LONGITUDINAL DATA ANALYSIS USING MULTILEVEL LINEAR MODELING (MLM): FITTING AN OPTIMAL VARIANCE-COVARIANCE STRUCTURE

A Dissertation

by

YUAN-HSUAN LEE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2010

Major Subject: Educational Psychology

Longitudinal Data Analysis Using Multilevel Linear Modeling (MLM): Fitting an Optimal Variance-Covariance Structure

LONGITUDINAL DATA ANALYSIS USING MULTILEVEL LINEAR MODELING (MLM): FITTING AN OPTIMAL VARIANCE-COVARIANCE STRUCTURE

A Dissertation

by

YUAN-HSUAN LEE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Victor L. Willson
	Oi-Man Kwok
Committee Members,	Robert Hall
	F. Michael Speed
Head of Department,	Victor L. Willson

August 2010

Major Subject: Educational Psychology

ABSTRACT

Longitudinal Data Analysis Using Multilevel Linear Modeling (MLM): Fitting an Optimal Variance-Covariance Structure. (August 2010)

Yuan-Hsuan Lee, B.A., National Tsing Hua University

Co-Chairs of Advisory Committee: Dr. Victor L. Willson
Dr. Oi-Man Kwok

This dissertation focuses on issues related to fitting an optimal variance-covariance structure in multilevel linear modeling framework with two Monte Carlo simulation studies.

In the first study, the author evaluated the performance of common fit statistics such as Likelihood Ratio Test (LRT), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) and a new proposed method, standardized root mean square residual (SRMR), for selecting the correct within-subject covariance structure. Results from the simulated data suggested SRMR had the best performance in selecting the optimal covariance structure. A pharmaceutical example was also used to evaluate the performance of these fit statistics empirically. The LRT failed to decide which is a better model because LRT can only be used for nested models. SRMR, on the other hand, had congruent result as AIC and BIC and chose ARMA(1,1) as the optimal variance-covariance structure.

In the second study, the author adopted a first-order autoregressive structure as the true within-subject V-C structure with variability in the intercept and slope

(estimating τ_{00} and τ_{11} only) and investigated the consequence of misspecifying different levels/types of the V-C matrices simultaneously on the estimation and test of significance for the growth/fixed-effect and random-effect parameters, considering the size of the autoregressive parameter, magnitude of the fixed effect parameters, number of cases, and number of waves. The result of the simulation study showed that the commonly-used identity within-subject structure with unstructured between-subject matrix performed equally well as the true model in the evaluation of the criterion variables. On the other hand, other misspecified conditions, such as Under G & Over R conditions and Generally misspecified G & R conditions had biased standard error estimates for the fixed effect and lead to inflated Type I error rate or lowered statistical power.

The two studies bridged the gap between the theory and practical application in the current literature. More research can be done to test the effectiveness of proposed SRMR in searching for the optimal V-C structure under different conditions and evaluate the impact of different types/levels of misspecification with various specifications of the within- and between- level V-C structures simultaneously.

DEDICATION

This dissertation is dedicated to Michelle and Katherine, my two beautiful and precious gifts from God.

ACKNOWLEDGEMENTS

I am proud of being a member of Research, Measurement, and Statistic Program at Texas A&M University. The training in the program equipped me with the professional knowledge of a quantitative methodologist. Particularly, I am grateful for my co-chairs, Dr. Oi-Man Kwok and Dr. Victor Willson, for their guidance through my doctoral study. Dr. Kwok is a wonderful adviser, who leads me to the world of quantitative and statistical research step by step and is always willing to share his experiences in academia. Dr. Willson has my deepest respect. I admire his insight and passion in quantitative research. More often than not, I am surprised by his brilliant and enlightening comments in research. He always reminds me of linking the simulations to real situations. I am also thankful for my committee members, Dr. Robert Hall and Dr. Mike Speed. They are very supportive in my doctoral study. I learned advanced mixed effects modeling from Dr. Speed, who helped me a lot in the data generation of study two. I also would like to thank Dr. Bruce Thompson and Dr. Myeongsun Yoon. The beginning statiscs courses I took from Dr. Thompson are sure to have a great impact on me for a life time. The measurement and IRT knowledge I learned from Dr. Yoon patched my knowledge in RMS. Of all the professionals, I also like to thank Dr. Hersh Waxman. It is my greatest honor to work for Dr. Waxman. Under his supervision, I learned how to work on research and complete a project as a team. He gave me great flexibility and opportunities to apply what I have learned in real research. Most important of all, I am impressed that he is such a great scholar who shows respect to his colleagues and is surprisingly humble.

I would like to thank my parents who brought me up and strived to provide me with a good education. I was pregnant with my second child when working on my dissertation. Without their help with the baby and everyday chores, I would never have finished my dissertation and passed my final defense successfully. Thank you, dad and mom! I love you!

Finally, I need to thank Jerry Jiun-Yu Wu, my dear husband. He is the one who makes me so productive such that I have two beautiful daughters and an earned doctorate.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Quantitative researchers have made extensive use of multilevel linear modeling (MLM) technique to analyze repeated measurement data. MLM has several advantages over traditional methods Univariate Analysis of Variance (UANOVA) or Multivariate Analysis of Variance (MANOVA) in analyzing repeated measures or longitudinal studies, such as allowing unbalanced data or missing data points, not requiring observations taken equidistantly, capturing the average growth trend of the outcome variable over time, and flexibly modeling the variance-covariance (V-C) structure (Diggle, 1988; Ferron, Dailey, & Yi, 2002; Laird & Ware, 1982; Luke, 2004; Wolfinger, 1993). The focus of this dissertation focuses on the last advantage, that V-C structure in MLM can be flexibly specified. Though MLM allow flexibly modeling of the V-C structure, the default V-C matrix in most of the commonly used statistical packages is still the identity structure, which assumes equal variance of each observation and no covariance between any pair of repeated measures. Careless or inexperienced researchers in performing MLM studies may just leave the choice of V-C structure to the computer software. Misspecification in the covariance structure in MLM or leaving the specification of V-C structure to the computer software generally causes no harm to the estimation of fixed effect /growth parameters (Ferron et al., 2002; Kwok, West, & Green, 2007; Murphy & Pituch, 2009); however, the corresponding estimates for the

---

This dissertation follows the style of *Psychological Methods.*

standard errors of the fixed effect/ growth parameters are biased, which will in turn lead to erroneous statistical inference of the hypothesis testing results (Davis, 2002; Diggle, Heagerty, Liang & Zeger, 2002; Kwok et al., 2007; Singer & Willett, 2003).

To motivate the use of MLM methodology, the second section of this dissertation reviews issues related to MLM: its advantages over traditional methods, MLM as a mixed effect model, effects of misspecifying the within-subject V-C Structure, types of misspecification, and existing methods in selecting an optimal V-C structure. Through the review of MLM, two research issues emerge and draw our attention. First, there is a lack of an optimal model selection method, and second, the effect of different types/levels of misspecification has not been investigated. As mentioned previously, misspecification of the within-subject V-C structure, although it may not have a negative influence on the fixed effect estimates, leads to biased estimation in the standard errors for the fixed effect. In other words, the statistical inferences drawn from the combination of unbiased fixed estimates and biased standard errors of the fixed effects will still be erroneous. This issue requires development of effective methods in selecting an optimal within-subject V-C structure, which is the third section of the dissertation. In section 3, the performance of LRT, AIC, BIC, and SRMR on searching for the correct covariance structure will be evaluated. The impact of several design factors, such as number of cases, number of repeated measurements, magnitude of the average growth model, and magnitude of the between-subject covariance matrix on the performance of these search methods, are also considered in the analysis.

A second consideration is that misspecification in both the between-subject (**G**-side) and the within-subject (**R**-side) covariance structure has rarely been examined simultaneously compared to misspecification of the within-subject V-C structure, which has been researched extensively (Ferron et al., 2002; Kwok et al., 2007; Murphy & Pituch, 2009; Vallejo, Ato, & Valdés, 2008). In a simple linear growth curve model, the G matrix is comprised of random effects, including variance of intercept (τ_{00}), variance of slope τ_{11} , and covariance between intercept and slope (τ_{01}), capturing the deviation of growth parameters from the population means for intercept and slope. Therefore, when fitting a mixed effect model for repeated measurements, researchers need to specify both the **G**- and **R**- side V-C structure for the data. In the best scenario, researchers will specify the V-C structures for the two sides correctly. In other cases, researchers may over-, under-, or generally misspecify the V-C structures. The fourth section of the dissertation will investigate the effects of different types and levels of misspecification in both the **G** and **R** side on the estimation of growth parameters, their corresponding standard errors, Type I error rate of the fixed effects, and the empirical statistical power for nonnull conditions.

1.1 Organization of Dissertation

The present dissertation is divided into five distinct sections. Sections 3 and 4 are written as individual manuscripts for potential publication in peer-reviewed journals. How each of the sections is conceptualized is presented below:

Section 1 serves as an introductory section that provides a brief overview of the topics to be examined along with a theoretical rationale for each of the individual studies. Section 2 provides a comprehensive literature review of issues related to MLM, including its advantages over traditional methods, MLM as a mixed effects model, the effects of misspecifying the within-subject V-C Structure and types of misspecification, and a review of existing methods in selecting an optimal V-C structure. Section 3 reports a Monte Carlo simulation study investigating the performance of commonly used fit statistics (i.e., AIC, BIC, Likelihood Ratio Test) in selecting the optimal V-C structure. A new index, Standardized Root Mean Square Residual (SRMR), is also proposed and evaluated. Annotated syntax is given to show how SRMR is calculated using Matlab. The third section is the first journal article. Section 4 reports results of a Monte Carlo simulation study that examines the effect of different types/levels of V-C misspecification in both the between- and within-subject matrices. The fourth section is the second journal article. Section 5, the last section, is the concluding section that connects the findings from the three manuscripts to provide overall and specific remarks for conclusions about MLM.

2. REVIEW OF ISSUES RELATED TO MULTILEVEL LINEAR MODELING

Multilevel linear modeling (MLM) for repeated measurement data has drawn increased attention in social and psychological studies over the past few decades. MLM is widely used for longitudinal studies because, for example, it can track the change of normal growth, identify risk factors, and assess the effect of intervention (Raudenbush, 2001). There are several advantages of modeling repeated measurement data using MLM over conventional statistical methods, such as (1) allowing unbalanced data or missing data points, (2) no requirement for observations to be taken equidistantly, (3) capturing the average growth trend of the outcome variable over time, and (4) flexibly modeling the variance-covariance (V-C) structure. Though MLM has many advantages over traditional methods, several issues remain unsolved and there are pitfalls that researchers may accidentally fall into when they do not have a thorough understanding of the MLM methodology. This section reviews the common issues related to MLM and is intended to function as a guide to introduce novice MLM researchers in the use of MLM to analyze repeated measurement data. It includes the advantages of MLM, MLM as a mixed effect model, effect of misspecifying a within-subject V-C structure, and methods in selecting an optimal V-C structure.

2.1 Advantages of MLM

2.1.1 Unbalanced Data or Missing Data Points

MLM does not require data to be balanced, where there are equal numbers of observations for all the combination of the classification factors, and allows analysis with missing data (Luke, 2004). Traditional Multivariate Analysis of Variance (MANOVA) deletes all the individuals or experimental units with missing data points (Hedeker & Gibbons, 2006); on the contrary, MLM uses all the available data, and the requirement of complete data is not necessary in the MLM analysis because the estimation method in MLM software packages such as PROC MIXED in SAS uses likelihood-based ignorable analysis, which assumes data to be missing at random (MAR), which can lead to valid analysis (Verbeke & Molenberghs, 2000).

2.1.2 No Requirement for Observations to Be Taken Equidistantly

Even if researchers can overcome the first constraints in traditional analysis methods and have complete data, equally-spaced observations will be required for both MANOVA and repeated measure ANOVA (Hedeker & Gibbons, 2006). In MLM, observations need not to be taken equidistantly. MLM can model pattern of change at unequally spaced time points as well as fixed time points.

2.1.3 Capturing the Average Growth Trend over Time

MLM allows the modeling of initial status and growth curve of each individual on an outcome variable. MLM has the capacity to depict the individual growth trend and the variation in the growth curve. The modeling of change is usually conducted in two levels, with level one being a function of time and level two examining individual difference in growth rate and initial status (Ferron et al., 2002). Repeated measure ANOVA, however, treat repeated measures as a within-subject factor on a nominal scale and can only test the difference in the response variable means at the different time points; similarly, the focus of MANOVA is on group mean comparison and gives no person-specific growth curves (Hedeker & Gibbons, 2006).

2.1.4 Flexibility in Modeling the V-C Structure

Most importantly, the focus of this paper is related to the advantage that the variance-covariance (V-C) structure can be flexibly modeled in MLM (Diggle, 1988; Laird & Ware, 1982; Wolfinger, 1993) while conserving degrees of freedom compared to unstructured modeling. Traditional UANOVA for repeated measurements requires the sphericity assumption or Huynh-Feldt (H-F) condition (with a compound symmetry V-C structure as the sufficient condition) which implies equal error variance for each measure within an individual and constant correlation between any pairs of repeated measures. This compound symmetry V-C structure may not be suitable for longitudinal data given that measures within a subject tend to correlate over time and the association diminishes as lags in time decreases (Hedeker & Gibbons, 2006). On the other hand, Multivariate Analysis of Variance (MANOVA) assumes an unconditional V-C structure by estimating all the unique elements in the V-C matrix which results in relatively low statistical power due to the large number of degrees of freedom required.

2.2 MLM as Mixed Effect Models

Alternative names for MLM-related modeling strategies includes "multilevel models" (Goldstein, 1995), "hierarchical linear models" (Bryk & Raudenbush, 1992), "random coefficient models" (Jennrich & Schluchter, 1986), "random effects models"(Laird & Ware, 1982), and "covariance component models" (Longford, 1993). Basically, MLM has so many synonymous names because it divides analysis into distinctive levels, allows level-specific parameters to vary across different experimental units, and accommodates various types of covariance structures. "The logical foundation for all longitudinal analysis is thus a statistical model defining parameters of change for the trajectory of a single participant. The task of comparing people then becomes the task of comparing the parameters of these personal trajectories" noted Raudenbush (2001, p. 502). For example, in the following linear growth curve model, the personal trajectory of change is a function of time (e.g. repeated measurement time points) in level one. Subject-specific parameters (π_{0i} and π_{1i} π_{1i}) are the level two outcome variables, varying around their grand means (β_{00} and β_{10} β_{00} *and* β_{10}) with variance (τ_{00} and τ_{11} τ_{00} *and* τ_{11}) and covariance (τ_{01} τ_{01}).

Level 1: $Y_{ti} = \pi_{0i} + \pi_{1i} time_{ti} + e_{ti}$, $e_{ti} \sim N(0, \sigma^2)$

Level 2: $\pi_{0i} = \beta_{00} + u_{0i}$, $\pi_{1i} = \beta_{10} + u_{1i}$ (1)

$$\text{with } \begin{pmatrix} u_{0i} \\ u_{1i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \; G = \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}\right]$$

Mathematically, MLM can be represented as a mixed effect model, with fixed effects defining the expected value of observations and random effects specifying the variance and covariance of the observations (Littell, Pendergast, & Natarajan, 2000). To have a clearer picture about how the between- and within- subject variance components are decomposed, we can take a simple linear growth curve model with *M* participants measured on *T* occasions in the same subject area for example.

$$\begin{bmatrix} y_{11} \\ \vdots \\ y_{T1} \\ \vdots \\ \vdots \\ \vdots \\ y_{1M} \\ \vdots \\ y_{TM} \end{bmatrix} = \begin{bmatrix} 1 & TIME_1 \\ \vdots & \vdots \\ 1 & TIME_T \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ 1 & TIME_1 \\ \vdots & \vdots \\ 1 & TIME_T \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} 1 & TIME_1 & 0 & & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & TIME_T & \vdots & & 0 & 0 \\ 0 & 0 & \vdots & & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \vdots & & 0 & 0 \\ 0 & 0 & \vdots & & 1 & TIME_1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & 1 & TIME_T \end{bmatrix} \begin{bmatrix} u_{01} \\ u_{11} \\ u_{02} \\ u_{12} \\ \vdots \\ \vdots \\ \vdots \\ u_{0M} \\ u_{1M} \end{bmatrix} + \begin{bmatrix} e_{11} \\ \vdots \\ e_{T1} \\ \vdots \\ \vdots \\ \vdots \\ e_{1M} \\ \vdots \\ e_{TM} \end{bmatrix} \quad (2)$$

Where y is a column vector with T repeated measures for M individuals. X is a [T*M by 2] matrix with intercept (i.e. 1) and the predictor variable TIME. β is a column vector with unknown growth parameters (i.e. β_0 $\boldsymbol{\beta_0}$ and $\beta_1$$\boldsymbol{\beta_1}$). Z is a [T*M] by [2M] design matrix, and $\boldsymbol{u}$ is a column vector with random effects representing between-subject variation in the intercept and slope. e is a column vector containing with-subject random errors for M individuals on T repeated measures.

Based on the above equation, the error structure can be divided into two parts, between-subject and within-subject error variance. The equation can be written for the general mixed effect model in matrix form according to Henderson (1975) as

$$y = X\beta + ZU + \varepsilon \quad (3)$$

Assuming $\boldsymbol{U}$ and $\boldsymbol{\varepsilon}$ are independently and normally distributed with

$$\varepsilon\begin{bmatrix} U \\ \varepsilon \end{bmatrix} = 0 \text{ and Var}\begin{bmatrix} U \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \quad (4)$$

where **y** is a vector of repeated measure outcome, **X** is the known design matrix of fixed effect, $\boldsymbol{\beta}$ is a vector of unknown fixed effect parameter estimates, **Z** is the known design matrix of the random effect, $\boldsymbol{U}$ is a vector of unknown random effect parameter estimates, and $\boldsymbol{\varepsilon}$ is the error associated with the measurement outcome. $\boldsymbol{\varepsilon}$ is assumed to be $N \sim (0, \boldsymbol{R})$ and U is assumed to be $N \sim (0, \boldsymbol{G})$. In repeated measurements, **R** corresponds to the within-subject error structure and **G** is the between-subject error matrix. The total variance in Y is $\boldsymbol{V}$, which is a function of $\boldsymbol{G}$ and $\boldsymbol{R}$**:**

$$Var(y) = ZGZ^T + R = V \quad (5)$$

Under the mixed model assumptions, (1) the means (expected values) of the responses are linearly related to the fixed-effects parameters (i.e. $E(\boldsymbol{y}) = \boldsymbol{X\beta}$), (2) random effect and residuals are normally distributed with mean zero and covariance matrices G and R respectively, and (3) random effect and residuals are independent of each other. Due to the independence of random effect and residuals, the **G** and **R** matrices can be flexibly modeled and conform to the structure of sample data.

2.3 Estimation Method

In the mixed effect model framework, the estimates for the fixed effects and random effects are calculated separately using different estimation methods.

2.3.1 Fixed Effect

For the estimation and hypothesis testing of the fixed effect parameters, Generalized Least Square[1] estimation method (GLS) is used. The GLS method is superior to the ordinary least square (OLS) method by taking into account the **G** and **R** covariance matrices or assuming an appropriate V-C structure (Tao, Littel, Patetta, Truxillo, & Wolfinger, 2002). The GLS method for the fixed effect takes into account of the covariance matrices for the random effect and residuals and contributes to more precise fixed effect parameter estimates (Littell, Milliken, Stroup, Wolfinger, & Schabenberber, 2006). In MLM, the fixed effects are estimated using GLS; therefore, the inferences directly incorporate the V-C structure the researcher specifies, while in OLS regression, ordinary least squares is used to estimate the fixed effects, and the inferences are made based on the fixed-effect only model (Tao et al., 2002). With the prediction of a random effect and the inclusion of random effect into any linear combination, the resulting fixed effect estimates are best linear unbiased prediction (BLUP), where the expected value of y given u is $E[\boldsymbol{Y}|\boldsymbol{u}] = \boldsymbol{X}\beta + \boldsymbol{Z}U$, a subject-specific (conditional) model; on the other hand, the expected value of y over entire population is $E[\boldsymbol{Y}] = E[\boldsymbol{X}\beta + \boldsymbol{Z}U + \boldsymbol{e}] = \boldsymbol{X}\beta$, which is a population-average (marginal) model (Littell et al., 2006). The estimated BLUP for a random effect shrinks toward the

---

[1] The GLS solutions for β are obtained by minimizing $(y - X\beta)'V^{-1}(y - X\beta)$ for β. The corresponding GLS solution estimate for β is $\hat{\beta} = (X'\hat{V}^{-1}X)^{-}X'\hat{V}^{-1}y$. The estimated GLS solutions for the random effects γ is obtained by $\hat{\gamma} = \hat{G}Z'\hat{V}^{-1}(y - X\hat{\beta})$. The fixed effect estimate in OLS is $\hat{\beta} = (X'X)^{-}X'y$, a special case for GLS if V=$\sigma^2 I_n$. $\hat{\beta}$ is the best linear unbiased estimator (BLUE). $\hat{\gamma}$ is the best linear unbiased predictor (BLUP).

population mean with a shrinkage factor equal $\frac{\sigma_b^2}{\sigma_b^2+\frac{1}{s}\sigma^2}$, where σ_b^2 is the variance of the random effect and σ^2 is the residual variance (Tao et al., 2002).

2.3.2 Random Effect

The common and popular estimation methods for the random effect parameters are maximum likelihood[2] based functions, such as maximum likelihood (ML) and restricted/residual maximum likelihood (REML). The two maximum likelihood estimation methods differ in the construction of likelihood function. REML uses the correct degrees of freedom by taking into account the degrees of freedom for the fixed effects in the model for the random effect and the residual likelihood function to obtain ML estimates for the variance components. Therefore, REML covariance component estimates are bias-free whereas ML covariance component estimates are biased downward when none of the covariance parameter estimates hit the non-negative boundary constraint (Tao et al., 2002). REML can be used to specify different V-C structures under the same mean model but ML should be used to account for the V-C structure the researcher specified when the researchers read the fit statistics for comparing the appropriateness of different V-C structures (Tao et al., 2002). REML received growing preference over ML for obtaining covariance parameter (McCulloch & Searle, 2001).

---

[2] In SAS Mixed Procedure, the log likelihood function for ML and REML are specified as:
ML: $l(\mathbf{G},\mathbf{R}) = -\frac{1}{2}\log|V| - \frac{1}{2}r'V^{-1}r - \frac{n}{2}\log 2\pi$
REML: $l_R(\mathbf{G},\mathbf{R}) = -\frac{1}{2}\log|V| - \frac{1}{2}\log|X'V^{-1}X| - \frac{1}{2}r'V^{-1}r - \frac{n-p}{2}\log 2\pi$
Where $r = y - X(X'V^{-1}X)^{-}X'V^{-1}y$ and p=rank(X)
In the default setting, a ridge-stabilized Newton-Raphson algorithm is used to minimize -2 times the log likelihood functions and obtain the parameter estimates.

2.4 Types of Variance-Covariance Structures

Though MLM has the flexibility in modeling different types of variance-covariance structure, the default V-C structure in popular statistical packages (e.g. HLM, SPSS Mixed, SAS Proc MIXED, and STATA XTMixed) is still the identity structure (i.e. $\sigma^2 I$) where the variance for all the repeated measures is the same and no covariance exists among repeated measures. Assuming no covariance among repeated measures is unrealistic as assuming static covariance in UANOVA with repeated measures. In the following section, commonly used variance-covariance structures in longitudinal data are introduced with the specification on the structures. According to Wolfinger (1993), a wide variety of V-C structures as an alternative to the identity structure can be used for repeated measures. These V-C structures include first-order autoregressive, banded, unstructured, Toeplitz, banded Toeplitz, and first-order autoregressive plus a diagonal. In this section, selected V-C structures are introduced and compared with an illustration of the V-C structures presented in Figure 1.

$$\begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}$$

Identity Structure

(ID)

$$\begin{bmatrix} \sigma^2+\sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2+\sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2+\sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2+\sigma_1 \end{bmatrix}$$

Compound Symmetry

(CS)

$$\sigma^2\begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

Autoregressive(1)

(AR(1))

$$\begin{bmatrix} \sigma^2 & \sigma_1 & 0 & 0 \\ \sigma_1 & \sigma^2 & \sigma_1 & 0 \\ 0 & \sigma_1 & \sigma^2 & \sigma_1 \\ 0 & 0 & \sigma_1 & \sigma^2 \end{bmatrix}$$

Toeplitz with 2 Bands

(TOEP(2))

$$\sigma^2\begin{bmatrix} 1 & \gamma & \gamma\rho & \gamma\rho^2 \\ \gamma & 1 & \gamma & \gamma\rho \\ \gamma\rho & \gamma & 1 & \gamma \\ \gamma\rho^2 & \gamma\rho & \gamma & 1 \end{bmatrix}$$

First-order Autoregressive Moving-Average

(ARMA(1,1))

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix}$$

Unstructured

(UN)

Figure 1. Commonly Used Within-Subject V-C Structure.

2.4.1 Identity Structure (ID)

The ID structure specifies that repeated measures are independent for each individual and have homogenous variance. The correlation function between all pairs of lags equals zero. Repeated measures under the ID structure assumption are unrealistic because it assumes no correlation among observation within an individual, though the default V-C structure for most of the popular statistical packages is the identity structure. In terms of equation (5), the ID structure says **G**=0 and **R**=$\sigma^2 \boldsymbol{I}$, where $\boldsymbol{I}$ is an identity matrix.

2.4.2 Compound Symmetric (CS)

The compound symmetric model specifies that individuals have homogeneous variance and homogeneous covariance among observations. The correlations were the same between any pairs of lags within an individual. There are two ways to specify a CS structure in terms of **G** and **R** in equation (5), either **G**=$\sigma^2_{CS,B}\boldsymbol{I}$ and **R**=$\sigma^2_{CS,W}\boldsymbol{I}$ or **G=0** and R=$\sigma^2_{CS,W}\boldsymbol{I} + \sigma^2_{CS,B}\boldsymbol{J}$, where J is a matrix of ones (Littell et al., 2000).

2.4.3 First-Order Autoregressive (AR(1))

AR(1) specifies the V-C structure to have homogeneous variance but covariance decreasing at an exponential rate with the increase of lags. The AR(1) can be presented as $y_t = \rho_1 y_{t-1} + \varepsilon_t$, where y_t is the predicted score at taken at time t, ε_t is the error associated with the measurement at time t, ρ is the autocorrelation coefficient, $0 \leq |\rho| \leq 1$. "AR models represent the most recent observation in a series as a function of previous observations within the same series." (Murphy & Pituch, 2009).

2.4.4 First-Order Autoregressive Moving Average Model (ARMA(1,1))

The ARMA(1,1) model is similar to the AR(1) model with the inclusion of an additional moving average parameter, θ . The ARMA(1,1) model with lag-1 process can be represented as $y_t = \rho_1 y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$. Like the AR(1) model,)). ARMA(1,1) specifies the V-C structure to have homogeneous variance and covariance decreasing at an exponential rate of the autocorrelation coefficient plus a multiplicative moving average constant with the increase of lags.

2.4.5 Toeplitz (TOEP)

"Toeplitz structure, sometimes called 'banded', specifies that covariance depends only on lag, but not as a mathematical function with a smaller number of parameters." (Littell et al., 2000). TOEP structure specifies the V-C structure to have homogenous variance σ^2_{TOEP} and mirrored equal covariance along the same band. In terms of equation (5), TOEP structure is specified with G = 0, elements in main diagonal of R are σ^2_{TOEP}, and $\sigma_{TOEP,|k-l|}$ for elements in the sub-diagonal, where $|k-l| = lag$ with *k* equal to the row number and *l* the column number (Littell et al., 2000).

2.4.6 Unstructured (UN)

Unstructured V-C matrix is the most general/unconditional form of V-C structure. Every unique element in the UN V-C structure is estimated with the upper triangle mirroring the lower triangle. In SAS PROC MIXED, the variance is constrained to be non-negative and the covariance is unconstrained (SAS Institute, 2008).

2.5 Effect of Misspecifying the Within-Subject V-C Structure and Types of Misspecification

Researchers have studied the effect of misspecifying the error structure in repeated measure data in the MLM context (Ferron et al., 2002; Kwok et al., 2007; Murphy & Pituch, 2009; Vallejo et al., 2008). Misspecification in the covariance structure generally reflected in negative influences on the estimates of standard errors for the fixed effects (Davis, 2002; Diggle, Heagerty, Liang, & Zeger, 2002; Kwok et al., 2007; Singer & Willett, 2003) and the associated hypothesis tests and caused biased statistical inferences or inflated type I error rate and lowered statistical power depending on the types of misspecification (Kwok et al., 2007; Murphy & Pituch, 2009; Vallejo et al., 2008). However, fixed effect estimate and its corresponding hypothesis test remained unbiased for most of the occasions (Ferron et al., 2002). Additionally, misspecification or no specification of the V-C structure may risk the potential of losing information of the change in the outcome variable over time that is reflected only in the covariance matrix of the within-subject residuals (Hedeker & Mermelstein, 2007).

Kwok et al. (2007) defined three types of misspecification in the covariance structure, over-specification, under-specification, and general-misspecification. Over-specification refers to misspecification of a simpler covariance structure to a more complex nested structure, for example, misspecifying an identity structure (ID) to a first-order autoregressive structure (AR(1)). On the contrary, under-specification means mis-identifying a more complex covariance structure to a simpler nested structure, for example, incorrectly specifying autoregressive moving average structure (ARMA(1,1)) to AR(1). General misspecification applies to misspecifying covariance structures between two non-nested covariance structure such as misspecifying ID to a banded toeplitz structure (TOEP(2)). Under-specification or general-misspecification often led to overestimation of the random effects and the corresponding standard errors while over-specification may lead to underestimation of the random effects and standard errors (Kwok et al., 2007). Though the effect of misspecification in the covariance structure has been researched, most of the studies only examined misspecification in the $\boldsymbol{R}$ side within-subject covariance structure (Ferron et al., 2002; Kwok et al., 2007; Murphy & Pituch, 2009; Vallejo et al., 2008). No research to date has examined the misspecification in both the between- and within-subject covariance structure simultaneously.

2.6 Selecting an Optimal V-C Structure

Littell et al. (2000) suggested four steps in modeling a mixed effect analysis.

Step 1: Model the mean structure by specifying the fixed effects to ensure unbiasedness of the fixed effect estimates

Step 2: Specify the covariance structure, between subjects as well as within subjects

Step 3: Use GLS estimation method to fit the mean model accounting for the covariance structure

Step 4: Make statistical inference based on the results of step 3 and make the mean model parsimonious

As a matter of fact, before proceeding from step 3 to step 4, there is an additional inexplicit step, that is, the model selection step for an optimal V-C structure. Researcher should choose among several competing V-C structures to determine which better accounts for the current data under the same mean model. Traditional model selection procedure involves using information criteria such as Akaike Information Criterion (AIC; Akaike, 1974) or Bayesian/Schwartz Information Criterion (BIC; Schwarz, 1978) and likelihood ratio test (LRT). A brief description about the traditional model selection methods was provided along with a potential alternative for searching the optimal V-C structure.

2.6.1 Likelihood Ratio Test (LRT)

Models with nested structures can be evaluated using the likelihood ratio test. The likelihood ratio test is the difference of the deviance statistics between one model nested in another. The deviance statistic is defined as -2 times the ratio of the log-likelihood statistic of the hypothesized model to the log-likelihood statistic of the saturated model:

$$\text{Deviance} = -2\left(\frac{\text{Log-likelihood}_{\text{Hypothesized model}}}{\text{Log-likelihood}_{\text{Saturated model}}}\right) \quad (6)$$

It quantifies the degree of badness of fit (to the data) of the current (hypothesized) model in comparison to the saturated model. Likewise, we can compute deviance statistics for nested competing models, one with simple covariance structure and the other with more complex structure (e.g. D_{simple} & $D_{complex}$), and obtained a change/difference in the deviance statistics ($\Delta D = D_{simple}-D_{complex}$) between the two models. This difference in the deviance statistics between nested models is termed the likelihood ratio test (LRT), which follows a chi-square distribution with degrees of freedom equal the difference between the total parameters estimated in each model. If the test statistic exceeds the value in the chi-square distribution associated with a specific alpha level of significance, we conclude that the simpler model (hypothesized model) is statistically worse than the more complex model and favor the complex model. Nevertheless, this hypothesis test can only reveal the difference but not magnitude of the difference between the two nested models. Moreover, model selection based on the standard significance test of the nested models is very sensitive to slight deviation between the nested models and tends to over-reject the parsimonious model when the sample size is large (Kuha, 2004).

2.6.2 Akaike Information Criteria (AIC)

AIC can be used for comparing non-nested covariance structures, for example, comparison between banded toeplitz (TOEP(2)) and first order autoregressive and moving average structure, (ARMA (1,1)). The formula for AIC can be written as the following expression:

$$AIC = d+2k \tag{7}$$

where d is the deviance statistic and k is the number of parameters estimated. Smaller AIC is the-better statistic because smaller values indicate better fit of the model to the data. AIC penalizes additional parameters to be estimated and the size of penalty is 2 multiplied by k.

2.6.3 Bayesian Information Criteria (BIC)

BIC is also used for non-nested models in comparing covariance structures. The formula for BIC can be represented as:

$$BIC = d+k*\ln(N) \tag{8}$$

where d is the deviance statistic, k is the number of estimated parameters, ln is the natural log, and N is the sample size. Like AIC, smaller BIC is also the-better statistic and penalizes for additional estimated parameters. BIC and AIC differ in the size of penalty, which is k multiplied by ln(N) for BIC. Therefore, BIC usually favors models with fewer parameters (Weakliem, 2004).

Unlike LRT, AIC and BIC quantify the degree of improvement for a given model over a comparison model (O'Connell & McCoach, 2008). However, the results of empirical studies showed that the accuracy of using these information criteria in search of an optimal covariance structure is not very promising. For example, AIC can accurately identify the true covariance structure 47% of the time while for BIC is only 35% (Keselman, Algina, Kowalchuk, & Wolfinger, 1998). Research is needed in finding an optimal fit statistic or index that can better determine the appropriate V-C structure for the data.

2.6.4 Standardized Root Mean Square Residual (SRMR)

A possible indicator that can be used for indentifying the true covariance structure is the standardized root mean square residual (SRMR; Bentler, 1995). SRMR is one of the absolute fit indices that evaluates how well a model reproduces the sample data under the SEM framework (Hu & Bentler, 1998). SRMR is defined as:

$$\text{SRMR} = \sqrt{\frac{\left\{2\sum_{i=1}^{t}\sum_{j=1}^{i}\left[\left(s_{ij}-\hat{\sigma}_{ij}\right)/\left(s_{ii}s_{jj}\right)\right]^{2}\right\}}{T(T+1)}} \quad (9)$$

where s_{ij} is an element (e.g., the covariance between the ith and jth time points) in the observed/unconditional covariance matrix, $\hat{\sigma}_{ij}$ is the corresponding element from the model-implied covariance matrix based on a specific model, s_{ii} and s_{jj} are the observed standard deviations of the ith and jth time points respectively. T is the number of repeated measures in longitudinal data analysis.

According to the algebraic definition, SRMR is a measure of the averaged difference of the standardized residuals between the observed/unconditional and model-implied covariance matrices (Bentler, 1995). SRMR is most sensitive to detecting misspecification in factor covariances (Hu & Bentler, 1998) and is a commonly reported fit index in SEM studies. We can obtain a similar SRMR under the MLM framework. Firstly, we need to confirm and estimate the fixed effect part or the mean model. Once we define a reasonable means model, the unstructured within-subject covariance structure can be treated as the unconditional covariance matrix because it represents the most general form of the covariance structure. With the same means model, we can specify the predicted/model-implied covariance matrix. The SRMR under the MLM framework can then be calculated based on these two covariance matrices. The effectiveness and performance of SRMR in searching for the optimal V-C structure has, however, not yet been evaluated.

2.7 Discussion

Through the review in the field of multilevel modeling, two research issues are emergent and draw our attention. First, there is a lack of an optimal model selection method, and second the effect of different types/levels of misspecification has not been investigated. Misspecification of the within-subject V-C structure, although it may not have a negative influence on the fixed effect estimates, it can be expected to lead to biased estimation of the standard errors for the fixed effects. In other words, the statistical inferences drawn from the combination of unbiased fixed estimates and biased standard error of the fixed effects will still be erroneous. The first issue is to develop effective methods for selecting an optimal within-subject V-C structure. Research can be conducted to assess efficacy of traditional fit statistics and evaluate in comparison the proposed SRMR index in selecting the optimal V-C structure.

Second, the misspecification on the R-side of the V-C structure has been extensively researched (Ferron et al., 2002; Kwok et al., 2007; Murphy & Pituch, 2009; Vallejo et al., 2008). Nevertheless, misspecification in both the G-side (between-subject) and the R-side (within-subject) covariance structures has rarely been examined simultaneously. In a simple linear growth curve model, the G matrix is comprised of random effects, including variance of intercept (τ_{00}), variance of slope τ_{11} , and covariance between intercept and slope (τ_{01}), capturing the deviation of growth parameters in individuals from the population means for intercept and slope as shown in equation (2). Therefore, whden fitting a mixed effect model for repeated measurements, researchers need to specify both the **G**- and **R**- side V-C structure for the data. In the best scenario, researchers will specify the V-C structures for the two sides correctly. In other cases, researchers may over-, under-, or generally misspecify the V-C structures. Research will thus be conducted to investigate the effect of different types/levels of misspecification in both the **G** and **R** side V-C matrices, considering the estimation of growth parameters, their standard errors, Type I error rate of the fixed effects, and the empirical statistical power so that we can have a general guideline as to the impact of different types/levels of misspecification on the interpretation of MLM studies.

3. SEARCHING FOR THE OPTIMAL WITHIN-SUBJECT COVARIANCE STRUCTURE IN LONGITUDINAL DATA ANALYSIS USING MULTILEVEL MODELING (MLM): A MONTE CARLO STUDY

3.1 Theoretical Framework

Multilevel linear modeling (MLM) is widely used in educational research because many educational data are in a multilevel structure (e.g., repeated measures nested within students and students nested with schools). MLM has also been adopted for analyzing longitudinal data (e.g. repeated measures nested within students) not only because it can capture the average growth trend of the outcome variable over time but also can flexibly model the within-subject covariance structure. However, when analyzing longitudinal data, researchers generally impose the simplest within-subject covariance structure, the identity structure: $\mathrm{R}=\sigma^2\mathrm{I}$, which is the default structure of the within-subject covariance matrix in many MLM related programs such as HLM, SAS PROC MIXED, SPSS MIXED, and STATA XTMIXED. This within-subject covariance structure is not realistic for longitudinal data because the repeated measures tend to correlate with each other over time and the correlations between measures tend to diminish as the lags in time increase (Hedeker & Gibbons, 2006). Failure to model an appropriate error covariance structure results in: 1) bias estimation of the standard errors of the fixed effects (Davis, 2002; Diggle et al., 2002; Kwok et al., 2007; Singer & Willett, 2003) and 2) the potential of losing information of the change in the outcome variable over time that is reflected only in the covariance matrix of the within-subject residuals (Hedeker & Mermelstein, 2007).

The impact of misspecifying the within-subject covariance matrix has been examined (Ferron et al., 2002; Kwok et al., 2007) and the importance of obtaining the correct covariance structure has been addressed (Singer & Willett, 2003). However, only a few studies (e.g., Keselman et al., 1998; Wolfinger, 1993) have examined the performance of the traditional methods including the likelihood ratio test (LRT) and the information criteria (e.g., Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC)) on searching for the correct covariance structure under the mixed model framework. The purpose of this study is to evaluate the performance of these traditional methods on searching for the correct within-subject covariance structure in longitudinal data analysis under the MLM framework. Additionally, a new alternative, standardized root mean square residual (SRMR), is proposed and its performance on searching for the correct covariance structure is compared with the traditional methods.

In this study, the performance of LRT, AIC, BIC, and SRMR on searching for the correct within-subject covariance structure were evaluated. The impact of several factors including sample size, magnitude of the average growth model, and magnitude of the between-subject covariance matrix on the performance of these search methods were also considered in the analysis.

3.2 Method

In this study, we focused on a common two-level growth model with level-1 modeling the repeated measures within individuals and level-2 modeling the differences of individual growth models between individuals. The level 1 and level 2 models can be written as the following expressions:

$$\text{Level 1: } Y_{ti} = \pi_{0i} + \pi_{1i} time_{ti} + e_{ti},\ e_{ti} \sim N(0, \sigma^2)$$

$$\text{Level 2: } \pi_{0i} = \beta_{00} + u_{0i},\ \pi_{1i} = \beta_{10} + u_{1i} \tag{10}$$

$$\text{with } \begin{pmatrix} u_{0i} \\ u_{1i} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix},\ G = \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}\right]$$

where Y_{ti} Y_{ti}, the personal trajectory of change, is a function of time ($time_{ti}$ $Time_{ti}$, e.g. $time_{ti}$ repeated measurement time points) in level one with *i* indicates individuals, *i* = 1, ..., *N*; and *t* indicates time points, 4 waves: *T* = -1.5, -0.5, 0.5, 1.5 or 8 waves: *T* = -3.5, -2.5, -1.5, -0.5, 0.5, 1.5, 2.5, 3.5 centered to have a mean of 0 and 1 unit between adjacent observations. The level two outcome variables, π_{0i} π_{0i} (intercept) and π_{1i} π_{1i} (slope) are the growth parameters in a linear growth model. π_{0i} π_{0i} and π_{1i} π_{1i} are multivariate normally distributed and vary around their grand means (β_{00} and β_{10} β_{00} *and* β_{10}) with variance (τ_{00} *and* τ_{11} τ_{00} and τ_{11})) and covariance (τ_{01} τ_{01}). We limited our focus to a simple *linear* growth model with correctly specified fixed effects collected in a balanced design.

3.2.1 Research Design and Model Parameterization

The simulation used a 2 (30 or 210 cases) x 2 (4 or 8 waves) x 3 (magnitude of growth parameter β_1: 0, .05 or .16) x 2 (**G** matrix: small or medium) x 4 (true **R** matrices for generating the data: ID, TOEP(2), AR(1), or ARMA(1,1)) factorial design to generate the data. A total of 500 replications were generated for each condition using the Mplus (V4.1) Monte Carlo procedure (Muthén & Muthén, 1998), yielding 48,000 total datasets. All data were generated under Mplus with a multivariate normal distribution (Muthén & Muthén, 1998). Each dataset was then analyzed using five separate specifications of the **R** matrix (ID, TOEP(2), AR(1), ARMA(1,1) and UN) using SAS PROC MIXED (Littell et al., 2006) yielding a total of 240,000 records (i.e., 48,000*5).

For the number of participants, we chose 30 as a “small” number of individuals and 210 as a “medium” number of individuals based on past simulation studies (Ferron et al., 2002; Keselman et al., 1998) and the review of the multiwave longitudinal studies published in *Developmental Psychology* by Khoo, West, Wu, & Kwok (2006). Additionally, we chose 4 waves as the small number of repeated measures and 8 waves as the medium number of measures based on the same review by Khoo and colleagues (Khoo et al., 2006). Three different magnitudes of the standardized effect size of the growth trajectory were examined in this study, no effect (i.e., $\beta_1 = .00$), small effect size (i.e., $\beta_1 = .05$) and medium effect size (i.e., $\beta_1 = .16$). The standardized effect size is calculated with the following equation (Raudenbush & Liu, 2001):

$$\delta = \frac{\beta_1}{\sqrt{\tau_{11}}} \tag{11}$$

where δ is the standardized effect size, β_1 is the average linear growth, and τ_{11} is the variance of the random effect associated with the growth, which captures the differences between individual growth trends and the average growth trend. The size of τ_{11} were set at .05 and .10, where $\tau_{11} = .05$ is recognized as small and $\tau_{11} =.10$ as medium according to Raudenbush and Liu (2001). Given the values of β_1 and τ_{11}, the corresponding δ could be easily computed and the resulting effect size in our simulated data is consistent with small effect size (i.e., $\delta = .20$) and medium effect size (i.e., $\delta = .50$) proposed by Cohen (1988). We also included the size of **G** matrix (small versus medium) as a design factor. According to the criteria provided by Raudenbush and Liu (2001), a medium **G** and a small **G** could be specified as:

$$R_{Medium} = \begin{bmatrix} .200 & .050 \\ .050 & .100 \end{bmatrix} \text{ and } R_{small} = \begin{bmatrix} .100 & .025 \\ .025 & .050 \end{bmatrix}.$$

The final design factor to be specified is the true within-subject covariance structure. Four covariance structures, namely, ID, AR(1), TOEP(2), and ARMA(1,1) were adopted in our study. The four within-subject covariance structures are commonly used when analyzing longitudinal data. AR(1) and ARMA(1,1) are commonly considered in time series analysis (Velicer & Fava, 2003; West & Hepworth, 1991). TOEP(2) is closely related to the moving average (1) structure which is also commonly used in time series analysis. σ^2 was set as 1 for all the ID, AR(1), TOEP(2), and ARMA(1,1) models, which is a common practice in power analysis under MLM studies (Bosker, Snijders, & Guldemond, 2003). The autoregressive parameter ρ was set as 0.8. The Toeplitz parameter, σ_1, for TOEP(2) and the moving average parameter, γ, for ARMA(1,1) were set as 0.5. These values were within the reasonable range of prior studies (Hamaker, Dolan, & Molenaar, 2002; Sivo & Willson, 2000).

3.2.2 Selection Criterion for the Optimal Covariance Matrix

The hit rate (i.e., the percentage of replications with correctly specified within-subject covariance structure) of each search method was used as the major outcome variable. For AIC and BIC, a correct hit in model selection was represented by an event that the smallest AIC or BIC value for the hypothesized covariance structure matches the true covariance structure. AIC and BIC hit rate for all investigated conditions and within-subject covariance structures were computed respectively. With respect to LRT, a correct identification in selecting a true covariance structure was determined in two stages. Firstly, LRT was conducted between the model with a hypothesized covariance structure and the model with the unstructured covariance (UN-structured) structure to examine whether the hypothesized model is statistically worse than the saturated model. If the results are statistically significant for the four hypothesized covariance structures (i.e., the UN-structured covariance fit the data best), there is no need to proceed to the second stage given that LRT fails to select the true covariance structure. Otherwise, cases that did not have a significant LRT proceeded to the second stage, indicating that the hypothesized models do not fit statistically worse than the saturated model. We computed the change in the deviance statistics between pairs of nested competing covariance structures and performed the chi-square differential test to determine whether the selected covariance structure matched the true covariance structure. After these steps, we calculated the overall hit rate for LRT.

To be selected as the optimal V-C structure for SRMR, two criteria must be met, (1) a SRMR value equal or less than 0.08 and (2) a matching target within-subject covariance matrix. For example, if ARMA(1,1) is the true within-subject covariance structure, one observation has a SRMR value less 0.08 and a target ARMA(1,1) covariance matrix, then SRMR succeeds in selecting the optimal within-subject V-C structure. The 0.08 criterion was selected as suggested by Hu and Bentler (1999) for the cutoff value of SRMR. The hit rate for LRT, AIC, BIC, and SRMR were compared as to their performance on searching for the correct covariance structure.

3.3 Result

The 2 (30 or 210 cases) x 2 (4 or 8 waves) x 3 (magnitude of growth parameter β_1: 0, .05 or .16) x 2 (**G** matrix: small or medium) factorial design yielded 24 simulation settings. The 24 settings were named through Model A to Model X. All condition investigated had an UN-structured T matrix with a simulated **R** matrix which is ID, TOEP(2), AR(1), or ARMA(1,1). The unconditional model for calculating SRMR had an unstructured within-subject variance covariance matrix with Null G so as to avoid model overparametization.

3.3.1 Convergence Rate

The average convergence rate across 24 models was 95%. The result from the ANOVA test suggested convergence rate was moderated by both number of case, number of wave, and their interaction effect ($F(3, 20) = 548.603$, $p < .001$). Models with 30 cases and 4 waves had the lowest mean convergence rate (mean = 0.90). Models with 210 cases and 8 waves had the highest mean convergence rate (mean = 0.99), followed by models with 30 cases and 8 waves (mean = 0.97) and models with 210 cases and 4 waves (mean = 0.96).

3.3.2 AIC and BIC Hit Rate

The AIC hit rate is shown in Table 1. The hit rate for a specific within-subject covariance matrix is shown by columns. The hit rate across all within-subject covariance matrices within a certain model is presented by rows. For ID covariance structure, the AIC statistic was able to correctly classify the covariance structure 68% of the time. For TOEP(2), the AIC hit rate was 72%. AR(1) covariance structure had an AIC hit rate of 65%. However, the AIC hit rate decreased to 41% for a true ARMA(1,1) covariance structure. The overall AIC hit rate across all conditions and within-subject covariance structures was about 62%. The BIC hit rate is presented in. For ID covariance structure, the BIC hit rate reached 84%, for TOEP(2) within-subject covariance 80%, for AR(1) 72%, and for ARMA(1,1) 28%. The overall BIC hit rate across all conditions and within-subject covariance structures was 66%. Analysis of variance (ANOVA) was conducted for both information criteria. The ANOVA test result is presented in Table 3. The η^2 statistic, the *semipartial* η^2 reported in SAS Proc GLM effect size option, was used to evaluate the impact of design factors on the hit rate. The semipartial η^2 is the proportion of total variation accounted for by the effect being tested, i.e. the ratio of observed sum of squares due to the effect being tested and the total corrected sample sum of squares. Number of cases ($F(5, 18) = 455.37$, $p < .01$) and number of waves ($F(5, 18) = 1094.39$, $p < .01$) were both significant factors for the hit rate of AIC. Number of waves alone accounted for 70% of the total variance in the AIC hit rate and number of cases explained the remaining 29% of the between factor variation. Models with 210 cases had a higher AIC hit rate (69.17%) than those with 30 cases (53.42%). Models with 8 waves (73.50%) had a higher AIC hit rate than those with 4 waves (49.08%).

Table 1. AIC Hit Rate

					Correct Classification (%)						
Model	# of cases	# of waves	Magnitude of growth parameter	**T** matrix	ID	TOEP(2)	AR(1)	ARMA(1,1)	Average	n	Convergence
A	30	4	0.00	Medium	58	46	39	9	38	1784	89
B	30	4	0.00	Small	55	52	40	12	40	1802	90
C	30	4	0.05	Medium	58	48	41	13	40	1795	90
D	30	4	0.05	Small	60	58	39	14	43	1802	90
E	30	4	0.16	Medium	57	51	37	15	40	1809	90
F	30	4	0.16	Small	54	55	40	14	41	1815	91
G	30	8	0.00	Medium	73	85	77	32	67	1947	97
H	30	8	0.00	Small	73	82	76	37	67	1953	98
I	30	8	0.05	Medium	73	82	76	36	66	1935	97
J	30	8	0.05	Small	73	79	77	38	67	1940	97
K	30	8	0.16	Medium	73	82	78	32	66	1947	97
L	30	8	0.16	Small	70	79	75	43	66	1947	97
M	210	4	0.00	Medium	71	74	60	31	59	1923	96
N	210	4	0.00	Small	67	65	64	29	59	1915	96
O	210	4	0.05	Medium	73	75	63	26	59	1925	96
P	210	4	0.05	Small	63	60	66	32	55	1936	97
Q	210	4	0.16	Medium	70	74	67	29	60	1915	96
R	210	4	0.16	Small	63	63	63	33	55	1913	96
S	210	8	0.00	Medium	78	82	79	83	81	1974	99
T	210	8	0.00	Small	74	81	81	85	80	1971	99
U	210	8	0.05	Medium	77	84	81	83	81	1973	99
V	210	8	0.05	Small	73	82	80	83	80	1978	99
W	210	8	0.16	Medium	73	83	79	83	80	1973	99
X	210	8	0.16	Small	74	85	80	86	81	1966	98
Total					68	72	65	41	62	45838	95

Table 2. BIC Hit Rate

Model	# of cases	# of waves	Magnitude of growth parameter	**T** matrix	Correct Classification (%) ID	TOEP(2)	AR(1)	ARMA(1,1)	Average	n	Convergence
A	30	4	0.00	Medium	63	46	43	5	39	1784	89
B	30	4	0.00	Small	64	54	45	9	43	1802	90
C	30	4	0.05	Medium	64	46	44	8	41	1795	90
D	30	4	0.05	Small	67	61	41	11	45	1802	90
E	30	4	0.16	Medium	66	52	41	9	42	1809	90
F	30	4	0.16	Small	63	57	44	11	44	1815	91
G	30	8	0.00	Medium	87	91	83	24	71	1947	97
H	30	8	0.00	Small	86	88	82	29	71	1953	98
I	30	8	0.05	Medium	89	90	82	29	72	1935	97
J	30	8	0.05	Small	86	86	81	30	70	1940	97
K	30	8	0.16	Medium	90	91	83	25	71	1947	97
L	30	8	0.16	Small	82	87	80	35	71	1947	97
M	210	4	0.00	Medium	93	84	68	12	64	1923	96
N	210	4	0.00	Small	81	76	71	12	59	1915	96
O	210	4	0.05	Medium	94	85	71	11	64	1925	96
P	210	4	0.05	Small	81	71	72	13	59	1936	97
Q	210	4	0.16	Medium	94	86	73	11	65	1915	96
R	210	4	0.16	Small	81	73	70	13	59	1913	96
S	210	8	0.00	Medium	96	97	88	61	86	1974	99
T	210	8	0.00	Small	95	96	89	59	85	1971	99
U	210	8	0.05	Medium	96	99	90	59	86	1973	99
V	210	8	0.05	Small	96	97	90	59	85	1978	99
W	210	8	0.16	Medium	97	98	89	56	85	1973	99
X	210	8	0.16	Small	96	98	92	63	87	1966	98
Total					84	80	72	28	66	45838	95

The magnitude of the growth parameter (η^2 = .0001) and **G** matrix (η^2 = .0001) rarely explained any variance in the hit rate of AIC. The hit rate for BIC had similar ANOVA result as the AIC hit rate. Only number of cases and number of waves were significant factors. They accounted for 98% of the variation in BIC hit rate (number of waves, 69%; number of cases, 29%). The 210-case models had a higher BIC hit rate (73.67%) than the 30-case models (56.67%). Likewise, 8-wave models (78.33%) had higher a BIC hit rate than the 4-wave models (52.00%).

3.3.3 SRMR Hit Rate

Table 4 presents the SRMR hit rate. The Matlab codes for calculating SRMR hit rate are presented in Appendix A. The overall SRMR hit rate was 81% across all investigated conditions and within-subject covariance structures. Correct classification of ID structure was 91%, while the TOEP(2) structure had the highest SRMR hit rate, 92%. AR(1) and ARMA(1,1) had lower SRMR hit rates, 79% and 61% respectively. ANOVA tests were conducted to examine the effect of number of cases, number of waves, magnitude of growth parameter, and **G** matrix on the hit rate of SRMR.

Table 3. Four-Way Analysis of Variance for AIC and BIC Hit Rate

	AIC hit rate				BIC hit rate				
Factor levels	M	SD	F (5, 18)	η^2	M	SD	F (5, 18)	η^2	n
# of cases			455.37**	0.2904			313.16**	0.289	
-30	53.42	13.71			56.67	15.05			22476
-210	69.17	11.94			73.67	12.7			23362
# of waves			1094.39**	0.698			751.42**	0.6935	
-4	49.08	9.33			52	10.39			22334
-8	73.5	7.33			78.33	7.69			23504
Magnitude of growth parameter			0.05	0.0001			0.21	0.0004	
0	61.38	16.08			64.75	17.36			15269
-0.05	61.38	15.24			65.25	15.59			15284
-0.16	61.13	15.53			65.5	16.73			15285
T matrix			0.11	0.0001			0.48	0.0004	
- Medium	61.42	15.56			65.5	17.01			22900
- Small	61.17	14.95			64.83	15.99			22938

**p < .0001 for two-tailed test.

Table 4. SRMR Hit Rate

Model	# of cases	# of waves	Magnitude of growth parameter	**T** matrix	Correct classification (%) ID	TOEP(2)	AR(1)	ARMA(1,1)	Average	n	Convergence
A	30	4	0.00	Medium	89	82	66	25	66	1784	89
B	30	4	0.00	Small	85	87	68	27	67	1802	90
C	30	4	0.05	Medium	86	82	66	24	65	1795	90
D	30	4	0.05	Small	82	86	68	29	67	1802	90
E	30	4	0.16	Medium	87	83	66	26	66	1809	90
F	30	4	0.16	Small	81	89	68	32	68	1815	91
G	30	8	0.00	Medium	88	78	63	56	71	1947	97
H	30	8	0.00	Small	86	87	64	60	74	1953	98
I	30	8	0.05	Medium	86	77	61	49	67	1935	97
J	30	8	0.05	Small	84	86	63	60	73	1940	97
K	30	8	0.16	Medium	86	77	65	56	70	1947	97
L	30	8	0.16	Small	86	89	67	60	75	1947	97
M	210	4	0.00	Medium	97	100	84	57	84	1923	96
N	210	4	0.00	Small	97	100	85	64	86	1915	96
O	210	4	0.05	Medium	97	100	86	58	85	1925	96
P	210	4	0.05	Small	97	100	84	63	86	1936	97
Q	210	4	0.16	Medium	96	100	85	63	86	1915	96
R	210	4	0.16	Small	96	100	85	66	87	1913	96
S	210	8	0.00	Medium	98	100	99	97	98	1974	99
T	210	8	0.00	Small	98	100	97	99	98	1971	99
U	210	8	0.05	Medium	98	100	99	97	99	1973	99
V	210	8	0.05	Small	97	100	98	96	98	1978	99
W	210	8	0.16	Medium	97	100	99	97	98	1973	99
X	210	8	0.16	Small	98	100	98	99	99	1966	98
Total					91	92	79	61	81	45838	95

The result of ANOVA test is shown in Table 5. Like the ANOVA test for AIC and BIC, magnitude of growth parameter and **G** matrix had small F values and hardly contributed to the variation in the SRMR hit rate. Number of cases ($F(5, 18) = 503.05$, $p < .01$) and number of waves cases ($F(5, 18) = 76.16$, $p < .01$) were significant factors in the hit rate for SRMR. However, unlike AIC and BIC, number of cases was the largest factor and explained 84% of the variance in SRMR hit rate while number of waves accounted for 13% of the variability in the SRMR hit rate. In the same vein, models with 210 cases (92%) had a higher hit rate than models with 30 cases (69.08%). Eight-wave models (85%) had higher hit rate than 4-wave models (76.08%).

3.3.4 Likelihood Ratio Test

The hit rate of LRT was investigated in four conditions only due to complexity of the two-stage process in selecting the optimal within-subject covariance structure. The overall correct classification was 69% as shown in the bottom of Table 6. The ID within-subject covariance structure had an average hit rate of 91% in LRT, TOEP(2) 93%, and AR(1) 75%. The drop in the LRT hit rate in AR(1) was due to the sharp decrease of the LRT hit rate in the condition with 30 cases and 4 waves, which was 34%. The ARMA(1,1) structure had an average LRT hit rate of 17%.

As we took a closer examination of the LRT hit rate for ARMA(1,1), the 210-case and 8-wave combination design had the highest LRT hit rate, 65%. The simulation design condition of 30 cases and 4 waves setting for ARMA(1,1) had only a 1% hit rate. The two remaining design combinations had 0% LRT hit rate. On average, the 210-case and 8-wave combination design had the highest average LRT hit rate (85%) across all within-subject covariance structures.

Table 5. Four-Way Analysis of Variance for SRMR Hit Rate

	SRMR hit rate				
Factor levels	M	SD	F (5, 18)	η^2	n
# of cases			503.50**	0. 8363	
-30	69.03	3.42			22476
-210	92.00	6.66			23362
# of waves			76.16**	0.1266	
-4	76.08	10.06			22334
-8	85.00	14.07			23504
Magnitude of growth parameter			0.41	0.0013	
0	80.50	12.98			15269
-0.05	80.00	13.93			15284
-0.16	81.13	13.23			15285
T matrix			3.52	0.0058	
- Medium	79.58	13.65			22900
- Small	81.50	12.41			22938

**p < .0001 for two-tailed test.

Table 6. LRT Hit Rate

			Correct classification (%)					
Target Σ Matrix	# of cases	# of waves	ID	TOEP(2)	AR(1)	ARMA(1,1)	Average	n
ID	30	4	92					2367
TOEP(2)	30	4		92				2996
AR(1)	30	4			34			2861
ARMA(1,1)	30	4				1	55	2583
ID	210	4	91					2567
TOEP(2)	210	4		96				3000
AR(1)	210	4			91			3000
ARMA(1,1)	210	4				0	69	2960
ID	30	8	93					2680
TOEP(2)	30	8		87				3000
AR(1)	30	8			83			3000
ARMA(1,1)	30	8				0	65	2989
ID	210	8	89					2835
TOEP(2)	210	8		96				3000
AR(1)	210	8			90			3000
ARMA(1,1)	210	8				65	85	3000
Total			91	93	75	17	69	45835

3.4 Demonstration with Empirical Data

We used a pharmaceutical example (Littell et al., 2000; Littell, Stroup, & Freund, 2002; Littell et al., 2006) to demonstrate the performance of AIC, BIC, LRT, and SRMR in an empirical dataset. The repeated measures data described the effect of three drugs (i.e. a standard drug (A), a test drug (C), and a placebo (P)) on respiratory capability of asthma patients. The dependent variable was respiratory capability termed FEV1 and was measured for 8 consecutive hours following treatment on 24 patients in a crossover design, yielding 576 observations. For demonstration purpose, only the time effect was considered in the means model in order to investigate the underlying time series profile. Four different within-subject structures (**R**) with an unstructured **G** matrix were fit to the pharmaceutical data. Additionally, an unstructured within-subject matrix with null G matrix was fit the empirical data, serving as the overall unconditional variance covariance structure. As in the simulated data, we fit the overall model (i.e. **R** = UN, **G** = Null) instead of an unstructured within-subject structure (**R** = UN) with unstructured T matrix (**G** = UN) so as to avoid model overparametization.

3.4.1 Result for Empirical Data

Table 7 presents the -2 log likelihood, AIC, BIC, and SRMR values for the four imposed within-subject structures. Nested models can be evaluated using LRTs. All the LRTs between nested models were statistically significant ($F(4, 5) = 7$, $p = .0082$ for ID versus AR(1); $F(5, 6) = 5.9$, $p = .0151$ for AR(1) versus ARMA(1,1); $F(4, 5) = 5.9$, $p = .0151$ for ID versus TOEP(2)). The result of LRTs indicates a preference for complex models over parsimonious models. In terms of performance of AIC and BIC, ARMA(1,1) had the smallest AIC (1731.3) and BIC (1738.4) values. The differences of BIC values between ARMA(1,1) and TOEP(2) (ΔBIC=3.8) and ARMA(1,1) and AR(1) (ΔBIC=2.7) suggested positive differences between the competing models while the change of BIC between ARMA(1,1) and ID (ΔBIC=6.5) corresponded to a strong model difference, according to Raftery's guideline in interpreting model comparison with BIC (Raftery, 1995). Regarding the performance of SRMR, three of the four covariance structures had SRMR values greater than the traditional .08 cutoff criterion (i.e. ID = .092; AR(1)= .087; TOEP(2) = .088) while the SRMR for ARMA(1,1) was the smallest among the four structures and was equal to .080. Therefore, SRMR had a congruent result as AIC and BIC and preferred ARMA(1,1) model over the other three models for the pharmaceutical data.

Table 7. Values of Fit Statistics on Four Imposed Σ with Empirical Data.

Fit Statistics	ID	AR(1)	ARMA(1,1)	TOEP(2)
-2LL	1732.2	1725.2	1719.3	1726.3
AIC	1740.2	1735.2	**1731.3**	1736.3
BIC	1744.9	1741.1	**1738.4**	1742.2
SRMR	0.0917	0.0874	**0.0801**	0.0880

3.5 Discussion

Our findings suggested the SRMR had the best average and individual performance in searching for the optimal within-subject variance covariance matrix. In the simulated data, the average SRMR hit rate was 81%. In particular, the SRMR correct identification for ARMA(1,1) outperformed all the other fit indices classification for ARMA(1,1). SRMR hit rate (61%) was higher than AIC (41%), BIC (28%), or LRT(17%). For an ID structure BIC had the best hit rate but the worst for an ARMA(1,1) structure, which corresponded to the fact that BIC penalizes additional parameters being estimated and favors parsimonious models. AIC also penalizes complex models but the size of the penalty (i.e. 2 times the number of parameter estimated) was not as striking as that for BIC, and thus AIC usually selects the less complex model such as TOEP(2). The ANOVA test revealed that number of cases and number of waves played a major role in the variability of the hit rate in these fit statistics. For AIC and BIC, number of waves was a more significant factor than number of cases while for SRMR number of cases accounted for more variation in the hit rate. In the analysis of the empirical dataset, LRTs favored more complex models; however, due to the fact that LRTs can only be used for nested models we could determine if ARMA(1,1) or TOEP(2) was a better model. On the other hand, the SRMR agreed with AIC and BIC in selecting the optimal variance structure as ARMA(1,1), which had the smallest value among the four structures.

The performance of these fit statistics is associated with their inherent natural fit categories. SRMR is categorized as the absolute fit index or more specifically an absolute misfit index (Browne, MacCallum, Kim, Andersen, & Glaser, 2002) because it decreases as the fit of the model to the data increases (Byrne, 2006).SRMR signifies the average discrepancy between the observed/unconditional sample and hypothesized correlation matrices with a value of zero indicating perfect fit (Byrne, 2006). On the other hand, AIC and BIC are relative fit statistics with lower information index value indicating a better fit among competing models. Following this vein, we learned that SRMR measures the difference between a hypothesized and an observed/unconditional covariance structures and reveals the extent of similarity between the two covariance structures while information criteria tell us what model fits better but fail to show the degree of similarity between the model-implied structure and the unconditional structure. The current study evaluated SRMR, a measure of absolute fit, in selecting the optimal variance covariance structure in MLM and found SRMR outperformed other fit statistics in the covariance structure selection.

4. EVALUATING THE IMPACT OF DIFFERENT TYPES/LEVEL OF MISSPECIFICATION IN THE WITHIN- AND BETWEEN-SUBJECT VARIANCE-COVARIANCE MATRICES IN MULTILEVEL MODELS WITH LONGITUDINAL DATA

4.1 Theoretical Framework

The technique of multi-level modeling (MLM) is commonly used to analyze repeated measurement data in various disciplines, where multiple observations are collected on the same participant over time (Littell et al., 2000). Using MLM for analyzing repeated measures has several advantages over the classical statistical methods. For example, MLM does not require data to be balanced and allows missing data or uneven data points (Luke, 2004). Additionally, observations need not to be taken equidistantly. In other words, data can be collected at various time points for different individuals. Most importantly, as the focus of this study, the error structure can be flexibly modeled in MLM. Traditional Repeated Measures Univariate Analysis of Variance (UANOVA) requires the sphericity assumption (with a compound symmetry V-C structure as the sufficient condition) which may not be suitable for longitudinal data given that measures within a subject tend to correlate over time and the association diminishes as lags over time increase (Hedeker & Gibbons, 2006). On the other hand, Multivariate Analysis of Variance (MANOVA) assumes an unconditional V-C structure by estimating all the unique elements in the V-C matrix, which results in relatively low statistical power. MLM represented as mixed effect models allows the error structure to be divided into two parts, a between-subject and a within-subject error variance. Due to the independence of random effects from residuals, the **G** and **R** matrices can be flexibly modeled and conform to the structure of sample data.

4.2 Purpose of This Study

The current study adopted a first-order autoregressive variance-covariance (V-C) structure as the true within-subject covariance structure and was intended to investigate the consequence of misspecifying simultaneously different levels of the variance-covariance matrices on the estimation and tests of significance of the growth/fixed-effect and random-effect parameters. Evaluation criteria include convergence rate, Type I error rate, statistical power, and relative bias of the fixed effects and their corresponding standard errors. Within the multilevel modeling framework, the total variance in the outcome variable is an additive function of the between-subject V-C structure and within-subject V-C structure. Given that the total V-C is the combination of the between-subject V-C and the within-subject V-C, a compensatory relation between the misspecifications in the between-V-C and the within-V-C may occur. For example, the impact of under-specification at one level may be balanced by over-specification at the other level. It is also possible that the impact of misspecification in V-C is not equally weighed across levels. In other words, the misspecification in the V-C at different level may have differential impact on the estimation and tests of significance of the fixed- and random-effect parameters. Based on the three possible types of misspecification proposed by Kwok and colleagues (2007), our research questions include:

1. What is the effect of an over-specified between-subject V-C matrix and an under-specified within-subject V-C matrix (Over G,

Under R) on the estimation and tests of significance of the fixed and random parameters?

2. What is the effect of an under-specified between-subject V-C matrix and an over-specified within-subject V-C matrix (Under G, Over R) on the estimation and tests of significance of the fixed and random parameters?
3. What is the effect of the generally misspecified between- and within-subject matrices (Generally misspecified G&R) on the estimation and tests of significance of the fixed and random parameters?

Design factors including number of waves of repeated measurement, number of participants, the magnitude of the fixed effect/growth parameters, and the magnitude of the autocorrelation parameter (ρ) were considered.

4.3 Method

The study employed the Monte Carlo Simulation approach in a 2 (number of cases: 30 and 210) by 2 (number of waves: 4 and 8) by 3 magnitude of growth parameters (β_1= 0, .05, 0.16) by 3 size of autocorrelation parameter (ρ = 0.2, 0.5 and 0.8) factorial design to examine the effect of different combination of under-specification, over-specification, or general misspecification in the between- and within-subject variance-covariance structure on the evaluation criteria. Selection of the levels of design factors are based on the past simulation studies (Ferron et al., 2002; Keselman et al., 1998; Kwok et al., 2007), the review of the multiwave longitudinal studies (Khoo et al., 2006), and the study by Raudenbush & Liu (2001). 500 replications were generated for each of the 36 simulation settings, yielding 18,000 datasets. The 18,000 datasets were then analyzed for three misspecified covariance structures and a correctly specified structure using SAS PROC MIXED (SAS Institute, 2002, 2008). Restricted maximum likelihood (REML) estimation method was used for generating the variance-covariance structures. For balanced data, REML solutions are the minimal variance unbiased estimators taking into account of the degrees of freedom lost for estimating the fixed effects and correcting the downward bias produced by full information maximum likelihood (ML) (Diggle, Liang, & Zeger, 1994; Smyth & Verbyla, 1996). A growing preference for REML over ML was observed for obtaining covariance parameter estimates (McCulloch & Searle, 2001). In terms of degrees of freedom, DDFM = KR(Firstorder) in SAS PROC MIXED is used. DDFM = KR(Firstorder) computes Satterthwaite-type degree of freedom based on the adjusted covariance matrix and eliminates the second derivatives from the calculation of the covariance matrix adjustment at the same time; thus, it is preferred for V-C structures that have nonzero second derivatives, such as AR(1) and ARMA (1,1). Other specification and evaluation of the simulation study are discussed below.

4.3.1 Model Specification

We considered a simple linear growth curve model with level-1 outcome variable as a function of time. For level-2 we modeled the variability in the common intercept and common slope. We focused our study on a correctly specified mean model with balanced design data only. The level-1 model can be represented as

$$y_{ti} = \beta_{0i} + \beta_{1i}Time_{ti} + e_{ti} \tag{12}$$

$$\text{with } e_{ti} \sim N\left[(0),\ \text{R}=\sigma^2\begin{bmatrix}1 & \rho & \rho^2 & \rho^2\\ \rho & 1 & \rho & \rho^2\\ \rho^2 & \rho & 1 & \rho\\ \rho^2 & \rho^2 & \rho & 1\end{bmatrix}\right]$$

where y_{ti} y_{ti} is the outcome of observation at time t for the ith individual, β_{0i} β_{0i} is the baseline status of the ith individual, β_{1i} β_{1i} is the linear growth rate of ith individual, $Time_{ti}$ $Time_{ti}$ is the time of the tth observation for the ith individual, and e_{ti} e_{ti} is the error term corresponding to the ith individual at time *t* with mean zero and variance $\sigma_e^2 = 1$. The within-subject structure of e_{ti} e_{ti} is denoted as **R**. In our study, the true within-subject covariance structure for our simulated data is an AR(1) structure, which is a commonly used covariance structure in longitudinal data analysis (Velicer & Fava, 2003; West & Hepworth, 1991). The AR(1) model was favored by Chi and Reinsel (1989) over other time series model due to its presentation of a more parsimonious correlation in addition to the random effects and a more appropriate display of the repeated measurement data.

The level-2 model is presented below with level-1 coefficient representing initial status (β_{0i}) and change of growth rate (β_{1i}) as the outcome variables in an unconditional random intercept and random slope model:

$$\beta_{0i} = \gamma_{00} + U_{0i} \qquad (13)$$

$$\beta_{1i} = \gamma_{10} + U_{1i} \qquad (14)$$

$$\text{with } \begin{pmatrix}U_{0i}\\ U_{1i}\end{pmatrix} \sim N\left[\begin{pmatrix}0\\ 0\end{pmatrix}, \boldsymbol{G} = \begin{bmatrix}\tau_{00} & 0\\ 0 & \tau_{11}\end{bmatrix}\right]$$

where γ_{00} is the average intercept with U_{0i} capturing the variability in initial status across all individuals. γ_{10} depicts the average rate of change of the outcome variable over time with U_{1i} quantifying the variation in the rate of change (or slope) across all individuals. U_{0i} has mean zero and variance τ_{00} = 0.2. U_{1i} has mean zero and variance τ_{11} = 0.1 . We adopted this medium between-subject V-C matrix from Raudenbush and Liu (2001) but assumed no covariance (i.e. $\tau_{01} = \tau_{10} = 0$) between U_{0i} and U_{1i}. The size of t τ_{11} is half of τ_{00} because the variation in the intercept is usually smaller than that in the slope (Kwok et al., 2007).

Based on our simulated data, we developed three combinations of misspecification in the between- and within- covariance structure for this study. As shown in Figure 2, the true model

True Model: True between- and within- subject covariance structure

G: UN(1) R: AR(1)

$$\boldsymbol{G} = \begin{matrix} \tau_{00} & 0 \\ 0 & \tau_{11} \end{matrix}, \mathbf{R} = \sigma^2 \begin{matrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{matrix}$$

Under G & Over R

Over G & Under R

Generally-misspecified G&R

$$\boldsymbol{G} = \begin{matrix} \tau_{00} & 0 \\ 0 & 0 \end{matrix}, \boldsymbol{R} = \sigma^2 \begin{matrix} 1 & \rho & \gamma\rho & \gamma\rho^2 \\ \rho & 1 & \rho & \gamma\rho \\ \gamma\rho & \rho & 1 & \rho \\ \gamma\rho^2 & \gamma\rho & \rho & 1 \end{matrix}$$

Under-specified G: variation in the intercept only
Over-specified R: ARMA (1, 1)

$$\boldsymbol{G} = \begin{matrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{matrix}, \boldsymbol{R} = \begin{matrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{matrix}$$

Over-specified G: UN-structured
Under-specified R: ID

$$\boldsymbol{G} = \begin{matrix} \sigma_U^2 & \sigma_{1U} \\ \sigma_{1U} & \sigma_U^2 \end{matrix}, \boldsymbol{R} = \begin{matrix} \sigma_e^2 & \sigma_{1e} & 0 & 0 \\ \sigma_{1e} & \sigma_e^2 & \sigma_{1e} & 0 \\ 0 & \sigma_{1e} & \sigma_e^2 & \sigma_{1e} \\ 0 & 0 & \sigma_{1e} & \sigma_e^2 \end{matrix}$$

Generally-misspecified G: TOEP(2)
Generally-misspecified R: TOEP(2)

Figure 2. Illustration of True Between- and Within-Subject Variance-Covariance Structure and 3 Misspecified Conditions in the Between- and Within-Subject Variance-Covariance Structures.

has an AR(1) within-subject structure with a first-band unstructured (UN(1)) between-subject structure (i.e. only estimating τ_{00}, and τ_{11} in the between-subject covariance structure). There are three misspecified conditions: 1) Over R & Under G, an over-specified ARMA(1,1) within-subject covariance with under-specified between-subject structure where there is only variation in the intercept (i.e. τ_{00}); 2) Under R & Over G, an under-specified identity (ID) within-subject covariance structure with an over-specified UN-structured between-subject covariance matrix (i.e. estimating all unique elements in the between-subject covariance structure including $\tau_{00}, \tau_{11,} and\ \tau_{01}$); and 3) Generally misspecified G&R, TOEP(2) structures in both the within- and between-subject covariance structure.

4.3.2 Evaluation Criterion

Several evaluation criteria were used to study the effect of misspecification in the between- and within-subject covariance structure. The criteria include: 1) rate of convergence of the replications, 2) bias of the estimates for the fixed effects (i.e., β_0, and β_1) and their corresponding standard errors (i.e., $SE_{\beta 0}$, and $SE_{\beta 1}$), and 3) Type I error rate and statistical power of the test of the fixed effects.

Relative bias for fixed effects (i.e. RB_θ) was calculated with population value not equal zero (i.e., β_1 = 0.05, 0.16, and β_0 = 0.1). RB_θ is defined as

$$\text{RB}=\frac{\theta-\theta}{\theta}$$

where θ was the population parameter value and θ is the sample estimate. Simple bias, defined as $\text{B}_{simple} = \theta - \theta$ was calculated for the fixed effect with parameter values equal zero (i.e., $\beta_0 = 0$). We also computed empirical relative bias for the standard errors of the regression coefficients. A zero value of RB reflected an unbiased estimate of the parameter. A negative value indicated an underestimation of the parameter; on the other hand, a positive value indicated an overestimation of the parameter. Type I error rate was evaluated using Bradley's conservative and liberal criterion (Bradley, 1978). The conservative criterion is $0.9\alpha \leq$ alpha $\leq 1.1\alpha$ (or $.045 \leq$ alpha $\leq .055$) and the liberal criterion is $0.5\alpha \leq$ alpha$\leq 1.5\alpha$ (or $.025 \leq$ alpha $\leq .075$).

4.4 Result

We present the result of the analyses in the following order: convergence rate of analysis, relative bias and simple bias of the fixed effects, empirical bias of the standard error of fixed effects, and finally Type I error rate and empirical statistical power for the test of fixed effects. The impact of design factors on the evaluation criteria was investigated using Univariate Analysis of Variance (UANOVA) with η^2 as the effect size indicator. We reported effects with $\eta^2 > .005$ to evaluate the influence of the design factors and avoided the use of significance test to rule out trivial factors.

4.4.1 Convergence of Analysis

The 18,000 simulated dataset were analyzed for a true model and 3 misspecified conditions, including i) Under G & Over R, ii) Over G & Under R, and iii) Generally misspecified G & R. The convergence rate was 100% for the true model, Over G & Under R, and Generally misspecified G & R. Six of 18,000 replications did not converge for the Under G & Over R condition (convergence rate = 99.97%). Of the six nonconvergent observations, three were in the condition of ρ = 0.2, number of cases = 30, number of waves = 4, and $\beta_1 = 0.05$. The other three were from the condition of ρ = 0.2, number of cases = 30, number of waves = 4, and $\beta_1 = 0.16$. We considered only the results from convergent datasets for further analyses.

4.4.2 Relative Bias and Simple Bias for the Fixed Effect

Simple bias was calculated for $\beta_1 = 0$. The UANOVA test of model specification (specification) on the variation of bias for the fixed effect, controlling for the number of cases (Ncases), number of waves (Nwaves), magnitude of β_1, and size of ρ (SzRho), revealed no design factors or their interaction terms with $\eta^2 > .005$.

The mean simple bias for detecting $\beta_1 = 0$ was -0.0004 for the True model, -0.0005 for the Under G & Over R, -0.0004 for the Over G & Under R, and -0.0005 for the Generally-misspecified G&R. Relative biases were computed for $\beta_0 = .01$ and $\beta_1 \neq 0$ (i.e., β_1 = 0.05 or 0.16). The mean relative bias for detecting $\beta_1 \neq 0$ was 0.0034 for the True model, 0.0053 for the Under G & Over R, 0.0019 for the Over G & Under R, and 0.0028 for Generally-misspecified G&R. On the other hand, the mean relative bias for detecting β_0 = 0.01 was 0.0026 for the True model, 0.0001 for the Under G & Over R, 0.0068 for the Over G & Under R, and 0.0058 for the Generally-misspecified G&R.

4.4.3 Empirical Relative Bias for the Standard Error of the Fixed Effect

We computed empirical relative bias for the standard errors of the regression coefficients. The impact of model specification on the variability of empirical RB was evaluated using UANOVA, controlling for number of cases, number of waves, magnitude of β_1, and size of ρ. The results are exhibited in Table 8. The UANOVA test showed that model specification (η^2 = .0367) had a differential effect on the empirical relative bias of standard error for β_1. The empirical RB for β_1 was near zero (-6.71475E-17) for the True model, -0.1181 for the Under G & Over R, 0.0135 for the Over G & Under R, and 0.1031 for the Generally-misspecified G&R.

In addition, the interactions between the specification main effect and SzRho (η^2 = .0283) and between specification and Nwaves (η^2 = .0196), also had an η^2 greater than .005 ($p < .0001$). A three-way interaction among specification, Nwaves, and SzRho (η^2 = .0142) was also found. Under G & Over R tended to underestimate the parameter while Over G & Under R tend to overestimate the true parameter especially as size of ρ increased to .8 and number of waves increased to 8.

Table 8. Impact of Model Specification on Empirical Relative Bias for the Standard Error of $\beta 1$ under MLM

parameter	Effect with η^2 >.0050	Mean Empirical RB			
		True model	Under G & Over R	Over G & Under R	Generally-misspecified G&R
RB of $SE_{\beta 1}$	**Specification** (η^2 = .0367, p<.0001)	0	-.1181	.0135	0.1031
	Specification*SzRho (η^2 = .0283, p<.0001)			**$\rho = 0.2$**	
		0	-0.0771	0.0049	0.0014
				$\rho = 0.5$	
		0	-0.0999	0.0147	0.0048
				$\rho = 0.8$	
		0	-0.1771	0.0209	0.3029
	Specification*Nwaves (η^2 = .0196, p<.0001)			**4 waves**	
		0	-0.0381	0.0156	0.0205
				8 waves	
		0	-0.1980	0.0114	0.1856
	Specification*SzRho*Nwaves (η^2 = .0142, p<.0001)			**4 waves, $\rho = 0.2$**	
		0	0.0066	0.0094	0.0029
				8 waves, $\rho = 0.2$	
		0	-0.1608	0.0004	-0.0001
				4 waves, $\rho = 0.5$	
		0	-0.0373	0.0171	-0.0066
				8 waves, $\rho = 0.5$	
		0	-0.1626	0.0123	0.0162
				4 waves, $\rho = 0.8$	
		0	-0.0836	0.0204	0.0652
				8 waves, $\rho = 0.8$	
		0	-0.2705	0.0215	0.5406

The mean empirical RB ranged from -0.2705 to 0.0066 for Under G & Over R, and from -0.0066 to 0.5406 for Generally-misspecified G&R. The true model and Over G & Under R had near unbiased estimates of the standard error for β_1 across conditions (mean empirical RB = 0 for Under G & Over R and mean empirical RB < 0.05 for Model 2).

The result of the effect of model specification on the empirical relative bias for the standard error of β_0 was presented in Table 9. Model specification (η^2 = .6660) was found to influence the empirical RB of standard error for β_0. The mean empirical RB for β_0 was also near zero for True model (3.584151E-17), 0.6153 for Under G & Over R, 0.0144 for Over G & Under R, and -0.1522 for Generally-misspecified G&R. We observed two similar two-way interaction effects, including specification*Nwaves(η^2 = .1841) and specification*SzRho (η^2 = .0140) and a three-way interaction effect, specification*Ncases*Nwaves (η^2 = .0062).

Table 9. Impact of Model Specification on Empirical Relative Bias for the Standard Error of $\beta 0$ under MLM

parameter	Effect with η^2 >.0050	Mean Empirical RB			
		True model	Under G & Over R	Over G & Under R	Generally-misspecified G&R
RB of $SE_{\beta 1}$	**Specification** (η^2 = .6660, p<.0001)	0	.6153	.0144	-.1522
	Specification*SzRho (η^2 = .0140, p<.0001)			$\boldsymbol{\rho = 0.2}$	
		0	.6917	-.0013	-.0474
				$\boldsymbol{\rho = 0.5}$	
		0	.5893	.0149	-.1396
				$\boldsymbol{\rho = 0.8}$	
		0	.5648	.0295	-.2695
	Specification*Nwaves (η^2 = .1841, p<.0001)			**4 waves**	
		0	.2729	.0113	-.1127
				8 waves	
		0	.9577	.0175	-.1916
	Specification*SzRho*Nwaves (η^2 = .0062, p<.0001)			**4 waves, $\boldsymbol{\rho = 0.2}$**	
		0	0.2625	0.0009	-0.0263
				8 waves, $\boldsymbol{\rho = 0.2}$	
		0	1.1209	-0.0035	-0.0686
				4 waves, $\boldsymbol{\rho = 0.5}$	
		0	0.2588	0.0117	-0.0906
				8 waves, $\boldsymbol{\rho = 0.5}$	
		0	.9198	.0182	-0.1886
				4 waves, $\boldsymbol{\rho = 0.8}$	
		0	0.2973	0.0212	-0.2213
				8 waves, $\boldsymbol{\rho = 0.8}$	
		0	0.8324	0.0379	-0.3177

The impact of these factors on the empirical RB of β_0, however, had a different pattern as that on the empirical RB of β_1. Under G & Over R tended to overestimate the parameter value as size of ρ decreased and number of waves increased while Generally-misspecified G&R tended to underestimate the parameter value as size of ρ increased and number of waves increased. The mean empirical RB ranged from 0.2588 to 1.1209 for Under G & Over R and from -0.3177 to -0.0263 for Generally-misspecified G&R. The true model and Under G & Over R had near unbiased estimates of SEs for β_0 (mean empirical RB = 0 for Under G & Over R and mean empirical RB < 0.05 for Model 2).

4.4.4 Type I Error Rate of Detecting β_0 and β_1

Type I error rate was computed for conditions whose true parameter value equal 0 (i.e., β_1=0). The impact of model specification on the variation of Type I error rate was evaluated using UANOVA, controlling for design factors including number of cases, number of waves, and size of ρ and their interaction terms. The result was shown in Table 10. Model specification (η^2 = .0072) was the only design factor that had η^2 greater than .005. True model and Over G & Under R maintained a nominal alpha rate close to .05, ($\bar{\alpha}$ = .0492 for True model and $\bar{\alpha}$ = .0485 for Over G & Under R). Under G & Over R had an inflated Type I error rate ($\bar{\alpha}$ = .0913) while Generally-misspecified G&R had slightly lower Type I error rate ($\bar{\alpha}$ = .0410) in detecting β_1=0. The mean Type I error rate for Generally-misspecified G&R was within Bradley's liberal criterion (i.e., .025 ≤ alpha ≤ .075) but fell out of his conservative criterion (i.e., .045 ≤ alpha ≤ .055). On the other hand, the Type I error rates for the True model and Over G & Under R were within Bradley's conservative criterion.

4.4.5 Statistical Power of Detecting β_0 and β_1

We examined statistical power for conditions whose population parameter values are greater than zero (i.e., β_0 = 0.10, and β_1 = 0.05 or 0.16). Number of cases, number of waves, magnitude of β_1, size of ρ and their interaction terms were controlled when we used UANOVA to evaluate the impact of model specification on the statistical power for the test of β_0. Specification (η^2 = .0455), specification*Ncases (η^2 = .0080), and specification*Nwaves (η^2 = .0054) were found to influence the variability in the power for detecting β_0. On average, Under G & Over R had the lowest statistical power (.0277) while Generally-misspecified G&R had the highest statistical power (.2361) in detecting β_0. True model (.1513) and Over G & Under R (.1492) had similar statistical power in detecting β_0. Compared to the true model, Under G & Over R tended to underestimate the statistical power in detecting β_0 as the number of waves increased and number of cases decreased. On the contrary, Generally-misspecified G&R tended to overestimate the statistical power in detecting β_0 as number of waves and number of cases increased. Over G & Under R maintained a similar statistical power as the true model across conditions.

Table 10. Impact of Model Specification on the Significance Test for Linear Growth Model under MLM

parameter	Effect with η^2 >.0050	Mean Empirical RB			
		True model	Under G & Over R	Over G & Under R	Generally-misspecified G&R
Type I error rate of β_1	**Specification** (η^2 = .0072, p<.0001)	0.0492	0.0913	0.0485	0.0410
Power of β_0	**Specification** (η^2 = .0455, p<.0001)	0.1533	0.0277	0.1492	0.2361
	Specification*Nwaves (η^2 = .0054, p<.0001)		**4 waves**		
		0.1284	0.0506	0.1251	0.1882
			8 waves		
		0.1742	0.0048	0.1733	0.2840
	Specification*Ncases (η^2 = .0080, p<.0001)		**N=30**		
		0.0722	0.0108	0.0710	0.1359
			N=210		
		0.2304	0.0446	0.2274	0.3363
Power of β_1	**Specification** (η^2 = .0055, p<.0001)	0.5099	0.5695	0.5007	0.4668

The impact of model specification was examined for the statistical power in testing β_1 = 0.05 and 0.16 using UANOVA, holding number of cases, number of waves, magnitude of β_1, size of ρ, and their interaction terms constant. We observed a differential effect on model specification (η^2 = .0055) for testing $\beta_1 \neq 0$, controlling for other factors. Compared to the true model, Under G & Over R had an overstatement of the power in testing $\beta_1 \neq 0$, while Generally-misspecified G&R had an understatement of power in testing $\beta_1 \neq 0$. The true model (.5099) and Over G & Under R (.5007) had similar power. The mean power in testing $\beta_1 \neq 0$ was .5099 for True model, .5695 for Under G & Over R, .5007 for Over G & Under R, and .4668 for Generally-misspecified G&R.

4.5 Discussion

This study was intended to examine the effect of different types/levels of misspecification in the between- and within-subject V-C structures simultaneously on the estimation and tests of significance for the growth/fixed-effects and their corresponding standard errors while considering the size of the autoregressive parameters, magnitude of the growth parameters, number of cases, and number of waves. Across all models, the estimates for the fixed/growth parameters were almost completely unbiased (RB<.05). This finding was consistent with previous research when the within-subject error structure is misspecified (Ferron et al., 2002; Kwok et al., 2007; Murphy & Pituch, 2009; Sivo, Fan, & Witta, 2005), a nesting in MLM is ignored

(Moerbeek, 2004), or a cross-classified structure is not considered (Luo & Kwok, 2009). Regarding the tests of the fixed effects and the estimation of the standard errors of the growth parameters, Under G & Over R usually underestimated the standard error of β_1 but overestimated the standard error of β_0, which in turn would lead to an inflated Type I error rate and power for the test of β_1 and lower statistical power for β_0. Generally-misspecified G&R had a reverse result compared to that for Under G & Over R. Generally misspecified G&R matrices tended to overestimate the standard error of β_1 but underestimate the standard error of β_0, which in turn would lead to inflated power for the test of β_0 and lowered Type I error rate and statistical power for β_1. On the other hand, Over G & Under R had nearly unbiased estimates of standard errors for the fixed effects, maintained a Type I error rate close to 0.05, and yielded comparable statistical power to the true model in testing the significance of growth parameters. In other words, Over G & Under R V-C structures performed equally well as the correctly specified model.

5. CONCLUSIONS AND LIMITATIONS

Researchers make extensive use of MLM when analyzing longitudinal data. Several studies showed the impact of mis-specifying the within-subject covariance (Ferron et al., 2002; Kwok et al., 2007) and articles addressed the importance of modeling the optimal covariance structure (Singer & Willett, 2003) when using MLM for longitudinal data analysis. Study 1 examined the performance of common fit statistics in selecting the optimal within-subject variance covariance matrix. The averaged overall hit rates for AIC, BIC, and LRT were below 70%. The worst concern as to these fit statistics was that their stability in searching for the optimal covariance structure aggravated as the target covariance structure became more complex. The SRMR had an averaged overall hit rate of 81%. The greatest advantage of SRMR over the common fit indices was that SRMR maintained its stability in selecting the optimal within-subject structure even when the target covariance structure was complex (e.g. ARMA(1,1) hit rate for SRMR was 61%). Based on the overall and steady performance of SRMR, we concluded SRMR had better discernment in the evaluation of optimal within-subject covariance structure. However, there were some limitations in Study 1. First of all, only balanced scenarios were considered. As shown in previous studies (Keselman et al., 1998; Wolfinger, 1993), the common fit statistics did not perform well on searching for the optimal covariance structure under the general mixed model framework. Nevertheless, the overall hit rates for AIC, BIC, and LRT were still high in the current study due to the research design consisting of balanced observations exclusively. Moreover, the size of **G** matrix, the element in between-subject covariance matrix, was known in the study. The observed covariance structure in the MLM framework is composed of the within-subject covariance matrix and the between-subject covariance matrix. The combination of different between- and within-subject covariance matrices may influence the performance of the fit statistics and test statistics under examination. Future research can be conducted to investigate the performance of SRMR with unknown **G** or different specifications of **G** and include cases with unbalanced observations. As more design factors are included, we are moving toward the direction of examining the robustness of SRMR as a possible alternative for selecting the optimal within-subject covariance structure in MLM.

Misspecification of covariance structures in MLM can produce bias in the statistical inference of the results. Previous studies have only examined the effect of misspecified within-subject error structure given a correctly specified between-subject covariance structure. Few studies to date have systematically examined the effect of misspecification in both between- and within-covariance matrices. Results of study 2 showed that an Over G& Under R model specification for a linear growth curve model performed as well as a correctly specified true model with an UN(1) G & AR(1) R structure in terms of unbiasedness of fixed effects, random effects, Type I error rate, and statistical power. The finding is consistent with the study by Kwok et al. (2007) in that there is a compensatory effect when we over-specify one side of the matrix and under-specify the other side of the matrix. However, misspecified conditions will cancel each other out only when all the essential elements or the diagonal terms in the G & R matrices have been estimated. The Under G & Over R condition did not form a compensatory effect because only the variation in the intercept was estimated in the G side while variation in slope was ignored. On the other hand, the generally misspecified

G & R condition also had biased estimation in the criterion variables because the variances for the intercept and slope in the G structure were forced to be identical, but in fact the variance in the slope was only half the variance in the intercept. The Over G & Under R condition had an unstructured between-subject structure which permits the G-side matrix to be estimated freely and thus can come closer to the true model. Though the R matrix was constrained to be an identity structure, as the size of ρ in the true model becomes small the off diagonal elements in R matrix approach zero since they are exponential multiples of the parameter. The findings of study 2, however, can only be applied to the current study design and specification and cannot be generalized to other model misspecifications. More research should be conducted to evaluate different types/levels of misspecification in both the between- and within- subject variance covariance structures.

Finally, the current dissertation only evaluated multilevel analysis with a single dependent variable; thus, the simulation results may not appropriately extend to Structural Equation Modeling (SEM) models, which permit multiple outcome variables. In addition, in real situation, there may be second level dependency in the research data, for example, repeated measures nested within students and students nested within schools. Future research can be conducted to address issues related to model selection with second level dependency data and the impact of ignoring the higher level V-C structure on the criterion variables.

REFERENCES

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

Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.

Bosker, R. J., Snijders, T. A. B., & Guldemond, H. (2003). PINT (Power IN Two-level Designs): Estimating standard errors of regression coefficients in hierarchical linear models for power calculations: User's manual. *Unpublished manuscript*.

Browne, M. W., MacCallum, R. C., Kim, C. T., Andersen, B. L., & Glaser, R. (2002). When fit indices and residuals are incompatible. *Psychological Methods*, *7*(4), 403–421.

Bryk, A., & Raudenbush, S. (1992). *Hierarchical linear models for social and behavioral research: Applications and data analysis methods*. Newbury Park, CA: Sage.

Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chi, E. M., & Reinsel, G. C. (1989). Models for longitudinal data with random effects and AR (1) errors. *Journal of the American Statistical Association*, *84*(406), 452–459.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Davis, C. S. (2002). *Statistical methods for the analysis of repeated measurements*. New York, NY: Springer Verlag.

Diggle, P. J. (1988). An approach to the analysis of repeated measurements. *Biometrics*, *44*(4), 959–971.

Diggle, P., Heagerty, P., Liang, K. Y., & Zeger, S. (2002). *Analysis of longitudinal data*. New York, NY: Oxford University Press.

Diggle, P. J., Liang, K. Y., & Zeger, S. L. (1994). *Analysis of longitudinal data*. Oxford, UK: Clarendon Press.

Ferron, J., Dailey, R., & Yi, Q. (2002). Effects of misspecifying the first-level error structure in two-level models of change. *Multivariate Behavioral Research*, *37*(3), 379–403.

Goldstein, H. (1995). *Multilevel statistical models* (2nd ed.). New York: Wiley & Sons.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. (2002). On the nature of SEM estimates of ARMA parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(3), 347–368.

Hedeker, D., & Mermelstein, R. J. (2007). Mixed-effects regression models with heterogeneous variance: Analyzing ecological momentary assessment (EMA) data of smoking. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 183–206). Mahwah, NJ: LEA.

Hedeker, D. R., & Gibbons, R. D. (2006). *Longitudinal data an*alysis. Malden, MA: Wiley-Interscience.

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, *31*(2), 423–447.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55.

Jennrich, R. I., & Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, *42*(4), 805-820.

Keselman, H. J., Algina, J., Kowalchuk, R. K., & Wolfinger, R. D. (1998). A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements. *Communications in Statistics - Simulation and Computation*, *27*(3), 591-604. doi:10.1080/03610919808813497

Khoo, S. T., West, S. G., Wu, W., & Kwok, O. (2006). Longitudinal methods. In M. Eid & E. Dienner (Eds.), *Handbook of psychological measurement: A multimethod perspective* (pp. 301-317). Washington, DC: APA.

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods & Research*, *33*(2), 188-229.

Kwok, O., West, S. G., & Green, S. B. (2007). The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. *Multivariate Behavioral Research*, *42*(3), 557–592.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, *38*(4), 963-974.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberber, O. (2006). *SAS for mixed models, Second Edition* (2nd ed.). Cary, NC: SAS Publishing.

Littell, R. C., Pendergast, J., & Natarajan, R. (2000). Modeling covariance structure in the analysis of repeated measures data. *Statistics in Medicine, 19*, 1793-1819.

Littell, R. C., Stroup, W. W., & Freund, R. J. (2002). *SAS for linear models*. Cary, NC: SAS Publishing.

Longford, N. T. (1993). *Random coefficient models*. Oxford, UK: Clarendon Press.

Luke, D. A. (2004). *Multilevel modeling*. Thousand Oaks, CA: Sage Publications.

Luo, W., & Kwok, O. (2009). The impacts of ignoring a crossed factor in analyzing cross-classified data. *Multivariate Behavioral Research*, *44*(2), 182-212. doi:10.1080/00273170902794214

McCulloch, C. E., & Searle, S. R. (2001). *Generalized, linear, and mixed models*. New York, NY: Wiley-Interscience.

Moerbeek, M. (2004). The consequence of ignoring a level of nesting in multilevel analysis. *Multivariate Behavioral Research*, *39*(1), 129-149.

Murphy, D. L., & Pituch, K. A. (2009). The performance of multilevel growth curve models under an autoregressive moving average process. *The Journal of Experimental Education*, *77*(3), 255–284.

Muth 255nal of Experimental Education m*Mplus user's guide*. Los Angeles, CA: Muthmental Educ.

O'Connell, A. A., & McCoach, D. B. (2008). *Multilevel modeling of educational data*. Charlotte, NC: Information Age Publishing.

Raudenbush, S. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology*, *52*(1), 501-525.

Raudenbush, S. W., & Liu, X. (2001). Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. *Psychological Methods*, *6*, 387-401.

SAS Institute. (2002). *SAS 9.2 [Computer software]*. Cary, NC: SAS Institute, Inc.

SAS Institute. (2008). *SAS OnlineDoc® for Windows 9.2.* Cary, NC: SAS Institute Inc.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 461–464.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Sivo, S. A., Fan, X., & Witta, L. (2005). The biasing effects of unmodeled ARMA time series processes on latent growth curve model estimates. *Structural Equation*

Modeling, *12*(2), 215-231.

Sivo, S. A., & Willson, V. L. (2000). Modeling causal error structures in longitudinal panel data: A Monte Carlo study. *Structural Equation Modeling*, *7*(2), 174-205.

Smyth, G. K., & Verbyla, A. P. (1996). A conditional likelihood approach to residual maximum likelihood estimation in generalized linear models. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(3), 565–572.

Tao, J., Littel, R., Patetta, M., Truxillo, C., & Wolfinger, R. (2002). *Mixed model analyses using the SAS system course notes.* Cary, NC:.SAS Institute Inc.

Vallejo, G., Ato, M., & Valdhe SAS system course notes.nger, R. (2002). to residual maximum likelihood estimation in generalized linear models. s in *Methodology*, *4*(1), 10–21.

Velicer, W. F., & Fava, J. L. (2003). Time series analysis. In J. A. Schinka, W. F. Velicer & I. B. Weiner (Ed.), *Handbook of psychology: Research Methods in psychology* (Vol. 2, pp. 581–606). Hoboken, NJ: Wiley.

Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York, NY: Springer Verlag.

Weakliem, D. L. (2004). Introduction to the special issue on model selection. *Sociological Methods & Research*, *33*(2), 167-187.

West, S. G., & Hepworth, J. T. (1991). Data analytic strategies for temporal data and daily events. *Journal of Personality*, *59*, 609–662.

Wolfinger, R. (1993). Covariance structure selection in general mixed models. *Communications in Statistics: Simulation and Computation*, *22*(4), 1079–1106.

APPENDIX A

ANNOTATED SYNTAX FOR CALCULATING SRMR

```
ID
% cov=Z*T*Z'+R
clear element11 element22 element33 element44 element21 element31 element41 element32
element42 element43 SRMR TargetR select_case N_case element_sum element_sum_square
element_sum_square_norm Descriptor Fullinfo_ID;
%Model selection, 4 waves. ID;
t=4; % # of waves;
Z=[1 -1.5; 1 -.5;1 .5; 1 1.5];
T_large=[.2 .05;.05 .10];
T_small=[.1 .025;.025 .05];
N=length(ID1);
% ID is assumed.
i=0;
for n=1:N,
   if nwaves(n)==4,
      if teffect(n) ==1 , % large t effect;
         i=i+1;
         TargetR(i)=targetg(n);
         % Design factors;
         Descriptor(i,:)=[nwaves(n) ncases(n) beffect1(n) targetg(n) teffect(n)];
         select_case(i)=n;
         G=Z*T_large*Z'; % Computation of G matrix = ZTZ';
         % elementij is the ith row and jth column element in computing SRMR index;
         % Computation of diagonal terms of SRMR index;
         element11(i)=(u11_5(n)-res1(n)-G(1,1))/(sqrt(u11_5(n))*sqrt(u11_5(n)));
         element22(i)=(u22_5(n)-res1(n)-G(2,2))/(sqrt(u22_5(n))*sqrt(u22_5(n)));
         element33(i)=(u33_5(n)-res1(n)-G(3,3))/(sqrt(u33_5(n))*sqrt(u33_5(n)));
         element44(i)=(u44_5(n)-res1(n)-G(4,4))/(sqrt(u44_5(n))*sqrt(u44_5(n)));
         % uij_5 is the unstructured var-cov structure estimates;
         %Computation of low triangle terms;
         element21(i)=(u21_5(n)-G(2,1))/(sqrt(u22_5(n))*sqrt(u11_5(n))); % u_21;
         element31(i)=(u31_5(n)-G(3,1))/(sqrt(u33_5(n))*sqrt(u11_5(n))); % u_31;
         element41(i)=(u41_5(n)-G(4,1))/(sqrt(u44_5(n))*sqrt(u11_5(n))); % u_41;
         element32(i)=(u32_5(n)-G(3,2))/(sqrt(u33_5(n))*sqrt(u22_5(n))); % u_32;
         element42(i)=(u42_5(n)-G(4,2))/(sqrt(u44_5(n))*sqrt(u22_5(n))); % u_42;
         element43(i)=(u43_5(n)-G(4,3))/(sqrt(u44_5(n))*sqrt(u33_5(n))); % u_43;
      else            % small t effect;
         i=i+1;
         TargetR(i)=targetg(n);
         % Design factors;
         Descriptor(i,:)=[nwaves(n) ncases(n) beffect1(n) targetg(n) teffect(n)];
         select_case(i)=n;
         G=Z*T_small*Z'; % Computation of G matrix = ZTZ';
         % elementij is the ith row and jth column element in computing SRMR index;
         % Computation of diagonal terms of SRMR index;
         element11(i)=(u11_5(n)-res1(n)-G(1,1))/(sqrt(u11_5(n))*sqrt(u11_5(n)));
         element22(i)=(u22_5(n)-res1(n)-G(2,2))/(sqrt(u22_5(n))*sqrt(u22_5(n)));
         element33(i)=(u33_5(n)-res1(n)-G(3,3))/(sqrt(u33_5(n))*sqrt(u33_5(n)));
```

```
            element44(i)=(u44_5(n)-res1(n)-G(4,4))/(sqrt(u44_5(n))*sqrt(u44_5(n)));
            % uij_5 is the unstructured var-cov structure estimates;
            %Computation of low triangle terms
            element21(i)=(u21_5(n)-G(2,1))/(sqrt(u22_5(n))*sqrt(u11_5(n)));
            element31(i)=(u31_5(n)-G(3,1))/(sqrt(u33_5(n))*sqrt(u11_5(n)));
            element41(i)=(u41_5(n)-G(4,1))/(sqrt(u44_5(n))*sqrt(u11_5(n)));
            element32(i)=(u32_5(n)-G(3,2))/(sqrt(u33_5(n))*sqrt(u22_5(n)));
            element42(i)=(u42_5(n)-G(4,2))/(sqrt(u44_5(n))*sqrt(u22_5(n)));
            element43(i)=(u43_5(n)-G(4,3))/(sqrt(u44_5(n))*sqrt(u33_5(n)));
        end
    end
end
% ALL the elements are combined to calculate the SRMR index;
element_sum_square =
element11.^2+element22.^2+element33.^2+element44.^2+2*element21.^2+2*element31.^2+2*e
lement41.^2+2*element32.^2+2*element42.^2+2*element43.^2;
element_sum_square_norm=element_sum_square/(t*(t+1));
SRMR=sqrt(element_sum_square_norm)'; % Resulted SRMR index value;

Fullinfo_ID=[Descriptor, SRMR];
N_case=sum(select_case~=0)
% computation of the hit rate of SRMR index;
hit_ID=0;
for iIDhit=1:length(Fullinfo_ID),
%    iARhit
    if SRMR(iIDhit)<=0.08 && Descriptor(iIDhit,4)==1
        hit_ID=hit_ID+1;
    end
end
hit_ID
% figure of SRMR values;
figure(1)
subplot(2,1,1)
stem(SRMR')
hold on
plot(1:length(SRMR'),0.05,'y');
plot(1:length(SRMR'),0.08,'g');
hold off
title('SRMR for ID')
AXIS([0 length(SRMR)+0.05*length(SRMR) -0.05 0.5]);
subplot(2,1,2)
stem(TargetR)
AXIS([0 length(SRMR)+0.05*length(SRMR) -inf inf]);
title('TargetR')

TOEP(2)
clear element11 element22 element33 element44 element21 element31 element41 element32
element42 element43 SRMR TargetR select_case N_case element_sum element_sum_square
element_sum_square_norm Descriptor Fullinfo_Toep;
%Model selection, 4 waves. Toep(2);
t=4;
Z=[1 -1.5; 1 -.5;1 .5; 1 1.5];
T_large=[.2 .05;.05 .10];
```

```
T_small=[.1 .025;.025 .05];
N=length(ID1);
% Toep is assumed.
i=0;
for n=1:N,
   if nwaves(n)==4  ,
      if teffect(n) ==1,
         i=i+1;
         TargetR(i)=targetg(n);  % large t effect;
         % Design factors;
         Descriptor(i,:)=[nwaves(n) ncases(n) beffect1(n) targetg(n) teffect(n)];
         select_case(i)=n;
         G=Z*T_large*Z';     % Computation of G matrix = ZTZ';
         % elementij is the ith row and jth column element in computing SRMR index;
         % Computation of diagonal terms of SRMR index;
         element11(i)=(u11_5(n)-res2(n)-G(1,1))/(sqrt(u11_5(n))*sqrt(u11_5(n)));
         element22(i)=(u22_5(n)-res2(n)-G(2,2))/(sqrt(u22_5(n))*sqrt(u22_5(n)));
         element33(i)=(u33_5(n)-res2(n)-G(3,3))/(sqrt(u33_5(n))*sqrt(u33_5(n)));
         element44(i)=(u44_5(n)-res2(n)-G(4,4))/(sqrt(u44_5(n))*sqrt(u44_5(n)));
         % uij_5 is the unstructured var-cov structure estimates;
         % Computation of low triangle terms;
         % toep(n) is the sigma_1e
         element21(i)=(u21_5(n)-toep(n)-G(2,1))/(sqrt(u22_5(n))*sqrt(u11_5(n)));
         element31(i)=(u31_5(n)-G(3,1))/(sqrt(u33_5(n))*sqrt(u11_5(n)));
         element41(i)=(u41_5(n)-G(4,1))/(sqrt(u44_5(n))*sqrt(u11_5(n)));
         element32(i)=(u32_5(n)-toep(n)-G(3,2))/(sqrt(u33_5(n))*sqrt(u22_5(n)));
         element42(i)=(u42_5(n)-G(4,2))/(sqrt(u44_5(n))*sqrt(u22_5(n)));
         element43(i)=(u43_5(n)-toep(n)-G(4,3))/(sqrt(u44_5(n))*sqrt(u33_5(n)));
      else                    % small t effect;
         i=i+1;
         TargetR(i)=targetg(n);
         Descriptor(i,:)=[nwaves(n) ncases(n) beffect1(n) targetg(n) teffect(n)];
         select_case(i)=n;
         G=Z*T_small*Z';     % Computation of G matrix = ZTZ';
         % elementij is the ith row and jth column element in computing SRMR index;
         % Computation of diagonal terms of SRMR index;
         element11(i)=(u11_5(n)-res2(n)-G(1,1))/(sqrt(u11_5(n))*sqrt(u11_5(n)));
         element22(i)=(u22_5(n)-res2(n)-G(2,2))/(sqrt(u22_5(n))*sqrt(u22_5(n)));
         element33(i)=(u33_5(n)-res2(n)-G(3,3))/(sqrt(u33_5(n))*sqrt(u33_5(n)));
         element44(i)=(u44_5(n)-res2(n)-G(4,4))/(sqrt(u44_5(n))*sqrt(u44_5(n)));
         % uij_5 is the unstructured var-cov structure estimates;
         % Computation of low triangle terms;
         % toep(n) is the sigma_1e;
         element21(i)=(u21_5(n)-toep(n)-G(2,1))/(sqrt(u22_5(n))*sqrt(u11_5(n)));
         element31(i)=(u31_5(n)-G(3,1))/(sqrt(u33_5(n))*sqrt(u11_5(n)));
         element41(i)=(u41_5(n)-G(4,1))/(sqrt(u44_5(n))*sqrt(u11_5(n)));
         element32(i)=(u32_5(n)-toep(n)-G(3,2))/(sqrt(u33_5(n))*sqrt(u22_5(n)));
         element42(i)=(u42_5(n)-G(4,2))/(sqrt(u44_5(n))*sqrt(u22_5(n)));
         element43(i)=(u43_5(n)-toep(n)-G(4,3))/(sqrt(u44_5(n))*sqrt(u33_5(n)));
      end
   end
end
% ALL the elements are combined to calculate the adjested SRMR index;
```

```
element_sum_square =
element11.^2+element22.^2+element33.^2+element44.^2+2*element21.^2+2*element31.^2+2*e
lement41.^2+2*element32.^2+2*element42.^2+2*element43.^2;
element_sum_square_norm=element_sum_square/(t*(t+1));
SRMR=sqrt(element_sum_square_norm)'/2; % Resulted SRMR index value;

Fullinfo_Toep=[Descriptor, SRMR];
N_case=sum(select_case~=0)
% computation of the hit rate of SRMR index;
hit_TOEP=0;
for iTOEPhit=1:length(Fullinfo_Toep),
   if SRMR(iTOEPhit)<=0.08 && Descriptor(iTOEPhit,4)==1
      hit_TOEP=hit_TOEP+1;
   end
end
hit_TOEP
% figure of SRMR values;
figure(1)
subplot(2,1,1)
stem(SRMR')
hold on
plot(1:length(SRMR'),0.05,'y');
plot(1:length(SRMR'),0.08,'g');
hold off
title('SRMR for Toep')
AXIS([0 length(SRMR)+0.05*length(SRMR) -0.05 0.5]);
subplot(2,1,2)
stem(TargetR)
AXIS([0 length(SRMR)+0.05*length(SRMR) -inf inf]);
title('TargetR')

AR(1)
%Model selection, 4 waves. AR(1);
t=4; % # of waves;
Z=[1 -1.5; 1 -.5;1 .5; 1 1.5];
T_large=[.2 .05;.05 .10];
T_small=[.1 .025;.025 .05];
N=length(ID1);
i=0;
for n=1:N,
   if nwaves(n)==4,
      if teffect(n) ==1,  % large t effect;
         i=i+1;
         TargetR(i)=targetg(n);
         Descriptor(i,:)=[nwaves(n) ncases(n) beffect1(n) targetg(n) teffect(n)]; % Design factors;
         select_case(i)=n;
         G=Z*T_large*Z'; % Computation of G matrix = ZTZ';
         % Estimates of elements in AR(1) var-cov matrix;
         AR=zeros(4,4);
         AR=[res3(n) ar1(n) ar1(n)^2 ar1(n)^3;
            ar1(n) res3(n) ar1(n) ar1(n)^2;
            ar1(n)^2 ar1(n) res3(n) ar1(n);
            ar1(n)^3 ar1(n)^2 ar1(n) res3(n)];
```

```
        % elementij is the ith row and jth column element in computing SRMR index;
        % AR(i,j) is ith row and jth column element estimate in AR matrix;
        % Computation of diagonal terms: AR(i,i) is the diagonal term;
        element11(i)=(u11_5(n)-AR(1,1)-G(1,1))/(sqrt(u11_5(n))*sqrt(u11_5(n)));
        element22(i)=(u22_5(n)-AR(2,2)-G(2,2))/(sqrt(u22_5(n))*sqrt(u22_5(n)));
        element33(i)=(u33_5(n)-AR(3,3)-G(3,3))/(sqrt(u33_5(n))*sqrt(u33_5(n)));
        element44(i)=(u44_5(n)-AR(4,4)-G(4,4))/(sqrt(u44_5(n))*sqrt(u44_5(n)));
        % uij_5 is the unstructured var-cov structure estimates;
        %Computation of low triangle terms;
        element21(i)=(u21_5(n)-AR(2,1)-G(2,1))/(sqrt(u22_5(n))*sqrt(u11_5(n)));
        element31(i)=(u31_5(n)-AR(3,1)-G(3,1))/(sqrt(u33_5(n))*sqrt(u11_5(n)));
        element41(i)=(u41_5(n)-AR(4,1)-G(4,1))/(sqrt(u44_5(n))*sqrt(u11_5(n)));
        element32(i)=(u32_5(n)-G(3,2)-AR(3,2))/(sqrt(u33_5(n))*sqrt(u22_5(n)));
        element42(i)=(u42_5(n)-G(4,2)-AR(4,2))/(sqrt(u44_5(n))*sqrt(u22_5(n)));
        element43(i)=(u43_5(n)-G(4,3)-AR(4,3))/(sqrt(u44_5(n))*sqrt(u33_5(n)));
      else
        i=i+1;
        TargetR(i)=targetg(n); % small t effect;
        Descriptor(i,:)=[nwaves(n) ncases(n) beffect1(n) targetg(n) teffect(n)];
        select_case(i)=n;
        G=Z*T_small*Z';
        AR=zeros(4,4);
        AR=[res3(n) ar1(n) ar1(n)^2 ar1(n)^3;
          ar1(n) res3(n) ar1(n) ar1(n)^2;
          ar1(n)^2 ar1(n) res3(n) ar1(n);
          ar1(n)^3 ar1(n)^2 ar1(n) res3(n)];
        % elementij is the ith row and jth column element in computing SRMR index;
        % AR(i,j) is ith row and jth column element estimate in AR matrix;
        % Computation of diagonal terms: AR(i,i) is the diagonal term;
        element11(i)=(u11_5(n)-AR(1,1)-G(1,1))/(sqrt(u11_5(n))*sqrt(u11_5(n)));
        element22(i)=(u22_5(n)-AR(2,2)-G(2,2))/(sqrt(u22_5(n))*sqrt(u22_5(n)));
        element33(i)=(u33_5(n)-AR(3,3)-G(3,3))/(sqrt(u33_5(n))*sqrt(u33_5(n)));
        element44(i)=(u44_5(n)-AR(4,4)-G(4,4))/(sqrt(u44_5(n))*sqrt(u44_5(n)));
        % uij_5 is the unstructured var-cov structure estimates;
        %Computation of low triangle terms;
        element21(i)=(u21_5(n)-AR(2,1)-G(2,1))/(sqrt(u22_5(n))*sqrt(u11_5(n)));
        element31(i)=(u31_5(n)-AR(3,1)-G(3,1))/(sqrt(u33_5(n))*sqrt(u11_5(n)));
        element41(i)=(u41_5(n)-AR(4,1)-G(4,1))/(sqrt(u44_5(n))*sqrt(u11_5(n)));
        element32(i)=(u32_5(n)-G(3,2)-AR(3,2))/(sqrt(u33_5(n))*sqrt(u22_5(n)));
        element42(i)=(u42_5(n)-G(4,2)-AR(4,2))/(sqrt(u44_5(n))*sqrt(u22_5(n)));
        element43(i)=(u43_5(n)-G(4,3)-AR(4,3))/(sqrt(u44_5(n))*sqrt(u33_5(n)));
      end
    end
end
% ALL the elements are combined to calculate the adjested SRMR index;
element_sum_square =
element11.^2+element22.^2+element33.^2+element44.^2+2*element21.^2+2*element31.^2+2*e
lement41.^2+2*element32.^2+2*element42.^2+2*element43.^2;
element_sum_square_norm=element_sum_square/(t*(t+1));
SRMR=sqrt(element_sum_square_norm)'/2;% Resulted SRMR index value;

Fullinfo_AR=[Descriptor, SRMR];
N_case=sum(select_case~=0)
```

```
N_case_AR=ncase_AR
% computation of the hit rate of SRMR index;
hit_AR=0;
for iARhit=1:length(Fullinfo_AR),
   if SRMR(iARhit)<=0.1 && Descriptor(iARhit,4)==3
      hit_AR=hit_AR+1;
   end
end
hit_AR
% figure of SRMR values;
figure(1)
subplot(2,1,1)
stem(SRMR')
hold on
plot(1:length(SRMR'),0.05,'y');
plot(1:length(SRMR'),0.08,'g');
hold off
title('SRMR for AR(1)')
axis([0 length(SRMR)+0.05*length(SRMR) -0.05 0.5]);
subplot(2,1,2)
stem(TargetR)
axis([0 length(SRMR)+0.05*length(SRMR) -inf inf]);
title('TargetR')

ARMA(1,1)
%Model selection, 4 waves.ARMA;
t=4;
Z=[1 -1.5; 1 -.5;1 .5; 1 1.5];
T_large=[.2 .05;.05 .10];
T_small=[.1 .025;.025 .05];
N=length(ID1);
% ARMA(1,1) is assumed.
i=0;
for n=1:N,
   if nwaves(n)==4  && ncases(n)==210 && beffect1(n)==.05,
      if teffect(n) ==1,  % large t effect;
         i=i+1;
         TargetR(i)=targetg(n);
         Descriptor(i,:)=[nwaves(n) ncases(n) beffect1(n) targetg(n) teffect(n)];
         select_case(i)=n;
         G=Z*T_large*Z'; % Computation of G matrix = ZTZ';
         % Estimates of elements in ARMA var-cov matrix;
         % gamma(n) is the gamma parameter;
         % rho(n) is the rho parameter;
         ARMA=[  res4(n) gamma(n) gamma(n)*rho(n) gamma(n)*rho(n)^2;
               gamma(n) res4(n) gamma(n) gamma(n)*rho(n);
               gamma(n)*rho(n) gamma(n) res4(n) gamma(n);
               gamma(n)*rho(n)^2 gamma(n)*rho(n) gamma(n) res4(n)];
         % elementij is the ith row and jth column element in computing SRMR index;
         % ARMA(i,j) is ith row and jth column element estimate in ARMA matrix;
         % Computation of diagonal terms AR(i,i) is the diagonal term;
         element11(i)=(u11_5(n)-ARMA(1,1)-G(1,1))/(sqrt(u11_5(n))*sqrt(u11_5(n)));
         element22(i)=(u22_5(n)-ARMA(2,2)-G(2,2))/(sqrt(u22_5(n))*sqrt(u22_5(n)));
```

```
            element33(i)=(u33_5(n)-ARMA(3,3)-G(3,3))/(sqrt(u33_5(n))*sqrt(u33_5(n)));
            element44(i)=(u44_5(n)-ARMA(4,4)-G(4,4))/(sqrt(u44_5(n))*sqrt(u44_5(n)));
            % uij_5 is the unstructured var-cov structure estimates;
            %Computation of low triangle terms;
            element21(i)=(u21_5(n)-ARMA(2,1)-G(2,1))/(sqrt(u22_5(n))*sqrt(u11_5(n)));
            element31(i)=(u31_5(n)-G(3,1)-ARMA(3,1))/(sqrt(u33_5(n))*sqrt(u11_5(n)));
            element41(i)=(u41_5(n)-G(4,1)-ARMA(4,1))/(sqrt(u44_5(n))*sqrt(u11_5(n)));
            element32(i)=(u32_5(n)-ARMA(3,2)-G(3,2))/(sqrt(u33_5(n))*sqrt(u22_5(n)));
            element42(i)=(u42_5(n)-G(4,2)-ARMA(4,2))/(sqrt(u44_5(n))*sqrt(u22_5(n)));
            element43(i)=(u43_5(n)-ARMA(4,3)-G(4,3))/(sqrt(u44_5(n))*sqrt(u33_5(n)));
        else      % small t effect;
            i=i+1;
            TargetR(i)=targetg(n);
            Descriptor(i,:)=[nwaves(n) ncases(n) beffect1(n) targetg(n) teffect(n)];
            select_case(i)=n;
            G=Z*T_small*Z';  % Computation of G matrix = ZTZ';
            % Estimates of elements in ARMA var-cov matrix;
            % gamma(n) is the gamma parameter;
            % rho(n) is the rho parameter;
            ARMA=[  res4(n) gamma(n) gamma(n)*rho(n) gamma(n)*rho(n)^2;
                    gamma(n) res4(n) gamma(n) gamma(n)*rho(n);
                    gamma(n)*rho(n) gamma(n) res4(n) gamma(n);
                    gamma(n)*rho(n)^2 gamma(n)*rho(n) gamma(n) res4(n)];
            % elementij is the ith row and jth column element in computing SRMR index;
            % ARMA(i,j) is ith row and jth column element estimate in ARMA matrix;
            % Computation of diagonal terms AR(i,i) is the diagonal term;
            element11(i)=(u11_5(n)-ARMA(1,1)-G(1,1))/(sqrt(u11_5(n))*sqrt(u11_5(n)));
            element22(i)=(u22_5(n)-ARMA(2,2)-G(2,2))/(sqrt(u22_5(n))*sqrt(u22_5(n)));
            element33(i)=(u33_5(n)-ARMA(3,3)-G(3,3))/(sqrt(u33_5(n))*sqrt(u33_5(n)));
            element44(i)=(u44_5(n)-ARMA(4,4)-G(4,4))/(sqrt(u44_5(n))*sqrt(u44_5(n)));
            % uij_5 is the unstructured var-cov structure estimates;
            %Computation of low triangle terms;
            element21(i)=(u21_5(n)-ARMA(2,1)-G(2,1))/(sqrt(u22_5(n))*sqrt(u11_5(n)));
            element31(i)=(u31_5(n)-G(3,1)-ARMA(3,1))/(sqrt(u33_5(n))*sqrt(u11_5(n)));
            element41(i)=(u41_5(n)-G(4,1)-ARMA(4,1))/(sqrt(u44_5(n))*sqrt(u11_5(n)));
            element32(i)=(u32_5(n)-ARMA(3,2)-G(3,2))/(sqrt(u33_5(n))*sqrt(u22_5(n)));
            element42(i)=(u42_5(n)-G(4,2)-ARMA(4,2))/(sqrt(u44_5(n))*sqrt(u22_5(n)));
            element43(i)=(u43_5(n)-ARMA(4,3)-G(4,3))/(sqrt(u44_5(n))*sqrt(u33_5(n)));
        end
    end
end
% ALL the elements are combined to calculate the adjested SRMR index;
element_sum_square =
element11.^2+element22.^2+element33.^2+element44.^2+2*element21.^2+2*element31.^2+2*e
lement41.^2+2*element32.^2+2*element42.^2+2*element43.^2;
element_sum_square_norm=element_sum_square/(t*(t+1));
SRMR=sqrt(element_sum_square_norm)'/3; % Resulted SRMR index value;

Fullinfo_ARMA=[Descriptor, SRMR];
N_case=sum(select_case~=0)
N_case_ARMA=sum(Descriptor(:,4)==4)
% computation of the hit rate of SRMR index;
hit_ARMA=0;
```

```
for iARMAhit=1:length(Fullinfo_ARMA),
%     iARMAhit
    if SRMR(iARMAhit)<=0.1 && Descriptor(iARMAhit,4)==4
        hit_ARMA=hit_ARMA+1;
    end
end
hit_ARMA
% figure of SRMR values;
figure(1)
subplot(2,1,1)
stem(SRMR')
hold on
plot(1:length(SRMR'),0.05,'y');
plot(1:length(SRMR'),0.08,'g');
hold off
title('SRMR for ARMA(1)')
axis([0 length(SRMR)+0.05*length(SRMR) -0.05 0.5]);
subplot(2,1,2)
stem(TargetR)
axis([0 length(SRMR)+0.05*length(SRMR) -inf inf]);
title('TargetR')
```

VITA

Yuan-Hsuan Lee received her Bachelor of Arts degree in foreign languages and literature from National Tsing Hua University, Taiwan in 2002. She entered the Research, Measurement, and Statistics program in the Department of Educational Psychology at Texas A&M University in August 2006. She is the data coordinator of the State of Texas Education Research Center at Texas A&M University. Her research interests include multilevel modeling and issues related to mixed effect models.

Ms. Lee may be reached at Dept. of Educational Psychology, c/o Dr. Oi-Man Kwok, Texas A&M University, College Station, TX 77843-4225. Her email is jasvi.rms@gmail.com.