

Two-Level Time Series Modelling With Autocorrelated Error Terms

By

AZEEZ, OLASUNKANMI ISIAKA

Matric. No: 89/020537

B.Sc. (Hons), M.Sc. Statistics (Ilorin), MBA (Ado-Ekiti)

**A Thesis Submitted to the Department of Statistics, Faculty of
Physical Sciences, University of Ilorin in Partial Fulfillment of the
Requirements for the Award of the Degree of Doctor of Philosophy
in Statistics.**

Supervisor: Prof. R.A. Ipinoyomi

CERTIFICATION

This is to certify that this thesis is an original work carried out by **Isiaka Olasunkanmi AZEEZ** with matriculation number 89/020537 in the department of statistics, Faculty of physical sciences, University of Ilorin. The thesis has been read and approved as meeting the requirements for the award of Doctor of philosophy in statistics of the University of Ilorin, Ilorin, Nigeria.

Prof. R.A. Ipinyomi
(Supervisor)

Date

Dr. G.M. Oyeyemi
(Departmental PG Coordinator)

Date

(Internal External from related Department)

Date

Prof. B.L. Adeleke
(Head of Department)

Date

External Examiner

Date

ACKNOWLEDGEMENTS

It was a long and tough journey to reach this point. However, now thinking back on all the experiences I have had in the doctorate program, it was the best journey I have ever had to make me grow as a person and a researcher. This work could not have been accomplished without the assistance of a number of persons and I would like to take these few lines to thank them warmly for their help.

First and foremost, I owe my deep gratitude to my supervisor, Prof. R. A. Ipinyomi without whom accomplishment of this work would not have been made possible. Baba is a very kind, generous and open person and it was a real pleasure working with him. I can only hope that our collaboration will keep on going in the future.

Many thanks to all the academic staff of the department especially Dr Oyeyemi, Dr Yahaya and Dr Oloyede, the success of my research is the result of our collaboration.

I am very grateful to my family, whose presence and support have always been important to me especially the woman who shares my life, Olabisi for her never-ending inspiration and support. You have made the past couple of years the best of my life so far.

Finally, I gratefully acknowledge the financial support from Tertiary Education Trust Fund TETFUND for my PhD programme.

DEDICATION

In memory of my late sister, Aminat with immense love and respect

ABSTRACT

The multilevel models, viewed as linear mixed-effects models, have been adapted as a standard method of conducting analyses in repeated measure data. This mixed model may still violate some of the assumptions, such as random error independence and homogeneity of variance, thereby leading to model misspecification. However, the literature has not adequately addressed these challenges. To this end, this work was motivated to examining the effects of the violation of these assumptions and their impacts on both standard and correlated modelling techniques. The objectives of this study were to: (i) investigate the consequences of violation of non-autocorrelation assumption on the parameter estimates in two-level time series modelling techniques; (ii) examine the asymptotic behaviour of two-level time series modelling techniques when the assumption of non-autocorrelated error terms is not met; (iii) determine the robustness of two-level correlated time series modelling technique to misspecification of the true dependency structure between observations; and (iv) validate the modelling techniques on suitable real-life data sets.

Data were simulated by injecting different levels of autocorrelation at varying sample sizes. The data generated were of first-order autoregressive errors with specified parameters. The maximum likelihood and Bayes estimation methods were used to estimate the parameters. The sensitivity of model fit criteria and the robustness of the mixed effect estimates of the two-level time series models for violations of non-autocorrelation assumptions were examined. The components of the covariance structure for the first-level model were mis-specified to determine their effects on parameter estimation. The measures of performance used were mean square error, bias, and variance. A real-life data set was used to validate the modelling techniques.

The findings of the study were that:

- i. violating non-autocorrelated errors assumption affected the accuracy of the parameter estimates of both standard and correlated modelling techniques, however, the severity of

- this bias depends on the degree to which errors are autocorrelated, but adding the autocorrelation parameter improved the fit to the data under correlated model;
- ii. though REML exhibited less bias when compared with MLE, however sensitivity of both level sample sizes to a non-autocorrelation violation is not consistent
 - iii. misspecifying the within-subject covariance structure has a negative effect on the parameter estimates; and
 - iv. with the observed variance-covariance matrices, there is low levels of inferential accuracy of mixed effect estimates and increase sample sizes at both levels has no effect on the bias rate.

The study concluded that violation of model assumptions has significant effects on the accuracy of the model parameter estimates regardless of sample sizes, while a correlated model provided a better fit. The study recommended that autocorrelation should be tested for, and if found should be modelled appropriately. This study, therefore, provides additional knowledge concerning the performance of two-level time series modelling techniques when a non-autocorrelation assumption is violated.

TABLE OF CONTENTS

TITLE PAGE	i
CERTIFICATION	ii
ACKNOWLEDGEMENT	iii
DEDICATION	iv
ABSTRACT	v
TABLE OF CONTENTS	vi
AND FIGURE	xi-xvii
	LIST OF TABLES

CHAPTER ONE

1.1 INTRODUCTION	1
1.1.1 Statistical modelling	2
1.1.2 Correlated data and multilevel models.....	4
1.1.3 Types of correlated data.....	6
1.1.4 Problems inherent in correlated data.....	7
1.2 STATEMENT OF THE PROBLEM.....	9
1.2.1 Disaggregation.....	9
1.2.2 Aggregation.....	9
1.2.3 Multilevel models.....	10
1.3 OVERALL AIMS OF THE STUDY.....	14
1.3.1 Research objectives.....	15

1.4 MOTIVATION FOR THE STUDY.....	15
1.5 SIGNIFICANCE OF THE STUDY.....	15
CHAPTER TWO	
2.0 INTRODUCTION.....	17
2.1 STRUCTURE OF THE TWO-LEVEL TIME SERIES MODEL.....	17
2.2 MODEL CONCEPTS.....	18
2.3 HETEROGENEITY.....	20
2.4 RANDOM INTERCEPT TTSM.....	21
2.5 RANDOM INTERCEPT AND SLOPE TTSM.....	23
2.6 CONDITIONAL AND MARGINAL PERSPECTIVES.....	31
2.7 A SINGLE REGRESSION OVER TIME.....	33
2.8 ESTIMATION OF PARAMETERS.....	35
2.9 PARAMETERIZATIONS FOR VARIANCE-COVARIANCE MATRICES.....	36
2.10 ESTIMATION AND INFERENCE.....	37
2.11 THE BASICS OF BAYESIAN INFERENCE FOR THE TWO-LEVEL MLM.....	47
2.12 HYPOTHESIS TESTING.....	50
2.12.1 Statistical testing of the presence of random effects.....	50
2.12.2 Profile-likelihood confidence intervals.....	51
2.13 CRITERIA FOR MODEL FIT AND MODEL CHOICE.....	53
2.14 PRESENCE OF CORRELATION.....	56
2.15 MAXIMUM LIKELIHOOD ESTIMATION AND BAYESIAN ANALYSIS.....	57
2.16 LIKELIHOOD RATIO TEST FOR VARIANCE COMPONENTS.....	61
2.17 NORMALITY OF ERRORS.....	62
2.18 CONSEQUENCES OF VIOLATING RESIDUAL NORMALITY ASSUMPTIONS.....	63
2.19 HETEROGENEITY OF ERROR VARIANCE AT LEVEL-1 AND LEVEL-2.....	64
2.20 INDEPENDENCE OF A LEVEL-1 PREDICTOR AND LEVEL-1 RESIDUALS.....	65
2.21 MODELING THE COVARIANCE STRUCTURES.....	66
2.22 OTHER CONDITIONS EXAMINED WITH MONTE CARLO STUDIES FOR A TWO-LEVEL TSM.....	67
2.23 METHODOLOGICAL ISSUES IN PREVIOUS SIMULATION STUDIES.....	71
CHAPTER THREE	
3.0. INTRODUCTION.....	75
3.1. MANIPULATION CHECK OF DATA GENERATION PROCESS.....	76
3.1.1 Normality assumption for the random effects.....	76
3.1.2 Residual diagnostics.....	76
3.2. THE MTS MODELS UTILIZED IN THIS STUDY.....	77
3.2.1 The correctly specified model (the correlated (c)- model).....	78
3.2.2. The misspecified model (the standard (s) multilevel-model).....	79

3.3. MANIPULATED CONDITIONS.....	80
3.4. SELECTION OF PARAMETER VALUES.....	82
3.4.1 Reisby et. al.;77 data.....	82
3.5. DATA GENERATION AND SIMULATION PROCEDURES.....	83
3.6. STATISTICAL ANALYSIS OF ACCURACY.....	86
CHAPTER FOUR	
4.0 INTRODUCTION	90
4.1 SAMPLING DISTRIBUTION OF MTSM PARAMETER ESTIMATES FROM THE SIMULATIONS.....	91
4.1.1 Convergence and inadmissible solutions.....	91
4.2 EVALUATION OF THE GENERATED DATA SETS.....	92
4.2.1 Existence of random effects.....	102
4.2.2 Statistical testing of the presence of random effects (examining slope variation).....	103
4.2.3 The likelihood ratio test for variance components.....	112
4.3 PERFORMANCE EVALUATIONS (CRITERIA FOR MODEL FIT AND MODEL CHOICE).....	114
4.4 THE EFFECT OF NON-AUTOCORRELATED ERROR ASSUMPTION VIOLATION ON THE BIAS OF FIXED PARAMETER ESTIMATES.....	116
4.4.1 The effect of non-autocorrelated error assumption violation on the mean square error (mse) of fixed parameter estimates.....	124
4.4.2 The effect of violations of non-autocorrelation assumption on the variance of the slope.....	127
4.4.3 Systematic versus random errors from mse components.....	129
4.4.4 Comparison of the estimators.....	129
4.4.5 Evaluating error indices under misspecification of autocorrelation coefficient.....	133
4.5 THE EFFECT OF NON-AUTO CORRELATION FIRST -LEVEL ASSUMPTION VIOLATION ON THE ACCURACY OF THE RANDOM PARAMETER ESTIMATES AS A METHOD OF MODEL VALIDATION.....	137
CHAPTER FIVE	
5.0 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS	148
5.1 Effects of autocorrelated errors of the level-1 residuals.....	149
5.2 CONCLUSIONS.....	157
5.3 LIMITATIONS OF THE PRESENT STUDY.....	160
5.4 RECOMMENDATIONS.....	161
5.5 CONTRIBUTION TO KNOWLEDGE.....	163
5.6 FUTURE RESEARCH.....	164
REFERENCES.....	166-179
Appendix A1	180-183
Appendix A2	184-193

LIST OF TABLES AND FIGURE

Table 3.1: Summary of design conditions utilized in the study	81
Table 4.1: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE with autocorrelated error	94
Table 4.2: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE with autocorrelated errors	94
Table 4.3: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE without autocorrelated error	94
Table 4.4: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE without autocorrelated errors	94
Table 4.5: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.9$, with the corresponding correlation matrix under FMLE with autocorrelated errors	94
Table 4.6: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.9$, with the corresponding correlation matrix under RMLE with autocorrelated errors	95
Table 4.7: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.9$, with the corresponding correlation matrix under FMLE without autocorrelated errors	95
Table 4.8: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.9$, with the corresponding correlation matrix under RMLE without autocorrelated errors	95

Table 4.9: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.7$, with the corresponding correlation matrix under FMLE with autocorrelated errors	95
Table 4.10: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.7$, with the corresponding correlation matrix under RMLE with autocorrelated errors	95
Table 4.11: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.7$, with the corresponding correlation matrix under FMLE without autocorrelated errors	96
Table 4.12: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.7$, with the corresponding correlation matrix under RMLE without autocorrelated errors	96
Table 4.13: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE with autocorrelated errors	96
Table 4.14: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE with autocorrelated errors	96
Table 4.15: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE without autocorrelated errors	97
Table 4.16: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE without autocorrelated errors	97
Table 4.17: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.7$, with the corresponding correlation matrix under FMLE with autocorrelated errors	97
Table 4.18: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.7$, with the corresponding correlation matrix under RMLE with autocorrelated errors	98
Table 19: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.7$, with the corresponding correlation matrix under FMLE without autocorrelated errors	98
Table 4.20: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.7$, with the corresponding correlation matrix under RMLE without autocorrelated errors	98
Table 4.21: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.9$, with the corresponding correlation matrix under FMLE with autocorrelated errors	99
Table 4.22: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.9$, with the corresponding correlation matrix under RMLE with autocorrelated errors	99
Table 4.23: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.9$, with the corresponding correlation matrix under FMLE without autocorrelated errors	99
Table 4.24: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.9$, with the corresponding correlation matrix under RMLE without autocorrelated errors	100

Table 4.25: Comparison of Estimated slope variances under maximum likelihood estimator with and without autocorrelated errors	101-102
Table 4.26: Full Maximum Likelihood Parameter Estimates and Wald 95% Confidence Intervals (CIs) Obtained by Fitting Model with Autocorrelated Errors	104-105
Table 4.27: Restricted Maximum Likelihood Parameter Estimates and Wald 95% Confidence Intervals (CIs) Obtained by Fitting Model with Autocorrelated Errors	105-106
Table 4.28: Full Maximum Likelihood Parameter Estimates and Wald 95% Confidence Intervals (CIs) Obtained by Fitting Standard Multilevel Model	106-107
Table 4.29: Restricted Maximum Likelihood Parameter Estimates and Wald 95% Confidence Intervals (CIs) Obtained by Fitting Standard Multilevel Model	107-108
Table 4.30: Slope Variances when t (Number of repeated measures) is 3 at varying level of ρ and Number of Subjects.....	110
Table 4.31: Slope Variances when t (Number of repeated measures) is 5 at varying level of ρ and Number of Subjects.....	110
Table 4.32: Slope Variances when t (Number of repeated measures) is 10 at varying level of ρ and Number of Subjects.....	111
Table 4.33: Random effect variance-covariance matrix.....	113-114
Table 4.34: Criterion for comparing the REML(Ω) and REML(I) models fit under different conditions and the corresponding p-value	115
Table 4.35: Criterion for comparing the ML(Ω) and ML(I) models fit under different conditions and the corresponding p-value.....	115
Table 4.36: Comparing the bias (upper) and percentage relative bias (below) of slopes for measurement occasion, $t = 3$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope.....	117
Table 4.37: Comparing the bias (upper) and percentage relative bias (below) of slopes for measurement occasion, $t = 5$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope	117
Table 4.38: Comparing the bias (upper) and percentage relative bias (below) of slopes for measurement occasion, $t = 10$ for two estimators under standard and autocorrelated models, for different numbers of	

subject N and different autocorrelation coefficient ρ as a function of the slope.....	118
Table 4.39: Comparing the MSE of slopes for measurement occasion, $t = 3$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope	125
Table 4.40: Comparing the MSE of slopes for measurement occasion, $t = 5$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope	125
Table 4.41: Comparing the MSE of slopes for measurement occasion, $t = 10$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope	125
Table 4.42: Comparing the variance of slopes for measurement occasion, $t = 3$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	127
Table 4.43: Comparing the variance of slopes for measurement occasion, $t = 5$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	127
Table 4.44: Comparing the variance of slopes for measurement occasion, $t = 10$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope	128
Table 4.45: Comparing the bias of slopes for measurement occasion, $t = 3$ for ML estimators under the correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	130
Table 4.46: Comparing the bias of slopes for measurement occasion, $t = 5$ for ML estimators under the correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	130
Table 4.47: Comparing the bias of slopes for measurement occasion, $t = 10$ for ML estimators under the correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	130
Table 4.48: Comparing the MSE of slopes for measurement occasion, $t = 3$ for ML estimators under the correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	131

Table 4.49: Comparing the MSE of slopes for measurement occasion, $t = 5$ for ML estimators under the correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	131
Table 4.50: Comparing the MSE of slopes for measurement occasion, $t = 10$ for ML estimators under the correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	131
Table 4.51 Comparing the Variance of slopes for measurement occasion, $t = 3$ for ML estimators under the correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	132
Table 4.52: Comparing the variance of slopes for measurement occasion, $t = 5$ for ML estimators under the correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slop.....	132
Table 4.53: Comparing the variance of slopes for measurement occasion, $t = 10$ for ML estimators the under correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.....	132
Table 4.54: Comparing the bias of slopes for measurement occasion, $t = 3$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N	133
Table 4.55: Comparing the bias of slopes for measurement occasion, $t = 5$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N	134
Table 4.56: Comparing the bias of slopes for measurement occasion, $t = 10$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N	134
Table 4.57: Comparing the MSE of slopes for measurement occasion, $t = 3$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject	135
Table 4.58: Comparing the MSE of slopes for measurement occasion, $t = 5$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N	135

Table 4.59: Comparing the MSE of slopes for measurement occasion, $t = 10$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N	136
Table 4.60: Comparing the variance of slopes for measurement occasion, $t = 3$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N	136
Table 4.61: Comparing the variance of slopes for measurement occasion, $t = 5$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N	137
Table 4.62: Comparing the variance of slopes for measurement occasion, $t = 10$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N	137
Table 4.63: Observed unstructured variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix	139
Table 4.64: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE with autocorrelated errors	139
Table 4.65: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE with autocorrelated errors	140
Table 4.66: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE without autocorrelated errors	140
Table 4.67: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE without autocorrelated errors	140
Table 4.68: Observed unstructured variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix	140
Table 4.69: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix under FMLE with autocorrelated errors	141
Table 4.70: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix under RMLE with autocorrelated errors	141
Table 4.71: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix under FMLE without autocorrelated errors	141
Table 4.72: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix under RMLE without autocorrelated errors	141

Table 4.73: Observed unstructured variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix	142
Table 4.74: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix under FMLE with autocorrelated errors	142
Table 4.75: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix under RMLE with autocorrelated errors	142
Table 4.76: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix under FMLE without autocorrelated errors	143
Table 4.77: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix under RMLE without autocorrelated errors	143
Table 4.78: Observed unstructured variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix	143
Table 4.79: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix under FMLE with autocorrelated errors	144
Table 4.80: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix under RMLE with autocorrelated errors	144
Table 4.81: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix under FMLE without autocorrelated errors	144
Table 4.82: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix under RMLE without autocorrelated errors	145
Figure 1. The data structure for the two-level hierarchical model	19

CHAPTER ONE

GENERAL INTRODUCTION

1.1. Introduction

The analysis of change is a fundamental component of so many research endeavours in almost every discipline. Our general idea inferred about change and its measurement, whether the developing traits under study are physiological, morphological, psychological or behavioural, remains a challenging area (Maria, Hopkins and Stephen, 2004). Although philosophical discussions of change have been traced back as far as Aristotle (Zeger and Harlow, 1987), various approaches used earlier to investigate change were very limited in that (a) only two occasions of measurement were addressed and (b) they focused exclusively either on measurement occasion level or on subject-level growth, resulting in data gap too scanty to allow examination of some of the most basic and interesting hypotheses about change over time to be tested (Kristopher J. P., Robert C. M., and Nancy E. B., 2008)

The foremost development of statistical analysis of data over time appears to be in the nineteenth century. The studies were oriented toward the change in aggregates or groups rather than in individuals. Every discipline in the social sciences seeks to understand the sources of stability and change in variables. A psychologist, for instance, maps the development of cognitive abilities in children. An economist traces the economic growth of nations or the sales of companies. A sociologist tracks crime rates across communities. The time interval for change might be annual, monthly, weekly, or daily and the unit of analysis might be children, businesses, or countries. Regardless of the subject area or the time interval, social and behavioural scientists have a keen interest in describing and explaining the time trajectories of their variables.

1.1.1. Statistical Modeling:

Statistics is about a set of values of qualitative or quantitative variables that deal with the data obtained by counting or measuring the properties of the population. Why not simply "statistics" but "statistical modelling" or "data analysis ?" As an objective process in which the researcher should, ideally, have no influence, data speak for themselves and the purpose of statistics is to extract the information that the data contained. But many questions are much more complicated, involving many variables and a researcher does not know what is doing what to what, and what conclusion is drawn from the data depends on the active and creative role in constructing and interrogating data by the researcher.

Modelling is about building a representation of a system in the real world which allows ideas to be investigated using general rules and concepts or the set of activities used to produce a model, models are comparable idealizations to the system being represented but less complicated. According to Anu Maria (1997) describing the behaviour of one or more elements over time in order to predict the result of changes to the system represents one of the reasons for constructing a model. It is expected that a model should closely approximate the salient features of the real system of interest and incorporate most of its salient features. A statistical model is a mathematical model specified by mathematical equations or mathematical ideal conceived to represent the structure of the data. The word "modelling" highlights that our goals, our beliefs, and our current state of knowledge all influence our analysis of data and with a model that is calibrated by data statements about the world can be made.

Conceptually, some data are difficult to model because the estimate of the mean is not constant and sometimes the variance, in such a case two statistical advances are required to properly analyze the data. Firstly, to detect the effect of change and its characteristics, more

appropriate statistical tools are necessary. For example, statistical methods are needed to identify trends, with one class of statistical methods needed for trends of stepwise and the second class of statistical methods needed for abrupt changes, Secondly, to account for the nonstationarities introduced by the change more sophisticated modelling tools are needed.

Time series analysis and regression analysis provide different ways to assess the robustness of research output and represent two important statistical tools used to extract information from research data. Regression analysis is a special type of predictive modelling technique which studies the relationship between a response (dependent) and predictor variable(s) (independent) usually applied to unordered data. The focus is on the relationship between dependent variable and one or more independent variables (or 'predictors') with interest in making statements about the response variable, controlling for the other variables. Time series modelling is the analysis of data set that comes with a temporal ordering, in many cases, time is not actually used to predict the magnitude of a random variable, but the actual times that observations are taken are important because the sequence of random variables are indexed by time.

Cross-sectional studies are distinct from time series and its analysis, because there is no natural ordering of the observations, and its data commonly analyzed by regression analysis. Contrarily, time series analysis, is a process which allows us to study relationships over time, the so-called dynamic aspect of a problem where one or more subjects are identified and observed over time. Actually, comparing time series with regression analysis may not be fair, because while time series analysis is much broader and refers to analysis techniques that handle series of data points indexed in time order in which the independent variable is time (or space), regression is the most frequently used statistical methods in calibrating the coefficients of an explicit

function. In this thesis, instead of the single (long) series of observations which characterizes a time series, the study shall assume that there are several (perhaps few) relatively short series of observations, one for each individual. Clearly, the problem we are addressing is intrinsically multivariate. However, it involves a restricted form of multivariate data. The existence of multiple series gives us the advantage that the study can more readily test (rather than assume) particular structures for the covariance matrix relating observations at subsequent times.

1.1.2. Correlated Data and Multilevel Models:

Time series exhibiting a large number of interacting components whose aggregate activity is nonlinear and typically exhibits hierarchical self-organization under selective pressures known as complex systems are always controlled by multi-level laws (Xuemei Han, Congfu Xu, Huifeng Shen and Yunhe, 2003). These multi-level laws bring on the combination of long-term correlated error of measurements and short-term correlated error of measurements. Traditional analysis and forecasting methods do not distinguish these multi-level influences and always make a single model for prediction by modelling each individual separately with its own model, which has to introduce a lot of parameters to describe the characteristics of complex systems and results in the loss of efficiency or accuracy (Xuemei Han et. al., 2003). When the data contain information at more than one level, or the unit of analysis differs from the unit of randomization in the study, then the major entity that is being analyzed (unit of analysis) becomes an issue (Eqbal Z. Darandari, 2004; Gelman and Jennifer 2007; Steenbergen and Bradford 2002). With classical approaches, one must conduct the analysis on the individual or group level by restricting the data set to eliminate hierarchy. An important feature of the longitudinal multilevel model that distinguishes it from its cross-sectional counterpart is that time

generally enters the level-1 specification. There are a number of ways that this can happen. One way is to let one or more of the explanatory variables be a function of time.

Particularly in the last decade, Multilevel models have become a standard method of conducting analyses of hierarchically structured data. The relevance of multilevel models to data is characterized by scale of the outcome variable (e.g., categorical, continuous), the type of design (e.g., longitudinal with repeated measures, cross-sectional and cross-classified), the number of outcomes (e.g., univariate, multivariate), and the number of levels (e.g., 2-level or 3-level) (Aitkin and Longford, 1986; Kreft and De Leeuw, 1986; Goldstein, 1986, Raudenbush and Bryk, 1986).

Different types of multilevel models designed to analyze data generated from a nested structure have been described under a variant of names in statistical literature, such as multilevel models, mixed models, random effects models, hierarchical linear models, variance component models, random coefficient models, empirical Bayes models and random regression models among others (Goldstein, 1995; Longford, 1987; Wolfinger, 1993; Laird and Ware, 1982; Raudenbush and Bryk, 2002; Dempster, Rubin and Tsutakawa, 1981; Kreft and De Leeuw, 1986; Hui and Berger, 1983; Stenio, Weisberg and Bryk 1983; Bock, 1983a,b; Gibbons, Hedeker, Watemaux and Davis, 1988; Gibbons, Hedeker, Elkin, Waternaux, Kraemer, Greenhouse, Shea, Imber, Sotsky and Watkins 1993). Though the nomenclature is different, the underlying idea is the same. A distinguishing feature of these models is the inclusion of one or more random components of cluster effects to account for the correlational structure and the influence of the subject on repeated measures data.

1.1.3. Types of Correlated Data:

The basic purpose in a longitudinal analysis as noted by Singer and Willett (2003) is in either the within-individual change, that is how does each subject or person change over time, or change in inter-individual differences, to predict differences among subjects in their changes. The longitudinal analysis involves multiple responses taken in sequence on the same subject over time and we generally refer to these observations as repeated measures. Longitudinal data can be considered as a particular case of repeatedly measured data, where a single measurement is collected repeatedly on each subject, which may or may not be distributed over time. It can consist of many time series measured on a sample of subjects. However, it differs from the traditional time series in the sense that they have fewer repeated measurements. This sort of data gives an appealing heuristic for framing for such research questions and underpins the statistical models.

Often in applied statistics, after some empirical data have been collected, the purpose of the analysis is to construct a statistical model. Otherwise said, we are interested in situations where the aim is to explain how an outcome or response, variable of particular interest is related to a set of explanatory variables, or covariates. Classically, a single observation on the response variable is obtained for each observational unit and one of the fundamental hypotheses of standard statistical modelling, in this case, is independence between observations. Hierarchically structure observations of data can be correlated with one another, and different forms of correlated data are: (i) Multivariate measurements on different individuals, for example, blood pressure, cholesterol, etc. Note that measurements on a particular individual are correlated, (ii) Clustered measurements, for example on a number of families we take cholesterol measurements. In this situation, the measurements on the different members of a family are

correlated, (iii) Repeated measurements, for instance, we could measure cholesterol over time for several experimental conditions. Here, the measurements on each individual would be correlated, (iv) Longitudinal data, for example, we could track the cholesterol levels for individuals over time. Again, the measurements on an individual would be correlated, and (v) Spatial data, for instance, we could measure the cholesterol level for individuals living in different counties. Note that the measurements which are near to one another in space would be correlated. This study emphasizes repeated observations over time although other types of clustering are of interest.

1.1.4. Problems Inherent In Nonindependent Observations/ Group Heterogeneity:

The inherent dependence or the possibility of dependence, that is associated with longitudinal data introduces extra complications into the analysis. The traditional statistical models' assumptions ("homogeneity of regression slopes", "independence of error") characterized by data which are identically and independently distributed can no longer be relied upon with correlated data. To guaranty a valid tests and confidence intervals, in which we can have confidence, we must somehow take the possible dependence into account.

When time series data are used in regression analysis, there would be similarities between multiple tracking data from the same individual, and often the error term is not independent through time, these systematic errors might exist in the observations at the same time point, issues which made the assumption of random error independence hard to meet. Ignoring the nonindependence of the observations can lead to inefficient effect-size estimates and biased estimates of the significance of the explanatory variables.

Traditional models also assume that explanatory variable effects are stable over the entire data set, meanwhile, with repeated measures over time, dependent variable occurred with a regular increase or decrease with time, which caused the increase or decrease of the variance.

Those changes had a great effect on the homogeneity of variance. Therefore, traditional statistical techniques might lead to unreasonable or even wrong conclusions. Moreover, the hierarchical linear model does not require the assumptions of independence of random error and the homogeneity of variance, so it is more suitable for longitudinal studies.

Also, the expected difference between effects for the group and effects for the individual which is known as aggregation bias can be eliminated in multilevel models but might occur in the standard linear regression model. Aggregation bias which indicates that the conclusion that what is true for the group must be true for the sub-group or individual implies that aggregated variables results analyzed at the subject level may be different from those at the original observation level (Kreft & De Leeuw, 1998). Wu (1995) indicates that multilevel models allow explicit modelling and analysis of between and within individual variation which comprises of the expected difference between effects for the group and effects for the individual into different levels, which can be interpreted as individual-level effects (i.e., within a cluster) and cluster-level effects (i.e., across clusters), respectively.

However, as multilevel time series model for handling repeated measures data over time becomes better known and used in research and the credibility of model applications depends on the plausibility of the model's assumption, in the current research environment there is need for an extended study design on handling multilevel time series models with applications to repeated measures data by focusing on the methodological dilemmas that arise from departure from AUTOCORRELATED ERRORS assumption. Specifically, the three study conditions manipulated in order to assess their effects on model estimation, include (i) Number of individuals, (ii) number of measurement occasion and (iii) the degree of autocorrelation.

1.2. Statement of the Problem:

With repeated measures data, irrespective of the nature of the repetition, the common theme is that the residuals of individuals within the unit will be correlated resulting from repeated measures from the same subject, or observational unit, which tend to be related to one another, a fact that poses particular challenges to the statistical procedures used for the analysis of such data. Generally, where variables exist at more than one level of analysis (e.g., a lower level outcome and both lower level and higher level predictors). there are at least three approaches to analyzing hierarchically structured data:

1.2.1. Disaggregation:

Disaggregation of data deals with hierarchical data issues by moving variables to a lower level thereby ignoring the presence of group differences, that is each unit at the lower level is allotted a variable that indicates the denomination representing the unit at the higher level within which it is nested. In data disaggregation, all relationships between variables are situated at level-1 of the hierarchy and consider to be context-free. Due to the presence of possible between-group variation that were ignored, the data analysis is based on the total number of units at the lower level included in the study.

1.2.2. Aggregation:

With the issues of data hierarchy and its analysis, aggregation of data adopts different approach when compared to disaggregation: By aggregation variables at a lower level are moved to a higher level (e.g., level-2 or level-3) thereby ignoring lower level individual differences instead of ignoring higher level group differences and information about individual variability is lost (Heather W. et al., 2012; Beaubien et al., 2001; Gill, 2003; Osborne, 2000). In aggregated statistical models, different data values from many sub-units, that is within-group variation are

combined into fewer values for fewer higher-level units. In summary, by tradition choice of approach has been between an aggregated model that discards potentially meaningful lower level variance or a disaggregated model that violates statistical assumptions and assesses the impact of higher level units based on the number of lower level units. Neither of these two options seems to be satisfactory.

1.2.3. Multilevel Models:

The third major approach to dealing with hierarchically nested data structures is multilevel models. Multilevel models (MLM) are designed specifically to surmount the weakness of the aggregated and disaggregated approaches discussed above. First, using multilevel, variations within an individual- and between-individual regressions are taken into account to describe the relationship between outcome variables and the predictors. In addition to Hierarchical Linear Models (HLM's) ability to assess variation at different levels and accurately partitioned the effects of between- and within-group variance, also it is a preferred method of data analysis when data are generated from a nested data structure because of its simplifying assumptions, accommodation of missing data and small and/or discrepant group sample sizes. Therefore, both individual and group level variations in individual outcomes can be modelled while utilizing measurements predictors at the first level (measurement level) and individual predictors at the second level (individual level). The estimates of the standard errors and effect size remain undistorted and the potentially meaningful variance overlooked resulting from the use of disaggregation or aggregation is retained (Beaubien, Hamman, Holt & Boehm-Davis, 2001; Gill, 2003; Osborne, 2000). According to Kreft and De Leeuw (1998), the generation of data based on hierarchical structures increases the popularity of multilevel models among researchers in social and behavioural sciences.

Expanded application of multilevel techniques is facilitated by an array of statistical, computational, and software advances as well as evolution in the world of data collection and management that has provided researchers with an abundance of data through which these models can be brought to life (Ronald H. Heck and Scott L. Thomas, 2015). Multilevel modelling for repeated measures data is most often discussed in the context of modelling change over time. The use of change models for longitudinal data analysis has exploded over the last several years. A survey of articles published in social science journals reveals the growing popularity of MLM as an analytical choice for data prevalent in social sciences and educational research. For instance, PSYCHINFO search with the “growth model” as the keywords produced only 18 articles for the publication year 1998. By 2008, a similar search netted 98 articles, and over 36,000 articles in 2017, while that of multilevel models netted over 44,000 in 2017.

Multilevel models (MLM) has a number of benefits. First, it can help to control for clustering of observations and heteroskedasticity. Secondly, it can improve the efficiency of estimated impacts, given that the assumptions of the MLM are correct. Third, even if the assumptions are violated MLM will still produce a best “MLM” fit, similar to the Best Linear Unbiased Estimate property of an OLS model (Goldberger, 1991). Fourth, a variation of the MLM, with group mean centring, does produce unbiased slope estimates under the same conditions that are normally used to justify a fixed effects model.

The price of such a powerful model for treating hierarchically structured data is the requirement of a set of strong functional form mathematical assumptions. In general, according to Raudenbush and Bryk (2002) data are assumed to satisfy the following conditions for MLM analysis utilizing maximum likelihood estimators:

1) the level-1 residuals, e_{ij} are identical, independently and normally-distributed with zero mean and a constant variance; 2) the level-1 predictors and level-1 residuals are independent; 3) random errors at level-2 (u_{ij}) exhibit a multivariate-normal distribution with zero means and a common covariance matrix; 4) level-2 residuals and level-2 predictors are independent; 5) level-2 residuals and level-1 residuals are independent; and 6) predictors at one level are independent of residuals at another level. In addition to these conditions, both the number of groups and the within-group sample size also need to be large enough for the accuracy of the parameter estimation and significance tests.

Statistical analyses assume that population(s) and sample(s) meet certain requirements called statistical assumptions which if violated, results of hypothesis testing and confidence intervals may be incorrectly interpreted and in particular, the probability of a test statistic may be inaccurate distorting Type I or Type II error rates (Kim F. Nimon, 2012). Consequently, the underlying assumptions of MLM govern by sampling technique and the characteristics of population data must be valid in order for the estimates and associated significance tests to have the desired properties.

Results of the covariance analysis in hypothesis tests and confidence intervals (CIs) are only as valid as the assumptions, failure of which may render the results invalid and may result in misleading inferences. The appropriate analytical methods are not as well developed, especially for more sophisticated models that permit more general forms of correlation among the repeated measurements. Researchers emphasized that key assumptions should be met prior to conducting any multilevel model analysis (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999 described in Eqbal Z. Darandari, 2004). In the same vein as contained in Burstein (1989) stated by Darandari (2004), researchers are reminded that the new and more powerful methods could

produce very complex, yet very inaccurate, results if data assumptions are not carefully considered. The goal of specification analysis involves assessing model assumptions in the context of available data. Practical solutions may be limited when the reliability of a model becomes suspect due to questions about the accuracy of the various estimation methods resulting from false models assumptions. The applied researcher who regularly adopts standard procedure as his traditional statistical methods without giving thought to its associated assumptions may inadvertently be filling the literature with non-replicable results (H. J. Keselman et al., 1998).

For the interpretation of most statistical techniques to be valid, a basic requirement requires that one or more assumptions be met. In published articles violations of assumptions are rarely checked for in the first place, and if at all they were, it was regularly by means of a statistical test, leading to inadequate information being reported on whether the data satisfy the assumptions underlying the statistical techniques used or not. In sum, these motivate the importance of studying the sensitivity of estimators and associated significance tests in cases that the assumption of non-autocorrelated errors does not hold.

Indeed, Monte Carlo studies of MLM began to appear in the late 1980s and have provided researchers with knowledge concerning the performance of this model under various data conditions. A review of these studies suggests that fixed effect parameters are usually the focus of scientific interest in the multilevel model. However, it is important to correctly specify the covariance structure to obtain valid statistical inferences for the fixed effects. As stated by Altham (1984) and described in Gumedze and Dunne (2011) over-parameterization of the covariance structure may lead to inefficient estimates and poor standard errors for the fixed effects whereas a too restrictive parameterization of the covariance structure renders the inferences about the fixed effect invalid.

It is known that the invalidity of the underlying assumptions of a statistical model can lead to serious problems of bias (in estimation) and low levels of inferential accuracy when sample sizes are small and asymptotic formulae poorly represent sampling behaviour. With this in mind, this study investigates the magnitude and pattern of the bias and the logical properties of the sampling scheme in order to examine how these conditions moderate the effect of the non-autocorrelation assumption violation on parameter estimation and the robustness of statistical tests.

1.3. Overall Aim of the Study:

Limited research has been published that provides applied researchers with the way to handle clustering in their data, leading to current lack of comprehensive knowledge about the robustness of Multilevel Time Series Linear Model (MTSM) parameter estimates to a level-1 non-autocorrelated error assumption violation, particularly with repeated measures data taken over time.

Understanding the dynamic correlation structure is important for achieving efficient parameter estimators. When data on individual history are limited and individual behaviours are similar conditional on certain variables, pooling the data could yield more accurate predictions of individual outcomes than generating predictions using the data on the individual in question. Here, the focus is on understanding relations among variables observed and unobserved, and to account for subject-level heterogeneity.

In sum, the goal of the study is to determine the extent to which violation of the first-level assumption of independence of residuals result in estimation bias and inaccurate apparent estimation precision. Furthermore, this study considered other conditional factors that can cause

the robustness of model estimates and estimation precision to vary, including the strength of autocorrelation among first-level residuals and the differing sample sizes at both levels.

1.3.1. Research Objectives:

The four main research objectives addressed by this thesis are to:

- i. investigate the consequences of violation of non-autocorrelation assumption on the parameter estimates in two-level time series modelling techniques.
- ii. examine the asymptotic behaviour of two-level time series modelling techniques when the assumption of non-autocorrelated error terms is not met
- iii. determine the robustness of two-level correlated time series modelling technique to misspecification of the true dependency structure between observations
- iv. validate the modelling techniques on suitable real-life data sets.

1.4. Motivation for the study

Researchers have increasingly applied the multilevel time series model to ever smaller sample sizes. The covariance structure for the first-level model, is often assumed to be simple, while becoming a common practice, in repeated measures contexts this assumption is more questionable since errors that are close in time may be correlated, the consequences of violation of independent errors at the first level provide a substantive motivation for this study.

1.5. The Significance of the Study

While several simulation studies have been conducted to examine how multilevel model parameter estimation is affected by the violation of its assumed conditions, few attempts have been made to investigate the extent of the robustness of MTSM to one of the key first level assumption violation that can occur in the data (i.e. non-autocorrelated error assumption).

In order to represent the impact of the serial correlation structure in models commonly used for repeated measures, this study utilized a form of MTSM that account for autocorrelated errors at first level, in comparison to MTSM that do not. Most previous simulation studies of MTSM used a model structure, with only autocorrelated error terms included in level-1 distributional assumptions. For these reasons, the results of the present study provide additional knowledge concerning the performance of MTSM under conditions with which the level-1 residuals are correlated. It also sets a new benchmark for investigating the robustness of statistical tests utilized in MTSM for repeated measure data.

By examining the efficiency of some multilevel time series modelling techniques in the presence of serial correlation in repeated measures data, this study adds information about the accuracy and precision of estimates and contributes to our understanding of the behaviour of multilevel time series models under less than ideal conditions. It will not only provide policymakers and academics with a better understanding of model properties, but will also assist researchers in making better decisions about the adequacy of data, and in meeting model assumptions. It also contributes to the research methods literature by explicitly showing the consequences of violation of non-autocorrelated error assumption at level 1 and by offering recommendations to address this issue.

CHAPTER TWO

THEORETICAL FRAMEWORK AND REVIEW OF RELATED LITERATURE

2.0. Introduction

This chapter begins with a formal explication of the model and theory underlying estimation techniques and hypothesis testing procedures of the two-level time series model. The past and recent studies that are related and relevant to this study will be reviewed, and the strength and weakness of Two-Level Time Series Model when the assumption of non-autocorrelated error terms is violated is discussed in detail. In an attempt to understand the trend, attention is paid to recommendations by methodologists which have encouraged this practice. Finally, an overview of the limitations of prior simulation work is provided in order to motivate the design of this study.

2.1. Structure of the Two-Level Time Series Model (TTSM):

In this section, ideas contained in (Laird & Ware, 1982) introduced by Harville (1977) to define a family of models for serial measurements that includes both growth models and repeated-measures models as special cases will be utilized. The two general types of designs from measurements of study participants are: (i) parallel design (Longitudinal studies), also called a parallel group study, where two or more treatments are compared by assigning participants randomly to either group, treatments are administered, and then the results are compared. The main objective is to compare how the outcomes variables trajectories vary between subjects. (ii) crossover design (Repeated measures design), in which subjects receive a sequence of different treatment (or exposures). The objective is to have a more efficient comparison of treatments and compare the responses of the same subjects to different conditions.

In longitudinal studies, the outcome variable can be continuous, binary and count. Similarly, the number and timing of observations can vary among subjects. This study focuses on parallel design case with continuous outcome variable, also it is assumed for individual k that the observations are taken at integer time points $t_{k,1}, \dots, t_{k,T_p}$ which are consecutive, in order to have a time series structure" missing data" situation, is not accommodated (i.e. balanced design: $t_{kj} = t_j$ for all subjects k). The general data structure is given by $y_{tj} = t^{th}$ observation on the j^{th} subject. To appropriately model measurement occasion-level and subject-level covariates simultaneously, a two-level time series model will be developed, and the model development strategy, estimation techniques and hypothesis testing procedures will be described in the following sections.

2.2. Model Concepts:

The simplest multilevel model that can be formulated considers only two levels of analysis. The first and most elementary of these levels will be referred to as level-1 and it is on this level that the analysis is focused. The remaining level is referred to as level-2 and provides the context for the level-1 units. For instance in longitudinal medical studies in which multiple measurements of patients cholesterol level are obtained from each individual, possibly under changing experimental conditions and at different times, as illustrated by (Laird & Ware, 1982), level-1 units could be measurement repetitions or occasions that are nested within patients (level-2 units). The dependent variable is measured for level-1 units since this is the primary level of analysis. This study shall denote the dependent variable y_{tj} as the sequence of observed measurements for the j^{th} subject measured at t^{th} measurements occasions.

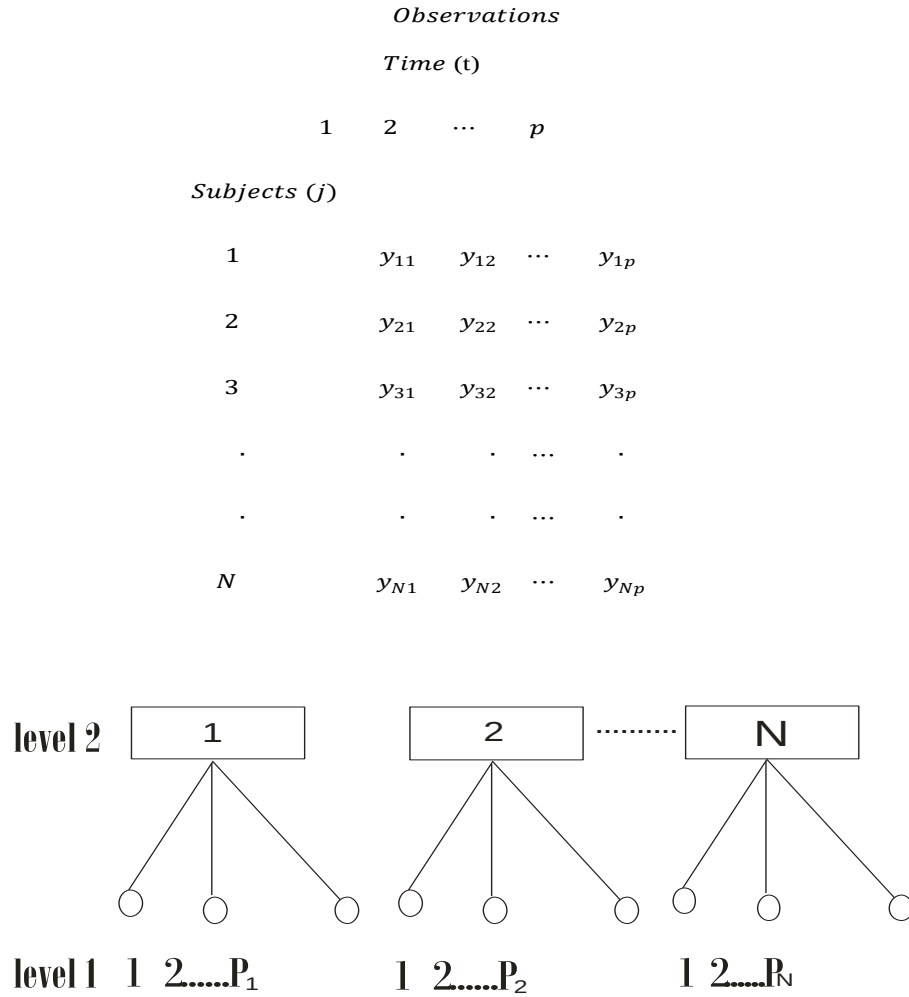


Figure 1 data structure for the two level hierarchical model.

Consider a conventional linear regression model illustrated by (Roberts, D. G., Donald, H., and Stephen D.; 2011) for the measurable dependent variable y of individual j ($j = 1, 2, \dots, N$ subjects) on occasion i ($i = 1, 2, \dots, p$ occasions), (where $i = t$)

$$y_{ij} = \beta_0 + \beta_1 X_{ij} + e_{ij} \quad \dots \quad \dots \quad \dots \quad \dots \quad (2.1)$$

The particulars of the data, namely whose observation it is (subscript j) and when was this observation made (the subscript i) are tracked by the subscripts which imply that both the outcome variable and the time variable are allowed to vary by individuals and occasions. With

the ignored subscripts, this model represents the regression of the response variable y on the independent variable, time (denoted X) and may represent time in weeks, months, etc.

In this linear regression models, like (2.1), e_{ij} are independent and identically distributed (i.i.d.) $N(0, \sigma^2)$ variates, an assumption which makes model given in equation (2.1) an unreasonable one for repeated measure data. According to Hedeker, D. and Gibbons, R. D. (2006) as pointed out by Azeez and Ipinyomi (2017), It is unreasonable because if data are dependent upon the higher-level unit then the residuals of measurements within the unit will be correlated, hence it is much more reasonable to assume that the residuals within an individual are correlated to some degree. In addition, the model given in equation (2.1) assumed that the time trajectory is the same for all individuals because the model parameters (the intercept or initial level, β_0 , and the linear change across time, β_1 ,) do not vary by individuals. Consequently, the need to incorporate individual-specific effects into the model for repeated measurements that will account for the data dependency and describe differential time trends for different individuals.

2.3. Heterogeneity:

By tracking subjects over time, it is possible to model subject behaviour in diverse forms. The two distinct approaches for modelling the quantities that represent heterogeneity (β_j) among subjects are the fixed and random effects model approach. Failure to include heterogeneity quantities in the model may introduce serious bias into the model estimators. If however, these subjects are regarded as a (random) sample from a population of subjects and we wish to make inferences about the variation between subjects in general, then a full multilevel approach is called for to describe simultaneously the relationships for several subjects and we can write, for subject j

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \cdots \cdots \cdots (2.2)$$

In particular, such a sample can provide estimates of the between-subjects variation and covariation in the slopes and intercept parameters which will allow the comparison of subjects with different characteristics.

2.4. Random Intercept TTSM:

A simple extension of the regression model given in (2.1) to allow for the influence of each individual on their repeated outcomes is provided by

$$y_{ij} = \gamma_{00} + \beta_1 X_{ij} + v_{0j} + e_{ij} \dots \dots \dots (2.3)$$

where y_{ij} is the dependent variable for observation i for the j^{th} individual, γ_{00} is the intercept, β_1 is the effect of the covariate X_{ij} , e_{ij} is the Level-1 random effect and v_{0j} represents the influence of individual j on his or her repeated observations (i.e. the Level-2 random effect). If individuals have no influence on their repeated outcomes, then all of the v_{0j} terms would equal 0. However, it is more likely that subjects will have positive or negative influences on their longitudinal data, and so the v_{0j} terms will deviate from zero. Representing the model in a multilevel form characterize the influence of subjects on their repeated observations, we have the model (2.3) partitioned into the following within subjects (or level-1) model,

$$y_{ij} = \beta_{0j} + \beta_1 X_{ij} + e_{ij} \dots \dots \dots (2.4)$$

and between-subjects (or level-2) model,

$$\begin{cases} \beta_{0j} = \gamma_{00} + v_{0j} \dots \dots \dots \\ \beta_1 = \beta_1 \end{cases} \quad (2.5)$$

Here in the first level model, individual j 's response at time i is influenced by his or her initial level γ_{00} and time trend, or slope β_1 , while in the second level model individual j 's initial level is determined by the population initial level γ_{00} , plus a unique contribution for that individual v_{0j} (the error term of all the observations in individual j). In (2.4), e_{ij} is the error term associated

with observation i in subject j . Therefore, among all observations within the same subject, the error term v_{0j} is the same, but the error term e_{ij} could be different. Thus, each individual is associated with his own distinct initial level. In the random intercept model, the intercept β_{0j} can vary across subjects, but the coefficient β_1 of the covariate X_{ij} does not differ across subjects. Consequently, there could be different intercept across subjects but the slope across subjects would be the same (Healy, 2001). Also, the random intercept model can be divided into two parts, namely the fixed effects and random effects. In Equation 2.3, regardless of observations and subjects, “ $\gamma_{00} + \beta_1 X_{ij}$ ” does not change, but “ $v_{0j} + e_{ij}$ ” could change according to subjects and observations. Due to these characteristics, the former and the latter can be referred to as fixed effects and random effects, respectively (Kreft & De Leeuw, 1998). Conversely, the present model indicates that each subject or individual’s slope is the same; all are equal to the population slope β_1

$$V(y_{ij}) = \sigma_v^2 + \sigma^2$$

$$\text{Cov}(y_{ij}, y_{ij}') = \sigma^2 \text{ where } j \neq j'$$

That is, both the variances and covariances across time are assumed to be the same. The level-2 model or between-subjects is sometimes referred to as a “slopes as outcomes” model (Burstein, 1980). The time-varying and time-invariant covariates can be considered for inclusion in the model, with the hierarchical representation showing that just as within subjects (level-1) covariates can be included in the model to explain variation in level-1 outcomes (y_{ij}), between-subjects (level-2) covariates can be included to explain variation in level-2 outcomes (the subject’s intercept β_{0j} and slope β_1).

In the model given by equation (2.4), the lower level residuals e_{ij} are now assumed to be conditionally independently and normally distributed in the population with zero mean and common variance σ^2 . According to Hedeker, D. and Gibbons, R. (2006), this assumption of conditional independence is much more reasonable than the ordinary independence assumption associated with (2.1).

2.5. Random Intercept and Slope TTSM:

According to Hedeker, D. and Gibbons, R. (2006) illustrated by Azeez & Ipinyomi (2017) it is more likely that individuals differ in their time trends, therefore very unlikely that the rate of change across time will be the same for all individuals. For these reasons, a more realistic TTSM allows both the intercept and time trend to vary by individuals. For this, the level-1 model now becomes

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \dots \dots \dots (2.6)$$

but the level-2 model is augmented as

$$\left\{ \begin{array}{l} \beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \dots \dots \dots (2.7) \\ \beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j} \dots \dots \dots (2.8) \end{array} \right.$$

$$\left\{ \begin{array}{l} \beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \dots \dots \dots (2.7) \\ \beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j} \dots \dots \dots (2.8) \end{array} \right.$$

where the β_j are modelled by the explanatory variables Z_j at the subject level. Any of the explanatory variables either the one time-level or one individual level explanatory variable may be measured at any of the levels, In the fixed part the variable can be measured at any levels, Note: we can use the same explanatory variable for both the fixed part and the random part of the model. These equations are used to model the contextual (i.e. subjects) variation in regression parameters. Both level-1 regression parameters constitute the left-hand side variables in these equations. The right-hand side variables consist of at least a minimum of a constant and typically also include at least one level-2 predictor and a disturbance.

In this model, γ_{10} is the overall population slope, γ_{00} is the overall population intercept, u_{1j} is the slope deviation for subject j and u_{0j} is the intercept deviation for subject j . As before, the residuals e_{ij} is an independent error term distributed normally with mean 0 and variance σ^2 and the assumption regarding the independence of e_{ij} is one of conditional independence, that is, they are independent conditional on u_{0j} and u_{1j} (Hedeker, D., and Gibbons, R., 2006). Since the intercept and slope coefficients are random variables that vary across the subjects, they are often referred to as random coefficients. Also, the subjects or individual-specific effects u_{ij} are treated as random effects since individuals in a sample are typically thought to be representative of a larger population of individuals (Hedeker, D., and Gibbons, R., 2006).

In equations (2.7) and (2.8), each parameter has two subscripts, if the first subscripts is "0", this is the parameter that relates to the intercept of (2.6), if the first subscript is "1", this is the parameter that relates to the slope of (2.6). Similarly, if the second subscript is "0", it means the intercept part of the second layer equation, if the second subscript is "1", it means the slope part of the second layer equation. According to Janet M. B., Henry E. B., and David C., (2008) summarizing multilevel as a single reduced-form equation, makes possible re-written of the model (2.6) as a single complex multilevel time series regression equation by substituting $\gamma_{00} + \gamma_{01}Z_j + u_{0j}$ for β_{0j} , and substituting $\gamma_{10} + \gamma_{11}Z_i + u_{1i}$ for β_{1j} , and redistributing, we have:

$$Y_{ij} = \gamma_{00} + \gamma_{01}Z_i + u_{0j} + \gamma_{10}X_{ij} + \gamma_{11}Z_iX_{ij} + u_{1j}X_{ij} + \varepsilon_{ij} \quad (2.9)$$

According to Gill (2003), rearranging is necessary, so that the fixed effects appear first, followed by the random effects, to allow for the variables classification and coefficients in terms of the hierarchical level they affect which leads us to our final mixed model, defined as

$$Y_{ij} = \gamma_{00} + \gamma_{01}Z_i + \gamma_{10}X_{ij} + \gamma_{11}X_{ij}Z_i + u_{0j} + u_{1j}X_{ij} + \varepsilon_{ij} \dots\dots\dots (2.10)$$

The term $X_{ij}Z_j$ is an interaction term that appears in the model because of modelling the varying regression slope β_{1j} of the time-level variable X_{ij} with the individual level variable Z_j .

The first four terms on the right-hand side of equation [2.10] indicate the fixed effects in the model. The first of these terms gives the intercept or constant, the second the effect of the level-2 predictor, the third the effect of the level-1 predictor, and the fourth the interactive effect between the level-1 and level-2 predictors (the so-called cross-level interaction). The latter term provides insight into how subjects (contexts) alter the impact of level-1 predictors. The last three terms in equation [2.10] are random effects – collectively they comprise the disturbance of the two-level time series model. Here, u_{0j} gives the residual contextual variance in the level-1 intercept after controlling for Z_j , $u_{1j}X_{ij}$ gives the residual individual variance in the slope for X_{ij} and e_{ij} is the usual level-1 disturbance term (capturing omitted level-1 predictors, measurement error in y_{ij} and any idiosyncratic sources of variation in y_{ij} that can be attributed to level-1 units). We can conceive of u_{0j} and $u_{1j}X_{ij}$ as *parameter noise*, and of e_{ij} as *level-1 unit noise*. Thus, prediction errors of the two-level time series model have two sources: (1) imperfect modelling of the dependent variable (e_{ij}); and (2) imperfect modelling of the level-1 parameters (u_{0j} and $u_{1j}X_{ij}$).

In equation (2.10), the terms u_{0j} and u_{1j} demonstrate that the level-1 units nested within each level-2 units are not independent and the errors are no longer independent of the level units. In addition, u_{0j} and u_{1j} may have different values within level-2 units, leading to heterogeneous variances of the error terms (Lisa, Kimberly and Elena, 1999). That is (2.10) shows that the composite error structure, $u_{1j}X_{ij} + u_{0j} + e_{ij}$ is now clearly heteroscedastic since it is conditioned on the level of the explanatory variable.

For models concerning continuous dependent variables, it is assumed that the errors in the level 1 models are normal random variables with mean zero and common variance σ^2 :

$$E(e_{ij}) = 0 \quad \text{var}(e_{ij}) = \sigma^2$$

The assumption of the level-2 model can be summarized as follows (Raudenbush & Bryk, 2002; Sullivan et al., 1999);

$$E(u_{0j}) = 0; \quad E(u_{1j}) = 0$$

$$\text{var}(u_{0j}) = \tau_{00} = \tau_0^2$$

$$\text{var}(u_{1j}) = \tau_{11} = \tau_1^2$$

$$E(\beta_{0j}) = \gamma_{00}; \quad E(\beta_{1j}) = \gamma_{01}$$

$$\text{var}(\beta_{0j}) = \text{var}(u_{0j}) = \tau_{00}; \quad \text{var}(\beta_{1j}) = \text{var}(u_{1j}) = \tau_{11}$$

$$\text{cov}(\beta_{0j}, \beta_{1j}) = \text{cov}(u_{0j}, u_{1j}) = \tau_{01} = \tau_{10}$$

$$\text{cov}(u_{0j}, \varepsilon_{ij}) = \text{cov}(u_{1j}, \varepsilon_{ij}) = 0$$

With two random individual specific effects, the population distribution of intercept and slope deviations is assumed to be bivariate normal $N(0, \Sigma_v I)$, with the random-effects variance-covariance matrix as

$$\Sigma_v = \begin{bmatrix} \tau_0^2 & \tau_{01} \\ \tau_{10} & \tau_1^2 \end{bmatrix} = D$$

equation (2.6) which described level-1 model is given by

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \dots \dots \dots (2.6 \text{ repeated})$$

while equations (2.7) and (2.8) which described level-2 models without explanatory variable are given by

$$\begin{cases} \beta_{0j} = \gamma_{00} + v_{0j} \\ \beta_{1j} = \gamma_{10} + v_{1j} \end{cases} \dots \dots \dots (2.11)$$

For estimation, we substitute (2.11) in (2.6 repeated) to yield

$$y_{ij} = (\gamma_{00} + v_{0j}) + (\gamma_{10} + v_{1j})X_{ij} + e_{ij} \dots \dots \dots (2.12)$$

$$= v_{0j} + v_{1j}X_{ij} + \gamma_{00} + \gamma_{10} X_{ij} + e_{ij} \dots \dots \dots (2.13)$$

Equation (2.13) shows that the two-level model may be written as a single linear mixed effects model. Specifically, we define $\mathbf{b}_i = (\mathbf{v}_{0j}, \mathbf{v}_{1j})'$, $X_{ij} = (1, X_{ij})'$, $\boldsymbol{\beta} = (\gamma_{00}, \gamma_{10})'$ and $X_{ij} = Z_{ij} \mathbf{v}_{0j} = \begin{bmatrix} \beta_{0j} - \gamma_{00} \\ \beta_{1j} - \gamma_{10} \end{bmatrix}$ contains random effects with expected value 0 and covariance matrix D.

Hence, in this case, the random effects $\mathbf{b}_i$ enter due to the random selection of units via the random coefficients. The overall model is

$$y_{ij} = Z_{ij}'\mathbf{b}_i + X_{ij}'\boldsymbol{\beta} + e_{ij} \dots \dots \dots (2.14)$$

$$\text{where } X_{ij} = Z_{ij} = \begin{bmatrix} 1 & x_{1j} \\ 1 & x_{2j} \\ \vdots & \vdots \\ 1 & x_{tj} \end{bmatrix}$$

is a fixed matrix which incorporates the regression model used. In this setting, where each of the coefficients in the fixed effect part of the model is random, X_{ij} and Z_{ij} are the same. The case where $X_{ij} = Z_{ij}$ is referred to as the random coefficients model (Demidenko, 2004) because all the fixed effects have a corresponding random effect.

The model (2.14), written in matrix notation, is given by

$$Y = X\beta + Zb + e \dots \dots \dots (2.15)$$

where $\mathbf{X}$ is a matrix of known coefficients of the fixed effect parameters γ_{00} , γ_{01} , γ_{10} , and γ_{11} . β is the vector of fixed effect parameters, $\mathbf{Z}$ is a matrix of coefficients (zeros and ones) of the random effects v_{ij} ; $\mathbf{b}$ is the vector of random effects v_{ij} , and $\mathbf{e}$ is the vector of the errors e_{ij} .

it is assumed that $\mathbf{b}$ and $\mathbf{e}$ are independent, and obtain

$$V(Y) = ZDZ' + R \dots \dots \dots (2.16)$$

The basic model for typical repeated measures data has been studied by several authors, among the earliest being Laird and Ware (1982). He proposed the so-called linear mixed-effects models in a biometric context, a linear regression models that include both variation that *is* explained by the independent variables of interest known as **fixed effects**, and normally distributed variation that *is not* explained by the independent variables of interest, **random effects**. A natural application is to longitudinal data where the random effects which represent unobserved factors that are common to all responses for a given subject vary between subjects and induce within-subject dependence among repeated measurements after conditioning on observed covariates. Introducing the Nan M. Laird; James H. Ware (1982) two-stage linear model

Let β denote a $p \times 1$ vector of unknown population parameters and $\mathbf{X}_i$ be a known $n_i \times p$ design matrix linking β to $\mathbf{y}_i$. Let $\mathbf{b}_i$ denote a $k \times 1$ vector of unknown individual effects and $\mathbf{Z}_i$ a known $n_i \times k$ design matrix linking $\mathbf{b}_i$ to $\mathbf{y}_i$. For measured, multivariate normal data, he proposed the following model:

Stage 1. For each individual unit, i ,

$$\mathbf{y}_i = \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i + \mathbf{e}_i, \dots \dots \dots (2.17)$$

where $\mathbf{e}_i$ is distributed as $N(\mathbf{0}, \mathbf{R}_i)$ (normal with mean $\mathbf{0}$ and covariance matrix $\mathbf{R}_i$).

$\mathbf{R}_i$ is a $n_i \times n_i$ positive-definite covariance matrix; it depends on i through its dimension n_i , but the set of unknown parameters in $\mathbf{R}_i$ will not depend upon i . At this stage, $\boldsymbol{\beta}$ and $\mathbf{b}_i$ are considered fixed, and the $\mathbf{e}_i$ are assumed to be independent (Laird and Ware, 1982).

Stage 2. The $\mathbf{b}_i$ are distributed as $N(\mathbf{0}, \mathbf{D})$, independently of each other and of the $\mathbf{e}_i$. Here $\mathbf{D}$ is a $k \times k$ positive-definite covariance matrix. The population parameters, $\boldsymbol{\beta}$, are treated as fixed effects, here the $\mathbf{b}_i$; are random effects, whose contribution to y_i are combined; through the design matrix $\mathbf{Z}_i$. Such models are sometimes called **two-stage models** because the $\mathbf{e}_i$; refer to within-subject variation (stage 1) and the $\mathbf{b}_i$; refer to between-subject variation (stage 2) (Laird and Ware, 1982). Marginally, the $\mathbf{y}_i$ are independent normal with mean $\mathbf{X}_i\boldsymbol{\beta}$ and variance-covariance matrix $\mathbf{R}_i + \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i'$. This expresses the structure of $V(\mathbf{Y})$ as a function of $\mathbf{D}$ and $\mathbf{R}_i$. The matrix $\mathbf{X}$ is assumed to be known, but the element of $\mathbf{D}$, $\mathbf{R}_i$, and possibly even $\mathbf{Z}_i$ may be functions of unobservable parameter vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)'$. $\mathbf{D}$, $\mathbf{R}$ and $\mathbf{Z}$ together imply the correlation structure for $\mathbf{Y}$, while $\mathbf{Z}\mathbf{b}$ with $\mathbf{D}$ models differences between subjects, $\mathbf{R}$ models the residual serial correlation not explained by $\mathbf{Z}\mathbf{b}$, as well as measurement error. In many repeated measures applications, $\mathbf{Z}_i\mathbf{D}\mathbf{Z}_i'$ represents the between-subject portion of the covariance structure, and $\mathbf{R}$ represents the within-subject portion.

Because (2.17) has been written in terms of the individual, this particular model formulation is referred to as the “unstacked” form. The model in (2.17) expresses the observations as a function of three terms. The term $\mathbf{X}_i\boldsymbol{\beta}$ describes the fixed contribution of the population, conditional on the predictor values, the term $\mathbf{Z}_i\mathbf{b}_i$ describes a random variation from the population value due to observing a particular subject and $\mathbf{e}_i$ describes an additional and distinct random deviation due to observing a particular subject on a particular occasion.

Further simplification of this model arises when $R_i = \sigma^2 I$, where I denotes an identity matrix. In that case, we call the model the 'conditional-independence model', since it implies that the n_i responses on Individual i are independent, conditional on b_i and β . Inference for this general linear model can be based either on least squares and maximum likelihood methods or on empirical Bayes methodology.

Many different types of linear mixed models can be constructed in which time is a variable, in this research work, the purpose of the longitudinal analysis is to see how the linear correlation of the dependent variable and time is influenced by subject-specific effects. Time may be modelled as a fixed effect to capture the linear correlation. Time can also be modelled as a random effect to capture the effect of time nested within subjects on the coefficient of time at level 1.

The linear mixed effects (LME) model are particularly used when there is no independence in the data, such as arises from a hierarchical structure and may be viewed as a generalization of the regression analysis models and variance component (VARCOMP). When the number of subjects is small and the number of observations per subject is large, subject-specific coefficients are treated as fixed, and as in the ANOVA model, ordinary regression analysis with dummy variables applies. Such a model is called a fixed effects model. Otherwise, when the number of subjects is large but the number of observations per subject is relatively small, a random effects model would be more adequate—then the subject-specific coefficients are random.

In fitting linear models to repeated measures data, the two modelling tasks are as follows:

(1) The choice of an appropriate covariance model is essential to obtain valid inference for the fixed effects so as to provide a good fit to the observed variances and covariances.

(2) We must fit a linear regression model that provides a good fit to the mean of the outcome variables

2.6. Conditional and Marginal Perspectives on the Mixed Model:

With the model in equation (2.17) for the measurements within a subject, two levels of correlation allowed are correlated errors resulting from the random effects and the within-subject variance-covariance matrix, R_i . Vonesh and Chinchilli (1997) noted that in some applications it makes sense to consider both levels and give some references where both levels are used. Chi and Reinsel (1989) also recommended the use of both levels of correlation where needed.

The model in (2.17) is tractable enough to allow the errors to be independent or correlated if the errors are correlated different correlated error structures for R_i include compound symmetry (CS), or autoregressive (AR), and Toeplitz (TOEP). For more details on the various types of correlated error structures that can be assumed for R_i , reference can be made to Ramon, C. L., George, A. M., Walter, W. S., Russell, W. (1996) or Schabenberger and Pierce (2002). Just as errors structures can be imposed on R_i , it can also be imposed on D , but in this study random effects are assumed to be uncorrelated with each other, meaning that D is restricted to be a diagonal matrix.

Also, we assume that $\text{cov}(e_i, b_i) = \emptyset$, where $\emptyset$ is a n by q matrix of zeros, that is, a distinction is made between the within-subject variation and the between-subject variation being the two sources of variation that account for differences in the response of subjects measured at the same occasion. By implication, the random effects and the random errors are uncorrelated, resulting in the conditional model given by

$$Y_i/b_i \sim N(X_i\beta + Z_ib_i, \sigma^2I), \quad b_i \sim N(0, D). \quad (2.18)$$

Interpretation: Random effects are subject-specific effects on the mean that vary within the population and are estimated subject to a regularization constraint.

Marginal perspective on the mixed model:

To simplify the estimation of model parameters in (2.17), the response vector may be expressed in terms of one purely fixed and one purely random vector.

If $e_i^* = \mathbf{Z}_i \mathbf{b}_i + \mathbf{e}_i$ in equation (2.17), then

$$\begin{aligned} \mathbf{y}_i &= \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \mathbf{e}_i \cdots \cdots \cdots \cdots \cdots (2.19) \\ &= \mathbf{X}_i \boldsymbol{\beta} + \mathbf{e}_i^* \\ &= \text{fixed} + \text{random} \\ &= E(\mathbf{y}_i) + V(\mathbf{y}_i) \end{aligned}$$

which implies that random effect $\mathbf{b}_i$ is marginalized to errors $\mathbf{e}_i$ to create $\mathbf{e}_i^*$ that is $N(\mathbf{0}, \mathbf{V})$, where

$\mathbf{V} = \mathbf{Z}_i \mathbf{D} \mathbf{Z}_i' + \mathbf{R}_i$. After marginalization process, (2.17) can be written as,

$$Y_i = X_i \boldsymbol{\beta} + e_i^* \text{ or } Y \sim N(X_i \boldsymbol{\beta}, V) \text{ where } V = \text{Cov}(Y_i) = Z_i D Z_i' + R_i \cdots \cdots \cdots (2.20)$$

The fixed component $X_i \boldsymbol{\beta}$ completely determines the mean, the first moment, of y_i , and has no effect on any variance or covariance. The random component e_i^* completely determines all variances and covariances, the second moment, and has no effect on the mean.

Equation (2.19) is called marginal model or sometimes called as the model of the population mean. The term 'marginal' is used here as an emphasizes to show that the mean response modelled is conditional only on predictor value and not on other responses (or random effects)

Interpretation: The random effects induce a correlation structure and thus enable a proper statistical analysis of the correlated data. An inference is usually based on the marginal model (2.19) for Y (unless Bayesian methods are used). Then, the unknown parameters are $(\boldsymbol{\beta}, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ contains the unknown parameters in $\mathbf{D}$ and $\mathbf{R}$ (and thus $\mathbf{V}$).

2.7. A Single Regression over Time:

One way of modelling heterogeneity is to use fixed parameters that vary by individual; this formulation is known as a *fixed effects* model, the fixed parameters induced similarities among different observations from the same subject through the mean function. Another way of modelling heterogeneity is to use parameters that vary by an individual that are represented as random quantities; these quantities are known as *random effects*, the random vectors α_i induced similarities through the covariance function. In each case, α_i is common to all observations within a subject and thus induces a dependency among these observations. Because models with random effects generally include fixed effects to account for the mean, models that incorporate both fixed and random quantities are known as *linear mixed effects models*. Just as a fixed effects model can be thought of in the linear model context, a linear mixed effects model can be expressed as a special case of the mixed linear.

In the linear mixed effects model of (2.14), i indexes time-order and many designed experiments with the regressors (X 's) controlled by the experimenter must be run over time, often with only one observation at a time. Here the definition of the expected value and variance of $Y | X$ is defined in terms of the expected behaviour of the response in the presence of X over some specified time frame of interest (the time over which the experiment is run). There may be some random quantities (such as unsuspected environmental factors) over time that influence “errors”, which creates correlated errors over time. In this research work, this problem is a remedy in the context of a first-order autoregressive, AR(1), model for the errors with observations equally spaced over time, the model is

$$e_t = \rho e_{t-1} + \varepsilon_t \cdots \cdots \cdots \quad (2.21)$$

where the ε_t are assumed to be independent and identically distributed with mean 0 and variance σ^2_ε . It depends on two parameters (*i.e.*, $q = 2$), and is often better suited to longitudinal data. This is also called a first-order Markov process and is extensively used in the time-series analysis (Gottman, 1981). Here, the (co)variance for time points j and j' equals

$$\sigma_{jj'} = \sigma^2 \rho^{|j-j'|}$$

where ρ is the AR(1) parameter and σ^2 is the error variance. For repeated measure data, AR(1) makes a logical sense in that the correlation of the errors clearly diminishes as time lag increases. Additional, it only depends on one additional parameter, also it represents a very parsimonious augmentation to the usual conditional independence assumption.

In work related to Model (2.17) & (2.21), La Vange and Helms (1983) used method of moments procedures to estimate the time series parameters for the model $Y_k = X_k + U_k$, with AR(1) and (moving average) MA(q) models for the process U_k , but without any random effects terms. Jennrich and Schluchter (1986) discussed various types of covariance structures, including random effects models and the AR(1) model separately, and used different algorithms to obtain the ML estimates. Ware (1985) overviewed some special covariance structures, including general multivariate, random effects, and AR(1) models.

The pattern of dependence among observations can have more effect on the validity and quality of a statistical analysis than any other feature. Therefore choosing an analysis must begin with determining the logical properties of the sampling scheme and thereby characterizing the patterns of independence and correlation among observations (Demidenko, 2004).

2.8. Estimation of Parameters

In terms of model matrices, the general two-level model (with cross-level interaction) is given by (following Hox's (2002) presentation)

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & z_j & z_j x_1 \\ 1 & x_2 & z_j & z_j x_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & z_j & z_j x_1 \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{01} \\ \gamma_{11} \end{bmatrix} + \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

and in more succinct form, this yields

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \mathbf{e} \cdots \cdots \cdots \quad (2.15) \text{ repeated}$$

$$Y \sim N(\mathbf{X}\boldsymbol{\beta}, V) \text{ where } V = \text{Cov}(Y_i) = \mathbf{ZDZ}' + \mathbf{R}$$

The first developments of linear mixed effects models were related to the so-called variance components models, defined as linear mixed effects models in which all random effects are independent (and hence no covariance components are present). The first known formulation of a variance components model seems to have been given by Airy (1861) while considering a standard measurement problem in astronomy. Fisher (1925) introduced the analysis of variance (ANOVA) method for estimating variance components whose use and its extension to 2-way crossed classification mixed effects models were clarified by Tippet (1931).

The general structure of the linear mixed effects model (2.15) seems to be accepted by most researchers today. The linear mixed effects models literature that has been published after (Harville, 1977) and Laird and Ware (1982) refers more to generalizations of the assumptions in the model (2.15) and/or to different estimation approaches than to reformulations of the basic model's structure.

Multilevel modelling requires estimation of three different types of parameters. They are the level-1 random coefficients (parameter estimates that are allowed to vary across subjects)

denoted by β_j , the level-2 fixed coefficients (parameter estimates that do not vary across subjects) denoted by γ , and the random effects, such as the variance-covariance matrix of the level-2 coefficients, $\mathbf{D}$, and the residual level-1 variance, σ^2 . The estimations of these parameters are mutually dependent in MLM. The estimation of level-1 random coefficients requires known values of random effects and level-2 fixed effects, whereas the estimation of level-2 fixed effects requires known values of random effects and level-1 random coefficients. Similarly, the estimation of random effects requires knowledge of the level-2 fixed effects. Generally, estimates of these parameters are obtained iteratively in most situations,

The most common method for estimating multilevel models is maximum likelihood estimation (MLE), whereby the fixed effects, level-1 coefficients, and variance components are estimated simultaneously. However, conceptually it is easier when we think of the estimation of the fixed effects, level-1 coefficients, and variance components as separate steps. When we do so, it becomes apparent that the estimation of multilevel models entails a mixture of generalized least squares (GLS), empirical Bayes (EB), and MLE methods.

2.9. Parameterizations for Variance-Covariance Matrices:

From (2.19) Marginally, the Y_i are independent normal with mean $X_i\beta$ and covariance matrix $Z_i D Z_i^T + R_i$. Further simplification of this model arises when $R_i = \sigma^2 I$ where I denotes an identity matrix.

Inference for this general linear model in (2.17) can be based either on least squares and maximum likelihood methods, or on empirical Bayes methodology. Estimation of the parameters in (2.17) is usually based on full maximum likelihood (FML) or restricted maximum likelihood (REML) estimation for the marginal distribution of Y_i . The (R)ML estimation of the variance-covariance components in the model (2.17) is usually a difficult numerical problem since the

resulting estimates should correspond to a positive semi-definite matrix. This difficulty has been pointed out by Harville (1977), Lindstrom and Bates (1988), & Searle, Casella and McCulloch (1992).

In the statistical literature, several estimation methods have been proposed for mixed effects models though full maximum likelihood and restricted maximum likelihood are generally adopted for linear mixed effects models (Harville, 1974, Longford, 1993). A maximum likelihood approach to the estimation of the variance components has some attractive features; The maximum likelihood estimators are functions of every sufficient statistics and are consistent and asymptotically normal and efficient (Miler,1973)

Harville (1977) remarks that:

Certain deficiencies of various other methods are not shared by maximum likelihood. In particular, the maximum likelihood approach is "always" well-defined even for many useful generalizations of the ordinary Analysis of variance (ANOVA) models, and, with maximum likelihood, nonnegative constraints on the variance components or other constraints on the parameter space cause no conceptual difficulties.

2.10. Estimation and Inference:

In repeated measures design, for the description of the multivariate normal density for the set of observations on the i^{th} subject, we need to introduce additional vector and matrix notation:

Let $Y_i = (Y_{i1}, Y_{i2}, \dots, Y_{ip})'$ denote the $(p \times 1)$ vector of responses, and

$$X_i = \begin{pmatrix} 1 & X_{i11} & X_{i12} & \cdots & X_{i1,k-1} \\ 1 & X_{i21} & X_{i22} & \cdots & X_{i2,k-1} \\ \vdots & \vdots & & & \vdots \\ 1 & X_{ip1} & X_{ip2} & \cdots & X_{ip,k-1} \end{pmatrix}$$

denote the $(p \times k)$ matrix of covariates.

Then the distribution of Y_i is assumed to have a multivariate normal distribution with mean

$$E(Y_i) = \mu_i = X_i\beta \text{ with covariance matrix } \Sigma = V.$$

Conditional on the X values in the model $\sigma_{ij} = \text{cov}(Y_i, Y_j) = \text{cov}(e_i, e_j)$ is the covariance between Y_i and Y_j and the correlation is $\text{corr}(Y_i, Y_j) = \sigma_{ij} / \sigma_i \sigma_j$. That is given vectors of observations $(Y_{i1}, Y_{i2}, Y_{i3}, \dots, Y_{iN},)$, the covariance matrix is defined as the following array of variances and covariances

$$\text{Cov}(Y/X's) = \Sigma = \text{Cov} \begin{bmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{iN} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{bmatrix}$$

Using least squares

For linear models, if we continue to use least squares $\hat{\beta} = (X'X)^{-1}X'Y$, our estimators will be unbiased but $\text{cov}(\hat{\beta}) = \Sigma_{\hat{\beta}} = (X'X)^{-1}X'\Sigma X(X'X)^{-1}$, so the naive estimate $MSE(X'X)^{-1}$ is clearly wrong. One way to modify this is to use $\hat{\Sigma}_{\hat{\beta}} = (X'X)^{-1}X'\hat{\Sigma}X(X'X)^{-1}$, where $\hat{\Sigma}$ is an estimate of Σ obtained in a manner particular to the model under consideration. This is the approach of traditional contextual analysis (Boyd and Iversen 1991; also see Hanushek 1974). An important drawback of this approach is, however, that it does not take into account that β_j is usually estimated with different levels of precision in different subjects, if only because the sample sizes in the different subjects differ. As a consequence, γ is not BLUE (best linear unbiased estimator), except in the special case of a balanced design in which the sample sizes for all level-2 units are identical.

Assuming multivariate Normality, we have a choice of algorithms for obtaining maximum likelihood estimators. Here we use the Iterative Generalized Least Squares (IGLS)

algorithm, one advantage of which is that it provides a conceptually straightforward procedure based upon generalized least squares regression as shown below:

The multivariate normal probability density function has the following representation:

$$f(Y_i) = f(Y_{i1}, Y_{i2}, \dots, Y_{ip}) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp[-(Y_i - X_i\beta)' \Sigma^{-1} (Y_i - X_i\beta)/2] \quad (2.22)$$

where $|\Sigma|$ is the determinant of Σ (*also known as generalized variance*).

Considering a framework for estimation of the unknown parameters β and Σ in the repeated measures setting, we first assume that Σ is known (later we will relax this unrealistic assumption)

Known Variance

Given that $Y_i = (Y_{i1}, Y_{i2}, \dots, Y_{ip})'$ are assumed to follow a multivariate normal distribution, then the following log-likelihood must be maximized.

$$\begin{aligned} L(y; \beta, \Sigma) &= \ln \left\{ (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} \exp \left[- \sum_{i=1}^n (Y_i - X_i\beta)' \Sigma^{-1} (Y_i - X_i\beta)/2 \right] \right\} \\ &= -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| - \frac{1}{2} \left[\sum_{i=1}^n (Y_i - X_i\beta)' \Sigma^{-1} (Y_i - X_i\beta) \right] \quad (2.23) \\ &= -\frac{1}{2} \sum_{i=1}^n [(Y_i - X_i\beta)' \Sigma^{-1} (Y_i - X_i\beta)] - \frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| \\ &= -\frac{1}{2} \sum_{i=1}^n [Y_i' \Sigma^{-1} Y_i - Y_i' \Sigma^{-1} X_i\beta - X_i' \beta' \Sigma^{-1} Y_i + X_i' \beta' \Sigma^{-1} X_i\beta] - \frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| \end{aligned}$$

maximizing the likelihood, we have $\partial L / \partial \beta' = 0$

$$\begin{aligned} \partial L / \partial \beta' &= -\frac{1}{2} \sum_{i=1}^n [-X_i' \Sigma^{-1} Y_i + X_i' \Sigma^{-1} X_i\beta] = 0 \\ \therefore \sum_{i=1}^n (X_i' \Sigma^{-1} Y_i) &= \sum_{i=1}^n (X_i' \Sigma^{-1} X_i\beta) \quad (2.24) \end{aligned}$$

This then yields an estimate of β

$$\hat{\beta} = \left[\sum_{i=1}^n (X_i' \Sigma^{-1} X_i) \right]^{-1} \sum_{i=1}^n (X_i' \Sigma^{-1} Y_i), \quad \text{with } Cov(\hat{\beta}) = \left[\sum_{i=1}^n (X_i' \Sigma^{-1} X_i) \right]^{-1}$$

This estimator differs from the least squares estimator, thus the least squares estimator will have different properties than the maximum likelihood estimator. Notice that if Σ is equal to $\sigma^2 I$, the estimators are the same. This estimate of β that maximizes the likelihood function is known as the generalized least squares estimate. Since we usually do not know the value of Σ , typically its value is estimated from the data. In general, it is difficult to write down simple expressions for the ML estimate of Σ but this has to be found using a numerical algorithm that maximizes the likelihood. Once the ML estimate of Σ has been obtained, we simply substitute the estimate of Σ , say $\hat{\Sigma}$, in the generalized least squares estimator to obtain the MLE estimate of β .

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

The resulting estimator of β in large samples will have all the same properties as when Σ is known.

Properties of GLS

- I. For any choice of Σ , GLS estimate of β is unbiased; that is $E(\hat{\beta}) = \beta$
- II. $Cov(\hat{\beta}) = (X' \Sigma^{-1} X)^{-1}$.
- III. Sampling distribution of $\hat{\beta} \sim (\beta, (X' \Sigma^{-1} X)^{-1})$.
- IV. If $\Sigma = \sigma^2 I$, GLS estimate reduces to the ordinary least square estimate.
- V. The most efficient generalized least square estimate is the one that uses the true value of Σ .

For situations in which autocorrelation or heteroskedasticity exists

$$Var(e) = \Sigma = \sigma^2 \Omega \neq \sigma^2 I$$

and the model can be written more generally as

- i. $y = X_i \beta + e$
- ii. $E(e/X) = 0$
- iii. $E(ee'/X) = \Sigma = \sigma^2 \Omega$
- iv. X may be fixed or stochastic, but is of rank k and generated by a process unrelated to e
- v. $e \sim N(0; \Sigma)$

Given that Σ is known up to a proportionality constant $\Sigma = \sigma^2 \Omega$, in which case

$\hat{\beta} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y$ with $cov(\hat{\beta}) = \sigma^2 (X' \Omega^{-1} X)^{-1}$, which would be estimated by $\hat{\sigma}^2 (X' \Omega^{-1} X)^{-1}$, where $\hat{\sigma}^2 = MSE$ is the estimated variance from the transformed fit.

The argument for this is that Σ (assumed to be a positive definite matrix, i.e., non-singular covariance matrix) can be partitioned or factored as $T'T$ where T is non-singular as shown below.

$$\begin{aligned} \hat{\beta} &= (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} y \\ &= (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} (X\beta + e) \\ &= \beta + (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} e \end{aligned} \tag{2.25}$$

$$E(\hat{\beta}/X) = \beta + (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} E(e/X)$$

Since $(X'X)^{-1}X'$ is either fixed or a function only of X if X is stochastic, it can be factored out of the expectation, leaving $E(e/X)$, which has an expectation of zero by the assumption of iid.

$$\therefore E(\hat{\beta}/X) = \beta$$

Similarly, for the variance of the GLS (*MLE and BLUE*) estimator

$$\hat{\beta} = \beta + (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} e$$

$$\hat{\beta} - \beta = (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} e$$

$$\text{var}(\hat{\beta} - \beta) = E \left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)' / X \right]$$

$$E \left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)' / X \right] = E[(X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} e e' \Sigma^{-1} X (X' \Sigma^{-1} X)^{-1} / X] \quad (2.26)$$

$$= (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} E(e e' / X) \Sigma^{-1} X (X' \Sigma^{-1} X)^{-1}$$

$$= (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} \Sigma \Sigma^{-1} X (X' \Sigma^{-1} X)^{-1}$$

$$= (X' \Sigma^{-1} X)^{-1}$$

$$= \sigma^2 (X' \Omega^{-1} X)^{-1}$$

$$= \frac{\sigma^2}{n} \left(\frac{1}{n} X' \Omega^{-1} X \right)^{-1}$$

If the regressors are non-stochastic, then this is also the unconditional variance of $\hat{\beta}$. If the regressors are stochastic, then the unconditional variance is given by $E_X[\text{Var}(\hat{\beta}/X)]$.

Residual Maximum Likelihood (RMLE) Estimation

In spite of the usual asymptotic properties of the MLEs, the MLE of Σ has a well-known bias in a small sample (e.g. the diagonal elements of Σ are underestimated). To address this problem, the theory of restricted maximum likelihood estimation was developed, and the main idea behind REML is to eliminate the parameter β from the likelihood so that it is defined only in terms of Σ .

Recall that the multivariate normal probability density has the following likelihood representation:

$$f(Y_i) = f(Y_{i1}, Y_{i2}, \dots, Y_{ip}) = (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} \exp[-(y - X\beta)' \Sigma^{-1} (y - X\beta)/2]$$

rather than maximizing

$$-\frac{1}{2} \sum_{i=1}^n [Y_i' \Sigma^{-1} Y_i - Y_i' \Sigma^{-1} X_i \beta - X_i' \beta' \Sigma^{-1} Y_i + X_i' \beta' \Sigma^{-1} X_i \beta] - \frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| \quad (2.27)$$

REML maximizes the following slightly modified log-likelihood

$$-\frac{1}{2} \sum_{i=1}^n [Y_i' \Sigma^{-1} Y_i - Y_i' \Sigma^{-1} X_i \beta - X_i' \beta' \Sigma^{-1} Y_i + X_i' \beta' \Sigma^{-1} X_i \beta] - \frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| - \frac{1}{2} \ln \sum_{i=1}^n (X_i' \Sigma^{-1} X_i) \quad (2.28)$$

When residual likelihood is maximized, we obtain an estimate of Σ whose degrees of freedom are correct for the reduction in degrees of freedom due to estimating β .

When REML estimation is used, we obtained the generalized least square estimate of β ,

$$\hat{\beta} = \left[\sum_{i=1}^n (X_i' \hat{\Sigma}^{-1} X_i) \right]^{-1} \sum_{i=1}^n (X_i' \hat{\Sigma}^{-1} Y_i), \quad \text{with } Cov(\hat{\beta}) = \left[\sum_{i=1}^n (X_i' \hat{\Sigma}^{-1} X_i) \right]^{-1}$$

where $\hat{\Sigma}$ is the REML estimate of Σ .

Estimating V

To this point, the concern has been with estimating the fixed and random effects from the multilevel model. These parameters will frequently be of primary interest; however, it is often useful to examine the level-1 coefficients directly. In the “two-step” approach discussed above, the “level-1” coefficients are explicitly derived from the first stage estimation, since these regressions are separate models for each j unit. The separate-models approach has the virtue of giving unbiased estimates of the level-1 coefficients, although the reliability (or precision) of these estimates will decline as the number of level-1 observations within each j unit decrease.

So far, all of these derivations have assumed that the variance-covariance matrix, V , and thus, R_i and D , are known. This is virtually not always the case. One possible solution would be

to obtain an estimate $\hat{V}$ for V and simply substitute this estimate into the forms for estimating $\underline{\beta}$ and $\underline{b}$, or using the distribution-free, iterative procedure from Peng and Lu (2012).

Two-Step Iterative Estimation Procedure

To describe the estimation procedure for the fixed and random effects using the two-step iterative procedure, as in Peng and Lu (2012), first Let $Y = (Y_1^T, Y_2^T, \dots, Y_m^T)^T$

$b = (b_1^T, b_2^T, \dots, b_m^T)^T$, $e = (e_1^T, e_2^T, \dots, e_m^T)^T$ be $n \times 1$ column vectors,

$X = (X_1^T, X_2^T, \dots, X_m^T)^T$ a $n \times p$ matrix, and let $Z = (Z_1, Z_2, \dots, Z_m)$ be $n \times qm$ block diagonal matrix, where $n = \sum_{i=1}^m n_i$. For an initial estimate, ignore the random effects and begin with the simple linear regression model. The initial estimate of the fixed effects coefficients is then given by ordinary least square (OLS), which can be expressed using the clustered structure of the data in the following form.

$$\hat{\beta} = \beta_0 = (X^T X)^{-1} X^T Y = \left(\sum_{i=1}^m X_i^T X_i \right)^{-1} \sum_{i=1}^m X_i^T Y_i$$

With this initial estimate, recall that

$$Y_i = X_i \beta + Z_i \mathbf{b}_i + \epsilon_i \xrightarrow{\text{yields}} \underbrace{Y_i - X_i \beta}_{=u_i} = Z_i \mathbf{b}_i + \epsilon_i \quad (2.29)$$

Then estimate the random effect for each i , $1 \leq i \leq m$ by least squares as

$$\hat{b}_i = (Z_i^T Z_i)^{-1} Z_i^T \hat{u}_i \quad \text{and} \quad \hat{\epsilon}_i = \hat{u}_i - Z_i \hat{b}_i.$$

With these estimates, Peng and Lu (2012) propose the following estimators of σ^2 and D .

$$\sigma^2 = \frac{\sum_{i=1}^m \hat{\epsilon}_i^T \hat{\epsilon}_i}{n - qm}$$

To obtain the estimator for D , observe that

$$D = E \left[\frac{1}{m\sigma^2} \sum_{i=1}^m b_i b_i^T \right] \approx \frac{1}{m\sigma^2} \sum_{i=1}^m b_i b_i^T$$

So for an estimator of D , use $\hat{b}_i$ as an estimator of b_i , $1 \leq i \leq m$. However, recall the equation

$$\hat{b}_i = (Z_i^T Z_i)^{-1} Z_i^T \hat{u}_i = b_i + (Z_i^T Z_i)^{-1} Z_i^T \epsilon_i.$$

So

$$\begin{aligned} \sum_{i=1}^m \hat{b}_i \hat{b}_i^T &= \sum_{i=1}^m (b_i + (Z_i^T Z_i)^{-1} Z_i^T \epsilon_i) (b_i + (Z_i^T Z_i)^{-1} Z_i^T \epsilon_i)^T \quad (2.30) \\ &= \sum_{i=1}^m [b_i b_i^T + (Z_i^T Z_i)^{-1} Z_i^T \epsilon_i b_i^T + b_i \epsilon_i^T Z_i (Z_i^T Z_i)^{-1} + (Z_i^T Z_i)^{-1} Z_i^T \epsilon_i \epsilon_i^T Z_i (Z_i^T Z_i)^{-1}] \\ &\approx \sum_{i=1}^m [b_i b_i^T + (Z_i^T Z_i)^{-1} Z_i^T \epsilon_i \epsilon_i^T Z_i (Z_i^T Z_i)^{-1}] \text{ Peng and Lu (2012)} \\ &\Rightarrow \sum_{i=1}^m b_i b_i^T \approx \sum_{i=1}^m [\hat{b}_i \hat{b}_i^T - (Z_i^T Z_i)^{-1} Z_i^T \epsilon_i \epsilon_i^T Z_i (Z_i^T Z_i)^{-1}] \\ &\approx \sum_{i=1}^m [\hat{b}_i \hat{b}_i^T - \sigma^2 (Z_i^T Z_i)^{-1}] \end{aligned}$$

Thus,

$$D \approx \hat{D} = \frac{\sum_{i=1}^m \hat{b}_i \hat{b}_i^T}{m\sigma^2} - \frac{\sum_{i=1}^m (Z_i^T Z_i)^{-1}}{m}.$$

With this estimate of D , re-estimate the fixed effects using weighted least squares. That is,

$$\hat{\beta} = (X^T V X)^{-1} X^T V Y,$$

where $V = \text{diag} \left\{ (Z_1^T \hat{D} Z_1^T + I_{n_1})^{-1}, \dots, (Z_m^T \hat{D} Z_m^T + I_{n_m})^{-1} \right\}$. Finally, iterate between (2) –

(6) until convergence in $\hat{\beta}$, $\hat{\sigma}^2$, and $\hat{D}$.

Unknown Variance

When the covariance matrices are unknown, but an estimate of θ , and thus of R_i and $\mathbf{D}$, is available, it is natural to set $\hat{V} = Z_i D Z_i' + R_i$ and estimate β and b_i by using the weighted least squares equations replacing each V , by $\hat{V}$. As stated by Demidenko (2004), in line with the large sample properties, a characteristic property of the linear mixed effects model with normally distributed errors is that the information matrix for beta coefficients and variance parameters is block diagonal, which means that MLEs for β and (σ^2, D) are asymptotically independent. The numerical consequence is that maximization of the log-likelihood function may be accomplished separately for β and the variance parameters (σ^2, D) , and the statistical consequence, in the line of the pseudo-likelihood theory (Gong and Saminiego, 1981; Parke, 1986) is that the use of any consistent estimate of D in the GLS formula of (2.22) would lead to an efficient estimator of β .

For the ML estimation, we must use an iterative procedure to find the solution with an initial guess for the regression coefficients taken from OLS regression estimates to generate a sequence of improving approximate solutions and with zeros for the variance components. In the first iteration, a complex iteration procedure is used to try to improve on the initial values. Then the likelihood function is evaluated and the second iteration is performed. This procedure continues until the process converges, i.e. until the changes in the estimates are below a small threshold. For some problems, there may be multiple values that maximize the likelihood. For other problems, no maximum likelihood estimate exists: either the log-likelihood function increases without ever reaching a supremum value, or the supremum does exist but is outside the bounds of the set of acceptable parameter values. Sometimes, however, the models do not converge. The most common cause of this is the inclusion of variance components that are close to zero. The remedy for this is to simplify the random part of the model.

Given the choice of estimation theory, we need a computational algorithm. If we have chosen "expectation maximization" (EM) algorithm as the method of estimation, we need to maximize the likelihood. In some cases, this can be accomplished by simply evaluating a formula. But, in most cases, there is no closed-form expression for the maximizer of the likelihood, and some iterative scheme is required. Popular choices include the "expectation maximization" (EM) algorithm (Dempster, Laird, and Rubin, 1977; Dempster, Rubin, & Tsutakawa, 1981), Fisher scoring (Longford, 1987), and iterative generalized least squares (IGLS) (Goldstein, 1986). In principle, any of these algorithm applied to the same data and the same model should produce identical results. However, the performance of an algorithm, including its rate of convergence and the reliability of convergence across difficult applications, can vary Raudenbush and Bryks (2002). In addition, Snijders and Bosker (1999) assert that different computational algorithms would yield the same estimates, but with regard to convergence and computational speed, their performances would differ.

Application of the EM algorithm to MLM is fully described in the literature (e.g. , Raudenbush, 1988; Raudenbush & Bryk, 2002; Braun, 1989), Application of Fisher's scoring and iterative GLS (IGLS) in MLM have been described in detail by Raudenbush and Bryk (2002) and by Goldstein (1986, 1995), respectively. Both of these methods also use iterative steps for the estimation, they typically employ OLS estimates as the starting values.

2.11. The Basics of Bayesian Inference for the Two-Level Time Series Model:

In the Bayesian view, the data to be observed arise from a probability distribution defined by certain unknown quantities. Prior distribution must be assigned to all these unknowns. Combining the data with the prior leads to a joint posterior distribution of all the unknown given the data. Inferences about specific unknowns (e.g., focal parameters) are achieved by averaging

over the possible values of other unknowns (e.g., auxiliary parameters) (Raudenbush & Bryk, 2002)

In Bayesian notation, the observed data, Y , are distributed according to $f(Y|\beta, \sigma^2)$, where f is the normal density. Here Y is a vector containing elements Y_{ij} for all level-1 units $i = 1 \dots n_j$ within each level-2 units $j = 1 \dots j$. Similarly, β is a vector containing the elements β_{qj} . The outcome Y_{ij} are assumed independently normally distributed with mean

$$E(Y_{ij}|\beta, \sigma^2) = \beta_{0j} + \sum_{q=1}^{Q-1} \beta_{qj} X_{qjk} \dots \dots \dots (2.31)$$

and variance σ^2 . We must specify a prior distribution for the unknown, β , and σ^2 . This entails a two-stage prior.

Framework

Let y_j be an observation and θ_j a parameter governing the data generating process for y_j . Assume further that the parameters $\theta_1, \theta_2, \theta_3, \dots, \theta_j$ are generated exchange ably from a common population, with distribution governed by a hyper-parameter $\emptyset$. The Bayesian multilevel model contains the following stages:

Stage 1: $y_j|\theta_j, \emptyset \sim P(y_j|\theta_j, \emptyset)$

Stage 11: $\theta_j|\emptyset, \sim P(\theta_j|\emptyset)$

The likelihood, as seen in stage I is $P(y_j|\theta_j, \emptyset)$, with $P(\theta_j, \emptyset)$ as its prior distribution. Note that the likelihood depends on $\emptyset$ only through θ_j .

The prior distribution from stage I can be broken down into:

$$P(\theta_j, \emptyset) = P(\theta_j|\emptyset)P(\emptyset) \text{ (from the definition of conditional probability)}$$

with $\emptyset$ as its hyper-parameter with hyperprior distribution, $P(\emptyset)$.

Thus, the posterior distribution is proportional to:

$$P(\emptyset, \theta_j | y) \propto P(y_j | \theta_j, \emptyset) P(\theta_j | \emptyset) \text{ (using Bayes theorem)} \quad \dots \dots \dots (2.32)$$

$$P(\emptyset, \theta_j | y) \propto P(y_j | \theta_j) P(\theta_j, \emptyset)$$

2-stage hierarchical model

In general, the joint posterior distribution of interest in 2-stage hierarchical models is:

$$P(\theta, \emptyset | Y) = \frac{P(Y | \theta, \emptyset) P(\theta, \emptyset)}{P(Y)} = \frac{P(Y | \emptyset) P(\theta | \emptyset) P(\emptyset)}{P(Y)} \dots \dots \dots (2.33)$$

$$P(\theta, \emptyset | Y) \propto P(Y | \theta) P(\theta | \emptyset) P(\emptyset)$$

Using the Bayesian approach, our unknown parameters are $\beta, \gamma, \sigma, \dots$, are assigned a prior distribution. Following the Bayes' theorem, the likelihood is multiplied by the prior distribution to derive estimates of the posterior distribution of the model parameters. In the multilevel models discussed above, level-1 parameters are conditioned by level-2 parameters. With the Bayesian approach, conceptually, inferences about all unknowns are derived from their joint posterior distribution given the data. In this case,

$$P(\beta, \gamma, \sigma^2, D | Y) = \frac{f(Y | \beta, \sigma^2) P_1(\beta | \gamma, D) P_2(\gamma, \sigma^2, D)}{h(Y)} \dots \dots \dots (2.34)$$

Here $h(Y)$, the marginal distribution of Y , is a normalizing constant that ensures that the posterior is a proper probability distribution. Inferences about focal parameters of interest are obtained by averaging over auxiliary parameters. if γ is of focal interest, we compute

$$f(\gamma | Y) = \iiint f(\beta, \gamma, \sigma^2, D | Y) \partial \beta \partial \sigma^2 \partial D \dots \dots \dots (2.35)$$

This integration will result in posterior distribution, also known as the updated probability estimate, as additional evidence on the prior distribution is acquired.

2.12. Hypothesis Testing:

The Existence of Random Effects

One question in multilevel modelling is how much variability there is across subjects or individuals in a given change component, if there is great variability in intercepts and slopes across subjects, we should explicitly model that variability. However, if there is little variability, then perhaps we do not need to model this variability. When the independence assumption is presumed to be violated in our simulated data, the random effect for the subject was added to our standard multilevel model which allows us to resolve this non-independence by assuming a different “baseline and slope” . For the chosen mixed model to be right for our needs, there should be an existence of random effects.

Apart from inference on the fixed effects, testing for the inclusion or exclusion of random effects may be of interest because estimating unnecessary random effects leads to a loss of power. Testing for random effects is the same as testing to see if the variance of the random effect distribution is equal to 0 under the null hypothesis. This is a difficult problem because 0 is on the boundary of the parameter space for variance components. As a result, the standard asymptotic hypothesis tests do not hold.

2.12.1. Statistical Testing of the Presence of Random Effects:

Testing for random effects is equivalent to testing to see if the variance of the random effect distribution under the null hypothesis is equal to 0 (Oliver E. L., and Thomas, M. B.; 2012). This is a difficult problem because 0 is on the boundary of the parameter space for variance components. It is worthwhile to mention that, since confidence intervals and hypothesis

testing are equivalent (Rao, 1973), the Profile Likelihood Confidence Interval (PCI) leads to a better test of

$$H_0: D = 0 \dots \dots \dots (2.36)$$

This test of $H_0: D = 0$, is based on the estimated standard error of D computed from the inverse of the information matrix, The ratio

$$Z = \hat{\tau}_{qq} / [\text{Var}(\hat{\tau}_{qq})]^{1/2}$$

is approximately normally distributed under the large-sample theory of maximum likelihood estimates. When $\hat{\tau}_{qq}$ is near zero, the normality approximation will be extremely poor and technically, the likelihood is unlikely to be symmetric about the mode (Raudenbush & Bryk, 2002), and a test based on symmetric confidence interval for τ_{qq} may be highly misleading. In terms of **D** the variance-covariance matrix, the accuracy of the estimation depends on the number of level-2 units, N. Standard errors for the variances in **D** based on the information matrix require large N for credibility, especially if the estimate of the variance is near zero. There is need for cautiousness in interpreting results of significance tests (tests for individual random effects and covariance components in particular) when the data are extremely unbalanced or number of level 2 units is small ($j < 30$) which calls for more research in order to determine the robustness of such tests.

2.12.2. Profile-Likelihood Confidence Intervals:

Confidence interval (CI) construction and statistical testing on beta parameters in the LME model usually are based on the asymptotic covariance matrix. For example, the statistical significance of the j^{th} component of the vector β uses the ratio $\frac{\hat{\beta}_j}{S_j}$, where S_j is the MLE standard error calculated as the square root of the $(j, j)^{th}$ diagonal element of the information matrix.

Approximately, or more precisely, asymptotically (when the number of clusters or subjects is large), $\frac{\hat{\beta}_j}{s_j} \sim N(0,1)$ under the null hypothesis, which will be referred to as Wald CI because of its association to Wald hypothesis testing. If the number of estimated parameters is relatively large (say, $N_T - m < 30$), one can use the t-ratio:

$$\frac{\hat{\beta}_j}{s_j} \sim t(N_T - m) \quad \dots \quad \dots \quad \dots \quad \dots \quad (2.37)$$

where $N_T = \sum n_i$ is the total number of observation and $N_T - m$ is the degrees of freedom (d.f.). It is worthwhile to note that the ratio in (2.37) does not have a *t – distribution* as in a standard linear regression model, because we use an estimate of matrix D and not the real value. This same approach can be taken to test linear hypotheses about the parameters in the covariance matrix. Wald-type tests may not work very well for small to moderate sample sizes, especially with lots of parameters present.

When the ML procedure converges, we obtain asymptotic estimates of the Standard Error (SE) for both fixed and random-effects. Thus, the random-effects *SE* can be used for hypothesis testing (Singer and Willet, 2003). For this model, the *SE* based tests would lead us to conclude that there is significant variability in each random-effect, suggesting that they are all needed. The problem is that these standard error based tests can be very sensitive to the normality assumptions of the random-effects, and in fact, some statisticians have suggested that “...they are so misleading that they should be abandoned completely” (Singer & Willet, 2003). Some methodologists recommend using the likelihood ratio test to assess the significance of variance estimates because doing so tends to yield more accurate test results than the Wald *z* statistic, particularly in small to moderate samples (Singer & Willet, 2003).

2.13. Criteria for Model Fit and Model Choice:

Assessing the fit of an estimated model relies on the following idea: When the data have been maximally grouped, the group-specific parameters β_{0j} and β_{1j} can be estimated using their mean values γ_{00} and γ_{10} respectively. The use of these mean values as estimators corresponds to the *saturated model*, i.e., the model which contains separate parameters for each group. Thus the saturated model provides the best fit to the data and serves as a benchmark when evaluating the fit of estimated regression models. Formally, test whether the departure between the estimated model and the saturated model is significant or not can now be tested. The Pearson statistic and the deviance are the most frequently used goodness-of-fit statistics used for testing such a departure, both requiring that the data have been grouped as much as possible. The *Pearson statistic* is given by the sum of the squared standardized residuals:

$$\sum_{i=1}^G \frac{(\beta_{0j} - \gamma_{00})^2}{\frac{\hat{\sigma}^2}{n_t}} \text{ for Intercept and } \sum_{i=1}^G \frac{(\beta_{1j} - \gamma_{10})^2}{\frac{\hat{\sigma}^2}{n_t}} \text{ for slope.}$$

where β_{0j} is the intercept for the j^{th} level 2 unit, G represents the number of groups or subjects, β_{1j} is the regression coefficient associated with level 1 predictor X for the j^{th} level 2 unit, Z_j is the level 2 predictor or covariate. In the level 2 models, γ_{00} and γ_{10} are the overall mean intercept and slope adjusted for Z , respectively. The regression coefficients (β_{0j} and β_{1j}) are considered as dependent variables and relate each to appropriate level 2 covariates.

The *deviance* is defined by

$$D = -2 \sum_{i=1}^G \{l(\hat{\beta}_{ij}) - l(\hat{\gamma}_{ij})\}$$

Choosing the Best Model

Several multilevel models can be tested (e.g., a two-level random intercept model and a two-level random coefficient model) depending on the researcher's purposes. The estimation of more than a two-level model is also possible, and there is no restriction regarding the number of random covariates (Hox, 2002; Kreft & De Leeuw, 1998; Eileen T. Lake, 2006). The choice of a relevant model is an important step, and it should be based on the necessity of parsimony in a model. Parsimony means that models should be as simple as possible (Hox, 2002).

Another way to examine the significance of variance and covariance components is to compare two models, one model including the variance components of interest and a second, simpler model that constraints certain components to zero. If the fit of the simpler model to the data is significantly worse than the fit of the more complex model, the simpler model is rejected as inadequately representing the variation in the data.

Testing $H_0: S^2 = 0$ Versus $H_1: S^2 > 0$

In this study, for the comparison of the quality of fits obtained with the two models under study (i.e. Full and Nested, where Full will be termed correlated multilevel model and Nested is the standard multilevel model), hypothesis tests will be carried out comparing the two models, one of which contains the other as a special case. That is, interest will centres on the comparison of quality of fit obtainable from models with some constraints on the random part to that without the constraints and not assessing parameter values as such. Typically, we allow the specification of constraints on the autocorrelated residuals of the more general model, this specification of constraints in some way, results in the restricted model, which we call the null hypothesis, and the general model with autocorrelated errors referred to as the alternative hypothesis. Our purpose is to determine whether the change in the quality of the fit is sufficient

to justify the greater complexity of model H_1 , because the more general model, H_1 , must provide a fit that is at least as good as the restricted model, H_0 . This comparison is often reduced to a p-value, which is the probability of seeing a difference in the model fits as large as we did, or even larger, when, in fact, H_0 is adequate. Like all probabilities, a p-value must be between 0 and 1. When the p-value for a test is small (close to zero) we prefer the more complex model, saying that we “reject H_0 in favour of H_1 ”. On the other hand, a non-small p-value implies that we “fail to reject H_0 ”, arguing that there is a non-negligible probability that the observed difference in the model fits could reasonably be the result of random chance, not the inherent superiority of the model H_1 . Under these circumstances, we prefer the simpler model, H_0 , according to the principle of parsimony.

For the application to the study simulated data, using Likelihood ratio tests for making inferences in the random errors of a model with random effects and that with random effects with auto-correlated error structure, consider the two specified models as M1 and M2 where M1 (model with conditionally independent errors) is nested in M2 (model with auto-correlated errors). That is, M2 is a general model of which M1 is a special case, and the same mean structure is present in each model. Note that M1 is a special case of M2 obtained by setting $\sigma^2\Omega_k = \sigma^2I$. Model M1 is simple covariance structure. We wish to test H_0 : the model is M1 versus H_1 : the model is M2. This would mean testing the hypothesis that the σ^2I , assumed conditional independent errors to start, is Auto-regressive of order 1. Because model M1 is nested within model M2, they are called nested models, and the likelihood-ratio test can be used for their comparison. For this, one compares the model deviance values (i.e., $-2 \log L$) to a chi-square distribution, where the degrees of freedom equals the number of parameters sets equal to zero in the more restrictive model. It should be noted that use of the likelihood ratio test for this

purpose also suffers from the variance boundary problem mentioned above (Verbeke and Molenberghs, 2000). Let $L1$ = value of the maximized log-likelihood under model $M1$ and $L2$ = value of the maximized log-likelihood under model $M2$. The likelihood ratio test statistic is $LRT = 2*(L2-L1) = -2L1-(-2L2)$. For suitable sample sizes, LRT is approximately chi-square with q degrees of freedom where q = number of distinct parameters in model $M2$ minus the number of distinct parameters in model $M1$. The test rejects H_0 if LRT is bigger than $\chi^2_{(\alpha,q)}$. According to Hox (2002) to select the model that best fits the data, the log-likelihoods of the two models are compared if one model is a special case of the other model. The log-likelihood is not usually of interest in itself, but it plays a pivotal role in choosing one model between two alternative models (Healy. 2001).

2.14. Presence of Correlation:

After centring and scaling the data, to determine the appropriate structure of the covariance matrix of errors for the LMM, Dawson, Gennings, and Carter (1997) proposed graphical methods by creating a draftsman's display and a parallel axis plot, which are complementary plots to determine the amount of correlation between successive observations and whether or not that amount is constant. Substantiated with examples the appearance of the two plots for independent data as well as data that follows an autoregressive (AR) or compound symmetry (CS) structure. The two plots give similar results and while these plots appear to be easily interpretable for smaller dimensions, they get more unwieldy as p (measurement occasion) increases.

Chi and Reinsel (1989) considered Laird and Ware (1982) model when the correlation structure is an autoregressive process of order one (AR(1)). Maximum likelihood estimators of the model parameter and a score test for the autocorrelation were derived. One of the main

conclusions is that the use of an AR(1) structure for the cluster-specific errors may have the effect of reducing the number of random effects needed in the model, but the investigation of ways to determine the best combination of time series error structure and number of random effects deserves further study. This issue is also considered by Jones (1990).

2.15. Maximum Likelihood Estimation and Bayesian Analysis

Maximum likelihood estimation for normal distribution variance components models seems to have been first considered by Crump (1947). The landmark paper on ML estimation for variance components models is Hartley and Rao (1967), in which, among other things, the first asymptotic results for the MLE were established. Miller (1977) corrected some problems in Hartley and Rao's results and established asymptotic results for a large class of variance component models, giving also conditions for them to hold. The restricted maximum likelihood was introduced by Thompson (1962) and later extended by Patterson and Thompson (1971). Harville (1977) presents a comprehensive review of maximum likelihood and restricted maximum likelihood estimation in linear mixed effects models and introduces the model formulation given in (2.15). Laird and Ware (1982) describe a general linear mixed effects model for repeated measures data and suggest the use of the EM algorithm for obtaining (R)MLEs of the variance-covariance components.

Jennrich and Schluchter (1986) consider ML estimation in linear mixed effects models for repeated measures with structured variance-covariance matrices. Their work was extended to the general linear mixed effects models by Wolfinger, Tobias, and Sall (1991), who also discuss restricted maximum likelihood. The use of structured matrices is very appealing in practice since many times it is known beforehand that the covariance structure of the random effects and/or the

errors follows a particular pattern, and substantial reductions in computing time can thus be achieved.

A Bayesian analysis of model (2.15) using the Gibbs sampler (Geman and Geman, 1984) was described in Gelfand, Hills, Racine-Poon and Smith (1990) and in Wakefield, Smith, Racine-Poon and Gelfand (1994). The Bayesian analysis is developed using a hierarchical model approach. In the second paper, the normal distribution of the random effects ($\mathbf{b}$) is replaced by a multivariate Student- t , enhancing the robustness of the fit and giving a method for detecting outlying random effects. The main advantage of this approach is its flexibility in handling complex situations, such as constrained parameters and non-Gaussian distributions for the random effects and/or error terms. The main drawbacks are the intensive computational effort required and the need for prior distributions for all the population parameters involved.

As stated by Kreft and De Leeuw (1986), The choices between FML and RML estimation method is a matter of interest, how and why one method is chosen over the other is still unclear. In Bryk and Raudenbush (1992) RML is presented as sometimes superior to FML, in particular for small data sets, where "small" means a small number of subjects rather than a small number measurement occasion per subject. However, unlike ML approach, RML estimate of the variance-covariance components does take into account the uncertainty about the unknown fixed effect.

The major distinction between RML and FML is that the former accounts for consumed degrees of freedom while the latter does not. FML, as is well known (Patterson and Thompson 1971; Longford 1993; Faraway 2006), gives biased estimates of the variances, though this bias decreases with increased sizes of level-2 units (Raudenbush and Bryk, 2002, Longford 1993). A variety of algorithms have been proposed for FML/RML. Among them include the EM

algorithm (Dempster, Rubin, and Tsutakawa, 1981; also for attention Raudenbush and Bryk 2002; Longford 1993; Mason, Wong, and Entwistle 1983), iterative generalized least squares for FML (Goldstein (1986, 1995), restricted iterative generalized least squares (for RML), the Fisher scoring algorithm (Longford ; 1987, 1993) for random coefficients models, and Newton-Raphson. As far as choices among these alternatives go, there is no agreed-upon “optimal” way to proceed. In the case of FML, Goldstein (1995) and Goldstein and Rasbash (1992) show that Fisher scoring and IGLS are equivalent (for the normal model; this is described in Kreft and De Leeuw 1998).

As noted by Bryks and Raudenbush (2002), In a two-level hierarchical model, the FMLE and REML approaches produce very similar results for, but there are some differences in estimating the level-2 variance-covariance components. FMLE and REML approaches have very similar results if the number of level-2 units (J) is large. On the other hand, if the number of level-2 units is small, the MLE variance estimates are smaller than the REML estimates. REML minimizes least squares residuals and is generally considered better than the maximum likelihood method based on the deviance of the data only. However, De Leeuw and Kreft (1995), noted that “the evidence of their superiority in complicated cases, and in the multilevel analysis, in particular, is not too convincing.”

There is lack of agreement about which estimation method with which algorithm should be used under particular sampling conditions. When the number of the units in group-level is large, any estimation method can be employed. However, parameter estimation is more problematic when the number of the units in group-level is small, because the group-level regression coefficients and individual effects are conditional upon the estimated values of group-

level variance, and the variance might be underestimated (Heck and Thomas, 2000; Morris, 1995).

Since questions such as stability of regression coefficients, the influence of multicollinearity, and testing of model assumptions, which are already problematic in large multiple regression models become even more complicated in multilevel models and correlation always create additional complexity, hence the need to examine the performance of these estimators under violation of non-autocorrelated errors assumption.

In contrast to classical estimation methods, Bayesian methods for multilevel modelling have been widely proposed. Bayesian methods can be used to obtain ‘empirical Bayes estimates’ for the subject-specific parameters in Laird and Ware models (1982). Indeed, the literature on Bayesian hierarchical models goes back more than 40 years (Lindley and Smith 1972, George, and Arnold (1964). Gill (2002) notes that "the expression in (2.33) can be expanded to account for more levels in the hierarchy, for example treating γ as conditioned by level-3 factors. Adding hierarchies adds substantial complication to the model, however (in either the Bayesian or classical case)". Goel and DeGroot (1981; cited in Gill 2002, p. 352) show that the quality of the posterior parameters “decrease at each progressive level of the hierarchical model” (Gill 2002, p. 352); this assertion is formally proved by Goel (1983), cited in Gill 2002, p. 352). Computationally, deriving Bayesian estimates of the model parameters entails specifying conditional distributions for the model parameters, specifying starting values for the parameters, and iterating over values for the model parameters. The iterations produce a chain of simulation draws from the posterior. If several chains based on different starting values are jointly run (i.e. multiple simulations), the Bayesian model is said to “converge” when the chains have mixed (Gelman and Hill 2007). Simulating chains and mixing, or evaluating convergence, is

computationally demanding. Gibbs samplers, or more generally Markov Chain Monte Carlo (MCMC) techniques are usually necessary to derive final parameter estimates (Gelman and Hill 2007, Gill 2002, Jackman 2000). Bayesian estimation is usually used in simple MLMs/LMEs because the numerical integration for complex models is computationally intensive and are not easy to program. However, when the sample size is large, FMLE and REML approaches work well and the Bayesian approach is not necessary (Bryk & Raudenbush, 2002).

2.16. Likelihood Ratio Test for Variance Components:

The linear mixed model (Laird and Ware, 1982) is a popular method for longitudinal data. It is often of interest to test for random effects that is whether certain random effects should be included, which is the same as testing to see if the variance of the random effect distribution is equal to 0. Hypothesis testing for random effects in linear mixed models began with the works of Self and Liang (1987) and Stram and Lee (1994). Although separated by seven years, these two works are almost universally cited together in the literature for testing random effects in linear mixed models. The work of Self and Liang focused on deriving the properties of the likelihood when the true parameter value may be on the boundary of its parameter space. This includes showing the existence of a maximum likelihood estimator, its large sample properties, and the asymptotic distribution of likelihood ratio statistics under these conditions. Stram and Lee applied the results of Self and Liang (1987) specifically to likelihood ratio tests for nonzero variance components in linear mixed models.

Stram and Lee (1994) showed that likelihood ratio statistics for nonzero variance components follow mixture χ^2 distributions. In their paper, they illustrated this through a number of specific hypotheses of the random effect covariance matrix that they have labelled as cases. For case 1, likelihood ratio statistics for tests of one random effect with no random effects

have an asymptotic distribution following a 50:50 mixture of χ^2_0 and χ^2_1 . In case 2 they test a model with 2 random effects that are potentially correlated against a model with just 1 random effect. The likelihood ratio test statistic for case 2 follows a 50:50 mixture of χ^2_1 and χ^2_2 . Case 3 is a generalization of case 2, when testing $q + 1$ random effects against q random effects. Here, the likelihood ratio statistic follows a 50:50 mixture of χ^2_q and χ^2_{q+1} . Lastly, case 4 addresses simultaneously testing multiple variance components. For the special case when the information matrix is equal to the identity under the null hypothesis then the asymptotic distribution follows a mixture χ^2 with binomial mixing probabilities. For all other situations, simulation methods are recommended. As is the case with standard asymptotic test statistics, the Wald and score statistics for hypothesis tests of nonzero variance components also follow the same mixture χ^2 distributions as the likelihood ratio test statistic (Verbeke and Molenberghs, 2003, Silvapulle, 1992).

2.17. Normality of Errors:

Pinheiro and Bates (2000) checked the assumption of normally distributed errors with mean zero and the variance equal to σ^2 for the LMM. To do so, they utilized the “within-group residuals” which are equivalent to our SS residuals. They considered various graphical methods such as boxplots of the residuals by profile, a scatter plot of residuals versus the fitted values, and normal probability plots of the residuals. Their approach is very similar to the classical approach of regression residual diagnostics.

Houseman, Ryan, and Coull (2004) proposed to assess the assumption of normally distributed errors via graphical methods. They noted that the approach of Pinheiro and Bates (2000) can be misleading because of the shrinkage of the estimated random effects. Houseman, Ryan, and Coull (2004) proposed to rotate the general residuals $(y_i - X_i\beta)$ by pre-multiplying

them by the Cholesky decomposition of the inverse of the estimated variance-covariance matrix. The rotated residuals can then be displayed on a Q-Q plot to determine normality.

2.18. Consequences of Violating Residual Normality Assumptions:

In multilevel model, where respondents are considered as the level-2 unit and occasions (time) within each respondent are considered as the level-1 unit, the normality of error assumption indicates univariate normality of residuals at level-1 and univariate or multivariate normality (if more than one parameter was considered as random) of random components at level-2. Failure of the normality assumption at level-1 will not bias estimation of the fixed effects, but it will introduce bias into standard errors at both levels, thereby affecting the validity of confidence intervals and hypothesis tests. Estimation of the level-2 fixed effects will not be biased by non-normality of the errors at level-2. However, the presence of skewness will affect inferences at level-2 (Raudenbush and Bryk, 2002)

Shieh, Fouladi and Pullum (2001) also found that the nonnormality of errors introduces a downward bias into standard errors as a function of the size of the kurtosis. In other words, the standard errors for both fixed and random effects tend to be smaller than they should be as the kurtosis increases. However, it was found that the bias in the standard errors of both fixed and random effects decreases as the level-2 sample size increases.

Like Raudenbush (2002), Goldstein (1995) also found that When the residual errors are not normally distributed, the parameter estimates produced by Maximum Likelihood Method (MLE) are still consistent and asymptotically unbiased. However, the asymptotic standard errors are incorrect so significant test, the confidence interval cannot be trusted. this problem does not get varnished as the sample size get larger. He emphasized further that when the sample size at level-1 is small with MLE, estimates of fixed regression coefficient appear generally unbiased,

but when the assumption of normality is not met in addition to small sample size, the standard error of the fixed parameter has a small downward bias, estimate of the residual at the lower level generally very accurate and the group variance underestimated.

2.19. Heterogeneity of Error Variance at Level-1 and Level-2:

Raudenbush and Bryk (2002) found that heteroscedasticity of the level-1 variances (σ_j^2) itself will not lead to bias in estimation of level-2 fixed effect parameters and associated standard errors. However, heteroscedasticity of level-1 variances might result from the misspecification of the level-1 model. Such misspecification due to the omission of an important level-1 predictor can introduce bias in the estimation of fixed effects. several Monte Carlo studies have estimated the impact of varying levels of heterogeneity.

In multilevel regression, the assumption of homoscedastic residuals relates to all levels. According to most authors (Bryk & Radenbush, 1992; Goldstein, 1995; Snidjer & Bosker, 1999), ignoring the violation of the homoscedasticity assumption, will not only give incorrect values for the residual variances but can also lead to biased standard errors and a worse model fit. It is strongly advocated to check for heteroscedasticity and model it, if found, in order to get correct estimates for the standard errors associated with these parameters (Snidjers and Bosker, 1999).

Maeda (2007) in his paper titled "Monte Carlo Evidence Regarding the Effects of Violating Assumed Conditions of Two-Level Hierarchical Models for Cross-Sectional Data" state that the purpose of studies by Kasim (1994), Kasim and Raudenbush (1998) was to examine the efficacy of the Bayesian approach in order to obtain unbiased estimates in the case that heterogeneous level-1 variance exists. In these studies, they compared the performance of parameter estimation with the Bayes estimation procedure, which models the heterogeneity of level-1 variances via Gibbs sampling, and the performance of parameter estimation with the

empirical Bayes-maximum likelihood (EBML) estimation procedure. Obtaining results consistent with the finding of Kasim, R. M., and Raudenbush, S. W. (1998), concluded that there is no significant difference between the quality of Bayes estimates and EB-ML estimates of fixed effects, regardless of the heterogeneity of level-1 variances. On the other hand, they found that the EB-ML estimator overestimates the level-1 variance under heterogeneity. However, its effect on the variance was found to be smaller when the level-2 sample size is large ($J = 100$).

As stated by Maeda (2007), Monte Carlo study that examined the effect of the non-normality of level-1 errors was conducted by Kim (1990). This study examined how outliers and the strength of the relationship between an outcome and predictors at level-1 relate to the accuracy and the efficiency of the fixed effect parameter estimation. One of the main findings of Kim's study is that the correlation between a level-1 predictor and an outcome is negatively related to the efficiency of the fixed effect estimation. Efficiency, which Kim estimated using the square root of the average squared difference between an estimate and a parameter (RMSE), was found to decrease as the magnitude of the correlation between a level-1 predictor increased from 0.2281 to 0.6318.

2.20. Independence of a Level-1 Predictor and Level-1 Residuals:

The effect of the dependence of a level-1 predictor on level-1 residuals has been observed only in the small-scale simulation study of Ebbes, Böckenhof and Wedel (2004). They discovered that for a moderate dependence of the level-1 predictor on the level-1 random component ($\rho = 0.3$), there are upward biases in the estimates of the fixed effects, in contrast to the downward bias observed in estimates of the level-1 random component in the random analysis of covariance (ANCOVA) model.

2.21. Modelling the Covariance Structures:

Modelling the population covariance matrix can be employed in conjunction with the analysis of longitudinal data, as illustrated by Potthoff and Roy (1964), in the analysis of growth curve data. In their illustration, four measurements on a subject were presumed to be serially correlated and the following model for Σ was applied:

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

where ρ is estimated from a separate but similar experiment.

Hedeker and Roberts (2006) state that "Including autocorrelated errors in regression models has been well-described in the econometrics literature, for example, to model longitudinal earnings data (MaCurdy,1982). For Mixed Regression Models (MRMs), a key reference is a paper by Chi and Reinsel (1989), which considered a MRM with AR(1) errors. Some other related references include Mansour, Nordheim and Rutledge (1985), Hedeker (1989), Jones and Boadi-Boateng (1991)".

An important consideration in the statistical modelling of correlated data concerns the type of outcome. Methods for continuous data are undeniably the best developed and the linear mixed model (Laird and Ware, 1982; Verbeke and Molenberghs, 1997, 2000) has played a prominent role in extending the general linear model to handle correlated continuous data, Goldstein et al (1995) proposed a multilevel time series model when the response variable follows a Normal distribution and the measurements are taken with unequal time intervals. Owing to the elegant properties of the multivariate normal distribution, its theory and implementation are greatly simplified. Software programs, such as the SAS procedure MIXED

(Littell et al., 1996), are therefore widely available to fit this kind of models and have facilitated dissemination of the methodology among the statistical community.

According to Jennifer S. K., Chan, and Boris Choy, S. T., (2008) " the strength of serial correlation and the length of time series determine the choice of robust covariance structure. For data of longer time series, modelling the covariance structure becomes more difficult because of the possibility of more complicated covariance structures while for data containing relatively shorter time series, AR1 which has a simple structure with only two parameters, the constant variance σ^2 and the auto-regressive parameter ρ is the choice of the most robust covariance structure irrespective of the strength of serial correlation. Similarly, EC and TOEP are also good choices for data of weak and strong serial correlation respectively". Simulation experiments show that EC is the most robust choice irrespective of the strength of serial correlation while AR1 and TOEP2 are good choices for data of strong and weak serial correlation respectively.

2.22. Other Conditions Examined with Monte Carlo Studies for a Two-level TSM:

Other design conditions that potentially influence the estimation of parameters and significance tests in MLM have been conducted under several simulation studies. The impacts of sample size selection, modelling the covariance structure and misspecification of random effect are reviewed in this section.

The use of generalized regression model in which errors are modelled by a covariance matrix with non-constant variances and nonzero covariances becomes imperative because of the use of repeated measure data commonly in clinical trials and stock markets. The importance of estimating such covariance matrix was pointed out by Lindsey (1993). However, the number of unknown elements in the matrix can increase quadratically with its dimension causing considerable difficulties in estimation. Among the various approaches suggested was that of

Yang and Berger (1994) who applied a spectral decomposition to a covariance matrix, Efron and Morris (1976) who estimated the inverse of a covariance matrix using a loss function, Barnard, Mcculloch and Meng (2000) who modelled the covariance matrix in terms of its standard deviations and correlation matrix and estimated the parameters using a Bayesian approach and Chiu, Tom and Kam-wah (1996) who applied a matrix exponential transformation to a covariance matrix and estimated the parameters in the transformed matrix as well as the mean model using an ML approach.

The accuracy of the various estimation methods with relatively small sample sizes due to asymptotic nature of MLE procedure used in multilevel is questionable when assumptions of normality and large samples are not met, the standard errors have a downward bias. In addition, the group-level variance components tend to be underestimated", also for highly accurate group-level variance estimates, many groups (more than 100) are needed (Van der Leeden and Busing, 1994; Van der Leeden, Busin and Meijer, 1997; Afshartous, 1995). Contrarily, Browne and Draper (2000) reported that with as few as six to twelve groups, Restricted ML (RML) estimation provides reasonable variance estimates and, with 48 groups, both RML and Full Information ML (FML) estimation produce reasonable variance estimates. Although the results of the available simulation studies are not completely in agreement with each other, they all conclude that the regression coefficients are estimated without bias while their standard errors tend to be biased downward with small sample sizes at the group level.

Studies have shown that assuming the within-subjects errors (general residuals) to be uncorrelated, when in fact they are not, can lead to inflated Type I error rates for tests of treatment effects. Just like in single-level regression models applied to the interrupted time-series data from a single case (Greenwood & Matyas, 1990; Toothaker, Banz, Noble, Camp, and

Davis,1983). If the autocorrelation is positive, standard errors tend to be smaller, and the results of the t or f tests will be inflated or biased in a positive manner. This inflation increases the Type I error rate (i.e., too often showing an effect when there actually is none). Conversely, if the correlation between the errors is negative, then the standard errors will be too large, causing the t or f test statistics to be smaller. This, in turn, will increase Type II error rates, or too often failing to show an effect when there actually is one (Kazdin, 2011).

Mok (1995) found that when the number of Independent sampling Units (ISUs) was equal to the number of observations or when the number of ISUs exceeded the number of observations fixed effects estimates were estimated with near identical precision. However, estimates in both conditions exhibited substantially smaller MSE and bias than did those obtained when the number of observations exceeded the number of ISUs. Random effect estimates, though consistent across all conditions, were more biased when the number of observations exceeded the number of ISUs than when the converse was true. There was no difference across conditions in the bias of residual variance estimates. Mok (1995) concluded that it was better to maximize ISU size in order to minimize estimation error.

As stated by Maeda (2007), sample size also plays a key role in determining the validity of HLM (Raudenbush & Liu, 2000). Obviously, having a large sample size at both levels is always desirable. However, it is not practical in some situations, because of the cost incurred in data collection (Snijders & Bosker, 1993). Having a large level-1 sample size will increase the precision of estimates at each level-2 unit, whereas a large level-2 sample size can be a decisive factor in determining the average effect of a level-2 predictor when the effect exhibits substantial variation among level-2 units (Raudenbush, Yang and Yosef, 2000). Thus, a particular choice of

sample size may be optimal for estimating certain parameters but not others in a model (Raudenbush et al, 2000).

Bassiri (1988) examined how changes in Intra Class Correlation (ICC) (.1 vs. .25), number of independent Sampling Units (ISUs) (10, 30, 60, 150) and number of observation per ISU (5, 25, 60, 150) affect the REML estimate. The model was composed of three main fixed effects, one fixed interaction, and two random effects. The covariance matrix of the random effects was a two-dimensional heterogeneous variance component structure. Consistent with the conclusions of Mok (1995), Bassiri found that the number of ISUs was a better explanatory variable of problems in REML estimation than the number of observations per ISU, with smaller numbers of ISUs resulting in the highest degrees of bias in both the fixed and random effects. Varying the number of observations within a fixed ISU size affected neither the estimation of fixed nor random effects. It is interesting to note that at a fixed sample size, estimates obtained under a large ICC (.25) exhibited less estimation error than when the ICC was small (.1). Bassiri (1988) suggested that because the ICC is a measure of the degree to which total variance is the result of between ISU variance, as ICC increases, so does the degree of variance associated with differences in ISUs. Such increases in variance parameters are tantamount to increases in the power to estimate parameters, for as parameters depart from zero they are likely to be more reliably estimated.

According to Charles E. McCulloch and John M. Neuhaus (2011), many authors have found that likelihood-based inference can be severely affected if the random-effect model is misspecified. For example, Heagerty and Kurland (2001) computed the asymptotic bias in the MLEs for the parameters in a logistic mixed model in four instances of random-effect model misspecification. They concluded that incorrect assumptions on the random effects can lead to

substantial bias in the MLEs for the fixed effects. Agresti, Caffo and Ohman- Strickland (2004) conducted empirical studies on the impact of model misspecification for the random effects in GLMMs, showing that the MLEs for the fixed effects can be very sensitive to the assumed random effect model. Finally, Litiere, Alonso and Molenberghs (2007) used simulation to show that the type I and type II errors of tests for the mean structure in a logistic mixed model can be seriously affected by violations of the random-effect model.

Kreft and De Leeuw (1998) remarks that " The statistical power for multilevel models differs depending on whether it is level 1 or level 2 effects that are being examined. Multilevel models required large sample sizes to conduct research with sufficient power. However, the number of measurement occasions or individual observations in groups is not as important as the number of subject or groups in a study. In order to detect cross-level interactions, given that the group sizes are not too small, recommendations have been made that at least 20 groups are needed. The fact that power varies as a function of effect size and intraclass correlations complicate the issue of statistical power in multilevel models, which differs for fixed effects versus random effects, and it changes depending on the number of groups and the number of individual observations per group."

2.23. Methodological Issues in Previous Simulation Studies:

Reviewing the above studies also reveals some common methodological issues inherent in multilevel model (MLM) simulation studies. First, conclusions from simulation studies are potentially sensitive to the choice of parameter values among others, all the reviewed simulation studies start with fitting a null model (a model contains only a response variable, and no explanatory variable other than an intercept). In these studies, the null models are used as a baseline for the estimation of 'explained' versus 'unexplained' variances in comparison to more

elaborate models. The null- model does not account for any variation in the response variable. It only separates the variances of the response variable into two parts; that is, the variance of clusters (i.e, subjects), σ_u^2 , and the variance of observations at Level 1, σ_e^2 . The null models afford the researchers the use of various values of intra-class correlations (ICC) in the response variable. As often found in educational research, ICC, written as the symbol “ ρ ,” can be computed on the basis of these two variance components, the intra-class correlation measures the proportion of variance in the outcome that is between groups i.e. level two units (which is equicorrelation from measurement to measurement).

The random intercept model (like our model 2.3) considered thus far in these reviewed simulation studies, has an important limitation: The so-called within- and between-subject effects of X are the same. The within-subject effect denotes the effect if X changes within the same subject on different occasions. The between-subject effect refers to different X values between different subjects. In either case, a difference of one unit of X in the model (2.3) induces a difference of β_1 in expected responses. This equality between the within- and between-subject effect might be questionable in applications.

Marginally, for the random intercept model, however, repeated measurements y_{ij} for subject i are correlated with the within-subject correlation coefficient

$$Corr(y_{ij}, y_{il}) = \frac{\tau_0^2}{\tau_0^2 + \sigma^2}, \quad j \neq l \dots \dots \dots (2.36).$$

The strength of the within-subject correlation Eq. (2.36) depends on the magnitude of the error variances τ_0^2 and σ^2 . The higher the random effects variance τ_0^2 relative to the error variance σ^2 , the stronger the within-subject correlation. Note also that the within-subject correlation is constant (equicorrelation) from measurement to measurement, that is random intercept models are still based on the assumption that the covariate effect of X is equal in size for each individual

or subject. However, this may be too restrictive and questionable for longitudinal data, where we might expect correlations dying off for measurements which are farther apart in time.

For repeated measures data, variances often change over time, where subjects are generally more similar at the start of the trial than at the end of the trial. Where some subjects are high responders and others are low. Secondly, covariances close in time are usually greater than covariances that are farther separated in time. As a result of these, the random coefficient model in equation (2.14) (where both random coefficients are mostly found to be away from zero) induces a certain (marginal) correlation structure on observations y_{ij} and y_{ir} from the same subject. This correlation structure is, however, more complicated as it depends on the covariate values X_{ij} . First of all, the marginal variances of y_{ij} are heteroscedastic with

$$Var(y_{ij}) = \tau_0^2 + 2\tau_{01}X_{ij} + \tau_1^2X_{ij}^2 + \sigma^2;$$

Thus, the marginal variances depend quadratically on the covariate values. The covariance between y_{ij} and y_{ir} can be shown to be

$$Cov(y_{ij}, y_{ir}) = \tau_0^2 + \tau_{01}X_{ij} + \tau_{01}X_{ir} + \tau_1^2X_{ij}X_{ir}$$

resulting in an intraclass correlation coefficient

$$Corr(y_{ij}, y_{ir}) = \frac{Cov(y_{ij}, y_{ir})}{Var(y_{ij})^{\frac{1}{2}}Var(y_{ir})^{\frac{1}{2}}},$$

which depends in a rather complicated way on the observed covariate values and is very difficult to interpret. As a remedy to the situations in the reviewed simulation studies, since no unique intra-class correlation can be used with the presence of the random slope, correlations between observations of the same individual or subject are taken into account and data simulated according to a first-order autoregressive process, so that ρ the degree of correlation between

subject data (i.e., average correlation of y from any two time points) can be expected to be dying off for measurements which are farther apart in time.

In addition, there must be quality control for all scientific tests, only one of the reviewed simulation studies (i.e., Bassiri, 1988) gave the rationale behind the selection of the particular parameter values he used. By contrast, however, most researchers do not explicitly describe the appropriateness of the selected parameter values. This is a serious matter because the selection of parameter values can significantly affect the result of simulation studies (e.g., Krull, 1997). For example, Kim (1990) used The Second International Mathematics Study data to obtain some guidance for selecting population parameters. However, he did not state the reason for choosing the data used to obtain the parameter values.

Finally, it should be noted that the reviewed simulation studies like many others do not consider carefully whether data and model assumptions are satisfied or not. Assessment of model assumptions in the context of available data is the objective of the specification analysis. The assumptions generate questions about the accuracy of the various estimation methods when these assumptions are false, practical solutions may be limited and the reliability of a model becomes suspect. According to Keseleman et. al. (1998), the applied researchers who traditionally adopts standard procedure without given thought to its associated assumption may inadvertently be filling the literature with non-replicable results.

While this literature provides insight for applied researchers, the limitations of these studies leave many questions unanswered. Specifically, how the effects of small sample sizes at either or both levels, complex covariance structure, and systematic variations in the degree of autocorrelation impact on fixed and random effects parameter estimation remain unclear. These are the focus of this thesis.

CHAPTER THREE

RESEARCH METHODOLOGY

3.0. Introduction

The present study is a Monte Carlo study designed to investigate the statistical performance of two-level time series models (TTSM) techniques for handling data with correlated errors under departure from non-autocorrelated errors assumption in the Level 1 residuals. As listed in chapter two, there are several statistical assumptions that are required to draw valid inferences from the multilevel model analysis. This study focuses on the non-tenability of independent of level-1 residuals assumption. Specifically, estimation of a two-level multilevel model with a standard assumption that the within-individual (level-1) residuals are uncorrelated was compared to that of a model allowing for modelling the additional correlation structure. When estimating both models, it was assumed that all residuals were independently distributed (as is typically assumed). Thus, the study assessed the robustness of the models' estimates to violations of non-autocorrelated residuals at level-1. Bias, Variance and mean square error were assessed for model evaluation. The details describing the investigated conditions, how data were generated, the models that were estimated, appear in the next sections.

It should be noted that the major goal of the present study is to provide researchers with a basis on which to judge whether or not the conditions for TTSM on independence of level-1 residuals are plausible with regard to repeated measures data. Therefore, this chapter presents the methodology of the study which describes the research design of the present study. It is divided into six sections: the first section present manipulation check of data generation process, the second introduces the model simulated; the third presents the manipulation condition; the

fourth addresses the selection of parameter values; the fifth addresses data generation procedures; and sixth the data analysis.

3.1. Manipulation Check of Data Generation Process:

Our simulated data involves interrelated set of specifications and assumptions at both hierarchical levels. The first step in any worthwhile statistical analysis is to become familiar with our data – describe it, get to know it. If we jump straight to fitting a model and looking at our p -value, missing important features of the data, it is quite likely that the desired p -value is meaningless. In our simulation studies before fitting parametric model to repeated measures data, it is important to know if many of the peculiarities of repeated measures data are embedded in the simulated data sets. Copt and Victoria-Feser (2006), noted that the likelihood estimator is not robust to departures from the assumption of normality of the response variable. If the simulated data sets were well generated and the model assumptions are met, then the parameter estimates obtained will be a good representation of the relationship between subjects' data and the estimates can then be used to examine the impact of autocorrelated errors.

3.1.1. Normality Assumption for the Random Effects:

Due to the shrinkage effect, the Best Linear Unbiased Predictors (BLUPs) $\hat{b}$ can look normal even if the true distribution for b is not normal (e.g. bimodal). Verbeke & Molenberghs, (2000) give a discussion on fitting mixture distributions for b instead to check for normality.

3.1.2. Residual Diagnostics:

Remember that $Cov(Y_i - X_i\beta) = D_i$. Thus, the residual vector $e_i = Y_i - X_i\hat{\beta}$ will have mean zero, but will be correlated. Diagnostics are typically based on standardized residuals $e_i^* = L_i^{-1}e_i$, where $\hat{D}_i = L_i L_i^l$ is the Cholesky decomposition. (E.g. residual plots, Mahalanobis distance between observed and predicted responses, semi-variogram.) e_i^* are approximately

uncorrelated with unit variance. One could also use the subject-specific residuals $y_i - X_i\hat{\beta} - Z_i\hat{b}_i$ to look at the serial correlation. However, $\hat{b}_i$ depend very much on the normality assumption for b_i and is also influenced by the assumed structure for D_i . Thus, these residuals are not well-suited to check assumptions on D_i

3.2. The Multi-Level Time Series Models Utilized in this Study:

An applications of multilevel models being considered is as follow; Assessment of change in a repeated measurements design, with the repeated measures assumed to be random samples within units, the samples being ordered in time. Intuitively, if there is a dynamic pattern that is common among subjects, then by observing this pattern over many subjects, we hope to estimate the pattern with fewer time series observations than required of conventional time series methods.

The general model used in this study has been introduced in chapter two. The idea in this study was to simulate multilevel data that have an error structure similar to repeated measures data taken over time in which the residuals are serially correlated. The 2 two-level multilevel time series models for repeated measures data utilized in the present study are the correctly specified model (hereafter the “Correlated-model”) and the misspecified model (hereafter the “Standard multilevel-model”). The latter was obtained by omitting correlated errors structures from the Correlated-model. The structures of the two multilevel time series models utilized in the present study are described in this section. The parameter values used for generating data and the details of the procedure employed to generate data for the modelled variables are described separately in subsequent sections of this chapter.

3.2.1. The Correctly Specified Model (the Correlated (C)- model):

The correctly specified TTS model, the Correlated-model, is regarded as the “true” model for repeated measures, (with additional correlation structure modelled): with within level-1 residuals (σ^2) that were normally distributed with a mean of zero and non-zero correlation of the errors. The level-1 residuals tend to correlate and the additional correlation structure was included in the random part of the model, so that completely independent observations may be contrasted with AR(1) patterns of nonindependent observations. The incorporation of correlated responses is crucial to obtain accurate results in the analysis. This model was used in the process of generating data for the present study.

The C-model described by (Azeez & Ipinyomi, 2017) is expressed as follows:

$$\text{Level-1: } Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \cdots \cdots \cdots (2.6 \text{ repeated})$$

$$\text{Level-2: } \beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + v_{0j} \cdots \cdots \cdots (2.7 \text{ repeated})$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + v_{1j} \cdots \cdots \cdots (2.8 \text{ repeated})$$

$$e_{ij} \sim N(0, \sigma^2 \Omega_k).$$

$$\text{Thus, } V(y_i) = Z_i D Z_i' + \sigma^2 \Omega_k \cdots \cdots \cdots (3.1)$$

where Ω_k depends on q autocorrelation parameters, with q varying depending on the type of autocorrelated error structure being considered.

The motivation lies in the need to allow for patterns of dependence, rather than complete independence among response values. The simplest way to allow such dependence is to assume $V(y_i) = Z_i D Z_i' + \sigma^2 \Omega_k$ with Ω_k of dimension $N \times N$, symmetric and positive definite or semi-positive definite (which allows any covariance matrix).

3.2.2. The Misspecified Model (the Standard (S) multilevel-model):

According to Daniel J. B., Nisha C. G., Danielle D., and Robert A. Z.,(2012) the modelling strategy termed the standard multilevel-model is perfectly reasonable in many instances, but it may not always conform optimally to our theoretical model of change over time. In particular, the modeling strategy described below assumes that the group is a stable entity that either exerts a constant effect on the individual over time (in the case that only a random intercept is included at the group level) or an effect that changes systematically with time (in the case that a random time slope is also included at the group level). In many cases, however, groups may undergo structural and/or functional changes over time that are more stochastic in nature, producing time-varying group-effects that are not captured well by conventional model specifications.

The misspecified model, i.e., the Standard multilevel-model, is created by omitting the correlation matrix (*i.e.*, Ω_k) in the error structure, replacing it with the identity matrix I . with omitted correlated error structure, the level-1 residuals (σ^2) were normally distributed with a mean of zero and all correlations of the errors equal zero. When this correlated error structure was omitted, the basic idea is that the model errors are assumed to be conditionally independent given the random effects, but the nature of the dependency has not been modelled correctly.

This model is given by the following:

$$\text{Level-1: } Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \dots\dots\dots (2.6 \text{ repeated})$$

$$\text{Level-2: } \beta_{0j} = \beta_0 + \beta_{01}Z_j + u_{0j} \dots\dots\dots (2.7 \text{ repeated})$$

$$\beta_{1j} = \beta_1 + \beta_{11}Z_j + u_{1j} \dots\dots\dots (2.8 \text{ repeated})$$

$$e_i \sim N(0, \sigma^2 I).$$

$$\text{Thus, } V(y_i) = Z_i D Z_i' + \sigma^2 I \quad \dots\dots\dots (3.2)$$

later we will contrast models that assume the errors are (conditionally) independent with the one that follows a nonstationary AR(1) process.

3.3. Manipulated Conditions:

A review of the literature reveals that the impact of violating the assumed conditions for TTSM analysis depends on the design conditions used, with the random intercepts and slopes model (or randomly varying coefficients for any functions of time), the variability of the response can change as a function of the times of measurement, and the magnitudes of the correlations between measurements from the same individual can depend on the time between them (also for attention, Azeez & Ipinyomi, 2017). Since the length of the time series and the strength of the serial correlation may affect the choice of robust covariance structures, noting this point, three conditions were manipulated in this study, and the manipulated conditions in the present study are the following: (1) the number of subjects or individuals ; (2) the length of time series or measurement occasions (i.e., time points) ; and (3) the strength of the serial correlation on the choice of robust covariance structure. The length of the time series (t) was set equal to 3 (small) or 5 (moderate) and 10 (long), while the numbers of the individual (N) were 30 (small) or 50 (moderate) and 100 (high). In repeated measures designs, it is not uncommon to have small time points; however, these small values reflect the measurement occasion conditions at the individual level. For example, In many clinical trials, the repeated measures are balanced in the sense that a common (and often limited) set of measurement times is considered for all individuals, which allows for a priori specification of a “saturated” model. The chosen values for these simulation conditions were based upon values reported in previous studies involving

longitudinal data and are further described in the following sections. The full set of combinations of conditions that will be explored in this simulation is shown in Table 3.1.

Table 3.1. Summary of design conditions utilized in the study

Auto-regressive parameter (ρ)	Level-2 Sample size								
	30			50			100		
	Average level-1 sample size								
	3	5	10	3	5	10	3	5	10
0.3	Covariance structure								
0.7	<i>SI</i> & <i>AR</i> (1)			<i>SI</i> & <i>AR</i> (1)			<i>SI</i> & <i>AR</i> (1)		
0.9	where <i>SI</i> & <i>AR</i> (1) are simple and autoregressive of order – 1 covariance structures								

In the present study, we do not allow the level-1 sample size to vary among level-2 units. In our data generation, all of the subjects have measurements that are equally spaced along the subject and at the same location for all subjects. We refer to this data type as the balanced, equally spaced data. This implies that $n_i = n$ for $i = 1, 2, \dots, p$, and that the values of the regressors (and consequently, X_i and Z_i) are the same for all subjects. If the data are unbalanced, estimates of the level-2 fixed effects will depend on weights that are functions of the maximum likelihood estimates of the variance-covariance matrix of the level-1 coefficients. Random errors in these weights will lead to uncertainty about the level-2 fixed effects, affecting their standard errors, thus, making their confidence intervals shorter and test of significance will be more liberal than they should be (Bryk & Raudenbush, 2002).

Selection of sample sizes at both levels is based on literature: the frequently occurring minimum number of groups (subjects) is 30, while 100 groups or subjects is seen as sufficient. In practice, the frequently occurring number is 50 groups. Similarly, for the highest group size (measurement occasion), 10 is chosen as being sufficient on the basis of the literature, while a group size of 3 can be considered minimal in clinical trials, and a group size of five is normal in

family research and in longitudinal research (Van Der Leeden et al., 1997; Kreft and De Leeuw, 1998). The degree of correlation between subject data or the proportion of unexplained variation in the dependent variable that is due to the subject (ρ) is totally based on practice. They span the customary level of autocorrelation coefficients found in the analysis of covariance structures in time series analysis.

Searching for patterns in observed sequential data is a problem well known and widely studied, transcending fields of research and types of data. There is a very wide variety of problems and correspondingly wide variety of solutions. In general, the quest to understand more about something that occurs over time create one's interest in modelling time series data, in order to make decisions, predictions, or simply to understand more about the world. A possible explanation for the choice of positive autocorrelation coefficient and autoregressive of order 1 is that the response shares a common, yet unobserved, subject-specific parameter that induces a positive correlation. Autoregressive models are the most popular time series model that can be fully estimated and tested within the framework of least square regression. Also, for repeated measures data, AR(1) makes a logical sense in that the correlation of the errors diminishes as the time lag increases. Additionally, it only depends on one additional parameter and represents a parsimonious augmentation to the usual conditionally independent assumption.

3.4. Selection of Parameter Values:

3.4.1. Reisby et. al. :77 data

One of the common weaknesses of the previous Monte Carlo studies is the lack of a rationale for selecting particular parameter values. The appropriate specification of parameter values is also related to the external validity of the simulation results. With this in mind, the current parameter values were specified on the basis of results of HLM analyses discussed at

great length by Hedeker and Robert Gibbons in their text, Longitudinal Data Analysis that focused on the longitudinal relationship between imipramine (IMI) and desipramine (DMI) plasma levels and clinical response in 66 depressed inpatients. The file REISBY5.TXT contains data from a study by Reisby, et al., in the 1977 Psychopharmacology 54, 263-272, which can be downloaded through the Vanderbilt library system.

The estimate of random effect variance-covariance matrix computed by Hedeker and Roberts (2006) from Reisby et. al., (1977) data set were used to generate data with realistic parameter values because this data set has been reported to provide a good fit of the variance-covariance matrix. The use of any consistent estimate of D in the GLS formula of (2.22) would lead to an efficient estimator of β (Demidenko, 2004). In addition, according to Burton et. al, (2006), the simulated data sets should have some resemblance to reality for the results to be generalizable to real situations and to have any credibility. They stated further that, a good approach is to use a real data set as the motivating example and hence the data can be simulated to closely represent the structure of this real data set.

3.5. Data Generation and Simulation Procedures:

In this research work, a simple two-level multilevel model, with one explanatory variable at the measurement occasion level and one explanatory variable at the subject level, is considered conforming to equation (2.9 repeated):

$$Y_{ij} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} + \gamma_{10}X_{ij} + \gamma_{11}Z_jX_{ij} + u_{1j}X_{ij} + \varepsilon_{ij} \quad (2.9 \text{ repeated})$$

where the model's components are as defined in section 2.5. of Chapter two.

This study employed steps for generating data that are similar to those utilized in several previous multilevel models simulation studies. (e.g., Maeda, 2007; Drandari, 2004; Busing, 1993; Mass & Hox, 2004). In this procedure, the level-1 residuals were generated separately

from known distributions. Further, a set of level-1 random coefficients (β_{0j} and β_{1j}) was computed with the values of the pre-specified fixed effects, the generated values of the level-2 predictor (Z_j), and the level-2 random effects for each level-2 unit (u_{0j} and u_{1j}). Finally, the level-1 outcome variable, Y_{ij} , was computed with the random coefficients obtained at the preceding step, the generated values of the level-1 predictors (X_{ij}) and the level-1 error (ε_{ij}). The details of the procedure for data generation used in the present study are described below.

Generating level-1 residuals.

The level-1 residuals, ε_{ij} , were assumed to follow a normal distribution with mean zero and constant variance σ^2 . To give a sense of what the addition of autocorrelated errors into a multilevel model represents, level-1 residuals were generated such that $Cov(e_t e_{t-s}) = \sigma_e^2 f(s)$ where $f(s) = \Omega$, and Ω is the autocorrelation matrix which depends on q autocorrelation parameters, with q varying depending on the type of autocorrelated error structure being considered. Thus the covariance between two measurements depends on the time difference between the measurements, and the function $f(s)$ is conveniently described by a negative exponential reflecting the common assumption that with increasing time difference covariance tends to a fixed value, $\alpha\sigma^2$, and typically this is assumed to be zero. Also, the first level residual variance σ^2 was fixed at 12.22

Generating level-2 residuals

In the multilevel time series analysis, it is assumed that the level-2 residuals exhibit a multivariate normal distribution. Therefore, the level-2 random components were randomly selected from normal distributions with zero means and a specified variance-covariance matrix, Σ_v . Thus, we have $u_j \sim N(\mathbf{0}, \Sigma_v)$. The assumption in the level-2 model (when errors are homogeneous at both levels) is that β_{0j} and β_{1j} have a multivariate normal distribution with

variances define by σ_{v0}^2 and σ_{v1}^2 and covariance as $\sigma_{v0\ v1}$. With two random individual specific effects, the population distribution of intercept and slope deviations is assumed to be bivariate normal, with the mean of zero and random-effects variance-covariance matrix Σ_v (*i.e.* $N \sim (0, \Sigma_v)$). Thus, we have $u_j \sim N(0, \Sigma_v)$. Hence the random effects of the C model have the property

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{v0}^2 & \sigma_{v0\ v1} \\ \sigma_{v0\ v1} & \sigma_{v1}^2 \end{bmatrix} \right)$$

The parameter values for the level-2 variances specified as;

σ_{v0}^2 , σ_{v1}^2 , and $\sigma_{v0\ v1}$ are randomly selected as 12.63, 2.08, and – 1.40 respectively.

Three conditions were varied in the simulation: (1) Number of Subjects (NS): three conditions, (NS= 30, 50, 100), (2) Time Points (TP): three conditions, (TP= 3, 5, 10), and (3) Strength of the Serial Correlation (SSC): three conditions, (SSC=0.3, 0.7, 0.9). The sizes of the conditions are partially based on literature and partially on practical experience.

There are $3 \times 3 \times 3 = 27$ conditions. Assuming normally distributed residuals we generated 1000 simulated data sets for each condition. The first-level (measurement occasion) predictor and second level predictor values (X and Z) are generated from a standard normal distribution and are fixed for each level of autocorrelation and sample size factors, just like its single-level counterpart. The regression coefficients are specified as follows: 1.00 for the intercept, and 0.3 (a medium effect size; see Cohen, 1988; Maas & Hox, 2005) for all regression slopes. To simplify the simulation model, without loss of generality, the covariance between the two u -terms was assumed to be equal to nonzero. Given the parameter values, the simulation procedure generates the residual errors ε_{ij} , u_{0j} and u_{1j} . This research work did not investigate

non-autocorrelated errors for the second level residuals, because autocorrelation of the within-individual (level 1) residuals does not affect the independence of errors assumption assumed at the second level.

Each of the above steps was checked thoroughly and the distributions of the fixed and random effects were checked to ensure that the program is working properly using generated values from the steps above. In order to specify the number of replications, with a replication size of 1000, stability of results was achieved, and this number was used for replications.

The multilevel analysis was conducted by writing R 3.2.5 codes, where estimates of the first- and second-level parameters and variance-covariance estimated values were saved as output for the second analysis. The researcher used the Maximum Likelihood estimation in order to have estimates of the fixed parameter and variance-covariance-components, their error variances were computed numerically with iterations via iterated numerical procedure algorithms in R package. The R program ran a number of iterations for each data set that resulted in a number of values of estimates. These parameter estimates of the simulated conditions were used to evaluate the accuracy of TTSM parameter estimates.

3.6. Statistical Analysis of Accuracy:

The purpose of the present study was to examine the consequences of violations of independence of errors assumption on the parameter estimates and model fit of the multilevel time series model. When data on individual history are limited and individual behaviours are similar conditional on certain variables, pooling the data could yield more accurate predictions of individual outcomes than generating predictions using the data on the individual in question (Cheng Hsiao, 2014). Maximum likelihood estimation procedure used in this study do not include shrinkage effects, although they were used in shrinkage estimation schemes. Shrinkage

estimator is optimal from one statistical perspective because it minimizes the expected mean square deviation also called mean square error (MSE). The absolute value of the mean error called Mean absolute error (MAE), mean square error (MSE), and the root mean square deviation otherwise known as the root mean square error (RMSE) are some of the several valuable indices used for model evaluation. The ability of these indices to indicate the error in the units (or squared units) of the constituent of interest expresses their values, which aids in the analysis of the results. RMSE, MAE, and MSE values of 0 indicate a perfect fit. Singh, Knapp and Demissie (2004), state that MAE and RMSE whose values are less than half the standard deviation of the measured data may be considered low and that either is suitable for model evaluation.

The ability to recover the generating parameter values under the condition that the assumed independence of level-1 residuals condition is not satisfied was evaluated in the multilevel time series model. The study confirmed the validity of the generated data set and for both models that were estimated, parameter, mean square errors and variance-covariance estimates were saved across each of the 1,000 simulated data sets for each combination of conditions. The study assessed the performance of the multilevel time series model fixed effects and variance components for accuracy. The estimated values of the fixed effects, the associated mean square errors and the variance-covariance matrix were recorded for subsequent analyses. The next section separately describes the data analysis methods used for investigating the sensitivity of estimators.

Bias: The extent to which the estimate of the parameter deviates from the true parameter value is referred to as the value of the bias. For each replication, bias in the estimator $\hat{\theta}_j$ was

computed, where θ_j is the true value of the j th parameter and $\bar{\hat{\theta}}_j$ is the average of the estimates for the j th parameter across the 1,000 simulated data sets.

The magnitude of the bias of the estimates for each parameter from the true value was estimated by using the bias index. Let $\hat{\theta}_j$ be the estimate of the population parameter θ_j , then the index was calculated from the sum over n replications of simulations of the difference between the estimated parameter values and its true, or population, value. The formula is as follows:

$$\text{The Bias of } (\hat{\theta}_j) = E(\hat{\theta}_j) - \theta. \quad (3.3)$$

In order to compare the bias magnitude across parameters and conditions, the percentage relative bias index was used in this study. The relative bias measures the tendency of the simulated data on the average to be smaller or larger than their observed counterparts, this index presented the bias in an estimate that was expressed in the percentage of the parameters true value. The percentage relative bias of the parameter estimates was calculated by dividing the absolute bias value by the parameter true value and multiplying it by 100 as follows:

$$\text{Relative Bias of } (\hat{\theta}_j) = \frac{|\bar{\hat{\theta}}_j - \theta|}{\theta_j} \times 100 \quad (3.4)$$

Criteria for substantial versus non-substantial relative parameter bias was assessed using the suggestions of Hoogland and Boomsma (1998), wherein they recommend if relative bias $|RB(\hat{\theta}_j)| \geq 0.05$, there is evidence for substantial bias.

Mean Square Error: The Mean Square Error (MSE) of parameter estimates was calculated for each parameter. The MSE is a measure of the expected error of parameter estimates and it indicates the recovery over the replications. If the MSE = 0, then the estimator always perfectly recovers the population parameter. Of course, that is not realistic. If the estimator is unbiased, then the observed squared errors capture only sampling variance—our

uncertainty about the parameter estimate. If the estimator is biased, then the observed squared errors capture both this bias and sampling variance. It is defined as the average squared distance from the estimator to the true value of the parameter. The formula is as follows:

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + Bias(\hat{\theta}, \theta)^2 \quad (3.5)$$

So, MSE is a method of comparison that considers both Bias and Inefficiency in evaluating performance, where smaller MSE is better.

Sample variance

The sample variance of a random variable demonstrates The two aspects of estimator bias demonstrated by the sample variance of a random variable are: the naive biased estimator, which can be corrected by a scale factor; and, the unbiased estimator that is not optimal in terms of mean square error (MSE), which can be minimized by using a different scale factor, resulting in an estimator biased with lower MSE than the unbiased estimator. In a concrete manner, the naive estimator sums the squared deviations and use n as the divisor, which is biased. Dividing instead by $n - 1$ yields an unbiased estimator. Conversely, minimizing MSE required division by a different number (depending on distribution), but this results in a biased estimator. This number is always larger than $n - 1$, so this is known as a shrinkage estimator, as the unbiased estimator is "shrinks" towards zero; the optimal value is $n + 1$ for the normal distribution. All else being equal, an estimator with smaller variance is preferable to one with greater variance.

CHAPTER FOUR

SIMULATION STUDY AND PRESENTATION OF RESULTS

4.0. Introduction

This chapter presents the results of the simulation study in which two-level time series models were fit to repeated measures data across different levels of violations to non-autocorrelated error assumption and sample sizes. More specifically, it summarizes the accuracy of the estimation of fixed effects and the random effects parameters and the robustness of the hypothesis testing utilized in the two-level time series analysis with repeated measures data in the case that there is a violation of non-autocorrelated error assumption. This chapter details the results of evaluating the robustness of the model's fixed and random parameter estimates, by examining the bias, Mean Square Error (MSE), and precision for each parameter under each of the study conditions. The TTSM parameter estimates and their mean square errors produced by the R program were obtained from 1000 replications for each of the study conditions. The bias, the MSE, and the variance for the estimates of each parameter under each of the 144 study conditions were presented and examined to evaluate and compare the performance of the TTSM parameter point estimates.

The results of the study are described in five sections: 1) Sampling distribution of TTSM parameter estimates from the simulations ; 2) the evaluation of generated data sets; 3) the models' behavior under various levels of violations of non autocorrelated error assumption to better understand the problem; 4) a summary of the performance of TTSM fixed effects when there exists autocorrelated error; and 5) a summary of the performance of TTSM random effects when there exists autocorrelated error.

4.1. Sampling Distribution of TTSM Parameter Estimates from the Simulations

Chapter three presents the process for producing the simulation data. In order to confirm the accuracy of the simulated data in representing the desired sampling distributions in the study, evidence was provided (in section 4.2) for the repeated measures data produced under the study simulation routine by examining their correlational structure and random effect.

In a multilevel model, the final reported coefficient is a weighted average of the within and between effects, and this type of hierarchical, partial pooling model known as a random effects model is a shrinkage estimation scheme. The amount of shrinkage under random effect is data-dependent, either end up shrinking too much or not enough thereby comparing sampling distribution characteristics for the simulated data (true values of means, variances, and correlations among variables) for the first- level variables and residuals with the expected distributional characteristics under study conditions becomes less useful and likely to make the recovery of the generating parameter values less successful in addition to the effect of Monte Carlo sampling error.

4.1.1 Convergence and Inadmissible Solutions

It is not that uncommon to get a negative variance in the Maximum Likelihood estimation of higher-level variance in a mixed or multilevel model, which can be fairly challenging to explain to someone whose understanding about multilevel models is cursory. Such solutions are inadmissible, and the common procedure is to constrain such estimates to the boundary value of zero. To overcome the infeasibility in situations where the non-convergence was more likely to be present, we used in our simulations the known parameter values to generate the data as starting values of the iterative algorithm as was done by Hartford and Davidian (2000) for nonlinear models. This reduces the frequency of non-convergence to zero. For all our simulation

studies containing a total of 144 different conditions, we tracked the frequency of non-convergence and found it to be zero. The estimation procedure converged in all 144,000 simulated data sets. However, all 144,000 simulated datasets produced only admissible solutions. This suggested that there was no problem in the estimation process and in the estimation conversion. The high rate of convergence was largely due to the adequate model, sample size, and the number of variables included in the analysis.

4.2. Evaluation of the generated data sets

Our simulated data involves an interrelated set of specifications and assumptions at level-1. The first step in any worthwhile statistical analysis is to become familiar with our data – describe it, get to know it. If we jump straight to fitting a model and looking at our p -value, missing important features of the data, it is quite likely that the desired p -value is meaningless. In our simulation studies before fitting a parametric model to repeated measures data, it is important to know if many of the peculiarities of repeated measures data are embedded in the simulated data sets. Recall that datasets were generated on the basis of the correctly specified model (i.e., the C-model) when the covariance structure is correctly specified. If datasets were appropriately generated, the average of the estimates over replications for a certain parameter should approach its corresponding parameter value. To test this, each of the generated datasets was analyzed with the corresponding C-model used to generate the data and S-model in order to examine the appropriateness of the data generation.

Simulated data sets were generated as a repeated measures data in which the errors are serially correlated according to a first-order autoregressive process. If the data sets were appropriately generated, there should be the presence of correlation between successive observations, the existence of random effects and Normality of random effects among others.

Because data generation does not vary by parameter values or model complexity, in order to determine whether the procedure generated valid data, below are the manipulation check adopted in this research work. The R statistical package was used to determine whether parameters governing data generation could be adequately recovered in estimation.

To account for correlation under standard and correlated model, We examine our simulated models under different data simulation conditions and find out that for our simulation studies, in all the 144 combinations, $\sigma_{v1}^2 < -2\sigma_{v0v1}$ an indication of increasing variance across time (Hedeker,2006), the largest value of the covariances and the least value of the variances of the slopes, are -1.421574 and 2.080286 respectively, which occurred at $t = 10$, $N = 100$ and $\rho = 0.9$, where t, N , and ρ are the number of measurement occasion, the number of subjects and the autocorrelation coefficient. Also, we examined the correlational structure of the residuals under the correlated model and find out that the correlations between pairs of responses often decrease as the time lag increases (there was a decreasing correlation among the time points). As noted by Singer and Willett (2003), the most important issue is whether correlation model that the multilevel model assumes is reasonable for our data. For the same examination under the standard multilevel model, it is an oversimplification, there is considerable variation in these correlations and so assuming that they are all the same does not appear to be reasonable. A sampled estimated variance-covariance matrix for the fitted models and pairwise correlations structure for repeated outcomes under Autoregressive of Order-1 (AR(1)) and simple covariance structure (SI) are given below in Table 4.1 to Table 4.24 below.

Table 4.1: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE with autocorrelated errors

86.28	34.99	19.97		1	0.3981	0.2180
34.99	89.54	40.47		0.3981	1	0.4337
19.97	40.47	97.24		0.2180	0.4337	1

Table 4.2: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE with autocorrelated errors

86.32	35.01	19.98		1	0.3981	0.2171
35.01	89.58	40.49		0.3981	1	0.4337
19.98	40.49	97.28		0.2171	0.4337	1

Table 4.3: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE without autocorrelated errors

86.21	12.67	13.12		1	0.1443	0.1434
12.67	89.41	18.15		0.1443	1	0.1947
13.12	18.15	97.17		0.1434	0.1947	1

Table 4.4: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE without autocorrelated errors

86.23	12.68	13.12		1	0.1444	0.1433
12.68	89.41	18.16		0.1444	1	0.1948
13.12	18.16	97.19		0.1433	0.1948	1

Table 4.5: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.9$, with the corresponding correlation matrix under FMLE with autocorrelated errors

93.23	85.96	79.48		1	0.7309	0.5367
85.96	96.62	91.44		0.7309	1	0.7450
79.48	91.44	104.19		0.5367	0.7450	1

Table 4.6: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.9$, with the corresponding correlation matrix under RMLE with autocorrelated errors

93.15	85.91	79.47		1	0.9060	0.8070
85.91	96.53	91.39		0.9060	1	0.9117
79.47	91.39	104.11		0.8070	0.9117	1

Table 4.7: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.9$, with the corresponding correlation matrix under FMLE without autocorrelated errors

88.06	12.67	13.12		1	0.1413	0.1405
12.67	91.26	18.15		0.1413	1	0.1909
13.12	18.15	99.02		0.1405	0.1909	1

Table 4.8: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.9$, with the corresponding correlation matrix under RMLE without autocorrelated errors

88.11	12.69	13.11		1	0.1415	0.1403
12.69	91.28	18.17		0.1415	1	0.1910
13.11	18.17	99.07		0.1403	0.1910	1

Table 4.9: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.7$, with the corresponding correlation matrix under FMLE with autocorrelated errors

100.61	74.76	56.86		1	0.7309	0.5367
74.76	103.98	80.24		0.7309	1	0.7450
56.86	80.24	111.57		0.5367	0.7450	1

Table 4.10: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.7$, with the corresponding correlation matrix under RMLE with autocorrelated errors

100.62	74.77	56.87		1	0.7310	0.5367
74.77	103.99	80.25		0.7310	1	0.7450
56.87	80.25	111.58		0.5367	0.7450	1

Table 4.11: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.7$, with the corresponding correlation matrix under FMLE without autocorrelated errors

100.28	12.61	13.15		1	0.1238	0.1245
12.61	103.56	18.09		0.1238	1	0.1686
13.15	18.09	111.24		0.1245	0.1686	1

Table 4.12: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.7$, with the corresponding correlation matrix under RMLE without autocorrelated errors

100.28	12.61	13.14		1	0.1238	0.1245
12.61	103.56	18.09		0.1238	1	0.1686
13.15	18.09	111.24		0.1245	0.1686	1

Table 4.13: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE with autocorrelated errors

216.41	74.02	31.72	19.45	16.18		1	0.3395	0.1431	0.0855	0.0689
74.02	219.70	79.44	39.25	29.09		0.3395	1	0.3555	0.1712	0.1229
31.72	79.44	227.23	89.08	51.00		0.1431	0.3555	1	0.3822	0.2119
19.45	39.25	89.08	238.98	102.94		0.0855	0.1712	0.3822	1	0.4170
16.18	29.09	51.00	102.94	254.97		0.0689	0.1229	0.2119	0.4170	1

Table 4.14: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE with autocorrelated errors

216.60	74.10	31.76	19.46	16.18		1	0.3396	0.1432	0.0855	0.0688
74.10	219.88	79.50	39.28	29.10		0.3396	1	0.3556	0.1713	0.1229
31.76	79.50	227.40	89.14	51.04		0.1432	0.3556	1	0.3823	0.2119
19.46	39.28	89.14	239.16	103.02		0.0855	0.1713	0.3823	1	0.4171
16.18	29.10	51.04	103.02	255.16		0.0688	0.1229	0.2119	0.4171	1

Table 4.15: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE without autocorrelated errors

216.31	12.69	13.27	13.85	14.43		1	0.0582	0.0599	0.0609	0.0615
12.69	219.59	18.09	20.79	23.49		0.0582	1	0.081	0.0908	0.0993
13.27	18.09	227.11	27.73	32.55		0.0599	0.081	1	0.1191	0.1353
13.85	20.79	27.73	238.87	41.61		0.0609	0.0908	0.1191	1	0.1687
14.43	23.49	32.55	41.61	254.87		0.0615	0.0993	0.1353	0.1687	1

Table 4.16: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE without autocorrelated errors

216.38	12.72	13.28	13.85	14.42		1	0.0584	0.0599	0.0609	0.0613
12.72	219.64	18.10	20.80	23.49		0.0584	1	0.0811	0.0908	0.0993
13.28	18.10	227.15	27.74	32.56		0.0599	0.0811	1	0.1191	0.1319
13.85	20.80	27.74	238.92	41.64		0.0609	0.0908	0.1191	1	0.1688
14.42	23.49	32.56	41.638	254.94		0.0613	0.0993	0.1319	0.1688	1

Table 4.17: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.7$, with the corresponding correlation matrix under FMLE with autocorrelated errors

217.21	156.39	114.01	84.53	64.08		1	0.7144	0.5122	0.3703	0.2719
156.39	220.57	161.84	121.55	94.16		0.7144	1	0.7215	0.5285	0.3965
114.01	161.84	228.12	171.48	133.29		0.5122	0.7215	1	0.7331	0.5518
84.53	121.55	171.48	239.85	185.31		0.3703	0.5285	0.7331	1	0.7482
64.08	94.17	51.04	185.31	255.77		0.2719	0.3965	0.5518	0.7482	1

Table 4.18: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.7$, with the corresponding correlation matrix under RMLE with autocorrelated errors

217.54	156.64	114.20	84.68	64.20		1	0.7146	0.5909	0.3705	0.2720
156.64	220.89	162.08	121.74	94.32		0.7146	1	0.7216	0.5286	0.3966
114.20	162.09	228.43	171.73	133.48		0.5909	0.7216	1	0.7332	0.5519
84.68	121.74	171.73	240.17	185.56		0.3705	0.5286	0.7332	1	0.7482
64.20	94.32	133.48	185.56	256.10		0.2720	0.3966	0.5519	0.7482	1

Table 4.19: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.7$, with the corresponding correlation matrix under FMLE without autocorrelated errors

216.76	12.69	13.27	13.85	14.43		1	0.0581	0.0598	0.0608	0.0613
12.69	220.04	18.09	20.79	23.49		0.0581	1	0.0809	0.0906	0.0991
13.27	18.09	227.56	27.73	32.55		0.0598	0.0809	1	0.1188	0.1350
13.85	20.79	27.73	239.32	41.61		0.0608	0.0906	0.1188	1	0.1683
14.43	23.49	32.55	41.61	255.32		0.0613	0.0991	0.1350	0.1683	1

Table 4.20: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.7$, with the corresponding correlation matrix under RMLE without autocorrelated errors

216.83	12.72	13.28	13.85	14.42		1	0.0582	0.0598	0.0608	0.0613
12.72	220.09	18.10	20.80	23.49		0.0582	1	0.0809	0.0906	0.0991
13.28	18.10	227.61	27.74	32.56		0.0598	0.0809	1	0.1189	0.1351
13.85	20.80	27.74	239.37	41.64		0.0608	0.0906	0.1189	1	0.1684
14.42	23.49	32.56	41.64	255.39		0.0613	0.0991	0.1351	0.1684	1

Table 4.21: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.9$, with the corresponding correlation matrix under FMLE with autocorrelated errors

230.24	209.14	190.24	173.26	158.06		1	0.9018	0.8072	0.7180	0.6378
209.14	233.63	214.62	197.78	182.90		0.9018	1	0.9041	0.7892	0.7299
190.22	214.62	241.19	224.26	209.50		0.8072	0.9041	1	0.9080	0.8228
173.26	197.78	224.26	252.91	238.06		0.7180	0.7892	0.9080	1	0.9131
158.06	182.90	209.50	238.06	268.80		0.6378	0.7299	0.8228	0.9131	1

Table 4.22: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.9$, with the corresponding correlation matrix under RMLE with autocorrelated errors

230.20	209.12	190.21	173.27	158.08		1	0.9018	0.8073	0.7181	0.6355
209.12	233.12	214.60	197.77	182.91		0.9018	1	0.9041	0.8137	0.7300
190.22	214.60	241.15	224.24	209.50		0.8073	0.9041	1	0.9081	0.8229
173.27	197.77	224.24	252.87	238.04		0.7181	0.8137	0.9081	1	0.9131
158.07	182.91	209.50	238.04	268.76		0.6355	0.7300	0.8228	0.9131	1

Table 4.23: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.9$, with the corresponding correlation matrix under FMLE without autocorrelated errors

228.81	12.63	13.24	13.85	14.46		1	0.0548	0.0565	0.0578	0.0585
12.63	232.14	18.06	20.77	23.49		0.0548	1	0.0766	0.0860	0.0943
13.24	18.06	239.67	27.70	32.52		0.0565	0.0766	1	0.1128	0.1285
13.85	20.77	27.70	251.42	41.55		0.0578	0.0860	0.1128	1	0.1603
14.46	23.49	32.52	41.55	267.37		0.0585	0.0943	0.1285	0.1603	1

Table 4.24: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.9$, with the corresponding correlation matrix under RMLE without autocorrelated errors

228.83	12.64	13.25	13.85	14.45		1	0.0549	0.0566	0.0577	0.0584
12.64	232.14	18.07	20.78	23.49		0.0549	1	0.0766	0.0860	0.0943
13.25	18.07	239.68	27.71	32.53		0.0566	0.0766	1	0.1129	0.1285
13.85	20.78	27.71	251.42	41.56		0.0577	0.0860	0.1129	1	0.1603
14.45	23.49	32.53	41.56	267.39		0.0584	0.0943	0.1285	0.1603	1

The model results, like others not shown is clearly picking up on the increasing variance across time and the diminishing covariance away from the diagonal. The correlations follow the commonly seen pattern of diminishing in value as one goes further away from the diagonal. Also, within a given time lag, it appears that the association may increase across time. For example in Table 4.13, for the time lag of one measurement occasion, the pairwise correlation equals 0.3395 at the beginning (*i.e.*, time 1 and time 2 correlation) and increases to 0.4170 at the end (*i.e.*, time 4 and time 5 correlation). Conclusively, We can see quite clearly that the correlation is quite high between adjacent time points and slowly decreases as the distance between time points increases.

Comparing the correlation matrices in Table 4.13 and Table 4.14 with the corresponding correlation matrices in Table 4.15 and Table 4.16, again suggests that standard multilevel model is an oversimplification, there is considerable variation in these correlations and so assuming that they are all the same does not appear to be reasonable. We do not quote estimates of covariances of vector-valued random effects because the covariance is a difficult scale to interpret, whereas a correlation has a fixed scale.

To further substantiate the evidence of autocorrelation in our simulated data, variances of the slope under S-model were compared with the slope variances of the C-model. In Table 4.25,

If we look at the random-effects of the S-model (i.e. model without autocorrelated errors), we will see that the level-2 random-effects were somewhat larger in that model except model fitted to a dataset with 10 measurement occasion and autocorrelation coefficient of 0.3. The random-effects for the slope in that model are in competition to capture the correlation over time. Thus, typically the reduction we see in the random-effects of C-models were due to the incorporated correlation structures.

Table: 4.25: Comparison of Estimated slope variances under maximum likelihood estimator with and without autocorrelated errors

S/NO	COMB.	ML Ω SLOPE VARIANCE	REML Ω SLOPE VARIANCE	MLI SLOPE VARIANCE	REMLI SLOPE VARIANCE
1.	3, 0.3, 30	2.214595	2.235301	2.283667	2.315
2.	5, 0.3, 30	2.114334	2.122918	2.120733	2.130917
3.	10, 0.3, 30	2.085541	2.088749	2.084937	2.087796
4.	3, 0.7, 30	2.128304	2.135736	2.283667	2.315
5.	5, 0.7, 30	2.094246	2.097807	2.120733	2.130917
6.	10, 0.7, 30	2.083019	2.084766	2.084937	2.087796
7.	3, 0.9, 30	2.094586	2.09683	2.283667	2.315
8.	5, 0.9, 30	2.080386	2.085483	2.120733	2.130917
9.	10, 0.9, 30	2.080973	2.081537	2.084937	2.087796
10.	3, 0.3, 50	2.160453	2.167449	2.2022	2.212826
11.	5, 0.3, 50	2.100522	2.10332	2.10444	2.107773

12.	10, 0.3, 50	2.083314	2.084249	2.082962	2.083798
13.	3, 0.7, 50	2.10877	2.111272	2.2022	2.212826
14.	5, 0.7, 50	2.088474	2.089629	2.10444	2.107773
15.	10, 0.7, 50	2.081794	2.082301	2.082962	2.083798
16.	3, 0.9, 50	2.088677	2.089432	2.2022	2.212826
17.	5, 0.9, 50	2.082605	2.08296	2.10444	2.107773
18.	10, 0.9, 50	2.080577	2.08074	2.082962	2.083798
19.	3, 0.3, 100	2.120113	2.121785	2.1411	2.143646
20.	5, 0.3, 100	2.090231	2.090884	2.09222	2.093
21.	10, 0.3, 100	2.081653	2.081857	2.081481	2.081664
22.	3, 0.7, 100	2.094307	2.094903	2.1411	2.143646
23.	5, 0.7, 100	2.08421	2.084478	2.09222	2.093
24.	10, 0.7, 100	2.080891	2.081001	2.081481	2.081664
25.	3, 0.99, 100	2.084311	2.084491	2.1411	2.143646
26.	5, 0.99, 100	2.081293	2.081375	2.09222	2.093
27.	10, 0.99, 100	2.080286	2.080322	2.081481	2.081664

where $ML\Omega$ and $REML\Omega$ are maximum likelihood estimates with autocorrelated errors, MLI and $REMLI$ are maximum likelihood estimates without autocorrelated errors.

4.2.1 Existence of Random Effects

One question in multilevel modelling is how much variability there is across subjects or individuals in a given change component, if there is great variability in intercepts and slopes across subjects, we should explicitly model that variability. However, if there is little variability,

then perhaps we do not need to model this variability. When the independence assumption is presumed to be violated in the simulated data, the random effect for the subject was added to our standard multilevel model which allows us to resolve this non-independence by assuming a different “baseline and slope”. For the chosen mixed model to be right for our needs, there should be the existence of random effects.

4.2.2 Statistical testing of the presence of random effects (Examining Slope Variation)

Rather than employing single-parameter or multiparameter hypothesis tests for our model's random level-1 coefficients, we can make use of hypothesis testing for variance components to ask whether random variation exists by testing a null hypothesis

$$H_0: \tau_{qq} = 0$$

According to Singer and Willet (2003), when the ML procedure converges, we obtain asymptotic estimates of the standard error (*SE*) for both fixed and random-effects, thus, we can use the random-effects *SE* for hypothesis testing. Because multilevel modelling assumes that subject-specific intercept (i.e., v_{0j}) and slope (i.e., v_{1j}) deviations from their respective regression coefficient estimates (γ_{00} and γ_{10}) are normally distributed. This means that 95% of plausible intercept and slope values fall between ± 1.96 standard deviation units from the regression coefficient estimate. According to Douglas (2010), the variance of a random effect is a poor scale on which to quote the estimate, because confidence intervals on the variance are so badly skewed. It is more sensible to assess the estimates of the standard deviations of random effects. Consequently, the degree of individual heterogeneity in both the intercepts and slopes can be revealed by the magnitude of the variances of the intercepts and slopes estimates from our simulated model. Finding the confidence intervals for the regression coefficients of the model simulated under different conditions, answers can be provided on whether the level of “between-subjects” variability is sufficient to warrant incorporating random effects in the model. The coverages of the 95% confidence intervals for the main fixed effects for all simulated conditions are presented in Tables 4.26–4.29 below.

**Table 4.26: Full Maximum Likelihood Parameter Estimates and Wald 95% Confidence Intervals (CIs)
Obtained by Fitting Model with Autocorrelated Errors.**

S/No	COMB.	MLQ SLOPES (Estimate)	VAR. SLOPES	STD DEV SLOPES	95% CI (LOWER)	95% CI (UPPER)
1.	3, 0.3, 30	-0.3267	2.2146	1.4882	-3.2434	2.5901
2.	5, 0.3, 30	-0.2107	2.1143	1.4541	-3.0606	2.6393
3.	10, 0.3, 30	0.0331	2.0855	1.4441	-2.7974	2.8636
4.	3, 0.7, 30	-0.4429	2.1283	1.4589	-3.3022	2.4165
5.	5, 0.7, 30	-0.3591	2.0942	1.4472	-3.1955	2.4773
6.	10, 0.7, 30	0.3363	2.0830	1.4433	-2.4925	3.1651
7.	3, 0.9, 30	-0.4745	2.0946	1.4473	-3.3111	2.3622
8.	5, 0.9, 30	-0.3867	2.0844	1.4437	-3.2164	2.4431
9.	10, 0.9, 30	0.4070	2.0810	1.4426	-2.4205	3.2344
10.	3, 0.3, 50	0.6721	2.1605	1.4698	-2.20879	3.5530
11.	5, 0.3, 50	0.4776	2.1005	1.4493	-2.3630	3.3183
12.	10, 0.3, 50	1.2309	2.0833	1.4434	-1.5981	4.0599
13.	3, 0.7, 50	0.7294	2.1088	1.4522	-2.1168	3.5757
14.	5, 0.7, 50	0.6169	2.0885	1.4451	-2.2156	3.4494
15.	10, 0.7, 50	1.6538	2.0818	1.4428	-1.1742	4.4818
16.	3, 0.9, 50	0.7547	2.0887	1.4452	-2.0779	3.5873
17.	5, 0.9, 50	0.6361	2.0826	1.4431	-2.1925	3.4646
18.	10, 0.9, 50	1.7345	2.0806	1.4424	-1.0927	4.5616
19.	3, 0.3, 100	0.8482	2.1201	1.4561	-2.0057	3.7021
20.	5, 0.3, 100	1.1363	2.0902	1.4458	-1.6974	3.9700
21.	10, 0.3, 100	0.5546	2.0817	1.4428	-2.2733	3.3824
22.	3, 0.7, 100	0.9212	2.0943	1.4472	-1.9153	3.7577

23.	5, 0.7, 100	1.2759	2.0842	1.4437	-1.5537	4.1055
24.	10, 0.7, 100	0.6573	2.0809	1.4425	-2.1701	3.4846
25.	3, 0.9, 100	0.9541	2.0843	1.4437	-1.8756	3.7838
26.	5, 0.9, 100	1.2959	2.0813	1.4427	-1.5318	4.1235
27.	10, 0.9, 100	0.6758	2.0803	1.4423	-2.1512	3.5027

Table 4.27: Restricted Maximum Likelihood Parameter Estimates and Wald 95% Confidence Intervals (CIs) Obtained by Fitting Model with Autocorrelated Errors.

S/No	COMB.	REML Ω SLOPES (Estimate)	VAR. SLOPES	STD DEV SLOPES	95% CI (LOWER)	95% CI (UPPER)
1.	3, 0.3, 30	-0.3172	2.2353	1.4951	-3.2475	2.6132
2.	5, 0.3, 30	-0.2183	2.1229	1.4570	-3.0741	2.6374
3.	10, 0.3, 30	0.0301	2.0887	1.4453	-2.8025	2.8628
4.	3, 0.7, 30	-0.4328	2.1357	1.4614	-3.2971	2.4316
5.	5, 0.7, 30	-0.3653	2.0978	1.4484	-3.2042	2.4735
6.	10, 0.7, 30	0.3314	2.0848	1.4439	-2.4986	3.1614
7.	3, 0.9, 30	-0.4642	2.0968	1.4480	-3.3024	2.3740
8.	5, 0.9, 30	-0.3928	2.0855	1.4441	-3.2233	2.4377
9.	10, 0.9, 30	0.4016	2.0815	1.4428	-2.4262	3.2294
10.	3, 0.3, 50	0.6838	2.1674	1.4722	-2.2018	3.5693
11.	5, 0.3, 50	0.4649	2.1033	1.4503	-2.3776	3.3075
12.	10, 0.3, 50	1.2259	2.0842	1.4437	-1.6037	4.0556
13.	3, 0.7, 50	0.7407	2.1113	1.4530	-2.1072	3.5886
14.	5, 0.7, 50	0.6050	2.0896	1.4456	-2.2283	3.4383
15.	10, 0.7, 50	1.6475	2.0823	1.4430	-1.1808	4.4758
16.	3, 0.9, 50	0.7658	2.0894	1.4455	-2.0673	3.5990

17.	5, 0.9, 50	0.6244	2.0830	1.4432	-2.2044	3.4531
18.	10, 0.9, 50	1.7278	2.08074	1.4425	-1.0994	4.5551
19.	3, 0.3, 100	0.8575	2.1218	1.4566	-1.9975	3.7125
20.	5, 0.3, 100	1.1347	2.0909	1.4460	-1.6994	3.9689
21.	10, 0.3, 100	0.5519	2.0819	1.4429	-2.2761	3.3799
22.	3, 0.7, 100	0.9304	2.0949	1.4474	-1.9064	3.7673
23.	5, 0.7, 100	1.2734	2.0845	1.4438	-1.5564	4.1032
24.	10, 0.7, 100	0.6531	2.0810	1.4426	-2.1744	3.4805
25.	3, 0.9, 100	0.9633	2.0845	1.4438	-1.8665	3.7931
26.	5, 0.9, 100	1.2931	2.0814	1.4427	-1.5345	4.1208
27.	10, 0.9, 100	0.67133	2.0803	1.4423	-2.15569	3.4982

Table 4.28: Full Maximum Likelihood Parameter Estimates and Wald 95% Confidence Intervals (CIs) Obtained by Fitting Standard Multilevel Model.

S/No	COMB.	MLI SLOPES (Estimate)	VAR. SLOPES	STD DEV SLOPES	95%CI (LOWER)	95% CI (UPPER)
1.	3, 0.3, 30	-0.1629	2.2837	1.5112	-3.1248	2.7990
2.	5, 0.3, 30	-0.0521	2.1207	1.4563	-2.9064	2.8022
3.	10, 0.3, 30	-0.1884	2.0849	1.4439	-3.0185	2.6417
4.	3, 0.7, 30	-0.1106	2.2837	1.5112	-3.0725	2.8513
5.	5, 0.7, 30	-0.0670	2.1207	1.4563	-2.9213	2.7873
6.	10, 0.7, 30	-0.1770	2.0849	1.4439	-3.0071	2.6531
7.	3, 0.9, 30	-0.0844	2.2837	1.5112	-3.0463	2.8775
8.	5, 0.9, 30	-0.0745	2.1207	1.4563	-2.9288	2.7798
9.	10, 0.9, 30	-0.1712	2.0849	1.4439	-2.0013	2.6589
10.	3, 0.3, 50	0.6542	2.2022	1.4840	-2.2544	3.5628

11.	5, 0.3, 50	0.2743	2.1044	1.4507	-2.5690	3.1176
12.	10, 0.3, 50	0.8867	2.0830	1.4432	-1.9421	3.7154
13.	3, 0.7, 50	0.6775	2.2022	1.4840	-2.2311	3.5861
14.	5, 0.7, 50	0.2710	2.1044	1.45067	-2.5723	3.1143
15.	10, 0.7, 50	0.8901	2.0830	1.4432	-1.9387	3.7188
16.	3, 0.9, 50	0.6891	2.2022	1.4840	-2.2195	3.5977
17.	5, 0.9, 50	0.2694	2.1044	1.4507	-2.5739	3.1127
18.	10, 0.9, 50	0.8918	2.0830	1.4432	-1.9370	3.7205
19.	3, 0.3, 100	0.8299	2.1411	1.4632	-2.0380	3.6979
20.	5, 0.3, 100	0.9397	2.0922	1.4465	-1.8953	3.7748
21.	10, 0.3, 100	0.4663	2.0815	1.4427	-2.3615	3.29407
22.	3, 0.7, 100	0.8781	2.1411	1.4632	-1.9899	3.7460
23.	5, 0.7, 100	0.9402	2.0922	1.4465	-1.8948	3.7752
24.	10, 0.7, 100	0.4656	2.0815	1.4427	-2.3622	3.2934
25.	3, 0.9, 100	0.90214	2.1411	1.4632	-1.9658	3.7701
26.	5, 0.9, 100	0.9404	2.0922	1.4465	-1.8946	3.7755
27.	10, 0.9, 100	0.4652	2.0815	1.4427	-2.3625	3.29301

Table 4.29: Restricted Maximum Likelihood Parameter Estimates and Wald 95% Confidence Intervals (CIs) Obtained by Fitting Standard Multilevel Model.

S/No	COMB.	REMLI SLOPES (Estimate)	VAR. SLOPES	STD DEV SLOPES	95% CI (LOWER)	95% CI (UPPER)
1.	3, 0.3, 30	-0.1257	2.3150	1.5215	-3.1079	2.8564
2.	5, 0.3, 30	-0.0414	2.1309	1.4598	-2.9026	2.8197
3.	10, 0.3, 30	-0.1930	2.0878	1.4449	-3.0251	2.6390
4.	3, 0.7, 30	-0.0734	2.3150	1.5215	-3.0555	2.9088

5.	5, 0.7, 30	-0.0563	2.1309	1.4598	-2.9175	2.8048
6.	10, 0.7, 30	-0.1816	2.0878	1.4449	-3.0136	2.6505
7.	3, 0.9, 30	-0.0472	2.3150	1.5215	-3.0294	2.9350
8.	5, 0.9, 30	-0.0638	2.1309	1.4598	-2.9249	2.7974
9.	10, 0.9, 30	-0.1759	2.0878	1.4449	-3.0079	2.6562
10.	3, 0.3, 50	0.6578	2.2128	1.4876	-2.2578	3.5734
11.	5, 0.3, 50	0.2707	2.1078	1.4518	-2.5749	3.1162
12.	10, 0.3, 50	0.8855	2.0838	1.4435	-1.9439	3.7148
13.	3, 0.7, 50	0.6811	2.2128	1.4876	-2.2345	3.5967
14.	5, 0.7, 50	0.2674	2.1078	1.4518	-2.5782	3.1129
15.	10, 0.7, 50	0.8889	2.0838	1.4435	-1.9405	3.7182
16.	3, 0.9, 50	0.6927	2.2128	1.4876	-2.2229	3.6083
17.	5, 0.9, 50	0.2657	2.1078	1.4518	-2.5799	3.1113
18.	10, 0.9, 50	0.8906	2.0838	1.4435	-1.9388	3.7199
19.	3, 0.3, 100	0.8229	2.1436	1.4641	-2.0468	3.69257
20.	5, 0.3, 100	0.9326	2.0930	1.4467	-1.9029	3.76821
21.	10, 0.3, 100	0.4671	2.08167	1.4428	-2.3608	3.2950
22.	3, 0.7, 100	0.8710	2.1436	1.4641	-1.9986	3.7407
23.	5, 0.7, 100	0.9331	2.0930	1.4467	-1.9025	3.7687
24.	10, 0.7, 100	0.4664	2.0817	1.4428	-2.3615	3.2943
25.	3, 0.9, 100	0.8951	2.1436	1.4641	-1.9746	3.7648
26.	5, 0.9, 100	0.9334	2.0930	1.4467	-1.9022	3.7689
27.	10, 0.9, 100	0.4660	2.0817	1.4428	-2.3619	3.2939

Observations: Our goal here is to just get a sense of the data: What do the trajectories look like? Supposing we fit individual regressions, how do the regression coefficients vary across subjects?. Because the slope varies for individuals, the fitted model allows the possibility that some individuals do not change across time, while others can exhibit dramatic change. zero is within all of the prediction intervals on the random effects, that the interval we have constructed includes zero reflects the fact that there is considerable heterogeneity in terms of subjects change across time. By the value of the dispersion of the coefficients, some subjects are above the line of the overall average slope while others are below it, the variation around the average line (as indicated by the value of variances) is large, this implies that individuals' slopes are not similar and cannot be modeled by the fixed-effect estimate.

The width of REML confidence interval estimates are wider than the ML confidence interval, but the differences are small. In the standard multi-level model, combinations with same numbers of subjects and measurements occasion appear to have the same width as their prediction interval, this adds quite a bit to our understanding of the variability in our random coefficients. Also observed from the constructed confidence interval, there is lopsidedness of coverage resulting in estimates falling more frequently to one side than the other of the true parameter. To substantiate our claim, standard normal distribution was used to estimate the expected percentage of regression coefficients that are less than 0.3 under correlated model (ML Ω) and found to be within the range of 0.15% to 100%. The lopsidedness of coverage is a direct consequence of the bias in the multilevel point estimator, on which the Wald interval is centred. Despite this problem, multilevel Wald 95% intervals appear to provide conservatively valid (i.e. at least 95%) average coverage for the parameter estimates (Sander, 2000).

In our constructed 95% confidence intervals, subjects (level-2 units) with low numbers of measurement occasion and low autocorrelation coefficients are predicted to have a wider confidence interval than subjects with high numbers of measurement occasion and high autocorrelation coefficient. Similarly, differences between the number of subjects indicate the relationship between the width of the confidence interval. Not surprising, the specific values for the intercept and slope coefficients are subjects characteristics, as shown in Tables 4.30–4.32 below.

Table 4.30 Slope Variances when t (Number of repeated measures) is 3 at varying level of ρ and Number of Subjects.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	2.2146	2.1283	2.0946	2.2353	2.1357	2.0968	2.2837	2.2837	2.2837	2.315	2.315	2.315
50	2.1605	2.1088	2.0887	2.1674	2.1113	2.0894	2.2022	2.2022	2.2022	2.2128	2.2128	2.2128
100	2.1201	2.0943	2.0843	2.1218	2.0949	2.0845	2.1411	2.1411	2.1411	2.1436	2.1436	2.1436

where ML Ω and REML Ω are maximum likelihood estimates with autocorrelated errors, ML I and REML I are maximum likelihood estimates without autocorrelated errors.

Table 4.31 Slope Variances when t (Number of repeated measures) is 5 at varying level of ρ and Number of Subjects.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	2.1143	2.0942	2.0804	2.1229	2.0978	2.0805	2.1207	2.1207	2.1207	2.1309	2.1309	2.1309
50	2.1005	2.0885	2.0806	2.1033	2.0896	2.0807	2.1044	2.1044	2.0830	2.1078	2.1078	2.0838
100	2.0902	2.0842	2.0801	2.0909	2.0845	2.0801	2.0922	2.0922	2.0922	2.0927	2.0927	2.0927

where ML Ω and REML Ω are maximum likelihood estimates with autocorrelated errors, ML I and REML I are maximum likelihood estimates without autocorrelated errors.

Table 4.32 Slope Variances when t (Number of repeated measures) is 10 at varying level of ρ and Number of Subjects.

	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
N	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	2.0855	2.0830	2.0809	2.0887	2.0848	2.0815	2.0849	2.0849	2.0849	2.0878	2.0878	2.0878
50	2.0833	2.0818	2.0801	2.0842	2.0823	2.0801	2.0833	2.0830	2.0830	2.0838	2.0838	2.0838
100	2.0817	2.0809	2.0803	2.0819	2.0810	2.0803	2.0815	2.0815	2.0815	2.0817	2.0817	2.0817

where ML Ω and REML Ω are maximum likelihood estimates with autocorrelated errors, ML I and REML I are maximum likelihood estimates without autocorrelated errors.

Similarly, confidence intervals were constructed for the intercepts and there is also considerable heterogeneity in terms of subjects' baseline though results not shown, because the intercept being significant is not particularly meaningful; it just indicates that repeated measures are different than zero at baseline. Hence the hypothesis that the regression intercepts and slopes vary across subjects is supported by the data.

Raudenbush (2002) noted that the predicted intercepts and slopes for the N subjects are not identical to the value we would obtain if we carried out N separate ordinary regression analyses in each of the N subjects, using standard ordinary least square techniques (OLS). If we were to compare results from N separate OLS regression analyses, to the values obtained from a multilevel regression analysis, we would find that the results from the separate analyses are more variable. This is because the multilevel estimates of the regression coefficients of the N subjects are weighted. They are so-called empirical Bayes (EB) or shrinkage estimates; a weighted average of the specific OLS estimates in each subject and the overall regression coefficients, estimated for all similar subjects.

Testing variances using their standard errors is not optimal because it assumes normality, and because the null hypothesis that a variance is zero is a test on the boundary of the permissible parameter space (variances cannot be negative), where standard likelihood theory is no longer valid. A variety of alternative approaches have been proposed (Berkhof & Snijders, 2001).

According to Singer & Willett (2003), the Standard Error (SE) based tests can be very sensitive to the normality assumptions of the random-effects, and in fact, some statisticians have suggested that they are so misleading that they should be abandoned completely. Since the

standard errors of the general linear model are based on asymptotic variance, they may not be a good estimator of the standard error for small samples. In particular, Wald Confidence Intervals may not perform very well. One should only use the Wald Confidence Interval if the likelihood function is symmetric about the MLE. Consequently, for our simulated model, the SE based tests might lead us to conclude that there is significant variability in each random-effect, suggesting that they are all needed. Some methodologists recommend using the likelihood ratio test to assess the significance of variance estimates because doing so tends to yield more accurate test results than the Wald z statistic, particularly in small to moderate samples.

4.2.3 The Likelihood Ratio Test for Variance Components

The most commonly used test in a regular hypothesis testing setting is likelihood ratio test (LRT) due to its desirable theoretical properties and the fact that it is easy to construct. The LRT statistic has an asymptotically χ^2 null distribution with the degrees of freedom equal to the number of independent parameters being tested under the null hypothesis under very general regularity conditions.

In the results of our simulation, the observed LRT statistic for $ML\Omega$, and MLI ranges from 17.54 to 55.26, and 23.07 to 55.23 respectively, while the corresponding $\chi^2_{2\alpha,k}$ are 4.61, 7.78, and 14.68. Then the level α likelihood ratio test reject $H_0: \tau_{qq} = 0$ since $T_{obs} \geq \chi^2_{2\alpha,k}$, where $\chi^2_{2\alpha,k}$ is the $(1 - 2\alpha)th$ quantile of the χ^2 distribution with one degree of freedom, hence the variance term associated with the slopes are significantly different than 0. Clearly, both of these interpretations under confidence interval and LRT are reasonable and consistent with each other.

Normality assumption allows us to express the variance-covariance matrix as a function of the random parameters. Under the two-level model, the variances and covariances will be functions of only four underlying parameters, expressing the variance-covariance matrix under different simulation conditions, as a function of random parameter results in the variance-covariance matrix for random effects shown in Table 4.33 below under $ML\Omega$, $REML\Omega$, MLI and $REMLI$

Table 4.33: Random effect variance-covariance matrix

S/NO	COMB.	<i>MLΩ</i>		<i>REMLΩ</i>		<i>MLI</i>		<i>REMLI</i>	
1	3, 0.3, 30	13.42	-1.69	13.47	-1.71	13.58	-1.83	13.65	-1.86
		-1.69	2.21	-1.71	2.22	-1.83	2.28	-1.86	2.99
2	5, 0.3, 30	13.09	-1.52	13.16	-1.54	13.08	-1.54	13.15	-1.56
		-1.52	2.11	-1.54	2.12	-1.54	2.12	-1.56	2.13
3	10, 0.3, 30	12.87	-1.45	12.98	-1.46	12.82	-1.45	12.90	-1.46
		-1.45	2.09	-1.46	2.09	-1.45	2.08	-1.46	2.09
4	3, 0.7, 30	13.55	-1.52	13.62	-1.52	13.58	-1.83	13.65	-1.86
		-1.52	2.13	-1.52	2.13	-1.83	2.28	-1.86	2.30
5	5, 0.7, 30	13.21	-1.46	13.29	-1.47	13.08	-1.54	13.15	-1.56
		-1.46	2.09	-1.47	2.09	-1.54	2.12	-1.56	2.13
6	10, 0.7, 30	12.95	-1.44	13.09	-1.44	12.82	-1.45	12.90	-1.46
		-1.44	2.08	-1.44	2.08	-1.45	2.08	-1.46	2.09
7	3, 0.9, 30	14.84	-1.45	15.18	-1.45	13.58	-1.83	13.73	-1.89
		-1.45	2.09	-1.45	2.10	-1.83	2.28	-1.89	2.32
8	5, 0.9, 30	14.05	-1.43	14.41	-1.44	13.08	-1.54	13.19	-1.57
		-1.43	2.08	-1.44	2.09	-1.54	2.12	-1.57	2.13
9	10, 0.9, 30	13.39	-1.43	13.83	-1.43	12.82	-1.45	12.93	-1.46
		-1.43	2.08	-1.43	2.08	-1.45	2.08	-1.46	2.09
10	3, 0.3, 50	13.10	-1.58	13.12	-1.59	13.20	-1.66	13.22	-1.67
		-1.58	2.16	-1.59	2.16	-1.66	2.20	-1.87	2.21
11	5, 0.3, 50	12.91	-1.48	12.93	-1.49	12.90	-1.49	12.92	-1.50
		-1.48	2.10	-1.49	2.10	-1.49	2.10	-1.50	2.11
12	10, 0.3, 50	12.78	-1.44	12.81	-1.44	12.74	-1.44	12.76	-1.43
		-1.44	2.08	-1.44	2.08	-1.44	2.08	-1.43	2.08
13	3, 0.7, 50	13.19	-1.48	13.22	-1.48	13.20	-1.66	13.22	-1.68
		-1.48	2.11	-1.48	2.11	-1.66	2.20	-1.68	2.21
14	5, 0.7, 50	12.98	-1.45	13.01	-1.45	12.90	-1.49	12.92	-1.50
		-1.45	2.09	-1.45	2.09	-1.49	2.10	-1.50	2.11

S/NO	COMB.	<i>MLΩ</i>		<i>REMLΩ</i>		<i>MLI</i>		<i>REMLI</i>	
15	10, 0.7, 50	12.82	-1.43	12.86	-1.43	12.74	-1.44	12.77	-1.44
		-1.43	2.08	-1.43	2.08	-1.44	2.08	-1.44	2.08
16	3, 0.9, 50	14.05	-1.44	14.17	-1.44	13.20	-1.66	13.25	-1.69
		-1.44	2.09	-1.44	2.09	-1.66	2.20	-1.69	2.21
17	5, 0.9, 50	13.52	-1.43	13.64	-1.43	12.90	-1.49	12.94	-1.50
		-1.43	2.08	-1.43	2.08	-1.49	2.10	-1.50	2.11

18	10,0.9,50	13.10	-1.42	13.23	-1.42	12.74	-1.44	12.78	-1.44
		-1.42	2.08	-1.42	2.08	-1.44	2.08	-1.44	2.08
19	3,0.3,100	12.87	-1.50	12.87	-1.50	12.92	-1.54	12.92	-1.54
		-1.50	2.12	-1.50	2.12	-1.54	2.14	-1.54	2.14
20	5,0.3,100	12.77	-1.45	12.77	-1.45	12.76	-1.45	12.77	-1.46
		-1.45	2.09	-1.45	2.09	-1.45	2.09	-1.46	2.09
21	10,0.3,100	12.70	-1.42	12.71	-1.43	12.69	-1.43	12.69	-1.43
		-1.42	2.08	-1.43	2.08	-1.43	2.08	-1.43	2.08
22	3,0.7,100	12.91	-1.45	12.92	-1.45	12.92	-1.54	12.92	-1.54
		-1.45	2.09	-1.45	2.09	-1.54	2.14	-1.54	2.14
23	5,0.7,100	12.81	-1.43	12.81	-1.43	12.76	-1.46	12.77	-1.46
		-1.43	2.08	-1.43	2.08	-1.46	2.09	-1.46	2.09
24	10,0.7,100	12.73	-1.42	12.73	-1.43	12.68	-1.43	12.69	-1.43
		-1.42	2.08	-1.43	2.08	-1.43	2.08	-1.43	2.08
25	3,0.9,100	13.37	-1.43	13.41	-1.43	12.92	-1.54	12.93	-1.55
		-1.43	2.08	-1.43	2.08	-1.54	2.14	-1.55	2.14
26	5,0.9,100	16.12	-1.42	16.26	-1.42	12.76	-1.46	12.77	-1.46
		-1.42	2.08	-1.42	2.08	-1.46	2.09	-1.46	2.09
27	10,0.9,100	12.87	-1.42	12.90	-1.42	12.69	-1.43	12.69	-1.43
		-1.42	2.08	-1.42	2.08	-1.43	2.09	-1.43	2.08

where $ML\Omega$ and $REML\Omega$ are maximum likelihood estimates with autocorrelated errors, MLI and $REMLI$ are maximum likelihood estimates without autocorrelated errors.

4.3 Performance Evaluations (Criteria for Model Fit and Model Choice)

At this point, the burden on our research work is to identify those characteristics that are appropriate and a way to measure the performance on that characteristic. An interesting question, at this point, is, should the assumption of compound symmetry be rejected or not, in other words, is the between-subjects variance in equation (3.1) necessary over that in equation (3.2). Fitting the model augmented with autocorrelated errors and the more restrictive compound symmetry model, yields the results contained in Tables 4.34–4.35 below. Because these are nested models, they can be compared using a likelihood-ratio test, albeit with the caveat that because the testing involves variance terms, it should be noted that use of the likelihood ratio test for this purpose suffers from the variance boundary problem (Verbeke and Molenberghs, 2000), and the p-value obtained from the likelihood ratio test should be divided by two. The difference between

two deviances is distributed as a χ^2 random variable with degrees of freedom equal to the difference in the number of parameters between the two models.

Table 4.34: Criterion for comparing the REML Ω and REML(I) models fit under different conditions and the corresponding p-value.

Comb.	3, 0.3, 30	5, 0.3, 30	10, 0.3, 30	3, 0.7, 30	5, 0.7, 30	10, 0.7, 30	3, 0.9, 30	5, 0.9, 30	10, 0.9, 30
Difference in model deviance values	-0.0838	-0.1202	-0.196	-0.39	-0.2894	-0.894	-3.3632	-0.3556	-1.19
P-Val.	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001
Comb.	3, 0.3, 50	5, 0.3, 50	10, 0.3, 50	3, 0.7, 50	5, 0.7, 50	10, 0.7, 50	3, 0.9, 50	5, 0.9, 50	10, 0.9, 50
Diff in model deviance values	-0.0704	-0.0956	-0.596	-0.1104	-0.312	-3.05	-0.2592	-0.5738	-3.816
P-Val.	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001
Comb.	3, 0.3, 100	5, 0.3, 100	10, 0.3, 100	3, 0.7, 100	5, 0.7, 100	10, 0.7, 100	3, 0.9, 100	5, 0.9, 100	10, 0.9, 100
Diff in model deviance values	-0.0058	-0.122	-0.112	-0.0114	-0.444	-0.496	-0.1616	-0.596	-0.9308
P-Val.	<0.939294	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001

Table 4.35: Criterion for comparing the ML Ω and ML(I) models fit under different conditions and the corresponding p-value

Comb.	3, 0.3, 30	5, 0.3, 30	10, 0.3, 30	3, 0.7, 30	5, 0.7, 30	10, 0.7, 30	3, 0.9, 30	5, 0.9, 30	10, 0.9, 30
Difference in model deviance values	-0.0224	-0.0372	-0.134	-0.4054	-0.139	-0.704	-3.8832	-0.3984	-1.016
P-Val.	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001
Comb.	3, 0.3, 50	5, 0.3, 50	10, 0.3, 50	3, 0.7, 50	5, 0.7, 50	10, 0.7, 50	3, 0.9, 50	5, 0.9, 50	10, 0.9, 50
Diff in model deviance values	-0.024	-0.132	-0.664	-0.0944	-0.346	-3.19	-0.5736	-0.804	-3.972
P-Val.	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001
Comb.	3, 0.3, 100	5, 0.3, 100	10, 0.3, 100	3, 0.7, 100	5, 0.7, 100	10, 0.7, 100	3, 0.9, 100	5, 0.9, 100	10, 0.9, 100
Diff in model deviance values	-0.016	-0.182	-0.058	-0.02	-0.5	-0.33	-0.176	-0.63	-0.418
P-Val.	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001

For all the 27 different conditions considered under FML estimator, the P-Values are less than 0.00001, indicating that the results are significant at $p < 0.05$, similarly under RML the P-Values are less than 0.00001, indicating that the results are significant at $p < 0.05$ except for the combination 3, 0.3, 100 with P-Value 0.939294, indicating that the result is not significant at 0.05.

Remarks: The models were fitted by maximum likelihood (ML & REML). As the models are nested and differ only in the random effects part, the ML & REML statistic also serve as the basis for assessing their relative goodness of fit. Generally, we seek the most parsimonious

model that adequately describes our data. In Tables 4.34 – 4.35, differences in deviances and the corresponding p-values are reported, allowing a separate level-1 variance for each time point in the random- slopes model marginally reduce the deviances. It appears that the differences in deviances under different combination are significant with one degree of freedom. Clearly, the correlated model is more appropriate, since adding the autocorrelation parameter improves the fit of the model to the data. In this likelihood ratio test the value of the parameter being tested, a correlation of zero, is not on the boundary of the parameter space. We can be confident that the p-value of the LRT adequately reflects the underlying situation.

As remarked by Raudenbush (2002), in general, a violation of the homogeneity assumption is not per se a serious problem for estimating either the level-2 coefficients or their standard errors. we are principally concerned about it because such heterogeneity may indicate a possible misspecification of the level-1 model. In particular, unidentified slope heterogeneity at level-1 would appear as heterogeneity of level-1 error variance, and such heterogeneity can bias estimates of the level-2 coefficients.

4.4 The Effect of Non-autocorrelated error Assumption Violation on the Bias of Fixed-Parameter Estimates

The parameter estimates in our simulated model represent our “best guess” at the chosen values of the model parameters and, as such, are important results in statistical modelling. However, they are not the whole story. Statistical models characterize the variability in the data and we must assess the effect of this variability on the parameter estimates using, bias, variance and mean square error. To examine the performance of the TTSM parameter estimates, bias measures are usually used. The statistical bias of parameter estimates is the difference between the expected value of the parameter estimate and its true value. Positive bias is identical to the overestimation of the true value, while negative bias is the underestimation of that value.

In this study, the absolute bias—the average difference between the estimated parameter values and the parameter true value—is used to evaluate the bias of the

parameter estimates. In addition, in an attempt to make bias values comparable across the entire range of parameters and conditions, a measure of the magnitude of the absolute bias in relation to the true parameter value was used. The percentage of bias of the estimates for each parameter under each condition was obtained by dividing the absolute value of the bias by the true parameter value and multiplying the value obtained by 100. The bias calculated by this measure was considered as a relative bias. Thus, the relative bias is the average ratio in percentage of the absolute bias values to the true value of the parameter.

The absolute and relative bias for each of the fixed parameter estimates were calculated from 1000 replications under each examined condition in the study. The results are displayed for each parameter across all conditions in Tables 4.36–4.38.

Table 4.36: Comparing the bias (upper) and percentage relative bias (below) of slopes for measurement occasion, $t = 3$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	-0.6267 -104.44	-0.7429 -123.81	-0.7745 -129.08	-0.6172 -102.86	-0.7328 -122.13	-0.7642 -127.36	-0.4629 -70.96	-0.4106 -62.22	-0.3844 -57.87	-0.4257 -77.16	-0.3734 -68.43	-0.3471 -64.07
50	0.3721 62.02	0.4295 51.58	0.4547 75.78	0.3838 63.96	0.4407 73.45	0.4658 77.64	0.3542 59.64	0.3775 63.51	0.3891 65.45	0.3578 59.04	0.3811 62.91	0.3927 64.85
100	0.5482 91.36	0.6212 103.53	0.6541 109.02	0.5575 92.92	0.6304 105.07	0.6633 110.55	0.5300 87.15	0.5781 95.17	0.6021 99.19	0.5229 88.32	0.5710 96.35	0.5951 100.36

Where Ω represents autocorrelation matrix and I is the identity matrix.

Table 4.37: Comparing the bias (upper) and percentage relative bias (below) of slopes for measurement occasion, $t = 5$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	-0.5107 -85.11	-0.6591 -109.84	-0.6867 -114.44	-0.5183 -86.39	-0.6653 -110.89	-0.6928 -115.47	-0.3521 -56.90	-0.3670 -59.39	-0.3745 -60.63	-0.3414 -58.69	-0.3563 -61.17	-0.3638 -62.41
50	0.1776 29.61	0.3169 52.81	0.3361 56.01	0.1649 27.49	0.3050 50.83	0.3244 54.06	-0.0257 -4.89	-0.0290 -5.44	-0.0306 -5.72	-0.0293 -4.28	-0.0326 -4.83	-0.0343 -5.10
100	0.8363 139.38	0.9759 162.65	0.9959 165.97	0.8347 139.12	0.9734 162.24	0.9931 165.52	0.6397 105.44	0.6402 105.52	0.6404 105.56	0.6326 106.62	0.6331 106.70	0.6334 106.74

Where Ω represents autocorrelation matrix and I is the identity matrix.

Table 4.38: Comparing the bias (upper) and percentage relative bias (below) of slopes for measurement occasion, $t = 10$ for two estimators under standard and autocorrelated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	-0.2669	0.0363	0.1070	-0.2699	0.0314	0.1016	-0.4884	-0.4770	-0.4712	-0.4930	-0.4816	-0.4759
	-44.48	6.05	17.83	-44.98	5.23	16.93	-82.17	-80.26	-79.31	-81.40	-79.49	-78.54
50	0.9309	1.3538	1.4345	0.9259	1.3475	1.4278	0.5867	0.5901	0.5918	0.5855	0.5889	0.5906
	155.15	225.63	239.08	154.32	224.58	237.97	97.58	98.14	98.42	97.78	98.34	98.63
100	0.2546	0.3573	0.3758	0.2519	0.3530	0.3713	0.1663	0.1656	0.1652	0.1671	0.1664	0.1659
	42.43	59.55	62.63	41.98	8	61.88	27.85	27.73	27.67	27.72	27.60	27.54

Where Ω represents autocorrelation matrix and I is the identity matrix.

Evaluating Bias

Correlated Model (C-Model)

The average of the computed relative biases over conditions (ARB) ranges from -129.08 to 239.08 and -127.36 to 237.97 for FML and REML respectively. As described in the previous chapter, the relative bias represents the deviation between the parameter value and its estimates, relative to the parameter value. For example, the value of -104.44% for ARB of the slope under ML (Ω) for the combination (3, 0.3,30) indicates that the estimation is negatively biased approximately by that value relative to its parameter values. In other words, the estimate of the fixed effect is more than twice the true size of the effect when autocorrelated error assumption is violated. Under Table 4.36, when the number of measurement occasion is 3, FML PRB values ranged from -129.08 to 109.02, while REML PRB values ranged from -127.36 to 110.55; both FML and REML estimates exceeded 10% bias. Under Table 4.36, with the autocorrelation coefficient, $\rho = 0.9$ bias was more extreme with ML and REML. Under Table 4.36, when N , the number of subjects is small (i.e., 30), both FML and REML were consistently negatively biased, with several values exceeding the threshold defining severe bias, with FML exhibited high bias when compared to REML. When the two estimators were contrasted, neither REML nor FML is more positively or negatively biased than each other.

Next, the accuracy of the fixed effect estimation was also examined with respect to design conditions. First, the effect of the autocorrelation coefficient (ρ) on the performance of the two-level time series model (TTSM) parameter estimate was examined, and bias was found to be increasing along with the values of autocorrelation coefficient, except for the combination (3,0.3,30). Second, the moderation effects of the level-1 and level-2 sample sizes on the bias of the fixed effect to the violation of non-autocorrelated error assumption were examined, Table 4.36 contains descriptive results of the bias estimates and percentage relative bias as a function of the sample sizes. However, the sensitivity of both level sample sizes to the non-autocorrelation assumption violation is difficult to evaluate because of the inconsistency of the fixed effect estimates over conditions. Both FML and REML estimates of the fixed effect exhibited problematic bias levels. Both FML and REML biases exceeded 10% whenever 30, 50 or 100 subjects were sampled irrespective of the level of autocorrelation.

Under Table 4.37, when the number of measurement occasion is 5, FML PRB values ranged from -114.44 to 165.97, while REML PRB values ranged from -115.47 to 165.52; both FML and REML estimates exceeded 10% bias. Under 4.37, at autocorrelation coefficient $\rho = 0.9$ bias was more extreme with ML and REML. Similarly, when N , the number of subjects is small (i.e., 30), both FML and REML were consistently negatively biased, with several values exceeding the threshold defining severe bias. FML exhibited substantial bias when compared to REML especially when the number of subjects was moderate and high. When the two estimators were contrasted, neither REML nor FML is more positively or negatively biased than each other.

Next, the accuracy of the fixed effect estimation was also examined with respect to design conditions. First, the effect of the autocorrelation coefficient (ρ) on the performance of the two-level time series model (TTSM) parameter estimate was examined, and biases were

found to be increasing along with the values of autocorrelation coefficient. Second, the moderation effects of the level-1 and level-2 sample sizes on the bias of the fixed effect to the violation of non-autocorrelated error assumption were examined, Table 4.37, contains descriptive results of the bias estimates and percentage relative bias as a function of the sample sizes. However, the sensitivity of both level sample sizes to the non-autocorrelation assumption violation is difficult to evaluate because of the inconsistency of the fixed effect estimates over conditions. Both FML and REML estimates of the fixed effect exhibited problematic bias levels. Both FML and REML biases exceeded 10% whenever 30, 50 or 100 subjects were sampled irrespective of the level of autocorrelation.

Under Table 4.38, when the number of measurement occasion is 10, FML PRB values ranged from -44.48 to 239.08, while REML PRB values ranged from -44.98 to 237.97; both FML and REML estimates exceeded 10% bias, except the combination with (10, 0.7, 30) with PRB values of 6.05 and 5.23 under ML and REML respectively. Unlike Tables 4.36 and 4.37, where bias was at the extreme at 0.9, under 4.38, bias was more extreme at 0.3 and 0.9 with ML and REML. FML exhibited substantial bias when compared to REML except with the combinations (10,0.3,30) and (10,0.7,30). Neither REML nor FML is more positively or negatively biased than each other.

Next, the accuracy of the fixed effect estimation was also examined with respect to design conditions. First, the effect of the autocorrelation coefficient (ρ) on the performance of the two-level time series model (TTSM) parameter estimate was examined, and biases were found to be increasing along with the values of autocorrelation coefficient except for the combination (10,0.3,30). Second, the moderation effects of the level-1 and level-2 sample sizes on the bias of the fixed effect to the violation of non-autocorrelated error assumption were

examined, just as noted in Table 4.36 and Table 4.37, the sensitivity of both level sample sizes to the non-autocorrelation assumption violation is difficult to evaluate because of the inconsistency of the fixed effect estimates over conditions. Both FML and REML estimates of the fixed effect exhibited problematic bias levels when the moderate and large number of subjects were sampled (i.e. 50 and 100).

Examination of raw biases (see Tables 4.36–4.38) revealed that the patterns observed were somewhat inconsistent with the increase in the number of measurement occasion (t) and the number of subjects (N). All the three Tables lead to the same conclusion regarding estimators performances. The high negative biases associated with the fixed effect estimate under the condition (5,30,0.7) and (5,30,0.9) observed in Table 4.37 was replaced with low positive biases in Table 4.38. Even with 100 subjects selected at varying numbers of repeated measurements both FML and REML were biased with more than 10%. Fixed effects are theoretically unbiased, with identical estimating equations for REML and FML, but results obtained in this study demonstrated that the asymptotic unbiased-ness of the fixed effect estimator was not robust to finite sample sizes. Though the MLEs of fixed effects differ negligibly between FML and REML.

Standard Model (S-Model)

The average of the computed relative biases over conditions (ARB) ranges from -82.16837 to 105.5589 and -81.3994 to 106.7405 for FML and REML respectively. Under Table 4.36, when the number of measurement occasion is 3, FML PRB values ranged from -70.96 to 99.19, while REML PRB values ranged from -77.16 to 100.36; both FML and REML estimates exceeded 10% bias. Under 4.36, at autocorrelation coefficient, $\rho = 0.9$ bias was more extreme with ML and REML. Similarly, when N , the number of subjects is small (i.e., 30), both

FML and REML were consistently negatively biased, with several values exceeding the threshold defining severe bias. FML exhibited substantial bias when compared to REML at all levels of correlation except when N , the number of subjects is 50. When the two estimators were contrasted, neither REML nor FML is more positively or negatively biased than each other.

Next, the accuracy of the fixed effect estimation was also examined with respect to design conditions. First, the effect of the autocorrelation coefficient (ρ) on the performance of the two-level time series model (TTSM) parameter estimate was examined, and bias was found to be increasing along with the values of autocorrelation coefficient, except for the combination (3,0.3,30). Second, the moderation effects of the level-1 and level-2 sample sizes on the bias of the fixed effect to the violation of non-autocorrelated error assumption were examined, Table 4.36, contain descriptive results of the estimates bias and percentage relative bias as a function of the sample sizes. However, the sensitivity of both level sample sizes to the non-autocorrelation assumption violation is difficult to evaluate because of the inconsistency of the fixed effect estimates over conditions. Both FML and REML estimates of the fixed effect exhibited problematic bias levels, though lower than that from the correlated model. Both FML and REML biases exceeded 10% whenever 30, 50 or 100 subjects were sampled irrespective of the level of autocorrelation.

Under Table 4.37, when the number of measurement occasion is 5, FML PRB values ranged from -60.63 to 105.56, while REML PRB values ranged from -62.41 to 106.74; both FML and REML estimates exceeded 10% bias. Under Table 4.36, with the autocorrelation coefficient, $\rho = 0.9$ bias was more extreme with ML and REML. Similarly, when N , the number of subjects is small and moderate, (i.e., 30 and 50), both FML and REML were consistently negatively biased, with several values exceeding the threshold of 10% defining severe bias

occurring at $N = 30$ and 100 . FML exhibited substantial bias when compared to REML except when the number of subjects is 50 . When the two estimators were contrasted. Neither REML nor FML is more positively or negatively biased than each other.

Next, the accuracy of the fixed effect estimation was also examined with respect to design conditions. First, the effect of the autocorrelation coefficient (ρ) on the performance of the two-level time series model (TTSM) parameter estimate was examined, and bias was found to be increasing along with the values of autocorrelation coefficient. Second, the moderation effects of the level-1 and level-2 sample sizes on the bias of the fixed effect to the violation of non-autocorrelated error assumption were examined. However, the sensitivity of both level sample sizes to the non-autocorrelation assumption violation is difficult to evaluate because of the inconsistency of the fixed effect estimates over conditions. Both FML and REML estimates of the fixed effect exhibited problematic bias levels only at $N=30$, and 100 .

Under Table 4.38, when the number of measurement occasion is 10 , FML PRB values ranged from -82.17 to 98.42 , while REML PRB values ranged from -81.40 to 98.63 ; both FML and REML estimates exceeded 10% bias. Just like Tables 3 and 5, bias was at the extreme at 0.9 with ML and REML. Bias trend of FML when compared to REML is not clearly defined. Neither REML nor FML is more positively or negatively biased than each other.

Next, the accuracy of the fixed effect estimation was also examined with respect to design conditions. First, the effect of the autocorrelation coefficient (ρ) on the performance of the two-level time series model (TTSM) parameter estimate was examined, and bias was found to be increasing along with the values of autocorrelation coefficient except for the combination $(10, 0.3, 30)$. Second, the moderation effects of the level-1 and level-2 sample sizes on the bias of the fixed effect to the violation of non-autocorrelated error assumption were examined, just as

noted in Table 4.36 and Table 4.37 the sensitivity of both level sample sizes to the non-autocorrelation assumption violation is difficult to evaluate because of the inconsistency of the fixed effect estimates over conditions. Both FML and REML estimates of the fixed effect exhibited problematic bias levels when 30, 50, and 100 subjects were sampled.

Examination of raw biases (see Table 4.36–4.38) revealed that the patterns observed were somewhat inconsistent with the increase in the number of measurement occasion (t) and the number of subjects (N). All The three Tables lead to a different conclusion regarding estimators performances. Even with 100 subjects selected at varying numbers of repeated measurements both FML and REML were biased with more than 10%. Fixed effects are theoretically unbiased, with identical estimating equations for REML and FML, but results obtained in this study demonstrated that the asymptotic unbiased-ness of the fixed effect estimator was not robust to finite sample sizes. Though the MLEs of fixed effects differ negligibly between FML and REML.

4.4.1 The Effect of Non-autocorrelated error Assumption Violation on the Mean Square Error (MSE) of Fixed-Parameter Estimates

The second measure used to evaluate the performance of the TTSM parameter estimates in this study was Mean Square Error (MSE). MSE of parameter estimates quantify the sum of the variance and squared bias associated with each parameter. For biased outcomes, MSE can be interpreted as a measure of overall accuracy, combining both bias and accuracy of precision into a single value. Thus, in cases where the parameter is unbiased, the MSE quantifies sampling variance. The best estimate should have the smallest possible MSE.

In this study, the MSE of the fixed parameter estimates was calculated from 1000 replications for each examined condition. Moreover, the shrinkage estimator is optimal from one

statistical perspective, it minimizes the expected mean square, and MSE comprises of two components, one due to any systematic bias, and other reflecting random variation about the expected value of the estimator. It is important to show the magnitude that each error source contributed to MSE estimates across parameters and conditions to help interpret the results, Singh et al. (2004) state that root mean square error (RMSE) and mean absolute error (MAE) values less than half the standard deviation of the measured data may be considered low and that either is appropriate for model evaluation.

Table 4.39: Comparing the MSE of slopes for measurement occasion, $t = 3$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	0.4082	0.5675	0.6155	0.3983	0.5546	0.6018	0.2330	0.1873	0.1665	0.1975	0.1556	0.1368
50	0.1484	0.1945	0.2168	0.1593	0.2063	0.2291	0.1335	0.1505	0.1594	0.1372	0.1544	0.1634
100	0.3067	0.3921	0.4341	0.3143	0.4010	0.4436	0.2851	0.3385	0.3669	0.2786	0.3313	0.3594

Where Ω represents autocorrelation matrix and I is the identity matrix.

Table 4.40: Comparing the MSE of slopes for measurement occasion, $t = 5$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	0.2640	0.4381	0.4753	0.2735	0.4483	0.4857	0.1273	0.1381	0.1436	0.1191	0.1295	0.1349
50	0.045	0.1037	0.1163	0.0293	0.0955	0.1077	0.0026	0.0028	0.0029	0.0033	0.0035	0.0036
100	0.7005	0.9537	0.9930	0.6978	0.9486	0.9875	0.4103	0.4109	0.4112	0.4014	0.4020	0.4023

Where Ω represents autocorrelation matrix and I is the identity matrix.

Table 4.41: Comparing the MSE of slopes for measurement occasion, $t = 10$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	0.0717	0.0021	0.0123	0.0733	0.0017	0.0112	0.2390	0.2279	0.2225	0.2434	0.2323	0.2268
50	0.8668	1.8333	2.0583	0.8576	1.8161	2.0391	0.3444	0.3484	0.3504	0.3430	0.3470	0.3490
100	0.0649	0.1279	0.1414	0.0636	0.1249	0.1381	0.0278	0.0275	0.0274	0.0280	0.0278	0.0277

Where Ω represents autocorrelation matrix and I is the identity matrix.

Evaluating MSE

C-model and S-model

The violation of the first-level residual independence assumption has a major effect, as expected, on MSE of the estimates of the fixed effect. From the illustrating Tables 4.39–4.41, it was noticed that when the model was correctly specified the MSE for the fixed effect increased substantially (between 0.0021 and 2.0583 for FML and between 0.0017 and 2.0391 for REML for slope). Similarly, by omitting correlation matrix (i.e., misspecified model), the MSE for the fixed effect increased substantially (between 0.0026 and 0.4112 for FML, and between 0.0033 and 0.4023 for REML respectively). For the C-Model, lowest MSE occurs at $\rho = 0.7$, while the highest occurs at $\rho = 0.9$ similarly, for S-Model, $\rho = 0.3$ for the lowest and 0.9 for the highest.

In general, increasing autocorrelation level led to an increase in the MSE of the fixed effect estimates of the slope under correctly specified model except when $N=30$ which is not clearly defined. Similarly, increasing autocorrelation level led to an increase in the MSE of the fixed effect estimates of the slope under misspecified model when $N=50$ but not clearly defined for $N=30$, and 100.

Moreover, as expected, increasing sample size decreased the MSE. However, this effect varied across conditions. To explain, when the model was correctly specified, increasing the sample sizes at both the first and second level decreased the MSE of the estimates for slope, particularly with increase measurement occasion (i.e., from $t = 3$ to 5 decreased MSE estimates from (0.4082, 0.5675, 0.6155) to (0.2640, 0.4381, 0.4753) respectively across the values of autocorrelation coefficients under FML, Similarly, (from $N=30$ to 50) MSE decreased from (0.4082, 0.5675, 0.6155) to (0.1484, 0.1945, 0.2168) under FML, $t = 3$, however, this effect varied across conditions under both models. There was an increase or decrease in MSE of the

estimates of the fixed effect when either the number of measurement occasion or the number of subjects changed.

As observed in bias results, the effect of non-autocorrelation assumption violation on the MSE of the estimates of some fixed effect parameters seemed to depend on the autocorrelation level and sample sizes at both levels. In general, the MSE for the estimates of the slope when $t=5$ were less than when $t=3$, and that of $t=10$ less than that of $t=5$. However, it is important to notice that the same pattern of change was observed for the bias and MSE.

4.4.2 The Effect of Violations of non-autocorrelation assumption on the Variance of the Slope

In this section, we will examine how violation of the homoscedasticity (constant variance) assumption at the first level affected the accuracy of the precision of parameter point estimates, particularly under small, medium to large violations. The effect of violating this assumption depended to some degree on the study's sample sizes.

Table 4.42: Comparing the variance of slopes for measurement occasion, $t = 3$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	0.0155	0.0156	0.0157	0.0175	0.0177	0.0178	0.0187	0.0187	0.0187	0.0162	0.0162	0.0162
50	0.0099	0.0100	0.0101	0.0120	0.0121	0.0121	0.0080	0.0080	0.0080	0.0092	0.0092	0.0092
100	0.0062	0.0062	0.0062	0.0035	0.0035	0.0036	0.0043	0.0043	0.0043	0.0052	0.0052	0.0052

Where Ω represents autocorrelation matrix and I is the identity matrix.

Table 4.43: Comparing the variance of slopes for measurement occasion, $t = 5$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9	0.3	0.7	0.9
30	0.0033	0.0037	0.0038	0.0048	0.0056	0.0058	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034
50	0.0029	0.0033	0.0034	0.0021	0.0024	0.0025	0.0019	0.0019	0.0019	0.0019	0.0019	0.0019
100	0.0011	0.0013	0.0013	0.0010	0.0011	0.0011	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010

Where Ω represents autocorrelation matrix and I is the identity matrix.

Table 4.44: Comparing the variance of slopes for measurement occasion, $t = 10$ for two estimators under standard and correlated models, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	ML (Ω)			REML (Ω)			ML (I)			REML (I)		
	ρ			ρ			ρ			ρ		
	0.3	0.7	0.9	0.3	0.7	0.99	0.3	0.7	0.99	0.3	0.7	0.90
30	0.0004	0.0007	0.0008	0.0005	0.0008	0.0008	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
50	0.0003	0.0005	0.0006	0.0002	0.0002	0.0004	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
100	0.0001	0.0002	0.0002	0.0001	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001

Where Ω represents autocorrelation matrix and I is the identity matrix

Evaluating Variance C-model and S-model

From the illustrating Tables 42–44, it was noticed that when the model has correctly specified the variance for the fixed effect increased substantially (between 0.0001 and 0.0157 for FML and between 0.0001 and 0.0178 for REML for slope). Similarly, by omitting correlation matrix (i.e., misspecified model), the variance for the fixed effect increased substantially (between 0.0001 and 0.0187 for FML, and between 0.0001 and 0.0162 for REML respectively). For the C-Model, the lowest variance occurs at $\rho = 0.3$, while the highest occurs at $\rho = 0.9$ similarly, for S-Model, $\rho = 0.3$ for the lowest and 0.9 for the highest. Although there were differences in random errors, with FML errors negligibly smaller than that of REML, they were at the fourth decimal places. For the fixed effect parameters, the variances of the estimates for the slope under S-model were smaller compared to the variance for the estimates of the slope under C-model.

As expected theoretically by the large sample theory, the inaccuracy of fixed effect precision will increase, indicating less accuracy, when sample sizes were decreased either at the second or first level. This increase in inaccuracy values was more notable with small samples at the first level. For example, Under ML(Ω) when $t=3$, the precision of the fixed effect at the first level under varying autocorrelation level, decreases from (0.0155, 0.0156, 0.0157) to (0.0033,

0.0037, 0.0038) when $t=5$, also decreases to (0.0004, 0.0007, 0.0008) when $t=10$, a decrease of (78.7%, 76.3%, 75.8%) and (87.9%, 87.1%,78.9%) respectively across correlation level.

Similarly, overall, the most accurate precision associated with the fixed effect estimate under C-model for FML and REML occur at $\rho = 0.3, t = 10, \text{ and } N = 100$, but under S-model, the effect of autocorrelation level was disregarded and the least precision is constant across autocorrelation level. Though both bias and MSE of fixed effect estimates exhibited inconsistent trends across sample sizes at both level, precision monotonically increases for all fixed effect estimates of the slope as the number of measurement occasion and the number of subjects sampled increased. Of interest is the constant variance across autocorrelation level for the fixed number of subjects under S-model, by implication precision will either be underestimated or overestimated depending on the level of autocorrelation or sample sizes.

4.4.3 Systematic Versus Random Errors From MSE Components

It is notable from the results presented previously, that the trend for MSE under the different conditions is quite similar to the trend in the bias, particularly in regard to assumption violations and small sample sizes. The results indicated a large magnitude of contribution by the systematic error (bias of parameter estimates) compared to the random error(sampling error) in most of the conditions—more specifically when non-autocorrelated error assumption was violated. This means that MSE estimates are dominated by the bias values and that they reflect more the systematic bias picture. The separate estimates of the systematic and random error computed will afford the researcher the opportunity to evaluate the robustness of the parameter estimates.

4.4.4 Comparison Of The Estimators

In Chapter three, we examine methods of parameter estimation for two-level time series models, including frequentist methods and Bayesian methods. Each of the estimators has its

advantages and disadvantages. Therefore, it is better to make a comparison of these estimators. The relative bias using the true generating parameter value of 0.3 for the slope as the reference, allows us to check whether each estimator's estimate is overestimated or underestimated. Furthermore, the MSE criterion is used to compare these estimators. We will compare these methods using different values of autocorrelation and sample size.

Table 4.45: Comparing the bias of slopes for measurement occasion, $t = 3$ for ML estimators under the correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	$\rho=0.3$			$\rho=0.7$			$\rho=0.9$		
	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES
30	-0.6267	-0.6172	1.1733	-0.7429	-0.7328	1.1985	-0.7745	-0.7642	1.2694
50	0.3721	0.3838	-0.0732	0.4295	0.4407	0.0801	0.4547	0.4658	0.1037
100	0.5482	0.5575	-1.4643	0.6212	0.6304	-1.5022	0.6541	0.6633	-1.5084

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient

Table 4.46: Comparing the bias of slopes for measurement occasion, $t = 5$ for ML estimators under correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	$\rho=0.3$			$\rho=0.7$			$\rho=0.9$		
	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES
30	-0.5107	-0.5183	-2.2616	-0.6591	-0.6653	-2.3060	-0.6867	-0.6928	-2.3862
50	0.1776	0.1649	-0.2179	0.3169	0.3050	-0.2991	0.3361	0.3244	-0.2769
100	0.8363	0.8347	1.4139	0.9759	0.9734	1.4377	0.9959	0.9931	1.4282

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient

Table 4.47: Comparing the bias of slopes for measurement occasion, $t = 10$ for ML estimators under correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	$\rho=0.3$			$\rho=0.7$			$\rho=0.9$		
	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES
30	-0.2669	-0.2699	-2.3698	0.0363	0.0314	-2.4063	0.1070	0.1016	-2.4052
50	0.9309	0.9259	0.6963	1.3538	1.3475	0.6659	1.4345	1.4278	0.6148
100	0.2546	0.2519	-1.0733	0.3573	0.35308	-1.0296	0.3758	0.3713	-1.0258

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient

Under Tables 4.45–4.47, when the number of measurement occasion is 3, 5, and 10 ML absolute raw biases values ranged from 0.0314 to 1.4345 while the corresponding biases values under Bayes ranged from 0.0732 to 2.4063, ; In general, the ML raw biases for the estimates of the slope were very small compared to that of Bayes. At autocorrelation coefficient, $\rho = 0.9$ bias was more extreme with ML and Bayes estimates. As t and N increases, the ML estimators

show some advantages due to a reduction in bias, except under a moderate number of subject $N=50$.

In addition, while the bias values seemed to increase across autocorrelation levels, the results indicate that both level-1 and level-2 sample sizes are related to the magnitude of the bias in the fixed effect estimates, but the moderation effect of the sample size is much weaker.

Evaluating MSE

Table 4.48: Comparing the MSE of slopes for measurement occasion, $t = 3$ for ML estimators under correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	$\rho=0.3$			$\rho=0.7$			$\rho=0.9$		
	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES
30	0.4082	0.3983	4.3298	0.5675	0.5546	4.4073	0.6155	0.6018	4.6245
50	0.1484	0.1593	2.2969	0.1945	0.2063	2.4504	0.2168	0.2291	2.3537
100	0.3067	0.3143	3.7541	0.3921	0.4010	3.9479	0.4341	0.4436	4.011

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient

Table 4.49: Comparing the MSE of slopes for measurement occasion, $t = 5$ for ML estimators under correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	$\rho=0.3$			$\rho=0.7$			$\rho=0.9$		
	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES
30	0.2640	0.2735	9.8024	0.4381	0.4483	10.1033	0.4753	0.4753	10.1920
50	0.0345	0.0293	1.8349	0.1037	0.0955	1.8959	0.1163	0.1163	1.8166
100	0.7005	0.6978	3.1762	0.9537	0.9486	3.2999	0.9930	0.9930	3.2500

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient

Table 4.50: Comparing the MSE of slopes for measurement occasion, $t = 10$ for ML estimators under correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	$\rho=0.3$			$\rho=0.7$			$\rho=0.9$		
	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES
30	0.0717	0.0733	12.485	0.0021	0.0017	12.2027	0.0123	0.0112	12.3150
50	0.8668	0.8576	3.0731	1.8333	1.8161	2.9901	2.0583	2.0391	2.9465
100	0.0649	0.0636	2.1621	0.1279	0.1249	2.0894	0.1414	0.1381	2.0360

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient

Tables 4.48–4.50 indicate that when the sample size gets larger, from 30 to 50 for $t = 3$ and 5, the ML and Bayes estimators have a reduction in their MSE values but rise again when $t = 10$, while Bayes show a steady decrease in MSE values across increased subject sizes,

the trend of ML MSE estimates are undefined. In general, the MSE estimates of the slope were very small compared to that of Bayes.

Tables 4.48–4.50 depict the simulation results of MSE with different autocorrelation levels and sample sizes. When the level-1 sample size is small and moderate (i.e., $t = 3$ and 5), Tables 4.48–4.50 show that the MSE of both Bayes and ML estimates increased across autocorrelation levels except for Bayes when $N=50$ and $\rho = 0.9$ but inconsistent across sample sizes. When $t = 10$, ML estimates of MSE increased along with the values of autocorrelation coefficient for $N = 50$, and 100, except when $N = 30$ which is not clearly defined.

Evaluating Variance

Table 4.51 Comparing the Variance of slopes for measurement occasion, $t = 3$ for ML estimators under correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	$\rho=0.3$			$\rho=0.7$			$\rho=0.9$		
	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES
30	0.0155	0.0175	2.9532	0.0156	0.0177	2.9708	0.0157	0.0178	3.0130
50	0.0099	0.0120	2.2916	0.0100	0.0121	2.4440	0.0101	0.0121	0.4037
100	0.0062	0.0035	1.6098	0.0062	0.0035	1.6913	0.0062	0.0036	1.7350

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient

Table 4.52: Comparing the variance of slopes for measurement occasion, $t = 5$ for ML estimators under correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient ρ as a function of the slope.

N	$\rho=0.3$			$\rho=0.7$			$\rho=0.9$		
	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES
30	0.0033	0.0048	4.6876	0.0037	0.0056	4.7858	0.0038	0.0058	4.4980
50	0.0029	0.0021	1.7874	0.0033	0.0024	1.8064	0.0034	0.0025	1.7399
100	0.0011	0.0010	1.1772	0.0013	0.0011	1.2329	0.0013	0.0011	1.7399

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient

Table 4.53: Comparing the variance of slopes for measurement occasion, $t = 10$ for ML estimators under correlated model and Bayes, for different numbers of subject N and different autocorrelation coefficient as a function of the slope.

N	$\rho=0.3$			$\rho=0.7$			$\rho=0.9$		
	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES	ML (Ω)	REML(Ω)	BAYES
30	0.0004	0.0005	6.8689	0.0007	0.0008	6.4122	0.0008	0.0008	6.5302
50	0.0003	0.0002	2.5882	0.0005	0.0002	2.5467	0.0006	0.0004	2.5686
100	0.0001	0.0001	1.0100	0.0002	0.0002	1.0293	0.0002	0.0002	0.9839

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient

In Tables 4.51–4.53 under Bayes, the variance of the residuals is far from stable; when $t = 3, \rho = 0.9$, it first decreases and then increases over time and similarly over the number of measurement occasion. In general, unlike the pattern of decreasing variance of the fixed effect estimates as a result of increasing, sample sizes at both level associated with ML, Bayes estimates of fixed effect decreases across the level-2 sample sizes, only for a fixed sample size at level-1. In sum, the variances of the fixed effects estimates under ML were generally smaller than that of Bayes.

4.4.5 Evaluating error indices under misspecification of the autocorrelation coefficient

Fixed effect parameters are usually the focus of scientific interest in the linear mixed model. However, it is important to correctly specify the covariance structure to obtain valid statistical inferences for the fixed effects. Our simulated model involves the interrelated set of specifications and assumptions, ρ the autocorrelation coefficient is one of the specifications when estimating the model's parameters for the correlated model, our interest centres on the effect and their tenability in the presence of misspecification.

Evaluating Bias

Table 4.54: Comparing the bias of slopes for measurement occasion, $t = 3$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N .

N	<i>ML</i> (Ω)		<i>REML</i> (Ω)		<i>ML</i> (Ω)		<i>REML</i> (Ω)	
	ρ		ρ		ρ		ρ	
	0.3/0.3	0.3/0.7	0.3/0.3	0.3/0.7	0.7/0.7	0.7/0.3	0.7/0.7	0.7/0.3
30	-0.6267	-0.5755	-0.6172	-0.5660	-0.7429	-0.7929	-0.7328	-0.7828
50	0.3721	0.3981	0.3838	0.4097	0.4295	0.4009	0.4407	0.4122
100	0.5482	0.5970	0.5575	0.6063	0.6212	0.5718	0.6304	0.5810

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient.

Table 4.55: Comparing the bias of slopes for measurement occasion, $t = 5$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N .

	$ML(\Omega)$		$REML(\Omega)$		$ML(\Omega)$		$REML(\Omega)$	
N	ρ		ρ		ρ		ρ	
	0.3/0.3	0.3/0.7	0.3/0.3	0.3/0.7	0.7/0.7	0.7/0.3	0.7/0.7	0.7/0.3
30	-0.5107	-0.5342	-0.5183	-0.5419	-0.6591	-0.6291	-0.6653	-0.6354
50	0.1776	0.1690	0.1649	0.1563	0.3169	0.3293	0.3050	0.3174
100	0.8363	0.8378	0.8347	0.8362	0.9759	0.9736	0.9734	0.9711

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient.

Table 4.56: Comparing the bias of slopes for measurement occasion, $t = 10$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N .

	$ML(\Omega)$		$REML(\Omega)$		$ML(\Omega)$		$REML(\Omega)$	
N	ρ		ρ		ρ		ρ	
	0.3/0.3	0.3/0.7	0.3/0.3	0.3/0.7	0.7/0.7	0.7/0.3	0.7/0.7	0.7/0.3
30	-0.2669	-0.2564	-0.2699	-0.2594	0.0363	0.0276	0.0314	0.0226
50	0.9309	0.9332	0.9259	0.9282	1.3538	1.3529	1.3475	1.3466
100	0.2546	0.2536	0.2519	0.2509	0.3573	0.3586	0.3531	0.3544

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient.

As reported in Tables 4.54–4.56, generally the bias of the fixed effect estimation examined under misspecified autocorrelation coefficient indicate that bias is either underestimated or overestimated. For example, in Table 4.54 when $t = 3$ and ρ the degree of correlation among the subject data which is 0.7 has been misspecified as 0.3 during parameter estimation, then the estimated biases across the various number of subjects were (-0.5755, 0.3981, and 0.5970) instead of expected (-0.7429, 0.4295, and 0.6212), a decrease of (22.53%, 7.31%, and 3.90%) under $ML(\Omega)$ in the result of our analysis. Similarly, in Table 4.55 when $t = 5$ and ρ the degree of correlation among the subject data which is 0.3 has been misspecified as 0.7 during parameter estimation, then the estimated biases were (-0.6291, 0.3293, and 0.9736) instead of expected (-0.5107, 0.1776, and 0.8363), an increase of (23.18%, 85.4% and 16.41%) respectively under $ML(\Omega)$.

Evaluating MSE

Next, the accuracy of the fixed effect estimation was also examined with respect to design condition under MSE in case misspecification of variance-covariance component occurred during parameter estimation. From the illustrating Tables 4.57–4.59, it was noticed that when the model variance-covariance component was misspecified generally by underestimating the size of the autocorrelation coefficient in the fixed effect estimation, the MSE estimated values were underestimated, a pattern of underestimation or overestimation similar to that of bias in Tables 4.54–4.56 was observed across study conditions, however, MSE values appeared to either being more dramatically underestimated or overestimated.

Table 4.57: Comparing the MSE of slopes for measurement occasion, $t = 3$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N .

	ML (Ω)		REML (Ω)		ML (Ω)		REML (Ω)	
N	ρ		ρ		ρ		ρ	
	0.3/0.3	0.3/0.7	0.3/0.3	0.3/0.7	0.7/0.7	0.7/0.3	0.7/0.7	0.7/0.3
30	0.4082	0.3467	0.3983	0.3378	0.5675	0.6444	0.5546	0.6305
50	0.1484	0.1684	0.1593	0.1799	0.1945	0.1708	0.2063	0.1820
100	0.3067	0.3626	0.3143	0.3711	0.3921	0.3332	0.4010	0.3411

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient.

Table 4.58: Comparing the MSE of slopes for measurement occasion, $t = 5$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N .

	ML (Ω)		REML (Ω)		ML (Ω)		REML (Ω)	
N	ρ		ρ		ρ		ρ	
	0.3/0.3	0.3/0.7	0.3/0.3	0.3/0.7	0.7/0.7	0.7/0.3	0.7/0.7	0.7/0.3
30	0.2640	0.2886	0.2735	0.2984	0.4381	0.3995	0.4483	0.4093
50	0.0345	0.0315	0.0293	0.0265	0.1037	0.1117	0.0955	0.1032
100	0.7005	0.7030	0.6978	0.7003	0.9537	0.9492	0.9486	0.9441

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient.

Table 4.59: Comparing the MSE of slopes for measurement occasion, $t = 10$ for ML estimators under C- model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N .

N	ML (Ω)		REML (Ω)		ML (Ω)		REML (Ω)	
	ρ		ρ		ρ		ρ	
	0.3/0.3	0.3/0.7	0.3/0.3	0.3/0.7	0.7/0.7	0.7/0.3	0.7/0.7	0.7/0.3
30	0.0717	0.0662	0.0733	0.0678	0.0021	0.0015	0.0017	0.0013
50	0.8668	0.8711	0.8576	0.8619	1.8333	1.8309	1.8161	1.8138
100	0.0649	0.0644	0.0636	0.0631	0.1279	0.1288	0.1249	0.1258

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient.

Next, the accuracy of the fixed effect estimation was also examined with respect to design condition under MSE in case misspecification of variance-covariance component occurred during parameter estimation. From the illustrating Tables 4.57–4.59, it is noticed that when the model variance-covariance component was misspecified generally by underestimating the size of the autocorrelation coefficient in the fixed effect estimation, the MSE estimated values were underestimated, a pattern of underestimation or overestimation similar to that of bias in Tables 4.54–4.56 was observed across study condition, however, MSE values appeared to either being more dramatically underestimated or overestimated.

Evaluating Variance

Table 4.60: Comparing the variance of slopes for measurement occasion, $t = 3$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N .

N	ML (Ω)		REML (Ω)		ML (Ω)		REML (Ω)	
	ρ		ρ		ρ		ρ	
	0.3/0.3	0.3/0.7	0.3/0.3	0.3/0.7	0.7/0.7	0.7/0.3	0.7/0.7	0.7/0.3
30	0.0155	0.0155	0.0175	0.0175	0.0156	0.0156	0.0177	0.0177
50	0.0099	0.0099	0.0120	0.0120	0.0100	0.0100	0.0121	0.0121
100	0.0062	0.0062	0.0035	0.0035	0.0062	0.0062	0.0035	0.0035

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient.

Table 4.61: Comparing the variance of slopes for measurement occasion, $t = 5$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N .

	ML (Ω)		REML (Ω)		ML (Ω)		REML (Ω)	
N	ρ		ρ		ρ		ρ	
	0.3/0.3	0.3/0.7	0.3/0.3	0.3/0.7	0.7/0.7	0.7/0.3	0.7/0.7	0.7/0.3
30	0.0033	0.0033	0.0048	0.0048	0.0037	0.0037	0.0056	0.0056
50	0.0029	0.0029	0.0021	0.0021	0.0033	0.0033	0.0024	0.0024
100	0.0011	0.0011	0.0010	0.0010	0.0013	0.0013	0.0011	0.0011

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient.

Table 4.62: Comparing the variance of slopes for measurement occasion, $t = 10$ for ML estimators under C-model in case $\rho = 0.7$ has been misspecified as $\rho = 0.3$ and vice versa, for different numbers of subject N .

	ML (Ω)		REML (Ω)		ML (Ω)		REML (Ω)	
N	ρ		ρ		ρ		ρ	
	0.3/0.3	0.3/0.7	0.3/0.3	0.3/0.7	0.7/0.7	0.7/0.3	0.7/0.7	0.7/0.3
30	0.0004	0.0004	0.0005	0.0005	0.0007	0.0007	0.0008	0.0008
50	0.0003	0.0003	0.0002	0.0002	0.0005	0.0005	0.0002	0.0002
100	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002	0.0002

Where Ω represents autocorrelation matrix and ρ is the autocorrelation coefficient.

As reported in Tables 4.60–4.62, generally the same pattern was observed for the fixed effect variance under variance-covariance component misspecification; however, the variances for the estimates for the slopes resulting from misspecification were very small when either being underestimated or overestimated compared to the results recorded under MSE.

4.5 The Effect of Non-Auto correlation First -level Assumption Violation on the Accuracy of the Random Parameter Estimates As a Method of Model Validation.

In this section, we present a real data example to illustrate the utility of our model. So far, we have focused on settings with the primary interest in inference about fixed effects by building a two-level time series model for a set of objectives and its validity will be determined for that purpose by examining the covariance structure. Indeed, just as model development will be influenced by the objectives of the performance study, so will model validation be. Even with a focus on fixed effects, accurate inference requires selecting an adequate approximation for the

true covariance structure. Unfortunately, although many model selection criteria have been suggested, none has been found to be clearly superior. Part of the difficulty stems from the interaction between fixed effect and covariance model selection (Mathew, Lloyd and Keith, 2011). Overall, one principle seems clear: controlling type I error for tests of fixed effects demands avoiding an under fitted covariance model.

In order to substantiate that our simulation models within their domain of applicability possess a satisfactory range of accuracy consistent with the intended application of the models, comparing the model input-output transformations to corresponding input-output transformations for the real-life data sets to ensure the validity of the exchangeability assumption seems necessary. To achieve this, we will consider the published psychiatric dataset from Bock (1983b), a real-life data from a psychiatric study described in Reisby et al. (1977) to validate our simulated model. As described by Hedeker and Gibbons (2006), the study focused on the longitudinal relationship between imipramine (IMI) and desipramine (DMI) plasma levels and clinical response in 75 depressed inpatients. Imipramine is the prototypic drug in the series of compounds known as tricyclic antidepressants and is commonly prescribed for the treatment of major depression (Seiden and Dykstra, 1977). Although the total number of subjects in the study was 75, only 54 subjects with all measures at each of the weeks were considered in our data analysis. If a comparison of system measurements and model results suggests that the results produced by the model are close to those obtained from the system, then the implemented model is assumed to be both a verified implementation of the assumptions and a valid representation of the system.

As a preliminary step, initial validation attempts were made in section 4.2 by examining data assumptions, next, we will employ the technique of a real system measurement of model

validation, by comparing the variance-covariance estimates. In order to facilitate comparison between observed unstructured variance-covariance and estimated variance-covariance structure from C and S models under two estimators, true covariance structures calculated from models adopting the j^{th} ($j = 1, 2$) estimators were fitted to models adopting each of the i^{th} ($i = 1, 2$) structures (i.e., simple and compound symmetry). To determine how representative the estimated variance-covariance output are, observed unstructured covariance matrices calculated were compared with the estimated ones as shown in Tables 4.63–4.82 below. With unstructured covariance matrix, there are no constraints, each variance and each covariance is estimated uniquely from the observed data, which will results in the best possible model fit, because each variance and covariance values are expected to be very close to what the data reflect.

Table 4.63: Observed unstructured variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix.

unstructured covariance matrix				Corresponding correlation matrix		
19.2456	12.7667	10.8533		1	0.5155	0.5006
12.7667	31.8667	14.6000		0.5155	1	0.5233
10.8533	14.6000	24.4267		0.5006	0.5233	1

Table 4.64: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE with autocorrelated errors

57.8381	30.0475	18.9309		1	0.5365	0.3170
30.0475	54.2388	31.9545		0.5365	1	0.5526
18.9309	31.9545	61.6520		0.3170	0.5526	1

Table 4.65: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE with autocorrelated errors

58.2696	30.3703	19.2124		1	0.5382	0.3196
30.3703	54.6420	32.2499		0.5382	1	0.5540
19.2124	32.2499	62.0289		0.3196	0.5540	1

Table 4.66: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under FMLE without autocorrelated errors

55.3827	18.6047	14.1610		1	0.3467	0.2465
18.6047	51.9943	20.7148		0.3467	1	0.3721
14.1610	20.7148	59.6032		0.2465	0.3721	1

Table 4.67: Estimated variance-covariance matrix for $t = 3$, $N = 30$, and $\rho = 0.3$, with the corresponding correlation matrix under RMLE without autocorrelated errors

55.5542	18.7836	14.3304		1	0.3490	0.2488
18.7836	52.1471	20.8759		0.3490	1	0.3740
14.3304	20.8759	59.7388		0.2488	0.3740	1

Table 4.68: Observed unstructured variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix.

unstructured covariance matrix				Corresponding correlation matrix		
18.2144	11.4984	11.2408		1	0.5562	0.5078
11.4984	23.4624	12.2888		0.5662	1	0.4891
11.2408	12.2888	26.9056		0.5078	0.4891	1

Table 4.69: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix under FMLE with autocorrelated errors

48.0178	22.1853	13.8732		1	0.4739	0.2719
22.1853	45.6428	25.2830		0.4739	1	0.5083
13.8732	25.2830	54.2132		0.2719	0.5083	1

Table 4.70: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix under RMLE with autocorrelated errors

48.0941	22.2646	13.9511		1	0.4748	0.2731
22.2646	45.7143	25.3576		0.4748	1	0.5091
13.9511	25.3576	54.2800		0.2731	0.5091	1

Table 4.71: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix under FMLE without autocorrelated errors

47.1226	15.8007	11.9036		1	0.3439	0.2373
15.8007	44.7999	18.9496		0.3439	1	0.3874
11.9036	18.9496	53.4204		0.2373	0.3874	1

Table 4.72: Estimated variance-covariance matrix for $t = 3$, $N = 50$, and $\rho = 0.2$, with the corresponding correlation matrix under RMLE without autocorrelated errors

47.1538	15.8568	11.9579		1	0.3240	0.2382
15.8568	44.8277	19.0024		0.3240	1	0.3882
11.9579	19.0024	53.4449		0.2382	0.3882	1

Table 4.73: Observed unstructured variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix.

18.0974	10.3258	9.5048	11.8786	9.8409		1	0.5090	0.4223	0.4105	0.3104
10.3258	22.7394	12.5110	15.0638	12.1866		0.5090	1	0.4959	0.4644	0.3429
9.5048	12.5110	27.9863	27.7675	28.4335		0.4223	0.4959	1	0.7716	0.7211
11.8786	15.0638	27.7675	46.2747	42.2510		0.4105	0.4644	0.7716	1	0.8333
9.8409	12.1866	28.4335	42.2510	55.5542		0.3104	0.3429	0.7211	0.8333	1

Table 4.74: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix under FMLE with autocorrelated errors

67.4762	37.7742	27.7012	18.1351	14.1152		1	0.5659	0.3563	0.2422	0.1705
37.7742	66.0301	39.6485	29.8958	26.6501		0.5659	1	0.5782	0.4037	0.3255
27.7012	39.6485	71.2248	48.1634	41.7311		0.3563	0.5782	1	0.6262	0.4907
18.1351	29.8958	48.1634	83.0600	63.3190		0.2422	0.4037	0.6262	1	0.6895
14.1152	26.6501	41.7311	63.3190	101.5358		0.1705	0.3255	0.4907	0.6895	1

Table 4.75: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix under RMLE with autocorrelated errors

67.6484	37.9376	24.8584	18.2871	14.2622		1	0.5669	0.3577	0.2438	0.1720
37.9376	66.1933	39.8030	30.0443	26.7934		0.5669	1	0.5791	0.4048	0.3265
24.8584	39.8030	71.3792	48.3094	41.8711		0.3577	0.5791	1	0.6269	0.4915
18.2871	30.0443	48.3094	83.2061	63.4567		0.2438	0.4048	0.6269	1	0.6899
14.2622	26.7934	41.8711	63.4567	101.6740		0.1720	0.3265	0.4915	0.6899	1

Table 4.76: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix under FMLE without autocorrelated errors

63.4278	16.2713	14.0013	11.7313	9.4613		1	0.2590	0.2138	0.1651	0.1198
16.2713	62.2038	18.3634	19.4095	20.4555		0.2590	1	0.2832	0.2757	0.2616
14.0013	18.3634	67.6119	27.0876	31.4497		0.2138	0.2832	1	0.3691	0.3857
11.7313	19.4095	27.0876	79.6521	42.4438		0.1651	0.2757	0.3691	1	0.4796
9.4613	20.4555	31.4497	42.4438	98.3244		0.1198	0.2616	0.3857	0.4796	1

Table 4.77: Estimated variance-covariance matrix for $t = 5$, $N = 50$, and $\rho = 0.4$, with the corresponding correlation matrix under RMLE without autocorrelated errors

63.5023	16.3425	14.0704	11.7984	9.5263		1	0.2599	0.2146	0.1658	0.1205
16.3425	62.2743	18.4306	19.4747	20.5188		0.2599	1	0.2839	0.2764	0.2621
14.0704	18.4306	67.6786	27.1510	31.5112		0.2146	0.2839	1	0.3697	0.3862
11.7984	19.4747	27.1510	79.7151	42.5036		0.1658	0.2764	0.3697	1	0.4800
9.5263	20.5188	31.5112	42.5036	98.3838		0.1205	0.2621	0.3862	0.4800	1

Table 4.78: Observed unstructured variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix.

18.2144	11.4984	11.2408	13.8304	11.7912		1	0.5562	0.5078	0.4898	0.3728
11.4984	23.4624	12.2888	15.0544	11.7712		0.5562	1	0.4891	0.4698	0.3299
11.2408	12.2888	26.9056	25.7328	26.7944		0.5078	0.4891	1	0.7499	0.7012
13.8304	15.0544	25.7328	43.7664	39.8272		0.4898	0.4698	0.7499	1	0.8172
11.7912	11.7712	26.7944	39.8272	54.2656		0.3728	0.3299	0.7012	0.8172	1

Table 4.79: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix under FMLE with autocorrelated errors

67.6106	37.6081	25.1163	19.1319	15.5657		1	0.5619	0.3631	0.2584	0.1919
37.6081	66.2608	39.1807	29.6112	26.5491		0.5619	1	0.5722	0.4040	0.3307
25.1163	39.1807	70.7557	46.5979	39.9508		0.3631	0.5722	1	0.6152	0.4815
19.1319	29.6112	46.5979	81.0953	59.8599		0.2584	0.4040	0.6152	1	0.6740
15.5657	26.5491	39.9508	59.8599	97.2797		0.1919	0.3307	0.4815	0.6740	1

Table 4.80: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix under RMLE with autocorrelated errors

67.8119	37.7917	25.2899	19.2982	15.7259		1	0.5630	0.3646	0.2600	0.1935
37.7917	66.4513	39.3537	29.7745	26.7054		0.5630	1	0.5732	0.4052	0.3319
25.2899	39.3537	70.9359	46.7609	40.1042		0.3646	0.5732	1	0.6159	0.4824
19.2982	29.7745	46.7609	81.2657	60.0132		0.2600	0.4052	0.6159	1	0.6744
15.7259	26.7054	40.1042	60.0132	97.4406		0.1935	0.3319	0.4824	0.6744	1

Table 4.81: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix under FMLE without autocorrelated errors

63.6415	17.2808	15.2644	13.2480	11.2315		1	0.2740	0.2333	0.1883	0.1451
17.2808	62.5261	19.0829	19.9839	20.8850		0.2740	1	0.2943	0.2865	0.2722
15.2644	19.0829	67.2457	26.7199	30.5384		0.2333	0.2943	1	0.3694	0.3837
13.2480	19.9839	26.7199	77.8002	40.1919		0.1883	0.2865	0.3694	1	0.4695
11.2315	20.8850	30.5384	40.1919	94.1896		0.1451	0.2722	0.3837	0.4695	1

Table 4.82: Estimated variance-covariance matrix for $t = 5$, $N = 30$, and $\rho = 0.4$, with the corresponding correlation matrix under RMLE without autocorrelated errors

63.7194	17.3626	15.3436	13.3246	11.3056		1	0.2749	0.2343	0.1892	0.1459
17.3626	62.5990	19.1598	20.0584	20.9569		0.2749	1	0.2952	0.2873	0.2728
15.3436	19.1598	67.3138	26.7921	30.6083		0.2343	0.2952	1	0.3701	0.3843
13.3246	20.0584	26.7921	77.8637	40.2596		0.1892	0.2873	0.3701	1	0.4700
11.3056	20.9569	30.6083	40.2596	94.2488		0.1459	0.2728	0.3843	0.4700	1

When $t = 3$, Table 4.63 and Table 4.68 are the observed variance-covariance matrices, the variances over the 3-time points for Table 4.63 vary from 19.2456 at time 0 to 24.4267; vaguely increasing with the variance at time $t = 2$ highest, while for Table 4.68 it varies from 18.2144 to 26.9056. Specifically, Tables 4.64–4.67 and Tables 4.69–4.72 are the variance-covariance matrices of the repeated measures given the model covariates X , which are the time trends. By comparing the observed matrices with the estimated variance-covariance matrices, presented in Tables 4.64–4.67 and Tables 4.69–4.72, the results show that the estimated variances are not close to the observed, though the models are clearly picking up the pattern of diminishing covariance away from the diagonal and at the earlier time points. Comparing the deviances of the C-model with that of the S-model using a likelihood ratio test under observed variance-covariance structure yields a significant results at 0.05, indicating that the more complex model (i.e., C-model) is preferred to S-model, arguing that there is a non-negligible probability that the observed difference in the model fits could reasonably not be the result of random chance, but the inherent superiority of the model. Clearly, both of these interpretations on model fits under observed variance-covariance matrix and under simulation conditions are reasonable and consistent with each other. Correlation structure is important for achieving

efficient parameter estimators; this aspect can be vital, especially for data sets with many observations over time. Also, considering the corresponding correlation matrices show that the correlations decrease as the time interval is increased, as can be seen, though the general pattern is similar, estimated correlation matrices under correlated models are a bit close to the observed one relative to estimated correlation matrices under standard models. Increasing the number of subjects from 30 in Table 4.63 to 50 in Table 4.68 indicate that recruiting more participants will not alleviate the bias rate associated with fixed effect inference.

In Table 4.63, taking the square root of the diagonal entries of the observed variance-covariance yields the standard deviations across time as 4.3870, 5.6451, and 4.9423 while the corresponding estimated standard deviations across time under ML (Ω), REML (Ω) ML (I), and REML (I) are (7.6051, 7.3647, and 7.8519), (7.6335, 7.3920, and 7.8758), (7.4420, 7.2107, and 7.7203) and (7.4535, 7.2213, and 7.7291), respectively. Similarly, in Table 4.68, taking the square root of the diagonal entries of the observed variance-covariance yields the standard deviations across time as 4.2678, 4.8438, and 5.1871 while the corresponding estimated standard deviations across time under ML (Ω), REML (Ω) ML (I), and REML (I) are (6.9295, 6.7559, and 7.3630), (6.3320, 6.7612, and 7.3675), (6.8646, 6.6933, and 7.3089) and (6.8669, 6.6953, and 7.3106), respectively. These estimated results did not clearly represent the observed one across the study time points, this gives some sense of why the estimated variance-covariance matrices did not fit the data well under small sample condition. Generally, with the observed variance-covariance matrices, results from the simulated models revealed a negatively biased inference for the fixed effects (which is consistent with the results obtained earlier) with biased covariance matrices in small samples. Even with a correctly specified covariance model, C-model has a high bias for smaller sample sizes at both levels. The underspecified S-model that

assumes simple random coefficients and conditional independence of the within-subject errors (i.e., compound symmetry) creates less bias in the variance-covariance matrix under small sample (*i.e.*, $N = 30, t = 3$) compared to our C-model.

Similarly, when $t = 5$, by comparing the estimated matrices of Tables 4.74 – 4.77 and Tables 4.79 – 4.82 with the observed variance-covariance matrices in Table 4.73 and Table 4.78, the results show that the estimated variances are not close to the observed. Just like the pattern in the observed matrices, for the estimated covariance matrices, the standard deviations increase across time indicating the growing spread in repeated measures across the study time points, and the correlations generally diminish as the time interval increases. Though the general patterns are similar, S-model's estimated correlations are a bit close to the observed relative to C-model, and pattern on the covariance structure may have important substantive implications (see Leckie & Goldstein, 2009). Apart from the issue of small sample sizes at both levels, another reason why these estimated standard deviations and correlations are not closer to the observed values given that the unstructured variance-covariance matrix was estimated, is that $\hat{\Sigma}$ is an estimate of the conditional variance-covariance matrix of y , given the covariates in $\mathbf{X}$, whereas the observed variance-covariance matrix consists of marginal statistics (Hedeker and Gibbons, 2006). Overall, our simulation results in support of the theoretical result and to compare the results in the scenario of the real data revealed that maximum likelihood estimators of variance parameters are biased especially under small to moderate sample sizes.

CHAPTER FIVE

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

5.0 Introduction

The present study and conclusions of the analysis results, as well as the study recommendations, are summarized and discussed in this chapter. This chapter contains four sections in which we consider the following; 1) the effects of autocorrelated errors of the level-1 residuals on the performance of the two-level time series models; 2) the general summary of the results; 3) conclusions of the present study; 4) limitations of the present study and suggestions for future research.

As with other statistical techniques, the validity of the estimates and associated significance tests of the multilevel time series model depend on how well data meet associated assumptions. These violations, seen in widespread applications in operations research, like in queuing theory, clinical trials and econometrics, where the usual assumption of independent error terms may not be plausible in most cases. Burstein (1989) reminds researchers that the new and more powerful methods could produce very complex, yet very inaccurate, results if data assumptions are not carefully considered. Hoekstra, Henk and Addie (2012) indicates that such assumptions are rarely reported in the literature and that some researchers might be unfamiliar with the techniques and remedies that are pertinent to the statistical tests they conduct.

Due to the lack of comprehensive knowledge about the robustness of TTSM parameter estimates to non-autocorrelation first-level assumption violations, this study was conducted to investigate the performance of TTSM analysis with repeated measures data under the condition that there is an interdependence between level-1 residuals and level-1 predictors and heteroscedasticity of the level-1 variance. Most of the available studies concerned with the

robustness of the HLM parameter estimates to the first-level assumption violations on error dependency were restricted to null model whose assumptions are more restrictive, more specifically other available studies are concerned with the non-normality of the error term at both levels, multicollinearity, etc under different design conditions as discussed extensively under review of related literature in chapter two.

Summary of Results

5.1. Effects of Autocorrelated Errors of the Level-1 Residuals:

Multilevel models share with many traditional models the assumption that the residuals at each level are homoscedastic, i.e., have constant variance and covariances, and do not depend on the particular values of the explanatory variable(s) included in the model. As earlier stated under literature review, the simulation study of Ebbes et al. (2004) discovered that for a moderate dependence of the level-1 predictor on the level-1 random component ($\rho = 0.3$), there are upward biases in the estimates of the fixed effects, in contrast to the downward bias observed in estimates of the level-1 random component in the random analysis of covariance (ANCOVA) model. Similarly, Raudenbush and Bryk (2002) found that heteroscedasticity of the level-1 variances (σ_j^2) itself will not lead to bias in estimation of level-2 fixed effect parameters and associated standard errors. However, heteroscedasticity of level-1 variances might result from the misspecification of the level-1 model. Such misspecification due to the omission of an important level-1 predictor can introduce bias in the estimation of fixed effects. The present study was conducted to investigate the effect of the interdependence of the level-1 residuals and level-1 predictors on the performance of TTSM.

While there are several conditions under which interdependence exists, the present study focused on that caused by autoregressive errors in the data generating process, though

autocorrelation is taken as a signal of misspecification, it is not rare that when multiple observations are taken on the same sampling unit over time, it is usually plausible to assume that observations on the same unit are correlated. It is hopeful that this study will provide new findings regarding the robustness of TTSM parameter estimates to the violations of the assumption of homoscedasticity of first-level residuals and independence of residuals.

The answers provided by the results of this study to each question below based on the objectives of the study are as follows:

1. *What are the consequences of violation of non-autocorrelation assumption on the parameter estimates and model fit of the two-level time series modelling technique?*

Violating the assumption of the non-autocorrelated errors affects the accuracy of the parameter estimation in TTSM. However, the severity of this bias depends on the degree to which errors are autocorrelated. The results suggested that estimates of both fixed and random parameters in TTSM are robust to the moderate and severe violations of this assumption. The present study found a severe bias in the estimation of level-2 fixed effects in the random coefficient model for highly correlated level-1 residuals. Similarly, correlated models exhibit higher bias compared to standard multilevel models, but standard multilevel models are more negatively biased than the correlated models. According to statistical literature, for random effect models, precision will decrease with increase heterogeneity and confidence intervals will widen correspondingly but in our simulation results with correlated errors, MLE has produced a corner solution with lost heterogeneity effect. The difference between FMLE and RMLE estimates were very small and inconsistent over conditions. FML percentage relative bias exceeded that of RMLE under C-model, similarly, FML exhibited higher PRB than that of REML under S-model. Though estimating equations for REML and FML are identical, thereby making the MLEs of

fixed effects differ negligibly between FML and REML. Their fixed effects are theoretically unbiased but result obtained in this study demonstrated that the asymptotic unbiased-ness of the fixed effect estimator was not robust to the selected sample size condition at both levels.

Similarly, the mean square error (MSE) of the estimators is contrasted across fixed effect estimates, however, the magnitude of MSE in the parameter estimates depends on the degree of correlation between subjects' data especially, under C-model compared to S-model. MSE of parameter estimates quantify the sum of the variance and squared bias associated with each parameter. For biased outcomes, MSE can be interpreted as a measure of overall accuracy, combining both bias and accuracy of precision into a single value. From the results of our analysis, FML exhibits greater MSE when compared with REML MSE estimates, contrary to prior simulation and theoretical work that demonstrated that for some models, REML, though uniformly less bias than FML may exhibit greater MSE than FML as model complexity increases (Corbeil and Searle,1976). A necessary consequence of these facts with respect to the results of the study is that for two equally biased estimates, FML with less MSE is a measure of efficiency upon which we would expect REML to outperform FML. Overall, the findings of the present study were consistent with those of previous studies with regard to the robustness of multilevel model to the violation of statistical assumptions. According to the results of previous studies, Demidenko (2004) state that, (i) for some (co)variance parameters, when few subjects are sampled, no matter how many observations are sampled per unit, problematic bias remains, (ii) Complex models exhibit greater bias than simpler models, and (iii) Full maximum likelihood (FML) estimates are more biased than Restricted maximum likelihood (REML) estimates. Raudenbush and Bryk (2002) reported that RML estimates variance components after removing

the fixed effects from the model, can lead theoretically to less bias than FML, especially when the number of groups is small.

Generally from the results of the study simulation, increasing the sample sizes at both levels show that the variances of the estimated fixed effect go to zero an indication that MSE is consistent. The fixed effect estimates are biased, an indication that a trade-off had occurred between variance and bias in such a way that a small increase in bias can be traded for a larger decrease in variance, resulting in an improvement in MSE. Here the accuracy issues appear to relate more to variability than to bias, the bias is under-estimated and overestimated under different simulation condition (as expected) but the variability falls rapidly with increasing sample sizes at both levels. This result is consistent with theory, we are using estimators that incorporate shrinkage effect and the biased estimator is expected to reduce some loss function, particularly MSE. Although many unbiased estimators are also reasonable from the standpoint of MSE, controlling bias does not guarantee that MSE is controlled.

For the comparison of the frequentist approach with that of Bayes, we perform a Monte Carlo simulation study to compare the estimates, obtained from maximum likelihood and Bayes estimation techniques under various conditions. Estimates based on a Bayes approach and a maximum likelihood approach were compared on the basis of their biases (differences between estimates and true values of the slope), mean square errors and the variance in each of the simulation experiments. The maximum likelihood estimation method for parameters gives very good results in terms of bias, MSEs, and variance by reducing the variability in recovered parameters. Under few conditions, little overall difference was observed between the ML and Bayesian methods, and the Bayes approaches gave more accurate estimates under bias for the

simulated participants, although this advantage is mostly restricted to medium numbers of subjects ($N = 50$).

Model fit

On the impact of autocorrelation on the efficiency of multilevel time series modelling techniques, TTSM_S are biased and inconsistent when there is autocorrelation but efficient. Our purpose on model fit is to determine whether the change in the quality of the fit from S-model to C-model is sufficient to justify the greater complexity of the C-model. The models were fitted by full maximum likelihood (FML) and restricted maximum likelihood (REML), a likelihood-ratio test comparing the two models was used to test the null hypothesis that the model is standard multilevel model and the results were significant at 0.05 α level, except under the condition $t = 3, \rho = 0.3, N = 100$, with P-Value 0.939294, indicating that the result is not significant at 0.05. Demidenko (2004), found that when the autocorrelation is small, little difference will exist for the estimates obtained by modelling the autocorrelation versus ignoring the autocorrelation, though consistent with the previous studies but under large N. According to Douglass (2010) when hypothesis tests are phrased in terms of the parameter restriction, it is important to realize that we are comparing the quality of fits obtained with two nested models. That is, we are not assessing parameter values per se; we are comparing the model fit obtainable with some constraints on parameter values to that without the constraints. A choice between the two models used in this study cannot be made based on the better model fit but has to be made based on theory and purpose of the study. This research work concentrate more on the effects of covariates at the second level, not only that random coefficients model is most appropriate by adding autocorrelation parameter but there is need to account for covariation between observations on the same individuals at different times. Based on the results of our analysis, there is a non-

negligible probability that the observed difference in the model fits could reasonably not be the result of random chance, but the inherent superiority of the model. Clearly, the autocorrelated model is more appropriate, since adding the autocorrelation parameter improves the fit of the model to the data.

2. What are the sufficient sample sizes for accurate estimation?

This study also found that across all combinations of conditions under sample sizes at both levels used for data simulation, the directions of the biases were inconsistent over conditions. Generally, the bias of MLE is larger than the bias of RMLE, and under small sample sizes at both levels, there is no convincing evidence that FML estimates are more negatively biased than REML estimates either under C-models or S-models. Under small sample sizes at both levels, bias was not only negative but substantial, and excessive bias was not restricted to the smallest sample sizes but could be observed in moderate to large sample sizes. No matter how many subjects were sampled and how many measurements were taken on each subject, problematic bias remains, in fact, the only condition where bias did not exceed the threshold of 10% occurred under S-model when $N=50$ and $t = 5$.

Similarly, the results of the present study indicate that generally, the magnitude of the autocorrelation coefficient within the range of the values under the design condition, do moderate the effect of the dependence on the accuracy of parameter estimation at a fixed subject size. However, it was not clear to what extent the relationship between the level-1 and level-2 sample sizes is related to the accuracy of the parameter estimation. The findings of the present study were consistent with those of previous studies with regard to the effect of sample sizes on the violation of associated statistical assumption. Mark and Stephen (2013) remarks that In general, the statistical properties of these estimators (MLE_s) are well-defined only when both the

number of subjects and the number of measurement occasion (numbers of repeated measures) are large, in which case estimates of parameters and their sampling variability are consistent (they converge to their true values with sufficiently large samples) and are asymptotically normally distributed.

3. *What is the effect of misspecification of covariance parameter in estimating fixed effect under Correlated model?*

The variance-covariance matrix of the repeated measures Y under our correlated model is of the form $V(y_i) = Z_i \Sigma_v Z_i' + R$ where $R = \sigma^2 \Omega$, where a structure model AR(1) was assumed. The results of our simulation show that, when the covariance parameters described in terms of variance of repeated measures was underspecified, the MSE and its components were underestimated, similarly when the covariance parameter was over specified, the resulting MSE and its components were overestimated. This result is consistent with our previous result on effects of autocorrelation on the fixed effect estimates as evaluated by the error indices

4. *Having determined parameters in our models that should have a random component, in this section we consider the questions of whether the scaled variance-covariance matrix of the random effects (**D**) structured in a simpler and complex form, i.e. with fewer parameters can be representative of the unstructured form using a real-life data.*

A Standard Error is meant to serve as a measure of the uncertainty of a parameter estimate, our methods of assessing bias and efficiency for parameter estimates could also be applied to estimates of standard errors (they too can be right or wrong, and they too can be widely dispersed or tightly clustered in repeated samples). Usually, the mean structure rather than the covariance model is of primary interest in practice, however, adequate covariance is essential to obtain valid model-based inferences for the mean model parameters. With respect to

the above question, the results of our simulation show that a small sample size at level two (meaning a sample of 50 or less) leads to biased estimates of the second-level standard errors

The correlated model with AR(1) covariance structure provided a better fit to the data. Thus, empirically, it better captures the nature of subject effects and how they change over time, but because of the biases of the estimated covariance structure, it is not likely to provide more accurate inferences about the fixed effects. The standard error estimates for both correlated and standard models are biased. For example, the standard error estimates obtained from C-Model from the AR(1) structure are 230% larger, on average, while the standard error estimates obtained from the S-model are 221% larger, on average (results not shown). To evaluate the performance of standard errors in our Monte Carlo simulation, we determine whether they meet their intended definition. In a large number of repeated samples, our confidence interval (CI) set at 95% include the true population parameter 95% of the time for both correlated and standard model. For the correlated model, a 95% CI was given by -0.046 ± 0.8761 , while that of standard model was given by 0.0136 ± 0.8755 , it seems both confidence intervals were too large, an indication of risk accepting a Null hypothesis when it is False. However, it is not accurate to say that the method with the smallest standard error is the most efficient as a general rule. We can have standard errors that are wrongly estimated to be small. A defensible inference from a set of data requires a scientifically credible model supported with diagnostic analysis, not merely a model that converges to a numerical solution (Keith & Paul, 2006). These results are consistent with the results of the previous studies, Demidenko (2004) proves that the number of subjects plays a larger role than the number of observations in achieving asymptotic properties of variance component estimators. Other studies revealed that misspecification of the random effects leads to biased estimates of the variance components, biased standard error estimates for

fixed effects, and confidence intervals with inaccurate coverage rates (Raudenbush and Bryk, 2002; Hox, 2010; Berkhof and Kampen, 2004).

5.2. Conclusions

Choosing an appropriate family of repeated measures data analytical linear models procedure is often a matter of concern to practitioners. With a two-level time series model, the analysis of repeated measures data can be conducted efficiently. Multilevel analyses are applied to data from many units of measurements with one or relatively few observations per unit (such as repeated measures data). However, as this method becomes better known and used in research, the credibility of model applications depends on the plausibility of the model assumptions. According to Keselman et al. (1998), "the applied researcher who routinely adopts a traditional procedure without giving thought to its associated assumptions may unwillingly be filling the literature with non-replicable results. Indeed, in their review paper on the effect of violating assumed condition for ANOVA, Glass, Peckham and Sanders (1972) stated the following: "The assumptions of most mathematical models are always false to a greater or lesser extent. The relevant question is not whether ANOVA assumptions are met exactly. But rather whether the plausible violations of the assumptions have serious consequences on the validity of probability statements based on the standard assumptions", a statement that can equally be extended to other statistical procedures.

The overall contribution of this study was to provide applied researchers with information on the behaviour of the three most commonly used estimators for multilevel model parameters to the violation of non-autocorrelation assumption under simulation and realistic data conditions. Before now, simulation work has focused exclusively on simple models whose assumptions are more restrictive, which limited the generalizability of previous findings. The

present study supports the previous findings with regard to the point that violation of non-autocorrelation assumption introduces bias into the fixed effect estimates, particularly at the level-2.

(Match conclusions with the objective) As the results of this study show, the estimation of level-2 fixed effects of subject characteristics is relatively robust with respect to the violation of non-autocorrelation assumption. This is an important finding, because the ability to model the covariance structure parsimoniously result in more efficient estimates of the fixed-effects parameters of the model and consequently more powerful tests of the effects of the repeated measures. Violating non-autocorrelation assumption led to seriously biased and inaccurate parameter estimates (high MSE values) for the second-level predictors and residuals. The precision of the fixed effects will be decreased with increase heterogeneity caused by autocorrelation which might lead to a widening confidence intervals. Further conclusions based on this study results are as follows:

The fixed and random parameter estimates were relatively robust under moderate and extreme violations of the non-autocorrelation assumption for first-level residuals. However, it should also be noted that using composite measures of estimator accuracy that combine summaries of bias and variability, our MLE in principle, is the right choice of method of estimation, though estimates are not unbiased, what we pay for in bias, we make up for in variance. This process which biased our estimates leaves us with values that tend to be closer to the parameter values (i.e., a smaller expected mean square error) than those based on OLS estimation as noted by Raudenbush & Bryk (2002). Higher precision of our shrinkage estimator is achieved at the expense of a certain bias, which is largest when we inspect subjects that are

both small and far removed from the overall mean. Hence MLE methods can be seen as improvements with the reduction in variance that outweigh the increased in squared bias.

The extent to which the results of the C-model can be compared to the results of the S-model should be examined. Although it is tempting to conclude that the impact of violations of non-autocorrelation on the fixed effect estimates is larger in the C-model than in the S-model, the C-model parsimoniously modeled the covariance structure and fits the data well. As noted by Daniel J. B. et. al. (2012) this modeling strategy is perfectly reasonable, it has conformed optimally to structural and/or functional changes over time that is more stochastic in nature, producing time-varying group-effects that are not captured well by conventional model specifications. Modelling of autocorrelations through use of time series models for the errors, such as AR(1) model, in addition to the inclusion of random effects terms, may be more appropriate and may lead to a more proper representation of the correlation structure. Simulation results make it clear that misspecification of the covariance model may introduce (additional) substantial inaccuracy in inference (Park, Park, and Davis, 2001; Muller, Edwards, Simpson, and Taylor, 2006).

Our Monte Carlo simulation analysis assesses the performance of estimators not only for MLE but Bayes estimator. The MLE and REML approaches produce very similar results, an attempt to use the combination of lack of bias but large variability as a feature of a better estimator, found no clear advantage for either parameter estimation method. FML that exhibited high bias, traded off its biases for low variance only at small and partly at moderate sample sizes at both levels, while REML that exhibited low bias traded off its own biases for low variance partly at moderate and large sample sizes at both levels. As noted by Kreft and De leeuw (1986), "the evidence of each estimator superiority in complicated cases, and in multilevel in particular,

is not too convincing." Maximum likelihood estimators under correlated model generally have a smaller bias, MSE and variance than the Bayes estimator whose estimates were conditioned on the ML estimator of the second-level variances, although this may be because of the selected prior which may have a large influence on the posterior. The superior performance of Bayesian estimation relies on the priors being correctly specified.

Generally, maximum likelihood estimators of variance-covariance parameters are biased in small to moderate sample sizes, however, when covariance parameters in the covariance models are either under or over specified, our results prove that accuracy of fixed effects inference cannot be guaranteed in small or large samples, Kackar and Harville (1984) proved that $\hat{\beta}$ from iterated approximate least squares is unbiased for a Gaussian (or any other symmetric) distribution. However, covariance parameters, at least in small samples, typically have a substantial bias.

5.3. Limitations of the Present Study:

It should be acknowledged that generalization of the results of Monte Carlo studies is limited to the conditions examined in that particular study. This study employed a two-level time series model that had first-level variables as the various measurements that were nested within the second-level variables, the individuals on which the measurements were made. The true parameter values of the variance-covariance parameters were used in the context of a multilevel model. They were specified on the basis of results of HLM analyses discussed at great length by Hedeker and Robert Gibbons (2006) in their text, Longitudinal Data Analysis that focused on the longitudinal relationship between imipramine (IMI) and desipramine (DMI) plasma levels and clinical response in 66 depressed inpatients to reflect a set of repeated measures data. The mean for the values was based on the medium effect size by Cohen(1988) while there were variations

only in their correlations or variances through the study conditions. Clearly, this selection of variables and measures could have affected parameter values. Other values for the parameters and other models are possible and could be adequate for other depressed settings and might lead to different conclusions.

Moreover, this study attempted to model conditions found in typical psychiatric data-set and multilevel models, where intercepts and slopes were allowed to vary and the cross-level interaction (The term that appears in the model because of modeling the varying regression slope β_{1j} of the time-level variable X_{ij} with the individual level variable Z_j) was not considered in the model included in the study. Even though it is expected that they may affect the estimates, it is difficult to speculate on the range of this effect. Kreft and DeLeeuw (1998) noted that adding a cross-level interaction to the multilevel model in general means adding instability which could remove many equivalencies in the analysis.

The present study was limited to balanced data because we want to model our covariance structure without complication. However, using unbalanced data is likely to give results not similar to those gotten by this study, with more bias in parameter estimates and less accuracy in the standard errors. This is because 2 level time series model is less attractive when the number and timing of observations vary among subjects because it does not take proper account of the weighting factor. Raudenbush and Bryk (2002) reported that ML bias with unbalanced data can be less with balanced data. This bias will further decrease as the number of groups increases or as the imbalance of the data decreases.

5.4. Recommendations:

The results of this study provide important information to researchers interested in making conclusions from the application of two-level time series model (TTSM) and several

recommendations can be drawn from them. The first recommendation is that since the critical number of subjects required for reliable estimation of subject effects is larger than is available in many existing datasets, a better approach is to explore methods that are more robust to small numbers of subjects, such as small sample corrections to test statistics, bootstrapping , and the two-step method; although some of these techniques require specialized knowledge and are available in only a few software packages.

When a simple covariance model (compound symmetry, unstructured) cannot be assumed, modeling of autocorrelations through the use of time series models for the errors, such as AR(1) model, may be more appropriate and may lead to a more proper representation of the correlation structure because with the presence of random effects it provides advantages to model the covariance structure which helps but cannot be guaranteed to be sufficient, nevertheless whether it is a general structure or one composed of random effects, under specification of the covariance model should be avoided. Inferences about change in the response and its relation to covariates are sensitive to the chosen model for the covariance among the repeated measures despite the fact that the main scientific interest is normally focused on the model for the mean response. Failure to properly account for the covariance results in confidence intervals and hypothesis tests that are invalid and may result in misleading inferences while parsimoniously modelling the covariance structure of the data should result in more efficient estimates of the fixed-effects parameters of the model and consequently more powerful tests of the effects of the repeated measures. Conscientious attention to good statistical practices in covariance model fitting which can provide confidence in fixed effect inference in a mixed model is recommended.

It is more often the rule than the exception that autocorrelation is violated in repeated measures data. For this reason, all repeated measures data should be exposed to tests of violations of non-autocorrelation, rather than assumed. The results of the simulation study add to the existing literature showing that when there are violations of autocorrelation, wrong conclusions can be made, especially where measurements are made close together in time. The appropriate statistical test should be used to test for autocorrelation. If it is found that the data have autocorrelated errors, then the nature of dependency should be modelled appropriately, and the choice of a covariance structure should be motivated by both theoretical expectations and empirical fit. Modelling the dependency appropriately is likely to result in smaller standard errors, that is more efficient model-based inferences.

5.5. Contribution to Knowledge:

In order to represent the impact of the serial correlation structure in models commonly used in repeated measures design research, this study utilized a form of TTSM that account for autocorrelated errors at first level, in comparison to TTSM that do not. Most previous simulation studies of TTSM used a model structure, with only autocorrelated error terms included in level-1 distributional assumptions. By examining the efficiency of some time series modelling techniques in the presence of serial correlation in repeated measures data, the results of the present study provide additional knowledge concerning the performance of TTSM under conditions with which the level-1 residuals are correlated, also this study adds information about the accuracy and precision of estimates and contributes to our understanding of the behavior of the two-level time series models under less than ideal conditions. It will not only provide policymakers and academics with a better understanding of model properties, but will also assist

researchers in making better decisions about the adequacy of data, and in meeting model assumptions.

5.6. Future Research:

The pattern of dependence among observations can have more effect on the validity and quality of a statistical analysis than any other feature. Therefore choosing an analysis must begin with determining the logical properties of the sampling scheme and thereby characterizing the patterns of independence and correlation among observations. Computer simulations can be used to compare the performance of different estimators and characterization of the effect of the design on the efficiency of the two-level time series models, by examining the sensitivity of estimators to model misspecification. In particular, how do estimators perform when the true errors are ARMA(p,q), and not an AR(1) or an AR(2), when it is MA and not AR ? and so on.

Although this study investigated the impact of violation of non-autocorrelation assumption, further research regarding unbalance data structures would benefit from assessing the impact of modelling autocorrelated errors in the residuals (e.g., under different methods of estimation). An additional limitation of the current study is that we confined our simulation to a relatively balanced TTSM, excluding interaction between both time and individual level. One important benefit of multilevel modelling is the possibility of explaining cross-level interaction.

Additional research is also needed to determine the sample sizes and other data characteristics under which correlated models will perform optimally, to determine which methods provide the most reliable statistical results and to compare their relative computational performance, one possibility is to construct the confidence intervals for all of these estimators. The confidence interval is another criterion for measuring the accuracy of an estimator. It can also be used to compare the coverage of different methods. Bootstrapping method has been

popular in calculating the confidence interval. It is a non-parametric method and easy to apply to almost any problems and any datasets. Therefore, bootstrapping method may be added for further comparison. There is usually a trade-off between the number of random effects incorporated in the model and the complexity of the cluster errors covariance structure (Jones, 1990). Further research is needed in that area, especially from a model building perspective.

REFERENCES

- Afshartous, S. D. (1995). Determining of sample size for multilevel design. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Agresti, A., Caffo, B., and Ohman-Strickland, P. (2004). Examples in which misspecification of a random effects distribution reduces efficiency, and possible remedies. *J. Comput. Graph. Statist* 47, 639–653.
- Airy, G. B. (1861). *On the algebraical and numerical theory of errors of observations and the combinations of observations*, London: MacMillan.
- Aitkin, M., & Longford, N. T. (1986). Statistical modelling issues in school effectiveness studies (with discussion). *Journal of the Royal Statistical Society, Series A* 144, 419 – 461.
- Allison, P.D. (1999). *Multiple Regression: A Primer*. California. London. New Delhi: Pine Forge Press, Thousand Oaks.
- Altham, P.M.E. (1984). Improving the precision of estimation by fitting a model. *Journal of the Royal Statistical Society Series B (Statistical Methodology)* 46(1), 118-119 .
- Anu Maria (1997). Introduction to modelling and simulation. Proceedings of the 1997 winter simulation Conference, State University of New York at Binghamton Department of Systems Science and Industrial Engineering Binghamton, NY 13902-6000, U.S.A.
- Azeez, O. I. and Ipinyomi, R. A. (2017). The influence of autocorrelated errors on the bias of multilevel time series parameter estimates. *International Journal of engineering and applied sciences (IJEAS)* ISSN: 2394-3661, 4(4), 98 – 102.
- Barnard, J., McCulloch, R., and Meng, X. L. (2000). Modelling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Statistical Sinica* 10, 1281–1312.
- Bassiri, D. (1988). Large and small sample properties of maximum likelihood estimates for the hierarchical linear model. Doctoral dissertation submitted to Michigan State University, East Lansing, MI.
- Beaubien, J. M., Hamman, W. R., Holt, R. W., and Boehm-Davis, D. A. (2001). The application of hierarchical linear modelling (HLM) techniques to commercial aviation research. Proceedings of the 11th Annual Symposium on Aviation Psychology, Columbus, OH: The Ohio State University Press.

- Berkhof, J. and Kampen, K. (2004). Asymptotic effect of misspecification in the random part of the multilevel model. *Journal of Educational and Behavioral statistics* 29(2), 201 – 218.
- Berkhof, J. and Snijders, T. A. B. (2001). Variance component testing in multilevel models. *Journal of Educational and Behavioral Statistics* 26, 133 – 152.
- Bock, R. D. (1983a). The discrete Bayesian. In H. Wainer and S. Messick, (eds). *Modern advances in psychometric research*. Hillsdale, NJ: Erlbaum, pp. 103 - 115.
- Bock, R. D. (1983b). Within-subject experimentation in psychiatric research. In R. D. Gibbons and M. W. Dysken, (eds). *Statistical and methodological advances in psychiatric Research*. New York: Spectrum, pp. 59 – 90.
- Boyd, L.H. and Iversen, G. R. (1991). *Contextual analysis: Concepts and statistical techniques*. Belmont, CA: Wadsworth.
- Braun, H. I. (1989). Empirical Bayes method: A tool for exploratory analysis. *Science Direct Multilevel Analysis of Educational Data*, Volume null, Issue null, pp. 19-55.
- Browne, W. J. and Draper, D. (2000). Implementation and performance issues in the Bayesian and likelihood fitting of multilevel models. *Computational Statistics* 15(3), 391–420.
- Bryk, A.S. and Raudenbush, S.W. (1992). *Hierarchical linear models in social and behavioural research: Applications and data analysis methods* (First Edition). Newbury Park, CA: Sage Publications.
- Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. In D. Berliner (ed). *Review of research in education* 8, 158 – 233. Washington, DC: American Educational Research Association.
- Burstein, L. (1989). Conceptual considerations in instructional sensitive assessment. Paper presented at the Annual Meeting of the American Education Research Association, San Francisco, CA.
- Burton, A., Althman, D. G., Royston, P. and Holder, R. L. (2006). The design of simulation studies in medical statistics. *Statistics in Medicine* 25, 4279–4292.
- Busing, F. (1993). Distribution characteristics of variance estimates in two-level models: A Monte Carlo study. *Issue 4 of Leiden psychological reports: Psychometrics and research methodology*.

- Charles, E., McCulloch and John, M. N. (2011). Misspecifying the shape of a random effects distribution: why getting it wrong may not matter. *Statistical Science* 26 (3), 388 – 402.
- Cheng, H. (2014). *Analysis of panel data*. California: Cambridge University Press.
- Chi, E.M. and Reinsel, G.C. (1989). Models for longitudinal data with random effects and AR(1) errors. *Journal of the American Statistical Association* 84, 452 – 459.
- Chiu, Y. M., Tom, L., and Kam-Wah, T. (1996). The Matrix-Logarithmic Covariance Model. *Journal of the American Statistical Association* 91 (433).
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. New York, N.J.: Psychology Press.
- Commenges, D. and Jacqmin-Gadda, H. (1997). Generalized score test of homogeneity based on correlated random effects models. *Journal of the Royal Statistical Society, Series B*. 59, 157–171.
- Copt, S. and Victoria-Feser, M. P. (2006). High-breakdown inference in the mixed linear model. *Journal of the American Statistical Association* 101(473), 292-300.
- Corbeil, R. R., and Searle, S. R. (1976). Restricted maximum likelihood (REML) estimation of variance components in the mixed model. *Technometrics American Statistical Association* 18(1), 31-38
- Coull, J. T., (2004). fMRI studies of temporal attention: allocating attention within or towards time. *Journal of Cognitive Brain Research* 21(2), 216 – 226.
- Crainiceanu, M. and Ruppert, D. (2004). Likelihood ratio tests for goodness-of-fit of a nonlinear regression model. *Journal of Multivariate Analysis* 91(1), 35-52.
- Crump, S. L. (1947). The Estimation of Variance in Multiple Classification. PhD Thesis. Department of Statistics, Iowa State University.
- Daniel, J. B., Nisher C. G., Danielle, D. and Roberts, A. Z. (2012). Analyzing Repeated Measures Data on Individuals Nested within Groups: Accounting for Dynamic Group Effects. *Journal of American Psychological Association* 18(1), 1-14.
- Darandari, E. Z. M. (2004). Robustness of hierarchical linear model parameter estimates under violation of second-level residual homoskedasticity and independence assumptions. Doctoral Dissertation. Florida State University.

- Dawson, K. S., Gennings C. and Carter W.H. (1997). Two graphical techniques useful in detecting correlation structure in repeated measures data. *The American Statistician* 51, 275–283.
- Demidenko, E. (2004). *Mixed models: Theory and applications*. New York: Wiley.
- Dempster, A., Laird, N. and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B.* 39, 1–38.
- Dempster, A. P., Rubin, D. B. and Tsutakawa, R. K. (1981). Estimation in covariance components models. *Journal of the American Statistical Association* 76(374).
- Douglas, B. (2010). *lme4: Mixed effect modelling with R*. Madison, WI, USA. Springer.
- Ebbes, P., Böckenholt, U. and Wedel, M. (2004). Regressor and random-effects dependencies in multilevel models. *Statistical Neerlandica* 58, 161–178.
- Efron, B., and Morris, C. (1976). Multivariate Empirical Bayes and Estimation of Covariance Matrices. *Ann Statistics* 4, 22-32.
- Eqbal Z. D. (2004). Robustness of hierarchical linear model parameter estimates under violations of second-level residual homoskedasticity and independence assumptions. A PhD. Dissertation Submitted to the Graduate School, Florida State University, Libraries Electronic Theses.
- Faraway, J. J. (2006). *Extending the linear model with R: Generalized linear, mixed effects and nonparametric regression models*. Boca Raton: Chapman & Hall.
- Feng, Z., and McCulloch, C. E.(1992). Statistical inference using maximum likelihood estimation and the generalized likelihood ratio when the true parameter lies on the boundary of the parameter space. *Statistics & Probability Letters* 11, 325–332.
- Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.
- Fitzmaurice, G. M., Laird, N. M. and Ware. J.H. (2004). *Applied Longitudinal Analysis*. New York: Wiley.
- Gelfand, A. E., Hills, S. E., Racine-Poon, A. and Smith, A. F.M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association* 85(412), 972–985.
- Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85, 398–409.

- Gelman, A. and Jennifer, H. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *Browse Journals and Magazines* 6(6), 721–741.
- George, C. T., and Arnold, Z. (1964). Bayes's theorem and the use of prior knowledge in regression analysis. *Biometrika* 51(1/2), 219–230.
- Gibbons, R. D., Hedeker, D. C., Watemaux, M. and Davis, J. M. (1988). Random regression models: A comprehensive approach to the analysis of longitudinal psychiatric data. *Psychopharmacology Bulletin* 24, 438–443.
- Gibbons, R.D., Hedeker, D., Elkin I., Waternaux, C., Kraemer, U.C., Greenhouse, J.B., Shea, M.T., Imber, S.D., Sotsky, S.M. and Watkins, J.T. (1993). Some conceptual and statistical issues in the analysis of longitudinal psychiatric data. *Archives of General Psychiatry* 50, 739–750.
- Gill, J. (2002). *Bayesian methods for the social and behavioural sciences*. New York: Chapman & Hall.
- Gill, J. (2003). Hierarchical linear models. In Kimberly Kempf-Leonard. *Encyclopedia of Social Measurement*. New York: Academic Press.
- Glass, G. V., Peckham, P. D., and Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance. *SAGE Journals* 42, 237-288.
- Goel, P. K., and DeGroot, M. H. (1981). Information about hyperparameters in hierarchical models. *Journal of the American Statistical Association* 76, 140–146.
- Goldberger, A. S. (1991). *A course in econometrics*, Cambridge, MA, Harvard University Press.
- Goldstein, H. (1995). *Multilevel statistical models, 2nd edition*, New York: Halstead Press.
- Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika* 73(1), 43–56.
- Goldstein, H., Micheal, J. R. H., and Jon, R. (1994). Multilevel time series model with application to repeated measures data. *Journal of Statistics in Medicine*. 13(16), 1643-1655.
- Goldstein, H., and Rasbash, J. (1992). Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalized least squares.

- Computational statistics and data analysis. *Journal of Computational Statistics and Data Analysis* 13(1), 63-71.
- Gong, G., and Saminiego, F. J. (1981). Pseudo maximum likelihood estimation: theory and applications. *Ann. Statistics* 9(4), 861--869.
- Gottman, J. M. (1981). *Time-Series Analysis: A comprehensive introduction for social scientists*, New York: Cambridge University Press.
- Greenwood, K. M., and Matyas, T. A. (1990). *Problems with the application of interrupted time series analysis for brief single-subject data*. *Journal of American Psychological Association* 12(5), 355-370.
- Gumedze, F. N. and Dunne, T. T. (2011). Parameter estimation and inference in the linear mixed model. *Journal of Linear Algebra and its Application* 435(8), 1813-2096
- Hanushek, E. A., and John E. J. (1977). *Statistical Methods for Social Scientists*. San Diego: Academic Press.
- Hartford and Davidian (2000). Consequences of misspecifying assumptions in nonlinear mixed effects models. *Journal of Computational Statistics and Data Analysis* 34(2), 139-164.
- Hartley, H. O., and Rao, J. N. K. (1967). Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika* 54(1/2), 93-108.
- Harville, D. A. (1974). Bayesian inference for variance components using only error contrasts. *Biometrika* 61(2), 383--385.
- Harville, D. A. (1976). Extension of the Gauss-Markov theorem to include the estimation of random effects. *Annals of Statistics* 4(2), 384-395.
- Harville, D. A. (1977). Maximum likelihood approaches to variance components estimation and to related problems, *Journal of the American Statistical Association* 72(358), 320--338.
- Heagerty, P., & Kurland, B. (2001). Misspecified maximum likelihood estimates and generalised linear mixed models. *Biometrika* 88(4), 973--985.
- Healy, M. (2001). Multilevel data and their analysis In: Leyland AH, Goldstein H, editors. *Multilevel modelling of health statistics*. Wiley; West Sussex. pp. 1--12.
- Heck, R. H., and Thomas S. L. (2000). *An introduction to multilevel modelling techniques*. Mahwah, NJ: Erlbaum Inc.
- Hedeker, D. Gibbons, R. D. (2006). *Longitudinal data analysis*. New York: Wiley.

- Henderson, C. R., Kempthorne, O., Searle, S. R. and Von Krosigk, C. N. (1959). Estimation of environmental and genetic trends from records subject to culling. *Biometrics* 15(2), 192-218.
- Hoeksta, R., Henk, A. L. K. and Addie, J. (2011). Are assumptions of well-known statistical techniques checked, and why (not)? *Front Psychol* 3(137).
- Hoogland, J.J. and Boomsma, A. (1998). Robustness studies in covariance structure modelling: An overview and a meta-analysis. *SAGE Journals, Sociological Methods & Research* 26, 329 – 367.
- Houseman, E. A., Ryan, L. M. and Coull, B. A. (2004). Cholesky residuals for assessing normal errors in a linear model with correlated outcomes. *Journal of the American Statistical Association* 99, 383–394.
- Hox, J. J. (2002). *Multilevel analysis: techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hox, J. J. (2010). *Multilevel analysis: techniques and applications* (2nd Ed) New York: Routledge.
- Hui, S. L. and Berger, J. O. (1983). Empirical Bayes estimation of rates in longitudinal studies. *Journal of the American Statistical Association* 78, 753–759.
- Jackman, S. (2000). Estimation and inference are missing data problems: unifying social science statistics via Bayesian simulation. *Oxford Journals* 8(4), 307–332.
- Janet M. B., Henry E. B. and David C., (2008). *The Oxford handbook of political methodology*. New York, USA: Oxford University Press.
- Jennifer S. K. Chan and Boris Choy, S. T. (2008). Analysis of covariance structures in time series. *Journal of Data Science* 6, 573-589.
- Jennifer Chan, Boris Choy, S.T. and Udi M. (2007). Robust Bayesian analysis of loss reserves data using the generalized-t distribution. *Econ Papers*. No. 196.
- Jennrich, R. I. and Schluchter, M. D. (1986). Unbalanced repeated measures models with structural covariance matrices. *Biometrics* 42(4), 805–820.
- Jones, R. H. (1990). Serial correlation or random subject effects. *Communications in Stat–Simulation and Computation* 19(3), 1105–1123.
- Jones, R. H. and Boadi-Boteng, F. (1991). Unequally spaced longitudinal data with serial correlation. *Biometrics* 47, 161 – 75.

- Kasim, R. M. (1994). The application of Gibbs sampling to nested variance components models with heterogeneous within-group variance. Doctoral Dissertation. Michigan State University, East Lansing, MI.
- Kasim, R. M. and Raudenbush, S. W. (1998). Application of Gibbs sampling to nested variance components models with heterogeneous within-group variance. *Journal of Educational and Behavioral Statistics* 23, 93–116.
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. (2nd ed.). New York: Oxford University Press.
- Keselman, H. J., Algina, J., Kowalchuck, R. K. and 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, 591–604.
- Keselman, H. J., Algina, J., Kowalchuk, R. K. and Wolfinger, R. D. (1999). A comparison of recent approaches to the analysis of repeated measurements. *British Journal of Mathematical and Statistical Psychology* 52, 63–78.
- Kim, F. N. (2012). Statistical assumptions of substantive analyses across the general linear model: A mini-review. *Frontiers in Psychology* 3, 322.
- Kim, K. (1990). *Multilevel data analysis: A comparative examination of analytical alternatives*. PhD Thesis, University of California, Los Angeles.
- Kreft, I. G. G. and De Leeuw, J. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics* 11, 57–85.
- Kreft, I. G. G. and de Leeuw, J. (1998). *Introducing multilevel modelling*. Thousand Oaks, CA.: SAGE Publications.
- Kristopher, J. P., Robert C. M. and Nancy, E. B. (2008). *Latent growth curve modelling*. Thousand Oaks, CA.: SAGE Publications.
- Laird, N. M. and Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics* 38, 963–974.
- Lavange, L. M., and Helms, R. W. (1983). The analysis of incomplete longitudinal data with modelled covariance structures. PhD Thesis, University of North Carolina,
- Leckie, G. and Goldstein, H. (2009). The limitations of using school league tables to inform school choice. *Journal of the Royal Statistical Society* 172, 835–851.

- Liang, K. Y. and Zeger, S. L. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42(1), 121-130.
- Lindley, D. V. and Smith, A. F. M. (1972). Bayes estimates for the linear models. *Journal of the Royal Statistical Society*, 34(1), 1–41.
- Lindsey, J. (1993). *Models for repeated measurements*. New York, USA: Oxford University Press.
- Lindstrom, M. J. and Bates, D. M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association* 83, 1014-1022.
- Litière, S., Alonso, A., and Molenbergh, G. (2007). Type I and type II error under random-effects misspecification in generalized linear mixed models. *Biometrics* 63, 1038–1044.
- Litiere, A. and Molenberghs, G. (2007). A note on Type II error under random effects misspecification in generalized linear mixed models. *Biometric* 67(2), 654 – 660.
- Lisa, M. S., Kimberly, A. D., and Elena, L. (1999). An introduction to hierarchical linear modelling. *Statistics in Medicine* 18(7), 855–888.
- Longford, N. T. (1993). *Random coefficient models*. New York: Oxford University Press.
- Longford, N. T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika* 74(4), 817–827.
- Longford, N. T. (1987). Logistic regression with random coefficients mixed models with nested random effects. *Biometrika* 1993(1), 1-50.
- Maas, C. J. M. and Hox, J. J. (2004a). The influence of violations of assumptions on multilevel parameter estimations and their standard errors. *Computation Statistics & Data Analysis* 46, 427–440.
- Maas. C. J. M. and Hox, J. J. (2004b). Robustness issues in multilevel regression analysis. *Statistical Neerlandica* 58, 127–137.
- Maas, C.J.M. and Hox, J. J. (2005). Sufficient sample sizes for multilevel modelling. *Hogrefe & Huber Publishers* 1(3), 86-92.
- Machin, D. (1975). On a design problem in growth studies. *Biometrics* 31, 749 – 753.
- Mackinnon, D. P. and Krull, J. L. (1977). Inconsistent mediation models. Paper presented at the annual convention of the American psychological association, New York, NY.

- MaCurdy, T. E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics* 18, 83 – 114.
- Maeda, Y. (2007). Monte Carlo evidence regarding the effects of violating assumed conditions of two-level hierarchical models for cross-sectional data. PhD Dissertation submitted to the Faculty of Graduate School, University of Minnesota, USA.
- Mansour, H., Nordheim, E. V. and Rutledge, J. J. (1985). Maximum likelihood estimation of variance components in repeated measures designs assuming autoregressive errors. *Biometrics* 41, 287 –294.
- Maria, V. G., Hopkins, M. and Stephen, W. (2004). Dietary choline supplementation improves behavioural, histological, and neurochemical outcomes in a rat model of traumatic brain injury. *Journal of Neurotrauma* (25) 8.
- Mark, L. B. and Stephen P. J. (2013). Regression analysis of country effects using multilevel data: A cautionary tale. Discussion Paper No. 7583 August 2013.
- Mason, W. M., Wong, G. Y., Entwisle, B. (1983). Contextual analysis through the multilevel linear model. *sociological methodology* 72–103.
- Matthew, J. G., Lloyd, J. E. and Keith E. M. (2011). Avoiding bias in mixed model inference for fixed effects. *Statistics in Medicine* 30(22), 2696–2707.
- McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association* 92(437), 162–170.
- McCulloch, C.E. and Neuhaus, J. M. (2011). Prediction of random effects in linear and generalized linear models under model misspecification. *Biometrics* 67(1), 270–279.
- Miller, J.J. (1973). Asymptotic properties and computation of maximum likelihood estimate in the mixed of the analysis of variance. Technical Reports (12)21, Nov 1973. <http://www.dtic.mil/dtic/tr/fulltext/u2/775399.pdf>.
- Miller, J.J. (1977). Asymptotic properties of maximum likelihood estimates in the mixed model of analysis of variance. *Annals of Statistics* 5, 746–762.
- Mok, M. (1995). Sample size requirements for 2-level designs in educational research. Unpublished manuscript, Macquarie University, Sydney Australia.
- Molenberghs, G. and Verbeke, G. (2000). *Linear mixed models for longitudinal data*. USA, NY: Springer-Verlag.

- Molenberghs, G. and Verbeke, G. (2007). Likelihood ratio, score, and Wald tests in a constrained parameter space. *The American Statistician* 61, 22-27.
- Morris, C. (1995). Hierarchical Models for Educational Data: An Overview. *Journal of Educational Statistics* 20, 190–200.
- Muller, E. K., Edward, L. J., Simpson, S. L., and Taylor, D. J. (2006). Statistical tests with accurate size and power for balance linear mixed models. *Statistics in Medicine* 26(19):3639-3660.
- Oliver, E. L. and Thomas, M. B. (2012). Permutation tests for random effects in linear mixed models. *Biometrics* 68(2), 486-493.
- Osborne, J. W. (2000). Advantages of hierarchical linear modelling. *Practical Assessment, Research and Evaluation* 7(1), 1–3.
- Park, T., Park, J. K. and Davis, C. S. (2001). Effects of covariance model assumptions on hypothesis tests for repeated measurements: analysis of ovarian hormone data and pituitary- pterygomaxillary distance data. *Statistics in Medicine* 20, 2441–2453.
- Parke, W.R. (1986). Pseudo maximum likelihood estimation: The asymptotic distribution. *The Annals of Statistics* 14, 355–357.
- Patterson, H. D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* 58, 545–554.
- Peng Zeng and Lu Lin. (2012). Finite mixture of heteroscedastic single-index models. *Open Journal of Statistics* 2, 12–20.
- Pinherio, J. C. and Bates, D. M. (2000). *Mixed-Effects Models in S And S-Plus*. New York: Springer-Verlag.
- Potthoff, R.F., and Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika* 51, 313–326.
- Ramon, C. L., George, A. M., Walter, W. S., Russell, W. (1996). *SAS System for Mixed Models*. NC, USA: SAS Institute Inc.
- Rao, C. R. (1973). *Linear statistical inference and its applications* (2nd ed). New York: Wiley.
- Raudenbush, S.W. (1988). Estimating change in the dispersion. *Journal of Educational Statistics*. 13(2), 148 – 171.
- Raudenbush, S. W. and Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education* 59, 1–17.

- Raudenbush, S.W., Yang, Meng-Li and Yosef, M. (2000). Maximum likelihood for generalized linear models with nested random effects via high-order, multivariate Laplace approximation. *Journal of Computational and Graphical Statistics* 9(1), 141–157.
- Raudenbush, S. W. and Anthony S. B. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: Sage Publications.
- Raudenbush, S.W. and Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. *Psychological Methods* 5(3), 199–213.
- Reisby, N. L., Gram, N. L. F., Bech, P., Nagy, A., Petersen, G. O., Ortmann, J., Ibsen, I., Dencker, S. J., Jacobsen, O., Krautwald, O., Sondergaard, I. and Christiansen, J. (1977). Imipramine: Clinical effects and pharmacokinetic variability. *Psychopharmacology* 54, 263–272.
- Robert, D. G., Donald, H., and Stephen, D. (2011). Advances in analysis of longitudinal data. *Annu Rev Clin Psychol* 6, 79–107.
- Ronald, H. H. and Scott L. T. (2015). *An Introduction to multilevel modelling techniques: MLM and SEM Approach using M Plus. Quantitative Methodology Series*. (Third Edition), New York and London: Routledge Taylor & Francis Group.
- Sander, G. (2000). Principle of multilevel modelling. *International Journal of Epidemiology* 29 (1), 158–167.
- Schabenberger, O. and Pierce, F. J. (2002). *Contemporary statistical model for the plant and soil science*. Boca Raton FL, USA: CRC Press.
- Searle, S. R., Casella, G. and McCulloch, C.E. (1992). *Variance components*. New York: John Wiley & Sons.
- Seiden, L. S. and Dykstra, L. A. (1977). *Psychopharmacology: A Biochemical and Behavioural Approach*. VanNostrand Reinhold, New York: Wiley & Sons.
- Self, S. G. and Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* 82, 605–610.
- Shapiro, A. (1988). Sensitivity analysis of non-linear programs and differentiability properties of metric projections. *Journal of control and optimization* 26(3), 628–645.
- Shieh, Y. Y., Fouladi, R.T. and Pullum, T. W. (2001). The effect of error term nonnormality on multilevel model parameter estimates and standard errors: A focus on estimation bias. *Multiple Linear Regression Viewpoints* 27, 12–37.

- Silvapulle, M.J. (1992). Robust Wald-type tests of one-sided hypotheses in the linear model. *Journal of the American Statistical Association* 87, 156–161.
- Singer, J. D. and Wille, J. B. (2003). *Applied longitudinal data analysis*. New York: Oxford University Press.
- Singh, J., Knapp, H. V. and Demissie, M. (2004). Hydrologic modelling of the Iroquois River watershed using HSPF and SWAT. Illinois State Water Survey Contract Report 2004-08.
- Snijders, T. A. and Bosker, R. J. (1993). Standard errors and sampling sizes for two-level research. *Journal of Educational Statistics* 18, 237–261.
- Snijders, T.A.B. and Bosker, R.J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modelling*. Thousand Oaks, CA: SAGE Publications.
- Steenbergen, M. R. and Bradford, S. J. (2002). Modelling multilevel data structures. *American Journal of Political Science* 46(1), 218–237.
- Stram, D. O. and Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics*. 50, 1171–1177.
- Strenio, J. L. F., Weisberg, H. I. and Bryk, A. S. (1983). Empirical Bayes estimation of individual growth curves parameters and their relationship to covariates. *Biometrics* 39(1), 71-86.
- Thompson, W. A. (1962). The problem of negative estimates of variance components. *Annals of Mathematical Statistics* 33, 273–289.
- Tippett, L. H. C. (1931). *The Methods of Statistics*. London: Williams and Norgate Ltd.
- Toothaker, L. E., Banz, M., Noble, C., Camp, J. and Davis, D. (1983). N=51 designs: The failure of ANOVA-based tests. *Journal of Educational Statistics* 4, 289–309.
- Van der Leeden, R. and Busing, F.(1994). First Iteration versus IGLS/ RIGLS estimates in two-level models: A Monte Carlo Study with ML3. Department of Psychometric and Research Methodology, Leiden University, Leiden.
- Van der Leeden, R., Busing, F. and Meijer, E. (1997). Bootstrap methods for two-level models. Leiden University, The Netherlands.
- Verbeke, G. and Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York: Springer.
- Verbeke G. and Molenberghs, G. (2003). The use of score tests for inference on variance components. *Biometrics* 59, 254–262.

- Verbeke, G. and Molenberghs, G. (2007). Likelihood ratio, score, and Wald tests in a constrained parameter space. *American Statistician* 61(1), 22-27.
- Vonesh, E. F. and Chinchilli, V. M. (1997). *Linear and nonlinear models for the analysis of repeated measurements*. New York: Marcel Dekker.
- Wakefield, J. C. and Kass, R. E. (1999). Bayes factors and approximations for variance component models. *Journal of the American Statistical Association* 94(448), 1242–1253.
- Wakefield, J. C., Smith A. F. M., Racine-Poon A. and Gelfand, A. E. (1994). Bayesian analysis of linear and non-linear population models using the Gibbs sampler. *Applied Statistics* 43(1), 201–221.
- Ware, J. H. (1985). Linear models for the analysis of longitudinal studies, *The American Statistician* 39, 95–101.
- Wolfinger, R. D. (1993). Covariance structure selection in general mixed models. *Communications in Statistics. Simulation and Computation*. 22(4), 1079–1106.
- Wolfinger, R.D., Tobias, R. D. and Sall, J. (1991). Mixed Models: A Future Direction: In *Proceedings of the 16th SAS Users Group Conference*, New Orleans, Louisiana. 1380–1388.
- Wu, M. (1995). An approximate generalized linear model with random effects for informative missing data. *Biometrics* 51(1), 151–168.
- Xuemei Han, Congfu Xu, Huifeng Shen and Yunhe Pan (2003). *A Novel Time Series Forecasting Approach with Multi-Level Data Decomposing and Modeling*. *IEEE journals*, Zhejiang University, Hangzhou, 310027, P.R. China.
- Yang, R. and Berger, J. O. (1994). Estimation of a covariance matrix using the reference prior. *Ann Statistics* 22(3), 1195–1211.
- Zeger, S. L. and Harlow, S. D. (1987). Mathematical models from laws of growth to tolls for biological analysis: Fifty years of growth 51(1), 1–21.
- Zeger, S. L., Liang, K.Y. and Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics* 44(4), 1049–1060.

Appendix A1: Psychiatric Dataset described in Reisby et al. (1977)

S/No	Subject. id	Time	Observations	S/No	Subject.id	Time	Observations
1	1	1	26	136	1	28	21
2	2	1	22	137	2	28	19
3	3	1	18	138	3	28	10
4	4	1	7	139	4	28	11
5	5	1	4	140	5	28	11
6	1	2	33	141	1	29	27
7	2	2	24	142	2	29	21
8	3	2	15	143	3	29	17
9	4	2	24	144	4	29	13
10	5	2	15	145	5	29	5
11	1	3	29	146	1	30	32
12	2	3	22	147	2	30	26
13	3	3	18	148	3	30	23
14	4	3	13	149	4	30	26
15	5	3	19	150	5	30	23
16	1	4	22	151	1	31	17
17	2	4	12	152	2	31	18
18	3	4	16	153	3	31	19
19	4	4	16	154	4	31	21
20	5	4	13	155	5	31	17
21	1	5	21	156	1	32	24
22	2	5	25	157	2	32	18
23	3	5	23	158	3	32	10
24	4	5	18	159	4	32	14
25	5	5	20	160	5	32	13
26	1	6	21	161	1	33	28
27	2	6	22	162	2	33	21
28	3	6	11	163	3	33	25
29	4	6	9	164	4	33	32
30	5	6	9	165	5	33	34
31	1	7	21	166	1	34	17
32	2	7	23	167	2	34	18
33	3	7	19	168	3	34	15
34	4	7	23	169	4	34	8
35	5	7	23	170	5	34	19
36	1	8	19	171	1	35	22
37	2	8	16	172	2	35	24

38	3	8	13	173	3	35	28
39	4	8	12	174	4	35	26
40	5	8	7	175	5	35	28
41	1	9	20	176	1	36	19
42	2	9	19	177	2	36	21
43	3	9	17	178	3	36	18
44	4	9	18	179	4	36	16
45	5	9	16	180	5	36	14
46	1	10	20	181	1	37	23
47	2	10	22	182	2	37	20
48	3	10	19	183	3	37	21
49	4	10	19	184	4	37	20
50	5	10	12	185	5	37	24
51	1	11	15	186	1	38	21
52	2	11	15	187	2	38	21
53	3	11	15	188	3	38	18
54	4	11	13	189	4	38	15
55	5	11	5	190	5	38	12
56	1	12	29	191	1	39	27
57	2	12	30	192	2	39	22
58	3	12	26	193	3	39	23
59	4	12	22	194	4	39	21
60	5	12	19	195	5	39	12
61	1	13	21	196	1	40	22
62	2	13	22	197	2	40	20
63	3	13	13	198	3	40	22
64	4	13	11	199	4	40	23
65	5	13	2	200	5	40	19
66	1	14	19	201	1	41	29
67	2	14	17	202	2	41	27
68	3	14	15	203	3	41	27
69	4	14	16	204	4	41	22
70	5	14	12	205	5	41	22
71	1	15	21	206	1	42	25
72	2	15	11	207	2	42	24
73	3	15	18	208	3	42	19
74	4	15	0	209	4	42	23
75	5	15	0	210	5	42	14
76	1	16	27	211	1	43	18
77	2	16	26	212	2	43	15

78	3	16	26	213	3	43	14
79	4	16	25	214	4	43	10
80	5	16	24	215	5	43	8
81	1	17	27	216	1	44	24
82	2	17	26	217	2	44	21
83	3	17	26	218	3	44	12
84	4	17	25	219	4	44	13
85	5	17	24	220	5	44	12
86	1	18	27	221	1	45	17
87	2	18	27	222	2	45	19
88	3	18	13	223	3	45	15
89	4	18	5	224	4	45	12
90	5	18	7	225	5	45	9
91	1	19	19	226	1	46	22
92	2	19	33	227	2	46	25
93	3	19	12	228	3	46	12
94	4	19	12	229	4	46	16
95	5	19	3	230	5	46	10
96	1	20	30	231	1	47	30
97	2	20	39	232	2	47	27
98	3	20	30	233	3	47	23
99	4	20	27	234	4	47	20
100	5	20	20	235	5	47	12
101	1	21	24	236	1	48	21
102	2	21	19	237	2	48	19
103	3	21	14	238	3	48	18
104	4	21	12	239	4	48	15
105	5	21	3	240	5	48	18
106	1	22	18	241	1	49	27
107	2	22	22	242	2	49	21
108	3	22	16	243	3	49	24
109	4	22	8	244	4	49	22
110	5	22	9	245	5	49	16
111	1	23	21	246	1	50	28
112	2	23	21	247	2	50	27
113	3	23	13	248	3	50	27
114	4	23	14	249	4	50	26
115	5	23	10	250	5	50	23
116	1	24	19	251	1	51	22
117	2	24	17	252	2	51	26

118	3	24	15	253	3	51	20
119	4	24	11	254	4	51	13
120	5	24	5	255	5	51	10
121	1	25	22	256	1	52	27
122	2	25	21	257	2	52	22
123	3	25	18	258	3	52	24
124	4	25	17	259	4	52	25
125	5	25	12	260	5	52	19
126	1	26	22	261	1	53	21
127	2	26	22	262	2	53	28
128	3	26	16	263	3	53	27
129	4	26	19	264	4	53	29
130	5	26	20	265	5	53	28
131	1	27	24	266	1	54	30
132	2	27	19	267	2	54	22
133	3	27	11	268	3	54	11
134	4	27	7	269	4	54	8
135	5	27	6	270	5	54	7

Appendix A2: Simulation code for data generating process, model estimation and error indices.

```
#####  
####R Code for Two-Level Time Series Modelling with Repeated Measures Data#####  
#####  
####DATA GENERATION PEOCESS#####  
library(MASS);library(mvtnorm);library(nlme);library(lme4);library(quantmod);library(rtf)  
#####  
  
rm(list=ls())  
  
set.seed(12345)  
  
N <- 30 ## 50,100  
  
t <- 3 ##5,10  
  
nsim<- 100#1000  
  
Nt=N*t  
  
dim<- Nt  
  
one<- diag(1,Nt)  
  
delta=1  
  
rho=0.3# 0.7,0.99  
  
beta=0.3  
  
sig2=12.22;#Z=cbind(1,c(1:t))  
  
zelta=matrix(c(12.63,-1.42,-1.42,2.08),2,2)  
  
sig <- diag(rep(1,t))  
  
#sig <- diag(rep(1,t))  
  
####Generation AR1 error  
  
eit=runif(Nt, -0.5,0.5)  
  
ei=eit
```

```

vit <- rmvnorm(N, sigma = sig)

vita<-Lag(vit,k=1)

vita[1,]<-0

AR1<- rho*vita+ei

####Generation MA1 error

vitm=Lag(eit, k=1)

vitm[1,]<-0

MA1=rho*vitm[,1]+ei #####MA1

####Generation ARMA11 error

ARMA1=AR1+MA1-ei #####ARMA1

###Generate Level two residuals

r2<- rmvnorm(Nt,c(0,0),zelta)

##Generating level 2 variables

w=0.2+runif(Nt,0,2)+r2[,1]+r2[,2]

###Generating level-1 random coefficients

u1=r2[,1]; u2=r2[,2]

B0=0.35+.2*w+u1

B1=0.5+.6*w+u2

W0=B0

W1=B1

W=cbind(W0,W1)

B0=mean(B0);B1=mean(B1)

##Intra correlation

ICC= 0.25/(0.25+.34)

wide <- data.frame(subject.id =factor(1:N),vit)

```

```

long <- reshape(wide, varying=list(2:(t+1)),
idvar='subject.id',v.names = "vit", direction = "long", timevar = "time")

aziz<-(long[order(long$subject.id),])

#Generate Outcome variable y

long$w <- w

long$AR1<-AR1 ##MA1, ARMA1

long$y <- delta*ICC + beta*aziz[,2] + beta*long$w +beta*long$w*aziz[,2]+
          u1*aziz[,2]+u2+long$AR1

###Variables List

y=long$y
y=unname(y)
y=y[,1]
time<- aziz[,2]
w<-long$w
subject.id<-aziz[,1]

Azeez=data.frame(cbind(subject.id,time,w,y))

#####

#Separable row col AUTOREGRESSIVE process

is = function(Nt, rho ) {
  m = diag(Nt)
  p=col(m)-row(m)
  d=abs(p)
  rho ^d
}

arcov=is(Nt, rho)

```

```

arcovi=diag(1,Nt)

sig2Reml=sig2*(N/(N-t-1))

Z=cbind(1,aziz[,2])

pcov=Z%*%zelta%*%t(Z)+sig2*arcov

Vcov=Z%*%zelta%*%t(Z)+sig2*arcovi

rcov<- Z%*%zelta%*%t(Z)+sig2Reml*arcov

VcovReml=Z%*%zelta%*%t(Z)+sig2Reml*arcovi

y=long$y

X=Z

Y=y

a=array(0,dim=c(nsim,2))

##MLE

for (i in 1:nsim) {

  Y=y+rnorm(y)

  timeModel <- solve(t(X)%*%solve(pcov)%*%X)%*%t(X)%*%solve(pcov)%*%Y

  a[i,] = timeModel

}

#####MLE Results#####

MeanML=apply(a,2,mean)

VarML=apply(a,2,var)

timeML <- solve(t(X)%*%solve(pcov)%*%X)%*%t(X)%*%solve(pcov)%*%Y

trueV=c(1,0.3)

biasML=MeanML-(trueV)

mseML=biasML^2+VarML

rmseML<-sqrt(mseML)

```

```

PRBml=100*((1/length(a))*sum(a[,2]-trueV[2]))/trueV[2]

ML<-(rbind(trueV,MeanML,VarML, biasML,mseML,rmseML ))

####

Bcov<-solve(t(X)%*%solve(pcov)%*%X)

sigma2ML<- (sum((y-X%*%timeML)^2))/(Nt-1)

e<-y-X%*%timeML

loglikeML <- ((-Nt/2)*log(2*pi)) - ((Nt/2)*log(sigma2ML)) - ((t(e)%*%e)/(2*sigma2ML))

llkML<-sum(loglikeML)

#####

#####

#####REML ESTIMATION#####

#REML

b=array(0,dim=c(nsim,2))

for (i in 1:nsim) {

Y=y+rnorm(y)

RModel <- solve(t(X)%*%solve(rcov)%*%X)%*%t(X)%*%solve(rcov)%*%Y

b[i,] = RModel

}

#####REML Results#####

MeanREML=apply(b,2,mean)

VarREML=apply(b,2,var)

timeREML <- solve(t(X)%*%solve(rcov)%*%X)%*%t(X)%*%solve(rcov)%*%Y

biasREML=MeanREML-(trueV)

mseREML=biasREML^2+VarREML

rmseREML<-sqrt(mseREML)

```

```

PRBrem1=100*((1/length(b))*sum(b[,2]-trueV[2]))/trueV[2]

REML<-(rbind(trueV,MeanREML,VarREML, biasREML,mseREML,rmseREML ))

BcovRM<-solve(t(X)%*%solve(rcov)%*%X)

sigma2REML<- (sum((y-X%*%timeREML)^2))/(Nt-1)

e<-y-X%*%timeREML

loglikeREML <- ((-Nt/2)*log(2*pi)) - ((Nt/2)*log(sigma2REML)) - ((t(e)%*%e)/(2*sigma2REML))

llkREML<-sum(loglikeREML)+((Nt+1)/2)*log(2*pi)-.5*(log(det(t(X)%*%solve(rcov)%*%X)))

#####ORDINARY#####

#####

##MLEOERDINARY

c=array(0,dim=c(nsim,2))

for (i in 1:nsim) {

  Y=y+rnorm(y)

  timeModelO <- solve(t(X)%*%solve(Vcov)%*%X)%*%t(X)%*%solve(Vcov)%*%Y

  c[i,] = timeModelO

}

#####MLEO Results#####

MeanMLO=apply(c,2,mean)

VarMLO=apply(c,2,var)

timeMLO <- solve(t(X)%*%solve(Vcov)%*%X)%*%t(X)%*%solve(Vcov)%*%Y

trueV=c(1,0.3)

biasMLO=MeanMLO-(trueV)

mseMLO=biasMLO^2+VarMLO

rmseMLO<-sqrt(mseMLO)

```

```

PRBmLO=100*((1/length(c))*sum(c[,2]-trueV[2]))/trueV[2]

MLO<-(rbind(trueV,MeanMLO,VarMLO, biasMLO,mseMLO,rmseMLO ))

####

BcovO<-solve(t(X)%*%solve(Vcov)%*%X)

sigma2MLO<- (sum((y-X%*%timeMLO)^2))/(Nt-1)

e<-y-X%*%timeMLO

loglikeMLO <- ((-Nt/2)*log(2*pi)) - ((Nt/2)*log(sigma2MLO)) - ((t(e)%*%e)/(2*sigma2MLO))

llkMLO<-sum(loglikeMLO)

#####REMLO

##REMLOERDINARY

d=array(0,dim=c(nsim,2))

for (i in 1:nsim) {

  Y=y+rnorm(y)

  timeModelREO <- solve(t(X)%*%solve(VcovReml)%*%X)%*%t(X)%*%solve(VcovReml)%*%Y

  d[i,] = timeModelREO

}

#####RMLEO Results#####

MeanRMLO=apply(d,2,mean)

VarRMLO=apply(d,2,var)

timeRMLO <- solve(t(X)%*%solve(VcovReml)%*%X)%*%t(X)%*%solve(VcovReml)%*%Y

trueV=c(1,0.3)

biasRMLO=MeanRMLO-(trueV)

mseRMLO=biasRMLO^2+VarRMLO

rmseRMLO<-sqrt(mseRMLO)

```

#####

```
RMLO<-(rbind(trueV,MeanRMLO,VarRMLO, biasRMLO,mseRMLO,rmseRMLO ))
```

####

```
BcovRO<-solve(t(X)%*%solve(VcovReml)%*%X)
```

```
sigma2RMLO<- (sum((y-X%*%timeRMLO)^2))/(Nt-1)
```

```
e<-y-X%*%timeRMLO
```

```
loglikeRMLO <- ((-Nt/2)*log(2*pi)) - ((Nt/2)*log(sigma2RMLO)) - ((t(e)%*%e)/(2*sigma2RMLO))
```

```
llkRMLO<-sum(loglikeRMLO)+((Nt+1)/2)*log(2*pi)-.5*(log(det(t(X)%*%solve(VcovReml)%*%X)))
```

#####April 2017 extension

```
PRBrmlO=100*((1/length(d))*sum(d[,2]-trueV[2]))/trueV[2]
```

```
elag= Lag(AR1,k=1)
```

```
elag[1,]<-0
```

```
RHO = sum(AR1*elag[,1])/sum(elag)
```

#####Chi-Squares for RMLO

```
l=diag(1,2)
```

```
D=zelta
```

```
bhat<- D%*%solve(l+D*t(Z)%*%Z)%*%t(Z)%*(y-X%*%timeRMLO)
```

```
kai2=(sum(bhat[2]-d[,2])^2)/sigma2RMLO
```

#####

#####

#####RESULTS

```
result1<-cbind(ML,REML,MLO,RMLO)
```

```

colnames(result1)<-
c("interceptML","slopeML","interceptREML","slopeREML","interceptMLO","slopeMLO","interceptRMLO",
"slopeRMLO")

result2= cbind(Bcov,BcovRM,BcovO,BcovRO)

result3<- c(sigma2ML,sigma2REML,sigma2MLO,sigma2RMLO)

result4=cbind(lIkML[1],lIkREML[1],lIkMLO[1],lIkRMLO[1])

#####

#####NEW COVARIANCE

####ZELTA

MLcov=Z%*%Bcov%*%t(Z)+sigma2ML*arcov

REMLcov=Z%*%BcovRM%*%t(Z)+sigma2REML*arcov

MLOcov<- Z%*%BcovO%*%t(Z)+sigma2MLO*arcovi

MLROcov<- Z%*%BcovRO%*%t(Z)+sigma2RMLO*arcovi

result5<-cbind(MLcov,REMLcov, MLOcov,MLROcov)

result6<- cbind(PRBml,PRBremI,PRBrmlO,PRBmlO,RHO)


result1

result2

result3

result4

result5

result6

#####3

#####Chi Square Test

D=result2[,1:2]

l=ncol(W)

```

```

#####ML

bhatML=array(0,dim=c(nsim,2))

for (j in 1:nsim){

    bhatML[j,]<- D%*%solve(I-t(W)%*%W%*%D)%*%t(W)%*%(Y-X%*%a[j,])

}

MLchi<- (sum((bhatML-trueV)^2))/(sigma2ML/t)

#####REML#####

bhatREML=array(0,dim=c(nsim,2))

for (j in 1:nsim){

    bhatREML[j,]<- D%*%solve(I-t(W)%*%W%*%D)%*%t(W)%*%(Y-X%*%b[j,])

}

REMLchi<- (sum((bhatREML-trueV)^2))/(sigma2REML/t)

####MLO

bhatMLO=array(0,dim=c(nsim,2))

for (j in 1:nsim){

    bhatMLO[j,]<- D%*%solve(I-t(W)%*%W%*%D)%*%t(W)%*%(Y-X%*%c[j,])

}

MLOchi<- (sum((bhatMLO-trueV)^2))/(sigma2MLO/t)

####RMLO

bhatRMLO=array(0,dim=c(nsim,2))

for (j in 1:nsim){

    bhatRMLO[j,]<- D%*%solve(I-t(W)%*%W%*%D)%*%t(W)%*%(Y-X%*%d[j,])

}

RMLOchi<- (sum((bhatRMLO-trueV)^2))/(sigma2RMLO/t)

result7=c(MLchi,REMLchi,MLOchi,RMLOchi)

```

