly different from zero and can be ignored. For data with random-effects error structure, small intraclass correlation (\(\rho = .2\)), \(r\) of .1, and sample size of 100 (Table 10), the growth curve model estimates \(\sigma^2_{e}, \sigma^2_{i}, \sigma^2_{s},\) and \(\sigma_{is}\) to be 79.21, 20.74, 1.99, and -0.09, respectively. These numbers are very close to the true values 80, 20, 2, and 0. The CS model gives an estimate of \(\sigma^2_{e} = 84.21\) and an estimate of \(\sigma = 27.35\), which are larger than the \(\sigma^2_{e}\) and \(\sigma^2_{i}\) in the GC model. Clearly, the CS model is trying to adjust for the increasing variance in the random-effects structure. The AR1 also gives enlarged estimates of \(\sigma^2_{e}\) and \(\rho\). To summarize, models with incorrect error structures at least recover the diagonal of the simulated covariance matrix successfully. In these models, however, elements off diagonal may not match those in the simulated covariance matrix, leading to misfit of the model. The issue of model fit is examined below.

**Fit statistics**

As expected, the repeated-measures ANOVA models always yield the lowest average AIC and BIC when the presumed error structure matches the true error structure. In other words, the CS model, on average, always fits data with CS error structure better than the AR1 and growth curve models (Table 3 to Table 5). Similarly, the AR1 model, on average, always fits data with AR1 error structure better than the CS and growth curve models (Table 6 to Table 8). In contrast, the growth curve model fails to yield the lowest average AIC and BIC: 1) most of the time when sample size is small (Table 9); and 2) when sample size is 100, intraclass correlation is small, and \(r\)
is small (Table 10).

The MIN AIC and MIN BIC reveal the same trend. These variables are created to count the number of times that a model has the smallest AIC and BIC among the three models, out of 100 times of comparison. They indicate the chance of being selected by AIC and BIC as the best fitting model. The numbers do not always add up to 100, because two models can have the smallest AIC and BIC in the same comparison. The results show that AIC and BIC perform fairly well when the true error structure is CS or AR1. For data with CS error structure, the fit indices suffer the most when intraclass correlation is small and sample size is small (Table 3). Yet, even in this case, AIC and BIC pick the correct model 80% of the time. As intraclass correlation and sample size increase, the rate of selecting the correct model quickly improves to more than 95% (see Figure 1 for sample size=20). For data with AR1 error structure, AIC and BIC also perform poorly when intraclass correlation and sample size are both small (Table 6). In the worst scenario (effect size is also small), these fit indices pick the correct model only 60% of the time. However, as intraclass correlation and sample size increase, AIC and BIC make almost perfect selection (see Figure 2 for sample size=20).

For data with random-effects error structure, the story is different. When sample size is small (Table 9), AIC and BIC rarely pick the correct model. For example, when $\rho=.2$ and $r=.1$, the growth curve model is considered the best fitting model only one out of 100 times! The rate of correct model selection increases as $\rho$ and $r$ get larger. Yet, even in the best scenario ($\rho=.8$, $r=.25$), AIC and BIC pick the growth curve model
only about 70% of the times (Figure 3). When sample size reaches 100 (Table 10), the performance of AIC and BIC improves substantially. However, when $\rho$ and $r$ are both small, AIC and BIC still make mistakes more than 80% of the time (Figure 4).

Having the percentage of correct selection in mind, it may be helpful to examine the type of mistake AIC and BIC make, namely, which model do AIC and BIC favor over the true model? Simply by looking at the MIN AIC and MIN BIC, it is clear that when the true error structure is CS or AR1, AIC and BIC tend to pick the other ANOVA model instead of the growth curve model (Table 3 to Table 8). When the true error structure is random-effects, AIC and BIC do not show obvious preference when sample size is small (Table 9). When sample size reaches 100, they lean towards the CS model (Table 10).
Chapter 5
Discussion

In this study, three models are compared on modeling longitudinal data with different error structures, intraclass correlation sizes, effect sizes, and sample sizes. These three models are: 1) repeated-measures ANOVA model with compound symmetry error structure; 2) repeated-measures ANOVA model with AR1 error structure; and 3) growth curve model. AIC and BIC are used to assess the fit of the model.

Results reveal that all three models successfully recover the fixed effects parameters. Somewhat counter-intuitively, the AR1 model tends to yield larger standard errors than the CS model, even when AR1 is the true error structure and the model fits the data better. It indicates that when the true error structure is AR1, the CS model underestimates the standard errors, and thus increase the risk of type I error. On the other hand, when the true error structure is compound symmetry, the AR1 model overestimates the standard errors, and thus may lead to type II error. A simple and convenient way to avoid making mistakes is to compare the fit statistics before looking at the significant tests. When the error structure is compound symmetry or AR1, AIC and BIC generally are good indicators for selecting the true model.

It should be noted that AIC and BIC perform surprisingly poorly when it comes to data with random-effects error structure, especially when sample size is small. When the sample size reaches 100, AIC and BIC still have difficulty picking out the true error structure when intraclass correlation is small and the proportion of slope variance to intercept variance is small. This is not hard to understand when we
compare the error structure of the models. The random-effects error structure is:

\[
\begin{bmatrix}
\sigma_i^2 + \sigma_e^2 & \sigma_i^2 & \sigma_i^2 & \sigma_i^2 & \sigma_i^2 \\
\sigma_i^2 & \sigma_i^2 + \sigma_s^2 + \sigma_e^2 & \sigma_i^2 + 2\sigma_s^2 & \sigma_i^2 + 3\sigma_s^2 & \sigma_i^2 + 4\sigma_s^2 \\
\sigma_i^2 & \sigma_i^2 + 2\sigma_s^2 & \sigma_i^2 + 4\sigma_s^2 + \sigma_e^2 & \sigma_i^2 + 6\sigma_s^2 & \sigma_i^2 + 8\sigma_s^2 \\
\sigma_i^2 & \sigma_i^2 + 3\sigma_s^2 & \sigma_i^2 + 6\sigma_s^2 & \sigma_i^2 + 9\sigma_s^2 + \sigma_e^2 & \sigma_i^2 + 12\sigma_s^2 \\
\sigma_i^2 & \sigma_i^2 + 4\sigma_s^2 & \sigma_i^2 + 8\sigma_s^2 & \sigma_i^2 + 12\sigma_s^2 & \sigma_i^2 + 16\sigma_s^2 + \sigma_e^2 \\
\end{bmatrix}
\]

And the compound symmetry error structure is:

\[
\begin{bmatrix}
\sigma_i^2 + \sigma & \sigma & \sigma & \sigma & \sigma \\
\sigma & \sigma^2 + \sigma & \sigma & \sigma & \sigma \\
\sigma & \sigma & \sigma^2 + \sigma & \sigma & \sigma \\
\sigma & \sigma & \sigma & \sigma^2 + \sigma & \sigma \\
\sigma & \sigma & \sigma & \sigma & \sigma^2 + \sigma \\
\end{bmatrix}
\]

In the simulation, \(\sigma_i^2\) is set to equal \(\sigma\). Hence, the random-effects structure only differs from the compound symmetry structure in that it has \(\sigma_s^2\) added to its elements. When intraclass correlation \(\rho = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_e^2}\) is small and \(r = \frac{\sigma_s^2}{\sigma_i^2}\) is small, \(\sigma_s^2\) is small compared to \(\sigma_i^2\) and \(\sigma_e^2\), and the two models become indistinguishable. Data with small sample size is more vulnerable to fluctuation errors, and thus makes it even harder for AIC and BIC to recognize the correct model. Another way to understand the difference between the two models is to consider the compound symmetry structure as a special case of the random-effects structure where \(\sigma_s^2\) equals 0. In a graph, data with compound symmetry error structure would form “growth curves” that are parallel lines. When the variance of slope is small, AIC and BIC may prefer the more parsimonious compound symmetry model over the growth curve model.

To some extent, the small chance that growth curve model is selected by fit statistics even when it is the true model suggests that the growth curve model is less robust than the repeated-measures ANOVA model. Whether it is the best fitting model...
depends heavily on the parameter values and the sample size. The bigger the $\sigma_s^2$ compared to $\sigma_i^2$ and $\sigma_e^2$, and the larger the sample size, the more likely that growth curve model fits the data better than the repeated-measures ANOVA models.

Unfortunately, the random-effects structure may not always be realistic in longitudinal studies. For a study with five equally spaced occasions, for example, the random-effects structure makes the following assumptions that are unlikely to be satisfied: 1) the total variance increases with time; 2) the covariance is the same between the first occasion and all later occasions; 3) the covariance between the second occasion and later occasions gradually increases as the interval gets larger; 4) the covariance between the third occasion and later occasions increases as the interval gets larger. What’s worse, the bigger the $\sigma_s^2$ compared to $\sigma_i^2$ and $\sigma_e^2$, the more unrealistic this pattern becomes. Consider, for example, the “threshold” covariance matrix for the growth curve model to be actually selected as the best fitting model.

Based on Table 10, this threshold covariance matrix for data with a sample size of 100 is:

$$
\begin{pmatrix}
100 & 50 & 50 & 50 & 50 \\
50 & 105 & 60 & 65 & 70 \\
50 & 60 & 120 & 80 & 90 \\
50 & 65 & 80 & 145 & 110 \\
50 & 70 & 90 & 110 & 180
\end{pmatrix}
$$

Transformed into a correlation matrix, it is:

$$
\begin{pmatrix}
1 & .49 & .46 & .42 & .37 \\
.49 & 1 & .53 & .53 & .51 \\
.46 & .53 & 1 & .61 & .61 \\
.42 & .53 & .61 & 1 & .68 \\
.37 & .51 & .61 & .68 & 1
\end{pmatrix}
$$
For data with a sample size of 20 (Table 9), the threshold covariance matrix becomes:

\[
\begin{pmatrix}
100 & 80 & 80 & 80 & 80 \\
80 & 120 & 120 & 140 & 160 \\
80 & 120 & 180 & 200 & 240 \\
80 & 140 & 200 & 280 & 280 \\
80 & 160 & 240 & 320 & 420
\end{pmatrix}
\]

which can be transformed into:

\[
\begin{pmatrix}
1 & .73 & .60 & .48 & .39 \\
.73 & 1 & .82 & .76 & .71 \\
.60 & .82 & 1 & .89 & .87 \\
.48 & .76 & .89 & 1 & .93 \\
.39 & .71 & .87 & .93 & 1
\end{pmatrix}
\]

It is clear that the patterns of the covariance matrices above are not what we would normally expect, and the latter one is even less likely than the first one. This creates a paradox, that is, the more likely the growth curve model is the best fitting model, the less likely the error structure can be observed in real life. The pattern of the correlation matrices seems more realistic after taking out the effect of the increasing total variance. However, it still assumes larger correlations between equally spaced occasions that are farther away form the origin (in this case, the first occasion). This pattern is not frequently observed.

This result provides evidence against the argument that growth curve model is generally a better choice than the repeated-measures ANOVA model in longitudinal data analysis (Llabre, Spitzer, Siegel, Saab, & Schneiderman, 2004; van der Leeden, 1998). It also puts into question the common practice of fitting growth curve models to longitudinal data without comparing the fit of different models. That being said, one may ask, what is the best strategy for analyzing repeated-measures data? Such
questions usually do not have a clear-cut answer, because the choice of model
depends on what the data is like. The current study suggests that repeated-measures
ANOVA is more appropriate when dealing with data where participants are measured
at identical points along the time axis. When there is no strong theatrical support for a
common growth pattern, researchers’ interest is often just the mean differences
between occasions. The best strategy, in this case, will be to find the best-fitting
model to get the most accurate statistical inference. Since longitudinal studies in
developmental psychology usually do not involve a large sample, the chance that a
growth curve model fits the data well may not be high.

Moreover, the repeated-measures ANOVA model has advantages over the growth
curve model in that it gives exact estimates of the means. This is more apparent when
the dependent variable does not show a linear change pattern. For example, consider
modeling the following set of arbitrary means:

\[
\begin{align*}
\bar{y}_1 &= 5 \\
\bar{y}_2 &= 10 \\
\bar{y}_3 &= 4 \\
\bar{y}_4 &= -3 \\
\bar{y}_5 &= 1
\end{align*}
\]

Using repeated-measures ANOVA, we will get exact estimates of the fixed effect
parameters:

\[
\begin{align*}
\beta_0 &= 1 \\
\beta_1 &= 4 \\
\beta_2 &= 9 \\
\beta_3 &= 3 \\
\beta_4 &= -4
\end{align*}
\]

They yield direct t-tests of the intercept and the difference in the means between the
last occasion and other occasions. If researchers are not interested in comparing the last occasion with other occasions, it is easy to construct contrasts to test any kind of time effect (e.g., linear, quadratic, cubic, $\bar{y}_2$ vs. $\bar{y}_3$, etc.). Using growth curve modeling, we can fit a polynomial growth function with time effects up to the power of 3, i.e., a cubic trend. Even with a cubic trend, however, the growth curve model will not recover the means perfectly, and researchers may lose information. Therefore, by setting the means to form a straight line, this study already puts the growth curve model in an advantaged situation in the comparison. When the change is not linear, which is often the case, repeated-measures ANOVA is even more appealing than is shown in this study.

On the other hand, when the interest of the study involves genuine growth, i.e., when individuals are expected to follow a universal growth pattern, growth curve modeling may be more appropriate in a conceptual sense. Moreover, growth curve modeling is generally more convenient when participants are measured on different occasions, or when dealing with a large amount of missing data.

In conclusion, this study contributes to the literature in several ways. First, it compares the repeated-measures ANOVA models and the growth curve model in the GLMM framework, thus helps researchers understand the similarity and difference between the two. Second, it explicitly compares the performance of these models under various conditions, and proposes a convenient procedure for model selection. Specifically, it shows that AIC and BIC are good indicators of the true model most of the time, and thus can be used to select the proper model. Third, it shows that the
growth curve model is not as robust as the repeated-measures ANOVA model. When sample size is small and the variance of slope is small compared to other covariance parameters, the repeated-measures ANOVA model often fits the data better than the growth curve model even when the error structure has a random-effects pattern. Hence, it provides evidence against the common notion of considering growth curve model as the first choice for longitudinal data analysis.

This study also has some shortcomings that limit its implications. First, it uses AIC and BIC as the fit indices, which do not provide information on the absolute fit of the model. Whereas we can select the best-fitting model out of the three, it is not possible to tell whether the best-fitting model is good enough, or, whether the other two models fit the data sufficiently well. Second, the parameters I use in the simulation consist of only a small proportion of all possible values. This is due to the large amount of time and effort required for the simulation. Since it is unrealistic to exhaust all the possibilities, I only select values that are commonly seen in empirical research and are representative of the spectrum. Third, all data is assumed to come from a study with five equally spaced occasions, and the means fall on a straight line. In reality, longitudinal studies vary on the number and spacing of measurements. It is also common to see change that shows a quadratic, cubic, or unstructured pattern. As discussed above, repeated-measures ANOVA may be even more favored when the change is not linear, but this argument remains to be examined.

In follow-up studies, a comparison will be made between a linear contrast in the repeated-measures ANOVA model and the slope in the linear growth model.
Demonstration of their equivalence will strengthen the arguments made in this work. Moreover, an important next step will be to estimate the models using a structural equations modeling (SEM) approach (e.g., LISREL) in order to get a measure of the absolute goodness-of-fit. To make a more comprehensive comparison, it is also possible to examine the performance of these models under other conditions, such as when the means do not fall onto a straight line, or the repeated measures are not equally spaced.


*Biometrics, 60*, 61-75.


Table 1

Values of parameters used in simulation of data with compound symmetry error structure

<table>
<thead>
<tr>
<th></th>
<th>Effect size=.5</th>
<th></th>
<th></th>
<th>Effect size=.8</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho=.2$</td>
<td>$\rho=.5$</td>
<td>$\rho=.8$</td>
<td>$\rho=.2$</td>
<td>$\rho=.5$</td>
<td>$\rho=.8$</td>
</tr>
<tr>
<td>$\sigma^2_c$</td>
<td>80</td>
<td>50</td>
<td>20</td>
<td>31.25</td>
<td>19.53125</td>
<td>7.8125</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>20</td>
<td>50</td>
<td>80</td>
<td>7.8125</td>
<td>19.53125</td>
<td>31.25</td>
</tr>
<tr>
<td>$\sigma^2_c+\sigma$</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>39.0625</td>
<td>39.0625</td>
<td>39.0625</td>
</tr>
</tbody>
</table>
Table 2

Values of parameters used in simulation of data with random-effects error structure

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$r = .1$</th>
<th>$r = .25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho = .2$</td>
<td>$\rho = .5$</td>
<td>$\rho = .8$</td>
</tr>
<tr>
<td>$\sigma_c^2$</td>
<td>80</td>
<td>50</td>
</tr>
<tr>
<td>$\sigma_t^2$</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>$\sigma_s^2$</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>$\sigma_{ts}$</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 3

Parameter estimates and fit statistics of CS, AR1, and GC models for data with CS error structure and sample size of 20.

<table>
<thead>
<tr>
<th></th>
<th>$\rho = .2$</th>
<th>$\rho = .5$</th>
<th>$\rho = .8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect size = .5</td>
<td>CS</td>
<td>AR1</td>
<td>GC</td>
</tr>
<tr>
<td>$\sigma^2_c$</td>
<td>79.17</td>
<td>99.03</td>
<td>76.64</td>
</tr>
<tr>
<td>$\sigma^2/\sigma^2_i$ *</td>
<td>(12.77)</td>
<td>(16.15)</td>
<td>(12.71)</td>
</tr>
</tbody>
</table>

$\sigma^2$        | 19.87 | 0.18 | 26.12 | 50.33 | 0.48 | 55.09 | 81.60 | 0.79 | 84.87 |

$\sigma^2/\sigma^2_i$ * | (11.22) | (0.13) | (21.05) | (20.44) | (0.12) | (25.66) | (25.83) | (0.07) | (29.91) |

$\beta_0$         | 24.67 | 24.67 | 4.89 | 25.25 | 25.25 | 5.66 | 24.91 | 24.91 | 4.86 |

$\beta_1$         | (2.17) | (2.17) | (1.82) | (1.95) | (1.95) | (1.95) | (2.33) | (2.33) | (2.02) |

$\beta_2$         | 2.22 | 2.22 | 1.88 | 2.23 | 2.23 | 2.03 | 2.24 | 2.23 | 2.17 |

$\beta_3$         | -19.88 | -19.88 | 4.98 | -19.61 | -19.61 | 4.94 | -20.02 | -20.02 | 5.02 |

$\beta_4$         | (2.65) | (2.65) | (0.60) | (1.91) | (1.91) | (0.44) | (1.55) | (1.55) | (0.32) |

$\beta_5$         | 2.80 | 3.13 | 0.65 | 2.23 | 3.03 | 0.52 | 1.42 | 2.40 | 0.34 |

$\beta_6$         | (2.73) | (2.73) | - | (2.12) | (2.12) | - | (1.40) | (1.40) | - |

$\beta_7$         | 2.80 | 3.11 | - | 2.23 | 2.93 | - | 1.42 | 2.19 | - |

$\beta_8$         | -9.80 | -9.80 | - | -9.86 | -9.86 | - | -10.05 | -10.05 | - |

$\beta_9$         | (2.59) | (2.59) | - | (2.25) | (2.25) | - | (1.37) | (1.37) | - |

$\beta_{10}$      | 2.80 | 3.05 | - | 2.23 | 2.72 | - | 1.42 | 1.88 | - |

$\beta_{11}$      | -4.73 | -4.73 | - | -4.70 | -4.70 | - | -4.92 | -4.92 | - |

$\beta_{12}$      | (2.98) | (2.98) | - | (2.01) | (2.01) | - | (1.51) | (1.51) | - |

$\beta_{13}$      | 2.80 | 2.82 | - | 2.23 | 2.24 | - | 1.42 | 1.41 | - |

$\beta_{14}$      | 2.80 | 2.82 | - | 2.23 | 2.24 | - | 1.42 | 1.41 | - |

AIC               | 717.36 | 719.97 | 735.31 | 692.31 | 702.17 | 708.55 | 630.15 | 646.38 | 644.21 |

BIC               | 719.36 | 721.96 | 738.65 | 694.30 | 704.16 | 711.93 | 632.14 | 648.35 | 647.68 |

MIN AIC           | 83 | 17 | 0 | 95 | 5 | 0 | 99 | 1 | 0 |

MIN BIC           | 84 | 18 | 0 | 95 | 5 | 0 | 99 | 1 | 0 |

* This parameter is $\sigma$ for the CS model, $\rho$ for the AR1 model, and $\sigma^2_i$ for the GC model.
Table 4
Parameter estimates and fit statistics of CS, AR1, and GC models for data with CS error structure and sample size of 100.

<table>
<thead>
<tr>
<th></th>
<th>Effect size = .5</th>
<th></th>
<th></th>
<th>Effect size = .8</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ρ = .2</td>
<td>ρ = .5</td>
<td>ρ = .8</td>
<td></td>
<td>ρ = .2</td>
<td>ρ = .5</td>
<td>ρ = .8</td>
</tr>
<tr>
<td>σ^2_c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>79.72</td>
<td>99.72</td>
<td>78.75</td>
<td>CS</td>
<td>20.10</td>
<td>99.24</td>
<td>19.77</td>
</tr>
<tr>
<td>AR1</td>
<td>(5.49)</td>
<td>(6.09)</td>
<td>(5.76)</td>
<td>AR1</td>
<td>(1.52)</td>
<td>(13.05)</td>
<td>(1.53)</td>
</tr>
<tr>
<td>GC</td>
<td></td>
<td></td>
<td></td>
<td>GC</td>
<td>(1.95)</td>
<td>(2.78)</td>
<td>(1.94)</td>
</tr>
<tr>
<td>σ/σ^2_i</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.20)</td>
<td>(1.79)</td>
<td>(2.78)</td>
</tr>
<tr>
<td></td>
<td>(1.88)</td>
<td>(1.95)</td>
<td>(1.48)</td>
<td></td>
<td>(1.84)</td>
<td>(1.8)</td>
<td>(1.8)</td>
</tr>
<tr>
<td>σ_m</td>
<td>-1.01</td>
<td></td>
<td></td>
<td></td>
<td>-1.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.88)</td>
<td>(1.95)</td>
<td>(1.48)</td>
<td></td>
<td>(1.84)</td>
<td>(1.8)</td>
<td>(1.8)</td>
</tr>
<tr>
<td>β_0</td>
<td>24.86</td>
<td>24.86</td>
<td>4.95</td>
<td>β_0</td>
<td>24.86</td>
<td>24.86</td>
<td>4.95</td>
</tr>
<tr>
<td></td>
<td>(1.12)</td>
<td>(1.12)</td>
<td>(0.75)</td>
<td>(1.12)</td>
<td>(1.12)</td>
<td>(0.75)</td>
<td>(1.12)</td>
</tr>
<tr>
<td>SE(β_0)</td>
<td>1.00</td>
<td>1.00</td>
<td>0.83</td>
<td>SE(β_0)</td>
<td>1.00</td>
<td>1.00</td>
<td>0.83</td>
</tr>
<tr>
<td>β_1</td>
<td>-19.94</td>
<td>-19.94</td>
<td>4.98</td>
<td>β_1</td>
<td>-19.94</td>
<td>-19.94</td>
<td>4.98</td>
</tr>
<tr>
<td>(1.28)</td>
<td>(1.28)</td>
<td>(0.29)</td>
<td></td>
<td>(1.28)</td>
<td>(1.28)</td>
<td>(0.29)</td>
<td></td>
</tr>
<tr>
<td>SE(β_1)</td>
<td>1.26</td>
<td>1.41</td>
<td>0.29</td>
<td>SE(β_1)</td>
<td>1.26</td>
<td>1.41</td>
<td>0.29</td>
</tr>
<tr>
<td>β_2</td>
<td>-14.90</td>
<td>-14.90</td>
<td></td>
<td>β_2</td>
<td>-14.90</td>
<td>-14.90</td>
<td></td>
</tr>
<tr>
<td>(1.33)</td>
<td>(1.33)</td>
<td></td>
<td></td>
<td>(1.33)</td>
<td>(1.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE(β_2)</td>
<td>1.26</td>
<td>1.41</td>
<td></td>
<td>SE(β_2)</td>
<td>1.26</td>
<td>1.41</td>
<td></td>
</tr>
<tr>
<td>β_3</td>
<td>-9.93</td>
<td>-9.93</td>
<td></td>
<td>β_3</td>
<td>-9.93</td>
<td>-9.93</td>
<td></td>
</tr>
<tr>
<td>(1.25)</td>
<td>(1.25)</td>
<td></td>
<td></td>
<td>(1.25)</td>
<td>(1.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE(β_3)</td>
<td>1.26</td>
<td>1.38</td>
<td></td>
<td>SE(β_3)</td>
<td>1.26</td>
<td>1.38</td>
<td></td>
</tr>
<tr>
<td>β_4</td>
<td>-4.98</td>
<td>-4.98</td>
<td></td>
<td>β_4</td>
<td>-4.98</td>
<td>-4.98</td>
<td></td>
</tr>
<tr>
<td>(1.19)</td>
<td>(1.19)</td>
<td></td>
<td></td>
<td>(1.19)</td>
<td>(1.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SE(β_4)</td>
<td>1.26</td>
<td>1.26</td>
<td></td>
<td>SE(β_4)</td>
<td>1.26</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>3677.76</td>
<td>3692.11</td>
<td>3690.99</td>
<td>AIC</td>
<td>3677.76</td>
<td>3692.11</td>
<td>3690.99</td>
</tr>
<tr>
<td>BIC</td>
<td>3682.96</td>
<td>3697.33</td>
<td>3699.92</td>
<td>BIC</td>
<td>3682.96</td>
<td>3697.33</td>
<td>3699.92</td>
</tr>
<tr>
<td>MIN AIC</td>
<td>99</td>
<td>1</td>
<td>0</td>
<td>MIN AIC</td>
<td>99</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MIN BIC</td>
<td>99</td>
<td>1</td>
<td>0</td>
<td>MIN BIC</td>
<td>99</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

*This parameter is σ for the CS model, ρ for the AR1 model, and σ^2_i for the GC model.*
Table 5
Parameter estimates and fit statistics of CS, AR1, and GC models for data with CS error structure and sample size of 200.

<table>
<thead>
<tr>
<th></th>
<th>Effect size = .5</th>
<th></th>
<th>Effect size = .8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \rho = .2 )</td>
<td>( \rho = .5 )</td>
<td>( \rho = .8 )</td>
</tr>
<tr>
<td>( \sigma_c^2 )</td>
<td>CS AR1 GC</td>
<td>CS AR1 GC</td>
<td>CS AR1 GC</td>
</tr>
<tr>
<td>(3.74)</td>
<td>(4.91)</td>
<td>(4.00)</td>
<td>(0.91) ( \hat{\rho} )</td>
</tr>
<tr>
<td>( \sigma_c^2 )</td>
<td>-15.08</td>
<td>-15.08</td>
<td>-15.08</td>
</tr>
<tr>
<td>( \sigma_m^2 )</td>
<td>-15.08</td>
<td>-15.08</td>
<td>-15.08</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>25.12</td>
<td>25.12</td>
<td>5.08</td>
</tr>
<tr>
<td>SE (( \beta_0 ))</td>
<td>0.77</td>
<td>0.77</td>
<td>0.59</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>-19.94</td>
<td>-19.94</td>
<td>4.98</td>
</tr>
<tr>
<td>SE (( \beta_1 ))</td>
<td>(0.89)</td>
<td>(0.89)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>-15.08</td>
<td>-15.08</td>
<td>-15.08</td>
</tr>
<tr>
<td>SE (( \beta_2 ))</td>
<td>(0.81)</td>
<td>(0.81)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>-10.18</td>
<td>-10.18</td>
<td>-10.18</td>
</tr>
<tr>
<td>SE (( \beta_3 ))</td>
<td>(0.89)</td>
<td>(0.89)</td>
<td>(0.89)</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>-5.20</td>
<td>-5.20</td>
<td>-5.20</td>
</tr>
<tr>
<td>SE (( \beta_4 ))</td>
<td>(0.93)</td>
<td>(0.93)</td>
<td>(0.93)</td>
</tr>
<tr>
<td>AIC</td>
<td>7378.75</td>
<td>7407.97</td>
<td>7389.57</td>
</tr>
<tr>
<td>BIC</td>
<td>7385.35</td>
<td>7414.57</td>
<td>7400.98</td>
</tr>
<tr>
<td>MIN AIC</td>
<td>99</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MIN BIC</td>
<td>99</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

*This parameter is \( \sigma \) for the CS model, \( \rho \) for the AR1 model, and \( \sigma_c^2 \) for the GC model.
<table>
<thead>
<tr>
<th></th>
<th>$\sigma_c^2$</th>
<th>$\sigma^2/\sigma_c^2$</th>
<th>$\sigma_c$</th>
<th>$\sigma$</th>
<th>$\rho$</th>
<th>$\rho^2$</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
<th>$\beta_4$</th>
<th>$\alpha$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>88.67</td>
<td>99.92</td>
<td>80.52</td>
<td>71.04</td>
<td>102.83</td>
<td>53.52</td>
<td>34.38</td>
<td>96.73</td>
<td>19.85</td>
<td>36.05</td>
<td>39.70</td>
<td>32.02</td>
<td>27.29</td>
</tr>
<tr>
<td>GC</td>
<td>11.32</td>
<td>0.20</td>
<td>23.86</td>
<td>31.31</td>
<td>0.50</td>
<td>63.86</td>
<td>62.94</td>
<td>0.78</td>
<td>89.58</td>
<td>3.66</td>
<td>0.20</td>
<td>11.05</td>
<td>11.74</td>
</tr>
<tr>
<td></td>
<td>(10.59)</td>
<td>(0.12)</td>
<td>(21.79)</td>
<td>(14.50)</td>
<td>(0.10)</td>
<td>(34.00)</td>
<td>(22.22)</td>
<td>(0.06)</td>
<td>(33.27)</td>
<td>(3.79)</td>
<td>(0.11)</td>
<td>(7.35)</td>
<td>(6.07)</td>
</tr>
<tr>
<td>$\sigma_c^2$</td>
<td>-</td>
<td>-</td>
<td>3.35</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2.93)</td>
<td>(3.97)</td>
<td>(2.46)</td>
<td>(1.25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{in}$</td>
<td>-</td>
<td>-</td>
<td>-14.20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(6.88)</td>
<td>(9.74)</td>
<td>(7.38)</td>
<td>(2.58)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>24.79</td>
<td>24.79</td>
<td>4.85</td>
<td>25.39</td>
<td>25.39</td>
<td>5.09</td>
<td>25.18</td>
<td>25.18</td>
<td>4.84</td>
<td>25.09</td>
<td>25.09</td>
<td>5.09</td>
<td>24.69</td>
</tr>
<tr>
<td></td>
<td>(2.42)</td>
<td>(2.42)</td>
<td>(2.00)</td>
<td>(2.21)</td>
<td>(2.21)</td>
<td>(2.39)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(2.20)</td>
<td>(1.46)</td>
<td>(1.46)</td>
<td>(1.16)</td>
<td>(1.38)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>2.23</td>
<td>2.23</td>
<td>1.88</td>
<td>2.25</td>
<td>2.26</td>
<td>2.16</td>
<td>2.19</td>
<td>2.18</td>
<td>2.22</td>
<td>1.40</td>
<td>1.40</td>
<td>1.22</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>(3.05)</td>
<td>(3.05)</td>
<td>(0.75)</td>
<td>(3.38)</td>
<td>(3.38)</td>
<td>(0.89)</td>
<td>(2.25)</td>
<td>(2.25)</td>
<td>(0.58)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(0.46)</td>
<td>(1.84)</td>
</tr>
<tr>
<td></td>
<td>(3.05)</td>
<td>(3.05)</td>
<td>(0.75)</td>
<td>(3.38)</td>
<td>(3.38)</td>
<td>(0.89)</td>
<td>(2.25)</td>
<td>(2.25)</td>
<td>(0.58)</td>
<td>(2.05)</td>
<td>(2.05)</td>
<td>(0.46)</td>
<td>(1.84)</td>
</tr>
<tr>
<td></td>
<td>(3.09)</td>
<td>(3.09)</td>
<td>(3.28)</td>
<td>(2.17)</td>
<td>(2.17)</td>
<td>-</td>
<td>(1.81)</td>
<td>(1.81)</td>
<td></td>
<td>(2.06)</td>
<td>(2.06)</td>
<td>(2.06)</td>
<td>(2.22)</td>
</tr>
<tr>
<td></td>
<td>(3.65)</td>
<td>(3.65)</td>
<td>(3.65)</td>
<td>(2.68)</td>
<td>(2.68)</td>
<td>(1.83)</td>
<td>(1.83)</td>
<td>(1.83)</td>
<td>(1.93)</td>
<td>(1.93)</td>
<td>(1.76)</td>
<td>(1.76)</td>
<td>(1.24)</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>2.97</td>
<td>3.14</td>
<td>0.75</td>
<td>2.65</td>
<td>2.96</td>
<td>-</td>
<td>1.85</td>
<td>2.18</td>
<td>-</td>
<td>1.89</td>
<td>1.97</td>
<td>-</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td>(2.97)</td>
<td>(3.09)</td>
<td>(3.28)</td>
<td>(2.17)</td>
<td>(2.17)</td>
<td>(1.83)</td>
<td>(1.83)</td>
<td>(1.83)</td>
<td>(1.93)</td>
<td>(1.93)</td>
<td>(1.76)</td>
<td>(1.76)</td>
<td>(1.24)</td>
</tr>
<tr>
<td></td>
<td>(2.87)</td>
<td>(2.87)</td>
<td>(2.87)</td>
<td>(2.07)</td>
<td>(2.07)</td>
<td>(1.40)</td>
<td>(1.40)</td>
<td>(1.40)</td>
<td>(1.90)</td>
<td>(1.90)</td>
<td>(1.43)</td>
<td>(1.43)</td>
<td>(0.80)</td>
</tr>
<tr>
<td>$\beta_7$</td>
<td>2.97</td>
<td>2.81</td>
<td>2.81</td>
<td>2.65</td>
<td>2.96</td>
<td>-</td>
<td>1.85</td>
<td>1.41</td>
<td>-</td>
<td>1.89</td>
<td>1.77</td>
<td>-</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td>(2.97)</td>
<td>(2.97)</td>
<td>(2.97)</td>
<td>(2.07)</td>
<td>(2.07)</td>
<td>(1.40)</td>
<td>(1.40)</td>
<td>(1.40)</td>
<td>(1.90)</td>
<td>(1.90)</td>
<td>(1.43)</td>
<td>(1.43)</td>
<td>(0.80)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>721.82</td>
<td>720.61</td>
<td>739.82</td>
<td>713.49</td>
<td>704.41</td>
<td>724.85</td>
<td>666.21</td>
<td>645.28</td>
<td>665.24</td>
<td>634.74</td>
<td>633.09</td>
<td>649.45</td>
<td>622.30</td>
</tr>
<tr>
<td></td>
<td>723.81</td>
<td>722.60</td>
<td>743.40</td>
<td>715.48</td>
<td>706.40</td>
<td>728.80</td>
<td>668.20</td>
<td>647.27</td>
<td>669.23</td>
<td>636.73</td>
<td>635.08</td>
<td>653.16</td>
<td>624.29</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>40</td>
<td>62</td>
<td>0</td>
<td>9</td>
<td>91</td>
<td>0</td>
<td>1</td>
<td>99</td>
<td>0</td>
<td>25</td>
<td>76</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>40</td>
<td>62</td>
<td>0</td>
<td>9</td>
<td>91</td>
<td>0</td>
<td>1</td>
<td>99</td>
<td>0</td>
<td>25</td>
<td>76</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 6

Parameter estimates and fit statistics of CS, AR1, and GC models for data with AR1 error structure and sample size of 20.

* This parameter is $\sigma$ for the CS model, $\rho$ for the AR1 model, and $\sigma_c^2$ for the GC model.
Table 7

Parameter estimates and fit statistics of CS, AR1, and GC models for data with AR1 error structure and sample size of 100.

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>Effect size = .5</th>
<th>Effect size = .8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS</td>
<td>AR1</td>
</tr>
<tr>
<td>$\sigma^2_c$</td>
<td>90.60</td>
<td>100.04</td>
</tr>
<tr>
<td>$\sigma^2_{\beta}/\sigma^2_c$</td>
<td>(6.54)</td>
<td>(6.26)</td>
</tr>
<tr>
<td>$\sigma^2_\sigma$</td>
<td>9.43</td>
<td>0.19</td>
</tr>
<tr>
<td>$\sigma^2_{\lambda}$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>24.77</td>
<td>24.77</td>
</tr>
<tr>
<td>SE ($\beta_0$)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>SE ($\beta_1$)</td>
<td>1.54</td>
<td>1.54</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-14.63</td>
<td>-14.63</td>
</tr>
<tr>
<td>SE ($\beta_2$)</td>
<td>1.35</td>
<td>1.35</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>-9.75</td>
<td>-9.75</td>
</tr>
<tr>
<td>SE ($\beta_3$)</td>
<td>1.46</td>
<td>1.46</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-4.85</td>
<td>-4.85</td>
</tr>
<tr>
<td>SE ($\beta_4$)</td>
<td>1.29</td>
<td>1.29</td>
</tr>
<tr>
<td>AIC</td>
<td>3702.01</td>
<td>3694.88</td>
</tr>
<tr>
<td>BIC</td>
<td>3707.23</td>
<td>3700.08</td>
</tr>
<tr>
<td>MIN AIC</td>
<td>6</td>
<td>94</td>
</tr>
<tr>
<td>MIN BIC</td>
<td>6</td>
<td>94</td>
</tr>
</tbody>
</table>

* This parameter is $\sigma$ for the CS model, $\rho$ for the AR1 model, and $\sigma^2_c$ for the GC model.
**Table 8**

Parameter estimates and fit statistics of CS, AR1, and GC models for data with AR1 error structure and sample size of 200.

<table>
<thead>
<tr>
<th>Effect size = .5</th>
<th>Effect size = .8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ρ = .2</strong></td>
<td><strong>ρ = .5</strong></td>
</tr>
<tr>
<td><strong>σ_ε^2</strong></td>
<td><strong>σ_ε^2</strong></td>
</tr>
<tr>
<td>CS</td>
<td>AR1</td>
</tr>
<tr>
<td>90.63</td>
<td>100.01</td>
</tr>
<tr>
<td>(4.66)</td>
<td>(4.33)</td>
</tr>
<tr>
<td><strong>σ/ρ/σ_ε^2</strong></td>
<td><strong>σ/ρ/σ_ε^2</strong></td>
</tr>
<tr>
<td>(9.37)</td>
<td>0.20</td>
</tr>
<tr>
<td>(2.99)</td>
<td>(0.04)</td>
</tr>
<tr>
<td><strong>σ_y^2</strong></td>
<td><strong>σ_y^2</strong></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(1.25)</td>
<td>(1.25)</td>
</tr>
<tr>
<td><strong>σ_μ</strong></td>
<td><strong>σ_μ</strong></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(2.98)</td>
<td>(3.12)</td>
</tr>
<tr>
<td><strong>β_0</strong></td>
<td><strong>β_0</strong></td>
</tr>
<tr>
<td>24.96</td>
<td>24.96</td>
</tr>
<tr>
<td>(0.70)</td>
<td>(0.70)</td>
</tr>
<tr>
<td><strong>SE (β_0)</strong></td>
<td><strong>SE (β_0)</strong></td>
</tr>
<tr>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>(0.82)</td>
<td>(0.82)</td>
</tr>
<tr>
<td><strong>β_1</strong></td>
<td><strong>β_1</strong></td>
</tr>
<tr>
<td>-19.98</td>
<td>-19.98</td>
</tr>
<tr>
<td>(0.92)</td>
<td>(0.92)</td>
</tr>
<tr>
<td><strong>SE (β_1)</strong></td>
<td><strong>SE (β_1)</strong></td>
</tr>
<tr>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>(1.02)</td>
<td>(1.02)</td>
</tr>
<tr>
<td><strong>β_2</strong></td>
<td><strong>β_2</strong></td>
</tr>
<tr>
<td>-14.79</td>
<td>-14.79</td>
</tr>
<tr>
<td>(1.02)</td>
<td>(1.02)</td>
</tr>
<tr>
<td><strong>SE (β_2)</strong></td>
<td><strong>SE (β_2)</strong></td>
</tr>
<tr>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>(0.97)</td>
<td>(0.97)</td>
</tr>
<tr>
<td><strong>β_3</strong></td>
<td><strong>β_3</strong></td>
</tr>
<tr>
<td>(0.97)</td>
<td>(0.97)</td>
</tr>
<tr>
<td><strong>SE (β_3)</strong></td>
<td><strong>SE (β_3)</strong></td>
</tr>
<tr>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>(0.77)</td>
<td>(0.77)</td>
</tr>
<tr>
<td><strong>β_4</strong></td>
<td><strong>β_4</strong></td>
</tr>
<tr>
<td>-4.91</td>
<td>-4.91</td>
</tr>
<tr>
<td>(0.77)</td>
<td>(0.77)</td>
</tr>
<tr>
<td><strong>SE (β_4)</strong></td>
<td><strong>SE (β_4)</strong></td>
</tr>
<tr>
<td>0.95</td>
<td>0.89</td>
</tr>
<tr>
<td><strong>AIC</strong></td>
<td><strong>AIC</strong></td>
</tr>
<tr>
<td>7419.33</td>
<td>7401.51</td>
</tr>
<tr>
<td>7305.72</td>
<td>7207.24</td>
</tr>
<tr>
<td>6847.15</td>
<td>6623.51</td>
</tr>
<tr>
<td>6480.88</td>
<td>6463.51</td>
</tr>
<tr>
<td>6365.16</td>
<td>6268.28</td>
</tr>
<tr>
<td>5915.59</td>
<td>5692.26</td>
</tr>
<tr>
<td><strong>BIC</strong></td>
<td><strong>BIC</strong></td>
</tr>
<tr>
<td>7425.93</td>
<td>7408.10</td>
</tr>
<tr>
<td>7312.32</td>
<td>7213.84</td>
</tr>
<tr>
<td>6853.74</td>
<td>6630.11</td>
</tr>
<tr>
<td>6487.48</td>
<td>6470.10</td>
</tr>
<tr>
<td>6371.75</td>
<td>6274.88</td>
</tr>
<tr>
<td>5922.18</td>
<td>5698.86</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MIN AIC</th>
<th><strong>MIN AIC</strong></th>
<th><strong>MIN AIC</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MIN BIC</th>
<th><strong>MIN BIC</strong></th>
<th><strong>MIN BIC</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>

*This parameter is σ for the CS model, ρ for the AR1 model, and σ_ε^2 for the GC model.*
Table 9
Parameter estimates and fit statistics of CS, AR1, and GC models for data with random-effects error structure, sample size of 20, and effect size of .5.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>r = .1</th>
<th>r = .5</th>
<th>r = .8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>σc^2</strong></td>
<td>CS</td>
<td>AR1</td>
<td>GC</td>
</tr>
<tr>
<td></td>
<td>83.21</td>
<td>108.16</td>
<td>77.98</td>
</tr>
<tr>
<td></td>
<td>(11.49)</td>
<td>(15.22)</td>
<td>(12.33)</td>
</tr>
<tr>
<td><strong>σ/ρ/σ^2</strong></td>
<td>CS</td>
<td>AR1</td>
<td>GC</td>
</tr>
<tr>
<td></td>
<td>24.91</td>
<td>0.25</td>
<td>20.35</td>
</tr>
<tr>
<td></td>
<td>(13.46)</td>
<td>(0.13)</td>
<td>(21.03)</td>
</tr>
<tr>
<td><strong>σs^2</strong></td>
<td>-</td>
<td>-</td>
<td>2.10</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>(2.40)</td>
</tr>
<tr>
<td><strong>σm</strong></td>
<td>-</td>
<td>-</td>
<td>-0.54</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>(5.97)</td>
</tr>
<tr>
<td><strong>β0</strong></td>
<td>25.61</td>
<td>25.61</td>
<td>5.25</td>
</tr>
<tr>
<td></td>
<td>(2.38)</td>
<td>(2.38)</td>
<td>(1.95)</td>
</tr>
<tr>
<td><strong>SE (β0)</strong></td>
<td>2.32</td>
<td>2.32</td>
<td>1.81</td>
</tr>
<tr>
<td><strong>β1</strong></td>
<td>-20.11</td>
<td>-20.11</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td>(3.22)</td>
<td>(3.22)</td>
<td>(0.69)</td>
</tr>
<tr>
<td><strong>SE (β1)</strong></td>
<td>2.88</td>
<td>3.26</td>
<td>0.70</td>
</tr>
<tr>
<td><strong>β2</strong></td>
<td>-15.60</td>
<td>-15.60</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2.88)</td>
<td>(2.88)</td>
<td>-</td>
</tr>
<tr>
<td><strong>SE (β2)</strong></td>
<td>2.88</td>
<td>3.23</td>
<td>-</td>
</tr>
<tr>
<td><strong>β3</strong></td>
<td>-10.28</td>
<td>-10.28</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2.94)</td>
<td>(2.94)</td>
<td>-</td>
</tr>
<tr>
<td><strong>SE (β3)</strong></td>
<td>2.88</td>
<td>3.15</td>
<td>-</td>
</tr>
<tr>
<td><strong>β4</strong></td>
<td>-5.82</td>
<td>-5.82</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2.94)</td>
<td>(2.94)</td>
<td>-</td>
</tr>
<tr>
<td><strong>SE (β4)</strong></td>
<td>2.88</td>
<td>2.83</td>
<td>-</td>
</tr>
</tbody>
</table>

AIC | 724.33 | 726.27 | 741.52 | 688.22 | 670.71 | 675.44 | 736.78 | 737.88 | 750.12 |
| BIC | 726.32 | 728.26 | 744.96 | 690.21 | 672.70 | 679.43 | 738.70 | 739.86 | 753.84 |
| MIN AIC | 73 | 27 | 1 | 62 | 37 | 1 | 63 | 36 | 1 |
| MIN BIC | 72 | 28 | 1 | 63 | 37 | 0 | 63 | 37 | 0 |

*This parameter is σ for the CS model, ρ for the AR1 model, and σ^2 for the GC model.
<table>
<thead>
<tr>
<th></th>
<th>( r = .2 )</th>
<th>( r = .5 )</th>
<th>( r = .8 )</th>
<th>( r = .25 )</th>
<th>( r = .5 )</th>
<th>( r = .8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS</td>
<td>AR1</td>
<td>GC</td>
<td>CS</td>
<td>AR1</td>
<td>GC</td>
</tr>
<tr>
<td>( \sigma^2_c )</td>
<td>84.21</td>
<td>111.65</td>
<td>79.21</td>
<td>61.99</td>
<td>131.86</td>
<td>49.90</td>
</tr>
<tr>
<td>( \sigma^2_c )</td>
<td>27.35</td>
<td>0.28</td>
<td>20.74</td>
<td>67.95</td>
<td>0.60</td>
<td>50.57</td>
</tr>
<tr>
<td>( \sigma^2_c )</td>
<td>(6.52)</td>
<td>(0.06)</td>
<td>(9.17)</td>
<td>(11.61)</td>
<td>(0.05)</td>
<td>(11.94)</td>
</tr>
<tr>
<td>( \sigma^2_c )</td>
<td>-</td>
<td>-</td>
<td>1.99</td>
<td>-</td>
<td>-</td>
<td>4.83</td>
</tr>
<tr>
<td>( \sigma^2_c )</td>
<td>(1.44)</td>
<td>(1.38)</td>
<td>(1.43)</td>
<td>(1.43)</td>
<td>(1.43)</td>
<td>(1.43)</td>
</tr>
<tr>
<td>( \sigma_{m} )</td>
<td>-</td>
<td>-</td>
<td>-0.09</td>
<td>-</td>
<td>-</td>
<td>0.12</td>
</tr>
<tr>
<td>( \sigma_{m} )</td>
<td>(2.70)</td>
<td>(2.73)</td>
<td>(2.73)</td>
<td>(2.73)</td>
<td>(2.73)</td>
<td>(2.73)</td>
</tr>
<tr>
<td>( \beta_{0} )</td>
<td>24.91</td>
<td>24.91</td>
<td>5.06</td>
<td>25.05</td>
<td>25.05</td>
<td>5.11</td>
</tr>
<tr>
<td>( \beta_{0} )</td>
<td>(1.00)</td>
<td>(1.00)</td>
<td>(0.96)</td>
<td>(1.49)</td>
<td>(1.49)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>( \beta_{1} )</td>
<td>-19.83</td>
<td>-19.83</td>
<td>4.97</td>
<td>-19.96</td>
<td>-19.96</td>
<td>4.99</td>
</tr>
<tr>
<td>( \beta_{1} )</td>
<td>(1.37)</td>
<td>(1.37)</td>
<td>(0.30)</td>
<td>(1.44)</td>
<td>(1.44)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>( \beta_{1} )</td>
<td>(2.70)</td>
<td>(2.73)</td>
<td>(2.73)</td>
<td>(2.73)</td>
<td>(2.73)</td>
<td>(2.73)</td>
</tr>
<tr>
<td>( \beta_{2} )</td>
<td>1.30</td>
<td>1.49</td>
<td>0.31</td>
<td>1.11</td>
<td>1.51</td>
<td>0.31</td>
</tr>
<tr>
<td>( \beta_{2} )</td>
<td>(1.34)</td>
<td>(1.34)</td>
<td>-</td>
<td>(1.30)</td>
<td>(1.30)</td>
<td>-</td>
</tr>
<tr>
<td>( \beta_{3} )</td>
<td>-9.97</td>
<td>-9.97</td>
<td>-</td>
<td>-9.94</td>
<td>-9.94</td>
<td>-</td>
</tr>
<tr>
<td>( \beta_{3} )</td>
<td>(1.20)</td>
<td>(1.20)</td>
<td>-</td>
<td>(1.06)</td>
<td>(1.06)</td>
<td>-</td>
</tr>
<tr>
<td>( \beta_{4} )</td>
<td>-4.84</td>
<td>-4.84</td>
<td>-</td>
<td>-4.93</td>
<td>-4.93</td>
<td>-</td>
</tr>
<tr>
<td>( \beta_{4} )</td>
<td>(1.23)</td>
<td>(1.23)</td>
<td>-</td>
<td>(0.95)</td>
<td>(0.95)</td>
<td>-</td>
</tr>
<tr>
<td>( \beta_{4} )</td>
<td>(1.30)</td>
<td>(1.27)</td>
<td>-</td>
<td>(1.11)</td>
<td>(1.03)</td>
<td>-</td>
</tr>
<tr>
<td>( \beta_{4} )</td>
<td>(1.68)</td>
<td>(1.39)</td>
<td>-</td>
<td>(1.85)</td>
<td>(1.40)</td>
<td>-</td>
</tr>
<tr>
<td>AIC</td>
<td>3719.61</td>
<td>3731.63</td>
<td>3728.05</td>
<td>3657.80</td>
<td>3669.44</td>
<td>3638.51</td>
</tr>
<tr>
<td>BIC</td>
<td>3724.82</td>
<td>3736.83</td>
<td>3738.20</td>
<td>3663.01</td>
<td>3674.65</td>
<td>3648.93</td>
</tr>
<tr>
<td>MIN AIC</td>
<td>83</td>
<td>6</td>
<td>11</td>
<td>6</td>
<td>0</td>
<td>94</td>
</tr>
<tr>
<td>MIN BIC</td>
<td>90</td>
<td>6</td>
<td>4</td>
<td>17</td>
<td>1</td>
<td>82</td>
</tr>
</tbody>
</table>

* This parameter is \( \sigma \) for the CS model, \( \rho \) for the AR1 model, and \( \sigma_c^2 \) for the GC model.
Figure 1. Percentage correct of AIC and BIC for data with CS error structure, sample size=20
Figure 2. Percentage correct of AIC and BIC for data with AR1 error structure, sample size=20
Figure 3. Percentage correct of AIC and BIC for data with random-effects error structure, sample size=20.
Figure 4. Percentage correct of AIC and BIC for data with random-effects error structure, sample size=100.
Appendix A

Simulation of data with compound symmetry error structure

Consider the simulation of 100 sets of data with compound symmetry error structure, medium effect size (.5), large intraclass correlation size (.8), and large sample size (200). The covariance matrix used for simulation is:

\[
\begin{bmatrix}
100 & 80 & 80 & 80 & 80 \\
80 & 100 & 80 & 80 & 80 \\
80 & 80 & 100 & 80 & 80 \\
80 & 80 & 80 & 100 & 80 \\
80 & 80 & 80 & 80 & 100 \\
\end{bmatrix}
\]

Generating seeds

Since there are five measurements points, we need to generate five columns of errors for each dataset. Each column requires a seed in the simulation. Therefore, each dataset requires 5 seeds, and 100 datasets require 500 seeds.

The seeds are generated using the array statement in SAS:

```sas
data;
array seed(i) seed1-seed500;
seedst = 49148716;
do i = 1 to 500;
   seed = ranuni(seedst);
   seed = seed * 1000000;
   seed = int(seed);
   file " C:\thesis\simulation\seed1.dat";
   put seed;
end;
run;
```

This program yields a 500×1 column of seeds, which is saved as a data file. The seed for the seeds (“seedst”) is changed every time a new round of simulation begins.
Simulating the data

Once we have the covariance matrix of errors and seeds, we can simulate data using SAS macro programs. Macro programs are iterative loops that are designed for carrying out the same commands repeatedly. They begin with %macro and end with %mend. Between the two statements we can put in macro-text which contains the commands that are to be carried out repeatedly. The macro program used in this simulation is:

```
%macro example;
  %do i = 1 %to 100;
  data one&i;
  infile "C:\thesis\simulation\seed1.dat";
  input seed1 #2 seed2 #3 seed3 #4 seed4 #5 seed5;
  iter=&i;
  if _n_ lt iter or _n_ gt iter then delete;
  file print;
  put iter seed1 seed2 seed3 seed4 seed5;
  run;

  data two&i;
  set one&i;
  do k=1 to 200;
    y1=1*rannor(seed1);
    y2=1*rannor(seed2);
    y3=1*rannor(seed3);
    y4=1*rannor(seed4);
    y5=1*rannor(seed5);
    output;
  keep y1 y2 y3 y4 y5;
  end;
  run;

(Continue on next page)```
proc iml;
start;
use two&i;
read all into y&i;
print "Raw Data for Iteration&i", y&i;

n=nrow(y&i);                          ***compute correlation matrix***;
sum=y&i[+,

ymeans=sum/n;
print "mean y for iteration&i", ymeans;

ypy=y&i*y&i-sum*sum/n;
s=diag(1/sqrt(vecdiag(ypy)));
corr=s*ypy*s;
print "Correlation Matrix for Iteration&i", corr;

cov={ 100 80 80 80 80,
     80 100 80 80 80,
     80 80 100 80 80,
     80 80 80 100 80,
     80 80 80 80 100};                 ***input covariance matrix***;
p=root(cov);                           ***calculate Cholesky***;
print "Cholesky", p;

ytrans=y&i*p;
print "Transformed Data for Iteration&i", ytrans;

n=nrow(ytrans);                        ***calculate covariance matrix on ytrans***;
sum=ytrans[+,

ytrmeans=sum/n;
print "Means ytrans for iteration&i", ytrmeans;
ccp=ytrans*ytrans-sum*sum/n;
covtrans=ccp/(n-1);
print "Covariance Matrix On Ytrans for iteration&i", covtrans;

means={5 10 15 20 25};                 ***add means***;
addmeans=shape(means, n, 5);
print addmeans;

ynew=ytrans+addmeans;
varname="c1"-"c5";
print "Transformed Data ynew for Iteration&i", ynew;

n=nrow(ynew);
sum=ynew[+,

ynewmns=sum/n;
print "Simulation Means for Iteration&i", means;
print "Means ynew for Iteration&i", ynewmns;
create three&i var {c1 c2 c3 c4 c5};
append from ynew;
finish;
run;                                                   (Continue on next page)
data four &i;
set three &i;
iter=&i;
file " C:\thesis\simulation\example.dat. " mod;
put iter c1 c2 c3 c4 c5;
run;

%end;

%mend example;

%example;
Appendix B

Estimating parameters in SAS PROC MIXED

*Estimating \( G \) and \( R \) in the mixed model*

SAS PROC MIXED estimates the parameters of the model using a generalized least squares approach. Estimating the fixed effects parameters requires minimizing:

\[
(y - X\beta)'V^{-1}(y - x\beta)
\]

where

\[
V = ZGZ' + R
\]

Since \( V^{-1} \) is unknown, reasonable estimates must be determined. This requires knowing \( G \) and \( R \) where \( G \) is the covariance matrix of \( \gamma \) and \( R \) is the covariance matrix of \( \epsilon \).

SAS PROC MIXED uses two different likelihood-based approaches under the assumption that \( \gamma \) and \( \epsilon \) are normally distributed (Harville, 1977, Laird and Ware, 1982; Jennrich & Schluchter, 1986). An objective function is constructed and maximized over all unknown parameters. The log-likelihood functions for maximum-likelihood (ML) and restricted maximum-likelihood (REML) are, respectively:

\[
ML : l(G, R) = -\frac{1}{2} \log |V| - \frac{1}{2} r^{-1} r - \frac{n}{2} \log(2\pi)
\]

\[
REML : l_r(G, R) = -\frac{1}{2} \log |V| - \frac{1}{2} \log |X'V^{-1}X| - \frac{1}{2} r'V^{-1}r - \frac{n-p}{2} \log(2\pi)
\]

where \( r = y - X(X'V^{-1}X)'X'V^{-1}y \) and \( p \) is the rank of \( X \). SAS PROC MIXED minimizes -2 times these functions.
Estimating $\beta$ and $\gamma$ in the mixed model

Once $G$ and $R$ have been obtained, we can obtain estimates of $\beta$ and $\gamma$ using the Henderson mixed model equations (Henderson, 1984):

$$
\begin{bmatrix}
X' \hat{R}^{-1} X & X' \hat{R}^{-1} Z \\
Z' \hat{R}^{-1} X & Z' \hat{R}^{-1} Z + \hat{G}^{-1}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta} \\
\hat{\gamma}
\end{bmatrix}
= 
\begin{bmatrix}
X' \hat{R}^{-1} y \\
Z' \hat{R}^{-1} y
\end{bmatrix}
$$

The solutions to the equations are:

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

and

$$
\hat{\gamma} = \hat{G}Z' \hat{V}^{-1} (y - X\hat{\beta})
$$

(Robinson, 1991) where ‘-‘ represents the generalized inverse. These estimates are equivalent to the empirical Bayes estimates (Bryk & Raudenbush, 1992; Carlin & Lewis, 1996; Laird & Ware, 1982).

Inferences for the fixed effects

The approximate covariance matrix of the parameters is given by:

$$
\hat{C} = (X' \hat{V}^{-1} X)^{-1}
$$

If $L$ represents an estimable contrast used to identify a fixed effect, then the t-statistic to test the fixed effect is given by:

$$
t = \frac{L\hat{\beta}}{\sqrt{LCL'}}
$$

If $\theta$ represents the vector of parameter estimates associated with $C$, then the t-statistic has Satterthwaite degrees of freedom:

$$
\nu = \frac{2(LCL')^2}{g'Ag}
$$

where $g$ is the gradient of $|C|$ evaluated at the parameters estimates for $V$ and $A$ is the
asymptotic covariance matrix of the parameter estimates, $\hat{\theta}$, obtained from the 2nd derivative of the likelihood equations.

**Model fit indices**

The model fit indices used to compare the models are the Akaike Information Criterion (AIC; Akaike, 1974) and the Bayesian Information Criterion (BIC; Schwarz, 1978) which are defined as:

\[
\text{AIC} = -2l + 2d
\]
\[
\text{BIC} = -2l + d \log(n)
\]

where $l$ is the log-likelihood, $n$ is the number of observations, and $d$ is the dimension of the model.